

ABCG2: The Lateral Slice Hypothesis as a Model for Multidrug Transport

Megan Hannah Cox, BSc Hons.

Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy

September 2018

Abstract

ABCG2 (also known as BCRP) is an efflux ABC transporter recognised to export a wide range of substances. ABCG2 is widely distributed in humans, predominantly residing within tissues which have protective roles, such as the intestine and blood brain barrier. ABCG2 is known for its role in the clinical phenomenon of multidrug resistance, particularly in limiting the exposure to chemotherapeutic agents.

Despite this significant role in tissue exposure to xenobiotics, exactly how ABCG2 binds and transports a wide range of structurally unrelated drug substrates remains unknown.

This project set out to determine the molecular basis of drug polyspecificity and transport in this pharmacologically important membrane protein. Previous research had identified a pair of residues (R482, P485) which influenced substrate selectivity. Based on the existing data, a 'lateral slice' hypothesis was proposed, identifying 10 residues in the transmembrane domain that were predicted through topological mapping, to be in a similar position in the plane of the membrane to R482 and P485 and would subsequently form part of a putative drug binding site. All 'lateral slice' residues investigated were mutated to alanine, as it is torsionally neutral and retains the propensity to form α -helices, so is unlikely to cause a structural perturbation. All 'lateral slice' ABCG2 mutants were able to express full length ABCG2 protein, that predominantly localised correctly to the plasma membrane. Functional investigations demonstrated the involvement of several of the 'lateral slice' residues in drug selectivity and efflux. For example, ABCG2_{F640A} showed just over a 2-fold increase in mitoxantrone efflux compared to the wild

type protein and a gain in the ability to efflux a non-native drug substrate; daunorubicin (an approximate increase of 7-fold compared to wild type protein). The interpretation of the 'lateral slice' experiment data was facilitated by the publication of structural data regarding the ABCG subfamily of ABC transporters. Structural mapping revealed that several of the functionally significant mutated 'lateral slice' residues were located to one patch on the surface of the protein. This observation led to a secondary working hypothesis, that the surface patch/pocket could be a site for drug recognition, so further residues lining this region were investigated in much the same way to the 'lateral slice' residues. All but one of the new 'binding pocket' residues had an implication in drug efflux, with one mutation, M548A, interestingly demonstrating an enhanced efflux for mitoxantrone (1.4-fold), an impaired efflux for pheophorbide A (2.4-fold), and a gain in efflux for daunorubicin (4.8-fold) compared to wild type protein, showing drug specificity for the mutated residue. The final 'binding pocket' residue, L633, demonstrated disrupted protein biogenesis, suggesting a role in protein stabilisation and/or folding.

All the experimental data from these hypotheses were interpreted in terms of a multisite drug binding model for ABCG2 and supported the proposal of a novel transport hypothesis, in which, residues at a cytoplasmic lipid-exposed surface site act as a recognition/binding site for drugs, and residues within a secondary site form part of a translocation pathway to the extracellular compartment. This model makes a number of testable predictions for future research.

Acknowledgements

I would like to express my deepest gratitude to Dr Ian Kerr, for the untold hours of support he has given me throughout the last 5 years, and for giving me the opportunities to develop into the scientist I am today.

I would also like to say a huge thank you to Debs (Deborah Briggs). I first met Deb as a bright-eyed undergraduate student doing my summer placement, I was taken under her wing to learn all the techniques required. As time has gone on our working relationship has blossomed into a true friendship, which I hope will last a lifetime.

To my fellow lab members Parth Kapoor (AKA Parsnip) and Aaron Horsey (AKA Horsey, Fairy or Aary Fairy) you have been a huge part of my life over the last four years and I value your friendships highly. I won't forget any of our "walk and talks", either it be round the lake, over a cup of coffee, or the occasional bowling evening.

I also offer my deepest gratitude to the University of Nottingham and Nottingham-Rothamsted BBSRC funded Doctoral Training Partnership. Their financial support and facilities have been instrumental in the success of this project.

I would like to extend my thanks to Dave Scott and Steve Carr for providing valuable support and guidance to me, during my professional placement at the Rutherford-Appleton Laboratory. Also, I would like thank David Onion, Nicola Croxall (Flow Cytometry Facility), and Tim Self (School of Life Sciences Imaging) for giving me their valuable time, advice and training on relevant equipment.

Thank you to my family, especially my step-dad Graeme (aka Father Fisher) for helping shape the person I have become, if he'd like to admit it or not, and Matthew for being there for me every step of the way and putting up with me through difficult phases of thesis writing.

Finally, a message to my mum (AKA Marmite) I could not have done this without you. Throughout my whole education you have gone above and beyond to seek help when I needed it and reassured me that I can do it when I doubt myself. You have taught me what it is to be a dedicated, hardworking, honest person. I love you so very much and dedicate this thesis to you for all love and support you have given me over the years, your little sweet pea.

Declaration

I declare that the thesis entitled “ABCG2: The Lateral Slice Hypothesis as a model for Multidrug Transport” is the result of my own work which has been undertaken during my period of registration for this degree at The University of Nottingham, under the guidance of Dr Ian Kerr.

Megan Hannah Cox
Student I.D. 4230851
September 2018

Notice to examiners



University of
Nottingham
UK | CHINA | MALAYSIA



The University of Nottingham BBSRC Doctoral Training Partnership (DTP) is undertaken in partnership with Rothamsted Research. Its aim is to develop an innovative, integrated, student-led programme of individualised research training in the Biosciences through an effective cohort-based programme. Students are not recruited to an individual project but, through a series of lab rotations, identify a PhD programme that they wish to pursue in one of three BBSRC priority research areas (Food Security; Industrial Biotechnology and Bioenergy; World Class Biosciences)

Research Training Structure

The research training structure is framed by individual development needs analysis and has the following core components:

- Generic and research skills development
- High-level modular training
- Laboratory rotations
- Experiential learning through public engagement and outreach activity

Generic and research skills development comprises the following: induction, introductory training programme, annual 'Spring School' conference-style training event, participation in Sysmic (online systems and mathematical biology course), specialist training days at partner organisations, twilight (after-lab) training which is arranged on an *ad hoc* basis to meet identified needs, and seminar series.

In Year One there is an initial period of intensive training which takes place over the course of the first six months. This comprises a two-week introductory training programme, 24 weeks of laboratory rotations, followed by a week-long 'Spring School' (conference-style training event for all DTP cohorts) and ending in a science outreach event. Each lab rotation is assessed focussing on a different form of scientific communication (Report, Poster, Presentation) and students receive feedback from supervisors. The PhD project begins in Month 7. Students are also encouraged to choose from a portfolio of relevant Master Courses which they can attend alongside their lab rotation projects.

In Years Two to Four training is mainly concentrated in the week long annual Spring School, although additional training and seminars are delivered where required throughout the year. A three month Professional Internship for PhD Students (PIPS) placement is usually taken in Year Three and, as BBSRC specifies, this is not in an area related to the students' PhD research programme.

The integrated programme of training is mapped to BBSRC priority areas, BBSRC Skills Statement: Scientific and Core Skills for UK Bioscience and the Researcher Development Framework (RDF) to ensure breadth and depth of provision, and to encompass vulnerable skills and strategic competencies. It builds on the best practice principles outlined in the Statement of Expectations for Doctoral Training, RCUK (2013) and the 'Good Practice Elements in Doctoral Training', LERU (2014).

It offers a wide range of skills development opportunities to prepare students for a successful bioscience-related career and supports the exchange of ideas and experiences within and across cohorts. It also generates a strong community of postgraduates and supervisors to stimulate the formation of cross-disciplinary links through cohort development, which prepares students for a career in an increasingly multidisciplinary environment.

Abbreviations

ABC	ATP-binding cassette
ABCG2	ATP-binding cassette transporter subfamily G member 2
ABCG5/G8	Heterodimeric ABCG5/ ABCG8 protein
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BCRP	Breast cancer resistance protein
BODIPY	Boron-dipyrromethene
bp	Base pair
BSA	Bovine serum albumin
CMV	Cytomegalovirus
dH₂O	Distilled water
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulphoxide
DMPC	1,2-Dimethyl-sn-glycero-3-phosphocholine
DNA	Deoxyribose nucleic acid
DNR	Daunorubicin
dNTPs	Deoxynucleotides
DTT	Dithiothreitol
EC₅₀	Concentration at which a drug evokes half maximal response
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(2-aminoethylether)- <i>N,N,N',N'</i> -tetraacetic acid
EM	Electron microscopy
ER	Endoplasmic reticulum
FSC	Forward scatter
FCS	Foetal calf serum
FCS	Fluorescence correlation spectroscopy
GFP	Green fluorescent protein
GlcNAc	N-Acetylglucosamine
HA	Haemagglutinin
HBSS	Hank's balanced salt solution
HEK293T	Human embryonic kidney 293T cell line
IOV	Inside out vesicle
kDa	Kilo Daltons
LB	Luria-Bertoni
LY	Lucifer yellow
MDR	Multidrug resistance
MRP	Multidrug resistance protein
MST	Microscale thermophoresis

MX	Mitoxantrone
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline-0.01% (v/v) Tween-20
PCH	Photon-counting histogram
PCR	Polymerase chain reaction
PEI	Polyethyleneimine
PhA	Pheophorbide A
Pi	Orthophosphate
PNGase	F Peptide: N-glycosidase F
PRA	Prazosin
ROV	Right-side out vesicle
SBE	Steroid binding element
SBP	Solute binding protein
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
Sf9	Spodoptera frugiperda
sfGFP	Super-folder green fluorescent protein
SLC	Solute carrier
SSC	Side scatter
SV40	Simian vacuolating virus 40
T_m	Melting temperature
TMD	Transmembrane domain
TMH	Transmembrane helix
TNF	Tumour necrosis factor
UV	Ultraviolet
VTA	Vesicular transport assay
WT	Wild type
3D	3-dimensional

Contents

Chapter 1 Introduction	1
1.1 Overview of transporters	1
1.1.1 Solute carrier (SLC) proteins	1
1.1.2 ATP-binding cassette (ABC) transport proteins	3
1.1.3 Four core ABC transporter domains	4
1.1.4 'ATP-switch' and 'constant contact' models of ABC exporters	17
1.1.5 ABC transporters in human disease.....	20
1.1.6 Multidrug resistance ABC transporters.....	24
1.2 ABCG2.....	26
1.2.1 Discovery of a third human MDR pump	26
1.2.2 The genetic locus and the genetic regulation of expression	27
1.2.3 The overall topology of the expressed protein.....	29
1.2.4 Tissue distribution and physiological roles of ABCG2.....	29
1.2.5 Polymorphisms of ABCG2	32
1.2.6 Site-directed mutagenesis	34
1.2.7 Oligomerisation of ABCG2	37
1.2.8 Pharmacology of ABCG2	38
1.2.9 Structure of ABCG2	44
1.3 Project aims.....	50
Chapter 2 Methods.....	52
2.1 Molecular Biology	52
2.1.1 Materials and reagents	52
2.1.2 Site-directed mutagenesis	52
2.1.3 Transformation	54

2.1.4 Plasmid DNA purification	54
2.1.5 Restriction Digest	55
2.1.6 DNA sequencing	56
2.2 Tissue culture	57
2.2.1 Materials and reagents	57
2.2.2 Cell passage maintenance.....	57
2.2.3 Transfection	58
2.2.4 Zeocin selection of stable cell lines.....	58
2.3 SDS-PAGE gel electrophoresis and western blotting.....	59
2.3.1 Cell harvesting.....	59
2.3.2 Protein concentration determination assay	59
2.3.3 SDS-polyacrylamide gel electrophoresis.....	59
2.3.4 N-linked glycanase digestion.....	60
2.3.5 Western blotting	60
2.4 Confocal microscopy	61
2.4.1 Materials and reagents	61
2.4.2 Poly-L-lysine coating.....	61
2.4.3 Cell seeding	62
2.4.4 Confocal Image Acquisition.....	62
2.5 Flow cytometry	63
2.5.1 Materials and reagents	63
2.5.2 Fluorescence based drug accumulation assay.....	63
2.5.3 Flow cytometry acquisition.....	64
2.5.4 Flow cytometry data analysis.....	65
2.6 Molecular visualisation and drug docking	69
2.7 Membrane based assays.....	70

2.7.1 Sodium butyrate analysis	70
2.7.2 Tissue culture	72
2.7.3 Membrane preparation	72
2.7.4 Membrane and protein stability analysis	73
2.7.5 Vesicular transport assay	73
2.7.6 ATPase assay	74
Chapter 3 Generation, expression and functional characterisation of ABCG2 'lateral slice' mutants	76
3.1 Development of the 'lateral slice' hypothesis	76
3.2 Construct information	81
3.3 Construct generation	82
3.4 Protein expression	87
3.5 Localisation.....	89
3.6 Functional characterisation	95
3.7 Flow cytometric analysis of ABCG2 mutant expression	96
3.8 Flow cytometric functional analysis of ABCG2 mutants.....	105
3.9 Summary	114
Chapter 4 Using the structural revolution to enhance the understanding of ABCG2 drug binding	116
4.1 The structural revolution.....	116
4.2 Structural mapping.....	117
4.2 Construct generation	120
4.3 Construct expression and localisation	121
4.4 Functional characterisation	125
4.5 Summary	130
Chapter 5 Enhancing the understanding of ABCG2-mediated allocrite recognition and translocation.....	132

5.1 ‘Gain in function’ mutation still requires the energy from ATP hydrolysis to drive the efflux of drug substrates.....	132
5.2 Sodium butyrate increases ABCG2 expression of <i>in vitro</i> investigation.....	137
5.3 Preservation of ABCG2 function in isolated membranes	140
5.4 Function of ABCG2 in isolated membrane validated by ATPase assays	145
5.5 Structural interpretation of ABCG2 mutants	153
5.6 Summary	159
Chapter 6 Discussion.....	161
6.1 Leucine to alanine mutation at position 633 may cause destabilisation of the re-entry helix.....	163
6.2 Methionine: aromatic interactions may be central to drug recognition and/or protein stabilisation	165
6.3 Hypothesis of an ABCG2 mediated allocrite transport cycle/mechanism.....	168
6.3.1 Current investigations and context.....	169
6.3.2 Recent structural advances.....	169
6.3.3 Structural evidence of multiple allocrite binding sites	171
6.3.4 Cox proposed drug binding sites.....	173
6.3.5 Cox hypothesis of the ABCG2 multi-drug transport mechanism.....	174
6.4 Directions for further research	177
6.4.1 Mechanism of ABCG2 regulation by cholesterol	177
6.4.2 Role of re-entry helix in structural stability and or transport.....	178
6.4.3 Role of methionine: aromatic interactions in allocrite efflux.....	178
6.4.4 Investigations into residues forming the translocation pathway/ the transport hypothesis	179
6.5 Summary	180
References	181
Appendices.....	203

List of figures

Figure 1.1 ABC transporter subfamilies.	4
Figure 1.2 The four core domains of an ABC transporter.....	7
Figure 1.3 ATP 'sandwich' dimer interface.	13
Figure 1.4 Simplified Schematic of the ATP-binding switch coupled to the alternate access model for ABC exporter function.	19
Figure 1.5 ABCG5/G8 structure resembles type II transporter fold.....	48
Figure 1.6 ABCG2 topology.	49
Figure 2.1 Flow cytometric data analysis of an ABCG2 variant	68
Figure 2.2 ABCG2 homology model with grid box assigned.....	70
Figure 3.1 The lateral slice hypothesis mutations.	78
Figure 3.2 An overview of systems for the expression of recombinant proteins.....	80
Figure 3.3 Plasmid map of pcDNA™-3.1/zeo(+)_sfGFP_ABCG2.....	82
Figure 3.4 Schematic representation of the generation of mutant constructs.....	83
Figure 3.5 A. PvuII restriction endonuclease digest.....	85
Figure 3.6 DNA mutagenesis confirmed by Sanger sequencing.....	86
Figure 3.7 Immunoblot analysis of ABCG2 mutant variants.....	88
Figure 3.8 Immunoblot analysis of ABCG2 mutant expression and glycosylation. ...	89
Figure 3.9 Confocal analysis of ABCG2 mutants.	94
Figure 3.10 Schematic representation of flow cytometry.....	96
Figure 3.11 Flow cytometric gating parameters.....	97
Figure 3.12 Relative expression of ABCG2 mutant variants stably expressed in HEK293T cell.....	98
Figure 3.13 Protein expression phenotype across total population.....	99
Figure 3.14 Implication of expression upon functional data interpretation.	101
Figure 3.15 Expression normalised drug efflux of WT and F640A variants.....	104

Figure 3.16 Inhibition of mitoxantrone (MX) efflux in the presence of various concentrations of Ko143.	106
Figure 3.17 ABCG2 mutants stably expressed in HEK293T cells demonstrate perturbed drug efflux.	109
Figure 3.18 Function of ABCG2 wild type and mutant variants stably expressed in HEK293T cells.	112
Figure 4.1 Structural mapping of lateral slice residues and proposal of a drug binding pocket.	119
Figure 4.2 Confirmation of ABCG2 “binding pocket” mutant construct generation.	120
Figure 4.3 Expression of ABCG2 _{WT} and mutant variants stably expressed in HEK293T cells.	122
Figure 4.4 Cell surface localisation of ABCG2 ‘binding pocket’ variants stably expressed in HEK293T cells.	124
Figure 4.5 Function of ABCG2 wild type and ‘binding pocket’ mutant variants stably expressed in HEK293T cells.	128
Figure 5.1 Generation and expression of E211Q F640A double mutant.	134
Figure 5.2 E211Q/F640A double mutant still requires the energy from ATP hydrolysis to drive efflux.	136
Figure 5.3 Sodium butyrate increases expression of ABCG2 in stably transfected HEK293T cells.	138
Figure 5.4 Sodium butyrate increases expression of ABCG2 in stably transfected HEK293T cells.	139
Figure 5.5 Vesicular transport assay overview.	141
Figure 5.6 Optimisation of buffering capacity and stabilisation of ABCG2 in isolated membranes.	143
Figure 5.7 ABCG2 function is conserved in isolated membrane vesicles.	144
Figure 5.8 Functional assessment of sfGFP-ABCG2 expressed in HEK293T membranes.	146
Figure 5.9 BODIPY-Prazosin is an allocrite of ABCG2.	147
Figure 5.10 Functional assessment of sfGFP-ABCG2 expressed in HEK293T membranes.	148

Figure 5.11 ABCG2 demonstrates an ATP concentration dependent increase in Pi release.	149
Figure 5.12 Dose response of ABCG2 _{WT} membranes with increasing concentrations of prazosin.	150
Figure 5.13 F640A demonstrates reduced ATPase activity.	152
Figure 5.14 ABCG2 homology model with grid box assigned.	154
Figure 5.15 The internal binding site exclusively acts as a docking site for mitoxantrone.	155
Figure 5.16 Multidrug binding site at the lipid interface of ABCG2.	157
Figure 5.17 Daunorubicin and pheophorbide A dock to a lipid exposed surface site.	158
Figure 6.1 Leucine at position 633 may form inter-helical interactions with isoleucine at position 573 to stabilise the re-entry helix.	164
Figure 6.2 Methionine: phenylalanine interaction between TMH4 and TMH6 is conserved.	167
Figure 6.3 Methionine: aromatic interactions may be crucial for either ligand: and/or protein: protein interactions/ stabilisation.	168
Figure 6.4 <i>In silico</i> docking of allocrites in an ABCG2 homology model reveals multiple potential allocrite interaction sites.	171
Figure 6.5 CryoEM structure of ABCG2 with molecules bound to the central cavity.	172
Figure 6.6 Schematic model of ABCG2 mediated allocrite efflux.	177

List of tables

Table 1.1 ABCG2 drug substrates and inhibitors.	41
Table 2.1 Mutagenic primers used in the generation of ABCG2 constructs.	53
Table 2.2 Polymerase chain reaction (PCR) cycling parameters for site-directed mutagenesis.	53
Table 2.3 Sequencing primers with a schematic representation of their binding location.....	56
Table 2.4 Drug fluorescent properties.	65
Table 3.1 Function of ABCG2 wild type and mutant variants stably expressed in HEK293T cells.	113
Table 4.1 Table of distances (below 5Å) between the residues in the lipid exposed pocket.....	118
Table 4.2 Function of ABCG2 wild type and 'binding pocket' mutant variants stably expressed in HEK293T cells.	129
Table 6.1 Cox proposed drug binding residues that overlap with other ABCG2 structural publications.	174

Chapter 1 Introduction

1.1 Overview of transporters

There are two main classes of proteins that govern vectorial transport of molecules across membrane lipid bilayers, these are solute carrier (SLC) and ATP-binding cassette (ABC) transporters. SLC and ABC transporters share a wide tissue distribution throughout the human body and are essential to cellular physiology. They are responsible for the transport of a diverse range of molecules, including endogenous metabolites, antioxidants, hormones, signalling molecules, nutrients, and neurotransmitters, as well as exogenous molecules including pharmacological agents. The endogenous function of many drug transporters is not well understood. However, the coordinated expression and activity of these transport proteins, especially at the intestinal, hepatic, and renal epithelial membranes, have a notable effect on the ADMETox (absorption, distribution, metabolism, elimination, and associated toxicity) of clinical agents and therefore are important to understand further (Glavinas et al., 2004; Nigam, 2014).

1.1.1 Solute carrier (SLC) proteins

The solute carrier (SLC) protein family consists of 384 unique sequences assigned to 48 subfamilies. Some SLC subfamilies demonstrate phylogenetic clustering, suggesting a common evolutionary origin, but the majority don't share any structural/sequence identities and therefore SLCs are predominantly classified based on their function (Fredriksson et al., 2008). SLCs facilitate the cellular uptake of molecules, either by facilitated diffusion down an electro-chemical gradient

acting as a uniporter, or secondary active transport coupled to the symport or antiport of ions as the driving force (Russel, 2010).

Of particular interest are SLC subfamilies: SLC15, SLC22, and SLC0 which are considered to have major implications in drug uptake, due to their tissue distribution across the intestine, liver, and kidneys, the primary organs/tissues for absorption and excretion of pharmacological agents. Coordinated expression and activities of subfamilies of both SLC and ABC transporters across the basolateral and apical side of the tissue epithelia contribute to the major limitations in the absorption, distribution, and elimination of pharmacological agents (Russel, 2010).

An example of this synergistic relationship is a member of the SLC0 subfamily that consists of organic anion transporting polypeptides (OATPs). OATPs are involved in the bidirectional transport of large and relatively hydrophobic organic anions in a pH-dependent manner, stimulated by an acidic extracellular environment (Russel, 2010). OATP1B1, OATP1B3, and OATP2B1 are localised to the sinusoidal membrane of the liver and transport amphipathic organic compounds into the cell via an organic ion (bicarbonate, HCO_3^-) antiport mechanism. These can include pharmacologically relevant agents such as; statins (rosuvastatin), anti-diabetic agents (glibenclamide), anti-rheumatic agents (methotrexate), antibiotics (benzylpenicillin), and anti-neoplastic agents (SN-38; active metabolite of irinotecan). These are then extruded into the blood or bile as either metabolites or parental compounds via ABC transporters; ABCC4, ABCC2, ABCB1, or ABCG2 (Russel, 2010).

1.1.2 ATP-binding cassette (ABC) transport proteins

Since their discovery in the 1980s, the ATP-binding cassette (ABC) proteins have been established as one of the largest, most diverse superfamilies of membrane transport proteins, present in both prokaryotes and eukaryotes (Higgins, 1992; Higgins et al., 1986). The superfamily utilises the energy derived from ATP hydrolysis to drive the movement of molecules across the membrane, which can be against a concentration gradient. These can be classified into type I and II ABC importers, present almost exclusively in prokaryotes, and ABC exporters, present in both prokaryotes and eukaryotes (Igarashi et al., 2004). In bacteria ABC transporters have been shown to encode 5% of the total genome, which hints towards their physiological importance (Higgins, 1992).

ABC transporter proteins are also prevalent in humans, with 49 proteins discovered to date. These can be classified into seven principal subfamilies: ABCA-G. Almost all subfamilies function as transporters. However, ABCE/F do not reside within the plasma membrane or transport ligands but rather have implications in translational control (Boël et al., 2014; Kerr et al., 2011; Su et al., 2018). The remaining 5 subfamilies (summarised in Figure 1.1 below) are involved in the cellular transport of a diverse array of substrates. These include endogenous metabolites, amino acids, peptides, sugars, nucleotides, ions, sterols, bile salts, lipids, toxins, antibiotics, and other clinically relevant agents (Sharom, 2008).

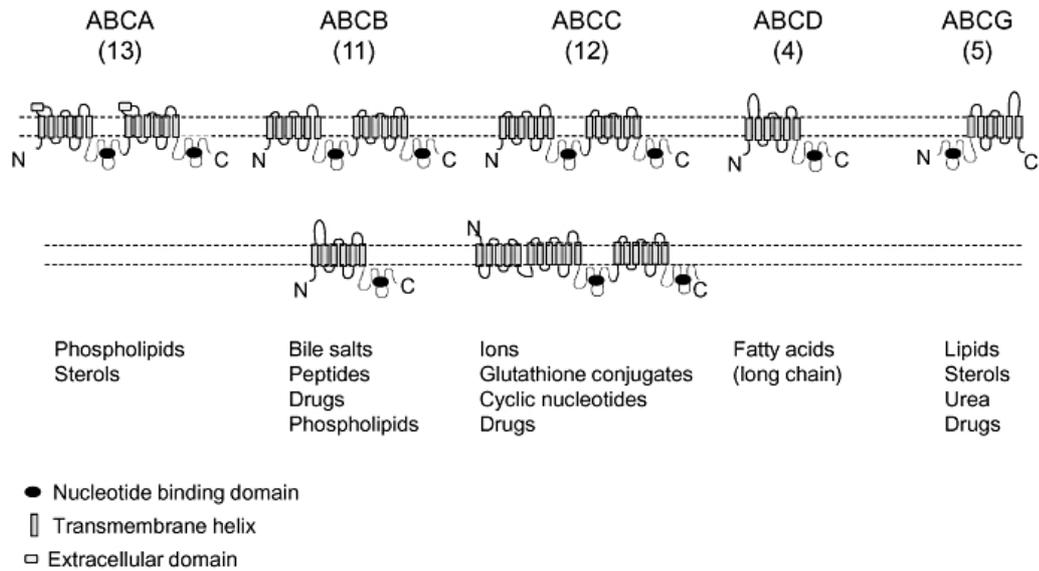


Figure 1.1 ABC transporter subfamilies. Figure depicts general protein topology and transport substrates. Figure taken from (Kerr et al., 2011). Number in brackets denotes the number of proteins within a sub-family.

1.1.3 Four core ABC transporter domains

The minimal functional configuration of an ABC transporter consists of four core domains. Characteristically; two cytoplasmic nucleotide binding domains (NBD₁ and NBD₂) and at least two hydrophobic transmembrane domains (TMD₁ and TMD₂), each one typically comprising between 5 and 8 TM spanning regions (alpha helices) (Higgins, 1992; Linton, 2007; Sharom, 2008). Human ABC transporters can be classified based on the structural arrangement of the NBD(s) and TMD(s) with the primary sequence as follows; ‘full transporters’, usually comprising of a single polypeptide chain with, two TMDs and two NBDs (in the order TMD₁-NBD₁-TMD₂-NBD₂), ‘half transporters’ which contain a single TMD and a single NBD in the polypeptide (TMD-NBD), and ‘reverse topology half transporters’ (NBD-TMD) (Higgins, 1992; Xu et al., 2004). The half-transporters as their name suggests only contain half the required domains, therefore to form a functioning transporter they need to associate with an identical (homo-) or closely related subfamily member

(hetero-), forming dimeric or higher ordered oligomeric structures (Wong et al. 2014). The individual domains of an ABC transporter can also be expressed as single polypeptides, occurring often in prokaryotic ABC transporters (e.g. histidine permease from *S. typhimurium*) (Higgins, 1992; Hung et al., 1998). The transmembrane spanning domains (TMD) of an ABC transporter protein form the structural basis of the translocation pathway (Dawson and Locher, 2006). The nucleotide binding domains couple the energy provided by the binding and hydrolysis of Mg-ATP and then release of ADP-Pi, to several possible outcomes: the direct movement of substrate ligand against its electrochemical gradient, referred to as primary active transport (e.g. ABCB1); the opening or closing of a specific membrane channel (ion-channel, e.g. ABCC7: cystic fibrosis transmembrane conductance regulator, CFTR), or the regulation of the permeability of protein channel complexes (receptors, e.g. ABCC8: SUR1 (sulfonylurea receptor)) (Fukuda et al., 2011; Kim and Chen, 2018; Sheppard and Welsh, 1999; Zolnerciks et al., 2007).

Although the four core domains are the minimum requirement for function, many ABC transporters have additional domains that serve regulatory or other peripheral roles (Higgins, 1992). For example, type I and II prokaryotic importers require a specific solute-binding protein (SBP), which acts to bind the transport substrate (allocrite) and deliver it to the transporter complex (Rice et al., 2014). The type I and type II importers vary in the number of SBP present, type I importers have an excess of SBPs present to efficiently capture any transport molecule, whereas the SBPs of type II importers seem to be stoichiometric with the transport protein (Berntsson et al., 2010). The human cystic fibrosis conductance regulator (CFTR,

ABCC7) also possesses an extra domain, this acts as a regulatory ('R') domain that separates the two halves of the protein. Phosphorylation of the 'R' domain serves to regulate the Cl⁻ channel activity of the CFTR (Cheng et al., 1991; Zhang and Chen, 2016). Other ABC transporters of the mammalian C subfamily have extra N-terminal membrane bound region/domain, termed TMD₀, which is often smaller than the core TMDs and connected to the core TMDs by a cytoplasmic linker (L₀). The additional TMD in ABCC8/9; the sulfonylurea receptor (SUR), has demonstrated modulatory activity of the receptor (Babenko and Bryan, 2003). Whereas, mutational studies demonstrated that removal of the TMD₀ of ABCC1 (multidrug resistance protein 1, MRP1) had little effect on function, but indicated a fundamental functional importance for the linker region between TMD₀ and TMD₁ (Bakos et al., 1999).

Nucleotide binding domains (NBDs)

The ability of ABC transporters to mediate vectorial transport of allocrites (transport substrates) requires an intricate coupling of signals and movements between the TMDs and NBDs (Kerr, 2002). The TMDs of ABC transporters are highly variable in terms of structure and composition, with limited sequence identity, suggestive of the diversity in transport substrates. Conversely, the NBDs share sequence identity across all ABC transporters irrespective of function or species origin (Holland and Blight, 1999). It was therefore hypothesised that all ABC transporter NBDs possessed common inter-domain communications (Jones and George, 1999).

The NBD sequence is most highly conserved in seven distinct regions (outlined in Figure 1.2), termed motifs. Two of the motifs, namely the Walker-A and -B motifs are common to most ATP-binding proteins (Kerr, 2002; Walker et al., 1982). The remaining five motifs; the signature motif, A-, H-, Q-, and D-loop, are thought to be exclusive to the ABC transporter NBDs (Kerr, 2002; Lawson et al., 2008; Linton and Higgins, 1998; Smith et al., 2002).

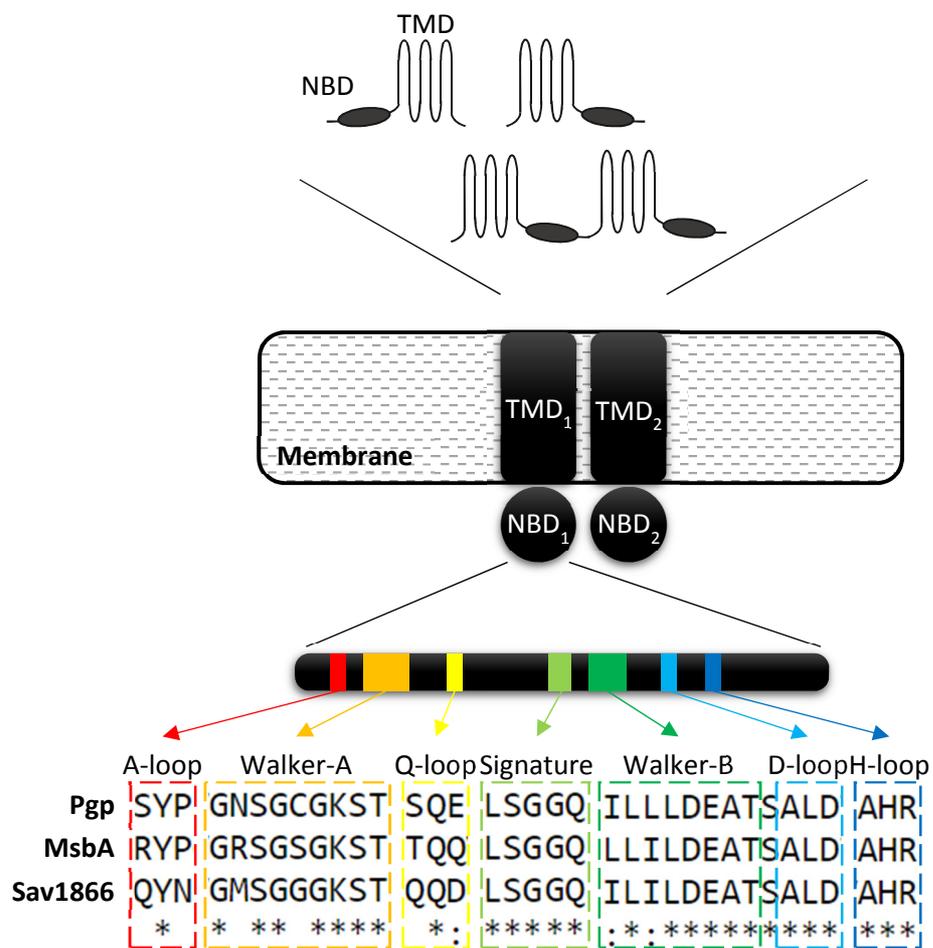


Figure 1.2 The four core domains of an ABC transporter. Schematic depicts the four core domains required for a functional transporter with examples of different topological arrangement (e.g. reverse topology (N-terminal NBD, C-terminal TMD) half transporter (NBD₁-TMD₁). The second part of the figure shows a sequence alignment of conserved NBD motifs across human Pgp (P08183), MsbA (*Pseudomonas aeruginosa*, Q9HUG8), and Sav1866 (*Staphylococcus aureus*, Q99T13).

The role of the NBD Motifs

The role of key residues within the conserved NBD motifs have been identified through mutagenesis studies and elucidated further with structural studies:

The Walker-A (GxxGxGK[S/T], where x is any residue) and Walker-B motifs (hhhhDE, where h is any hydrophobic residue) form multiple interactions with associated nucleotide (ATP) and/or cofactor (Mg^{2+}). Mutations to the glycine, lysine, and serine of the Walker-A motif (GxxGx**GK**[S/T]) resulted in the retention of ATP binding, but disruption of nucleotide hydrolysis, suggesting a role for the lysine and serine in metal ion coordination (Urbatsch et al., 2000).

The acidic glutamate residue of the Walker-B motif (hhhh**DE**), often referred to as the 'catalytic carboxylate', is proposed to form a hydrogen bond with an 'attacking' water molecule to coordinate the hydrolytic attack of the γ -phosphate of ATP (Hrycyna et al., 1999; Lawson et al., 2008; Moody et al., 2002; Smith et al., 2002; Tomblin et al., 2004b; Urbatsch et al., 2001). The H-loop (bhbu**H**, where b is a branched aliphatic residue, and u is a hydrophilic residue) located C-terminal to the Walker-B motif is also thought to coordinate the hydrolytic attack by the formation of H-bonds with the γ -phosphate itself (Kerr, 2002; Lawson et al., 2008; Linton and Higgins, 2007, 1998; Smith et al., 2002). The D-loop (S**ALD**) is proposed to aid in the formation of cross-dimer interactions, potentially forming a charge:charge interaction with the H-loop histidine residue (Lawson et al., 2008).

The A-loop containing any aromatic residue, most commonly a tyrosine residue, is located approximately 25 residues upstream of Walker-A motif, the aromatic

(benzene) ring has been shown to form π - π stacking interactions against the adenine ring of ATP. Mutagenesis studies of the aromatic loop in Pgp (human ABCB1 and mouse (Mdr3/ABCB1a) demonstrated the requirement of an aromatic residue for full transporter ATPase activity/function, suggesting a role in orientating the ATP in the correct position for nucleophilic/ hydrolytic attack (Carrier et al., 2007; Hung et al., 1998; Kim et al., 2006; Lawson et al., 2008).

The Q-loop (conserved glutamine residue) is located between the Walker-A motif (of the 'core' domain) and signature motif (of the α -helical subdomain, discussed below) (Diederichs et al., 2000). There is evidence to suggest that the highly conserved glutamine residue residing within the Q-loop is involved in NBD:TMD inter-domain communication, potentially through connections with the coupling helices of the TMD (intracellular loops of TMDs). Mutagenesis studies support this, where by the glutamate residue in mammalian Pgp was substituted to either a glutamine or alanine. The mutants were able to demonstrate a normal binding capability for transport substrate and nucleotide (ATP), but showed reduced ATPase activity, suggesting that the NBD:TMD communication was disrupted (Urbatsch et al., 2000). In a recent structural study of the E556Q/E1201Q human Pgp mutant (i.e. a mutant with both Walker-B glutamates neutralised to glutamine) resolved in the ATP bound state, the Q-loop has been shown to possess multiple roles; it forms part of the interface between the NBD and TMD, with interactions from intracellular coupling helices 2 and 4. The Q-loop also contributes to NBD dimer formation by directly contacting the opposing NBD (Kim and Chen, 2018; Smith et al., 2002). Furthermore, the glutamine residues of the Q-loop are oriented to

coordinate with a Mg^{2+} ion and the γ -phosphate of ATP. Mutations of the Q-loop glutamine may destabilise the ATPase site and NBD dimer formation (Kim and Chen, 2018).

The ABC signature (LSGGQ) motif is proposed to form part of a cross-dimer ATP binding pocket (see below for more details). The serine residue (LSGGQ) lies at the dimer interface with potential interactions from residues residing in the opposing NBD Walker-A motif (Hopfner et al., 2000; Lawson et al., 2008; Linton and Higgins, 1998; Tomblin et al., 2004a).

NBD structure and dimerisation

The first crystal structure of an ABC transporter's nucleotide binding domains was of the isolated NBDs of a histidine permease (HisP) from *Salmonella typhimurium* co-crystallised with ATP published in 1998 by Hung et al.. The 1.5Å structure revealed a somewhat stubby 'L-shaped' topology of the NBD (monomer), which comprised two sub-domains. The α/β subdomain which showed structural homology to the F1-type ATPase and RecA structures, termed the 'core' domain. This core domain contained six of the seven conserved motifs (Walker A/B, A-, Q-, D-, and H-loops). The second subdomain, a primarily α -helical structure, contained the characteristic ABC signature motif (or C-loop) (Hung et al., 1998). The crystallographic unit of HisP contained a dimer of two monomers in a 'back-to-back' association of the two opposing α -helical subdomains (Hung et al., 1998; Kerr, 2002).

Subsequent structural studies of bacterial ABCs confirmed the 'L-shaped' topology of the NBD monomer, but provided contrasting data for the dimeric association of the nucleotide-binding domains (Diederichs et al., 2000; Hopfner et al., 2000; Lamers et al., 2000). The differences observed in the NBD association were converse to the hypothesis that all ABC transporter NBDs possessed common inter-domain communications, so investigations proceeded to evaluate the validity of the structural models. The entropy ratio of residues at the NBD interface was calculated, an entropy ratio can discriminate between artefactual dimers and true protein:protein contacts (Elcock and McCammon, 2001; Kerr, 2002). According to this criterion the dimeric interface represented by the ABC bacterial protein, Rad50, involved in DNA double-strand break repair, represented a true protein: protein interaction (described below) (Kerr, 2002).

The Rad50 NBDs are suggested to dimerise upon ATP binding forming an interlocking of the 'L-shaped' monomers in a 'head-to-tail' arrangement, such that the Walker-A/-B motif of one monomer forms an ATP binding pocket with the signature motif of the opposing monomer, 'sandwiching' the ATP between the two opposing NBDs (Hopfner et al., 2000; Jones and George, 1999; Kerr, 2002).

A later study by Smith et al., 2002 built upon the structural studies of MJ0796, from *Methanococcus jannaschii* by Yuan et al., 2001 and Moody et al., 2002 to determine the mechanism by which the NBDs dimerise and the contacts required for hydrolytic cleavage of the γ -phosphate of ATP.

Mutations of the catalytic base (E of Walker-B motif) of MJ0796, glutamate (E) → glutamine (Q), reportedly produced a cassette that was hydrolysis deficient but able to form stable dimers upon association with ATP, whereas ADP was unable to drive dimerisation (Moody et al., 2002). This showed a symmetrical ATP 'sandwich' similar to that seen for the Rad50 dimer (Figure 1.3).

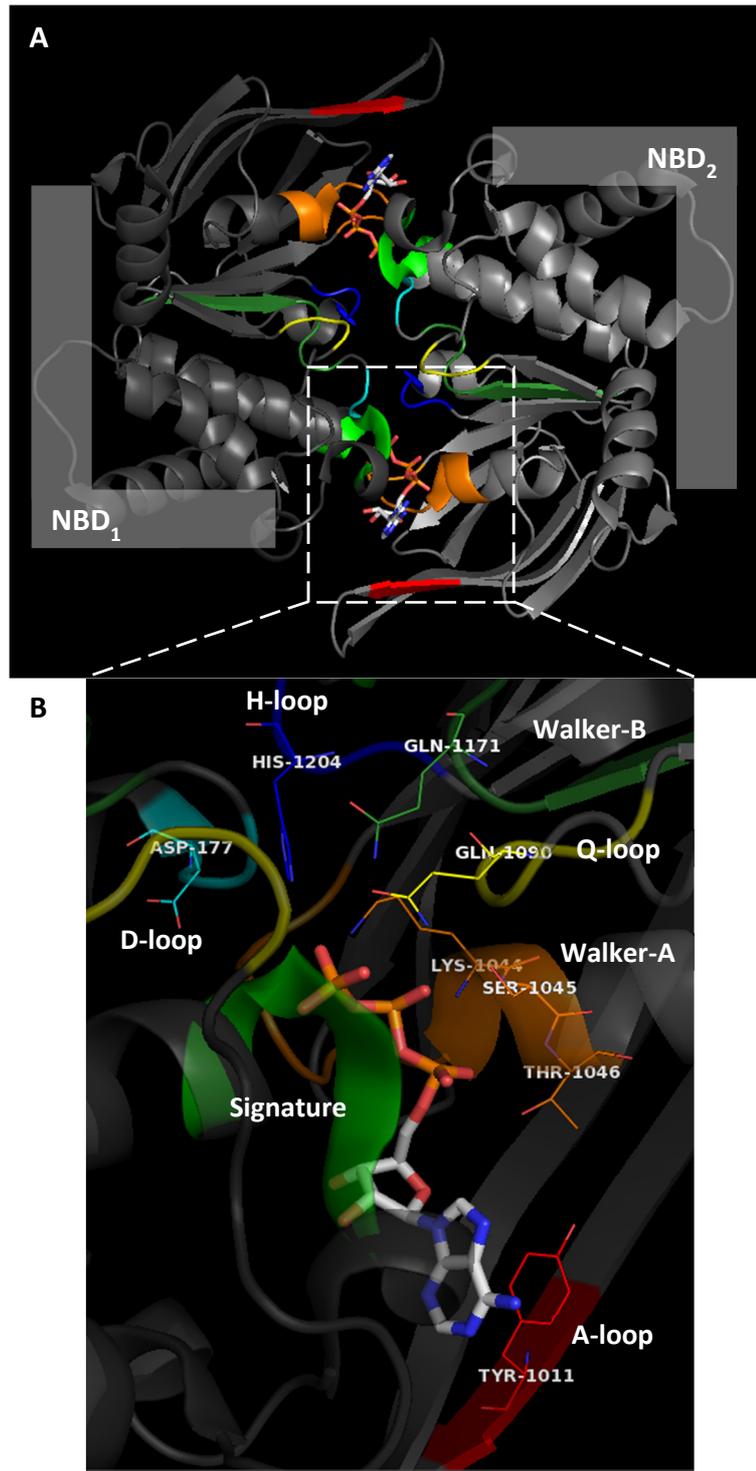


Figure 1.3 ATP 'sandwich' dimer interface. **A.** The NBD dimer interface formed by the E→Q MJ0796 mutant with ATP bound. The structure shows the head-to-tail arrangement of the two L-shaped NBD monomers with the Walker-A/B motifs (orange and dark green, respectively) of one monomer forming an ATP binding site with the signature motif (light green) of the opposing NBD. **B.** closer inspection of the ATP binding sites, highlights residues which are thought to contribute to nucleotide coordination and/or hydrolysis. Figure created with PyMol using PDB file 1l2t.

The E → Q mutant structure shows a water molecule within the NBD active site in an ideal geometry for hydrolytic attack on the γ -phosphate of ATP. The paper describes the formation of potential interactions between the cofactor (Na^+ in the case of the mutant, Mg^{2+} metal ion in wild type protein) and the oxygen on the amide sidechain of the glutamine residue (Q-loop) positions the amide to donate a hydrogen bond to the hydrolytic water molecule (Smith et al., 2002). The side chain amide of the mutated residue is suggested to donate a second hydrogen bond to the hydrolytic water molecule, which in turn donates a bifurcated (split) hydrogen bond to two oxygen atoms of the γ -phosphate of ATP, and a second hydrogen bond to the backbone carbonyl of an alanine in the dimer related subunit (Smith et al., 2002). In the wild type protein, i.e. the presence of the glutamate residue will result in four acceptors competing for two potential donor protons on the water molecule (with H-bond donor only from glutamine in Q-loop).

Hydrolysis of ATP is presumed to occur upon the rotation of the 'attacking' water molecule so that it donates hydrogen bonds to the acceptor carboxylate of glutamate (Walker-B), while donating a lone pair of electrons to the γ -phosphate of ATP in a nucleophilic attack (Smith et al., 2002).

The structural difference between the Mg.ADP-bound and Na.ATP-bound demonstrate a reorientation (rigid-body movement/rotation) of the α -helical subdomain (residues within the signature motif) relative to the 'core' subdomain upon ATP binding, causing predominantly the side chains of glutamine (Walker-B, in the glutamate to glutamine mutant) and histidine (H-loop) residues to undergo

'induced-fit' rotamer changes, creating a contact between the hydrolytic water and γ -phosphate (of ATP), respectively (Hopfner et al., 2000; Smith et al., 2002). This causes the propensity for an ATP to bind in the second binding site promoting a fully closed NBD formation, via electrostatic interactions. It is proposed that this rigid body movement of the α -helical subdomain may potentially mediate the binding/ translocation of transport substrate in the transmembrane domains (Smith et al., 2002).

It is then surmised that the presence of ADP.Pi in the NBD active site disrupts the inter-domain electrostatic interactions (due to an increase in negative charge of 'free' phosphate) promoting the separation of the NBD dimer (Smith et al., 2002).

Whilst Smith et al. provide mechanistic detailing on the potential interactions required for NBD dimerisation and catalysis of ATP, there is limited information to describe how this is coupled to the vectorial transport of allocrites (i.e. via the coupling to the TMDs), i.e what drives allocrite transport, ATP binding or hydrolysis? And can the energy from the hydrolysis of one ATP drive the transport cycle (in an asymmetric fashion) or are two required (in a sequential manner)?

TMD structural topology

The full structures of ABC transporters revealed a divergence in the TMD structure, resulting in their classification into two categories; type I and type II, based on transmembrane helix topology and proposed mechanism of transport. These were originally proposed and described for ABC importers. Generally, Type I importers contain 12 TM helices, including a core of 10TMHs and the addition of an N-

terminal TMH wrapping around the TMD partner (exemplified by the maltose transporter: MalFGK₂ (Chen et al., 2013; Oldham et al., 2007a)) (ter Beek et al., 2014). Type II importers consist of two identical TMDs, each comprising 10 TMHs. The TMDs line up next to each other without cross overs of TMHs between the two TMDs (exemplified by ButCD an *E.coli* vitamin B₁₂ importer (Locher et al., 2002) and HmuUV involved in the bacterial uptake of haem (Woo et al., 2012)) (ter Beek et al., 2014). Both type I and type II importers contain a loop that extends into the NBD to couple allocrite (transport substrate) transport with ATP hydrolysis (Dawson et al., 2007).

Type I ABC importers are associated with a domain-wide rigid-body mechanism of transport, displaying an alternating access model, cycling between an open inward (resting state) and an open outward (allocrite binding, nucleotide associated) state. Type II importers are thought to have a more restricted movement between these states (Rice et al., 2014).

The type I and II structural folds of ABC importers can be extended generally to ABC exporters. Sav1866 and Pgp adopt a type I fold, demonstrating TMD domain swapping/cross-over, with two TMHs from each TMD crossing over forming interactions with the NBD via the intracellular loops (Dawson and Locher, 2007; Kim and Chen, 2018; Zolnerciks et al., 2007).

A mechanism of transport i.e. the coupling of ATP hydrolysis to allocrite transport of ABC exporters has been suggested by two models, discussed below.

1.1.4 'ATP-switch' and 'constant contact' models of ABC exporters

How ABC exporters couple nucleotide hydrolysis to allocrite efflux remains an area of hot debate, with no one definitive answer. Currently, there are two 'well-accepted' models for exporter efflux: the 'ATP-switch' and 'constant contact', that go some way to describing a coupled mechanism for allocrite transport.

Both models of transport are built on the structural description of ABC exporters. Sav1866 (*Staphylococcus aureus*) was one of the first accurate transporter crystal structures to be reported and remains one of the highest resolution (approx. 3.0Å) structural models for interpreting the binding and hydrolysis of ATP (Dawson and Locher, 2007, 2006). The protein was resolved in an 'open-inward' conformation (also described as a 'teepee') usually associated to a nucleotide free or ADP bound state, or an 'open-outward' conformation, associated with the ATP bound structure (AMP:PNP, non-hydrolysable ATP) (Dawson and Locher, 2007). Several lower resolution structural models of the mammalian ABC exporter ABCB1/ Pgp also demonstrated this conformation (Rosenberg et al., 2003, 2001; Zolnerciks et al., 2007).

The models for allocrite translocation within ABC transporters work on the premise that the association of allocrite to a high-affinity binding site(s) in the TMD is intrinsically coupled to the catalytic cycle within the NBDs. It is crucial that the mechanism is coupled in this order i.e. the ABC transporter in its basal state possessing a low affinity for ATP, otherwise ATP binding and hydrolysis would occur

independently of transport and the process rendered useless (Linton and Higgins, 2007).

The 'ATP-switch' model for transport originally described by Higgins and Linton in 2004 and updated in 2007 describes a multi-step transport process, whereby allocrites associate in the aforementioned high-affinity binding state causing a transmembrane helical conformational change, which is propagated to the NBD (most likely associated with conformational changes to the intracellular loop(s) (coupling helices) of the TMD that extend into the NBD (Kim and Chen, 2018; Lawson et al., 2008)), that brings about the rigid-body movement of the α -helical subdomain of the NBD, 'switching' the affinity of ATP in the nucleotide binding domain from a low state to a high state. This results in the association of a second ATP and the formation of a closed NBD dimer. The dimer formation is coupled to further conformational changes in the TMHs which leads to allocrite translocation. The sequential hydrolysis of ATP to ADP.Pi results in the NBD dissociation (via electrostatic repulsion as described previously in section 1.1.3 (NBD structure and dimerization)) and resetting of the transporter to its basal state (Higgins and Linton, 2004; Linton and Higgins, 2007; Smith et al., 2002) (overview of the ATP-switch model is represented in Figure 1.4).

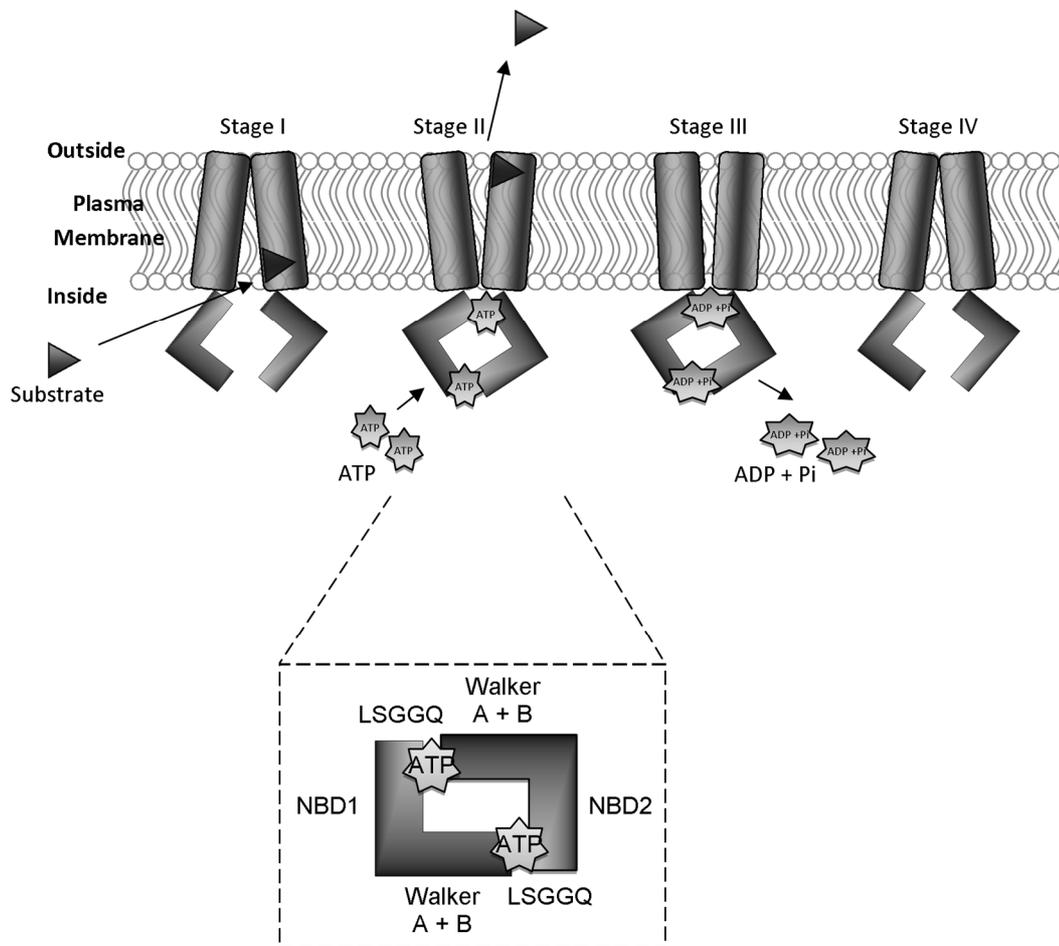


Figure 1.4 Simplified Schematic of the ATP-binding switch coupled to the alternate access model for ABC exporter function. Stage I, the ligand binds to the TMDs in the high-affinity open (separated) NBD conformation, inducing increased affinity for ATP. **Stage II**, ATP binding triggers the formation of the closed NBD dimer, which in turn stimulates a large conformational change in the TMDs which is sufficient to translocate the substrate ligand. The enlarged segment of the diagram shows a simplified version of the interactions taking place in the ‘ATP-sandwich’ **Stage III**, ATP hydrolysis initiates separation of the closed NBD dimer. **Stage IV**, transporter reset; Pi then ADP is released to complete the transport cycle and restore the protein to a high affinity state for another cycle of ligand binding (Linton, 2007). (Schematic was generated using Microsoft Office Publisher).

The major differences between the ‘ATP-switch’ model for transport and the ‘constant contact’ model are; that the ‘ATP-switch’ suggests the requirement of two hydrolytic events to drive a single transport cycle, this does not take into account transporters containing a degenerate nucleotide binding site (such as ABCG5/G8 (Wang et al., 2011)). It also shows that in the absence of ATP, the NBDs in the nucleotide ‘free’ state are completely separated (by up to 30Å) thus the switch

model implies significant conformational changes to accommodate the NBD dimerisation to facilitate allocrite transport (George and Jones, 2012; Linton and Higgins, 2007). The significant separation of the NBDs is currently a topic of discussion within the field, with leading thoughts suggesting that the significant separation is an artefact of detergent solubilisation (Wright et al., 2018).

The 'constant contact' model is derived from studies of the mammalian Pgp transporter by Senior et al. in 1995, that established an alternating hydrolysis (asymmetric) model of transport (Senior et al., 1995). Where one nucleotide binding site opens at the point of hydrolysis (presumably by electrostatic repulsion), enough to allow the release of ADP.Pi, whilst the second site remains occupied with ATP in an occluded/ dimeric state, the binding of ATP to the first site forms a 'fully closed' NBD dimer interface priming the second ATP for hydrolysis (George and Jones, 2012; Senior et al., 1995).

Despite the advances within the structural field of research, a comprehensive understanding of the mechanism underpinning allocrite transport remains elusive (Robey et al., 2018). Such an understanding is necessary to fully comprehend the physiological and pathological roles of ABC transporters.

1.1.5 ABC transporters in human disease

A number of heritable diseases are linked to germline mutations associated with the loss of ABC transporter physiology (Robey et al., 2018). The management of these diseases presents a considerable challenge, described below are three examples of ABC transporter disease association and the current therapies.

ABCC7

CFTR is a polytopic membrane protein of the ABCC subfamily, ABCC7. It consists of two transmembrane spanning domains and two nucleotide binding domains comprising the four core required domains for a functioning ABC transporter, as previously mentioned CFTR also contains an extra cytosolic domain termed the 'R' (regulatory) domain. The CFTR protein is not a typical ABC transporter, instead of participating in active transport, the protein is a chloride ion channel. The protein utilises the energy derived from ATP binding and subsequent relatively slow hydrolysis to open and close the chloride ion channel, respectively (Sheppard and Welsh, 1999). The opening and closing of the channel is also regulated by the phosphorylation of the regulatory domain (Sheppard and Welsh, 1999). Mutations within ABCC7 are the resultant cause of cystic fibrosis, an illness that is characterised by impaired bronchial and pancreatic secretions. A deletion of phenylalanine at position 508 (F508del) within the NBD1 of ABCC7 is the most common mutation found within at least one allele of 90% of patients (Rabeh et al., 2012). The deletion leads to a decreased folding efficiency and protein stability, resulting in increased ubiquitin mediated protein degradation (Rabeh et al., 2012). A combination therapy is currently available for the F508del, containing a corrector (lumacaftor), that improves the processing of the protein and its trafficking to the plasma membrane, and a potentiator (ivacaftor), that increases the open probability of the channel and therefore the transport of chloride (Deeks, 2016). This drug combination is not a perfect fix but does help reduce some of the symptoms associated with cystic fibrosis.

ABCA4

ABCA4, expressed in the photoreceptor cells, functions as an important transporter of vitamin A derivatives in the visual cycle (Auricchio et al., 2015). Mutations within the ABCA4 gene are the major cause of autosomal recessive retinal degeneration and was first characterised in 1997 as the gene responsible for autosomal recessive Stargardt disease (Auricchio et al., 2015). A large number (>900) of disease associated sequence variants have been identified, and result in the progressive loss of retinal function and structure leading to/ the main cause of vision loss (Auricchio et al., 2015; Kjellström, 2015; Tanna et al., 2017). The highly polymorphic nature of the ABCA4 gene makes it difficult to assign a single disease causation, and therefore difficult to assign a protein-specific target for treatment. There are clinical agents available that neutralise toxic by-products of the visual cycle, thereby slowing disease progression (Tanna et al., 2017). Gene replacement therapy using adeno-associated virus (AAV) derived vectors is currently being investigated as a curative method for retinal disorders (Auricchio et al., 2015).

ABCB4

The MDR3 gene, also known as ABCB4, was originally cloned from a liver cDNA library shortly after the discovery of MDR1 (ABCB1/Pgp, discussed later) (Van der Bliek et al., 1987). The gene encodes a 140kDa polytopic transporter with a typical MDR configuration of TDM₁:NBD₁:TMD₂:NBD₂ contained within a single polypeptide chain. MRD3 shares approximately 77% sequence similarity to Pgp and was thought to arise from MDR1 gene duplication, and therefore initially thought to share

extensive tissue distribution and allocrite diversity (Van der Blik et al., 1987). On the contrary, MDR3 protein is predominantly constrained to the apical side of canalicular membrane within the hepatocytes, colocalised with MDR1 and BSEP (ABCB11) (Smit et al., 1994). Its physiological role is to translocate phospholipids with a choline head group into the bile canaliculus (Sarkadi et al., 2006; Smit et al., 1994; Smith et al., 1994). Due to its restricted tissue distribution, and physiological function, ABCB4 is thought to play a key role in bile formation (Sarkadi et al., 2006). Single nucleotide polymorphisms within ABCB4 have been associated to cholestatic liver disease, including type three progressive familial intrahepatic cholestasis (PFIC), which is an autosomal recessive disease in which mutations lead to a frame shift resulting in nonsense and missense, often truncated, protein (Andress et al., 2017, 2014). This results in early onset cholestasis (bile flow disruption), progressing to cirrhosis and liver failure before adulthood (de Vree et al., 1998; Jacquemin et al., 2001). Intrahepatic cholestasis can also occur during pregnancy (ICP) and has serious consequences for both mother and foetus, often associated with foetal distress and spontaneous immature delivery (Sarkadi et al., 2006). ICP often occurs in women with no known family history of PFIC and the genetic basis of this disorder is unknown, although there is evidence of heterozygote missense mutations causing a disruption to protein trafficking and subsequent function (Dixon et al., 2000). ABCB4 mutations can also associate to low phospholipid-associated cholelithiasis (LPAC), where defects in ABCB4 function results in the production of bile with low phospholipid content and bile that is prone to gallstone production. Prophylactic therapy with ursodeoxycholic acid (UDCA) is given at the

diagnosis of LPAC to prevent the occurrence/ reoccurrence of the disorder and its associated symptoms (Rosmorduc and Poupon, 2007).

1.1.6 Multidrug resistance ABC transporters

As is exemplified by CFTR, ABCA4, and ABCB4, ABC transporters have a significant role in influencing human disease, making the understanding of their transport mechanisms of great importance. Another important aspect of ABC transporters is their ability to transport a diverse range of transport substrates limiting the bioavailability of pharmacological agents, including antineoplastic agents. The export of pharmacological agents (in humans) is mediated predominantly by ABCB1 (multidrug resistance transporter 1 (MDR1), P-glycoprotein (Pgp)), ABCC1 (multidrug resistance protein 1 (MRP1)), and ABCG2 (breast cancer resistance protein (BCRP)). These transporters are often overexpressed in certain cancers and limit the bioavailability of pharmacological agents, by active efflux out of the target cell. This leads to the characteristics of cellular resistance, termed multidrug resistance (MDR), which is thought to be the main cause for suboptimal disease related outcomes (Leonard et al., 2003). Therefore it is vital to understand the molecular mechanisms of these exporters for future disease management (Huang, 2009; Leonard et al., 2003).

Out of the three ABC transporters associated with MDR, ABCB1 and ABCC1 are the most extensively investigated. ABCB1 is the first member of the ABCB subfamily of human ABC transporters. The MDR gene, located on chromosome 7 q21.12, encodes a 140kDa protein that is comprised of an N-terminal transmembrane

domain and a C-terminal nucleotide binding domain in tandem repeat (i.e. TMD₁-NBD₁-TMD₂-NBD₂). The protein was first discovered in 1976 by Juliano and Ling, in a drug resistant Chinese hamster ovary cell line. It was reported that a single surface expressed glycoprotein was responsible for drug impermeability of the cell line, therefore they termed the protein P-glycoprotein (Pgp). The gene encoding hamster Pgp was subsequently cloned and soon after the human gene homologue was discovered, termed MDR1 (Gros et al., 1986; Roninson et al., 1986; Ueda et al., 1986). ABCB1 is expressed highly at the blood:brain barrier, intestine, placenta, kidney proximal tubules, and the liver bile ducts, and provides a protective mechanism by removal of xenobiotic compounds (Cascorbi, 2011; Cordon-Cardo et al., 1989; Leslie et al., 2005; Schinkel, 1999). The expression of ABCB1 has been attributed to clinical resistance to chemotherapy and a poor clinical outcome in adult patients with acute leukaemia (Marie et al., 1991). It has also been shown to be over expressed in certain breast cancers (Trock et al., 1997) and soft tissue sarcomas (Abolhoda et al., 1999). Because of this, ever since the first discovery of Pgp in 1976, research has been conducted to better understand its mechanisms of action in an attempt to create protein modulators for the reversal of MDR and influence drug design to avoid transport (Krishna and Mayer, 2001; Maki et al., 2003; Mlejnek et al., 2017; Sikic et al., 1997).

ABCC1 (also known as MRP1), located on chromosome 16 p13.11 was first isolated from a doxorubicin-selected human lung cancer cell line (H69AR) in 1992 by Cole et al. it encodes a 140-150kDa protein transporter with a topological arrangement (TMD₀-TMD₁-NBD₁-TMD₂-NBD₂). ABCC1 is expressed in most tissues/ ubiquitously

throughout the body with relatively high levels of expression residing within the blood:brain barrier, placenta, lung, kidneys (Hipfner et al., 1997; Leslie et al., 2005). It has been implicated in the transport of a structurally diverse array of conjugate organic anions, including cysteinyl leukotriene (LTC₄) and oxidised glutathione (Leier et al., 1994). MRP1 has shown to be expressed in several cancers, including glioblastomas, breast sarcomas, lung cancer, and Ewing's sarcoma of the bone, but there is a lack of convincing evidence that it affects the pharmacokinetics of clinical drugs and therefore has proved to be an unlikely target for cancer therapy (Filipits et al., 1996; Galimberti et al., 1998; Oda et al., 1997; Robey et al., 2018; Tews et al., 2000).

Comparatively, ABCG2 is less well understood and therefore presents itself as an interesting, important area for study.

1.2 ABCG2

1.2.1 Discovery of a third human MDR pump

ABCG2 was first discovered by Doyle et al. in 1998 from a multidrug-resistant breast cancer derived cell line, MCF-7 Adr/Vp. The cell line displayed an ATP-dependent reduction in the accumulation of anthracycline agents (mitoxantrone, doxorubicin, and daunorubicin) in the absence of ABCB1 or ABCC1 expression. RNA fingerprinting identified the overexpression of a 2.4kb mRNA, which encoded a 663 amino acid ABC transporter that was termed the breast cancer resistance protein (BCRP). Simultaneous to this discovery, Allikmets et al. presented a placenta-specific ATP-binding cassette gene, ABCP, encoding a 655 amino acid protein. The

gene was found to express in very high levels in the placenta and was thought to have an important role in protecting the foetus from harmful compounds (Allikmets et al., 1998). Finally, in January of 1999, only one month after the first reports of a new MDR ABC transporter a third group reported the molecular cloning of cDNAs which were highly overexpressed in a mitoxantrone-resistant human colon carcinoma cells (S1-M1-80) (Miyake et al., 1999). Once again mitoxantrone resistance was ATP dependent and therefore attributed to an ABC transporter that was termed MXR (mitoxantrone resistant) (Miyake et al., 1999). The sequence presented by Doyle et al., 1998 isolated from a drug-resistant breast cancer cell line showed a sequence variation at position 482 (glycine in comparison to arginine in the other sequences), thought to arise from repeat exposure to drug, which was later identified to alter substrate specificity (Honjo et al., 2001; Robey et al., 2003).

1.2.2 The genetic locus and the genetic regulation of expression

The ABCG2 gene is located on chromosome 4q21-22. The gene spans 66kb comprising of 16 exons, which extend from 60 to 532pb, and 15 introns (Bailey-Dell et al., 2001). The translation of the ABCG2 gene starts in exon 2 with the majority of the 5' UTR (untranslated region) in exon 1. ABCG2 promoter activity was characterised by luciferase reporter assays using cell lines with high endogenous expression of ABCG2 (including carcinoma cell lines, JAR, BeWo, and JEG-3). ABCG2 gene expression is under the regulation of a TATA-less promoter containing several putative Sp1, AP1, and AP2 binding sites. The promoter is located downstream from a putative CpG island which itself is located 312bp upstream of the transcriptional start site (Bailey-Dell et al., 2001).

The ABCG2 promoter contains several response elements including; estrogen response element (ERE) (Ee et al., 2004), hypoxia response element (Krishnamurthy et al., 2004), aryl hydrocarbon response element (Ebert et al., 2005), and an NFκB response element (Zhang et al., 2011). As a result, there are numerous mechanisms by which ABCG2 protein expression can be regulated. ABCG2 expression has also been reported to be suppressed by DNA methylation within the CpG island (Turner et al., 2006).

ABCG2 protein expression can also be regulated post-transcriptionally by microRNAs. MicroRNAs are small non-coding RNA molecules, several miRNAs have been characterised for ABCG2 including miR-519c, miRNA-328, and miR-655-3p. These microRNAs results in a decrease in ABCG2 mRNA (messenger RNA) and subsequently reduced protein expression, predicted to interact within a putative binding site within the 3' untranslated region (Awortwe et al., 2018; Pan et al., 2009; To et al., 2009). It is proposed that both 5' and 3' UTR putative binding sites are absent within some drug selected cell lines leading to an increase in ABCG2 protein expression (Nakanishi et al., 2006; To et al., 2009).

ABCG2 has also been reported to be regulated at the post-translational level by phosphorylation of threonine and position 362 by Pim-1 kinase. Knockdown of Pim1L (44kDa isoform) and mutagenesis of T362 to alanine resulted in compromised ABCG2 oligomerisation and localisation (Xie et al., 2008).

1.2.3 The overall topology of the expressed protein

The ABCG2 gene encodes a 655 amino acid polypeptide chain. The polypeptide chain forms a polytopic reverse half-transporter topology of approximately 72kDa in size. It is localised to the plasma membrane of epithelial cells, and consists of an intracellular, N-terminal nucleotide binding domain (NBD) and a C-terminal, transmembrane spanning domain (TMD) (NBD-TMD). The TMD was predicted to have 6 membrane spanning alpha helices, interconnected via intra/extra cellular looped regions (Allikmets et al., 1998; Wang et al., 2008).

1.2.4 Tissue distribution and physiological roles of ABCG2

The exact physiological role of ABCG2 is still to be established. However, it is being increasingly recognised for its role in the disposition and tissue exposure of drugs and xenobiotics. In its native form ABCG2 predominantly localises to the apical membrane of cells, frequently in tissues which have secretory or barrier functions, this includes tissues such as the blood-brain barrier, placenta, gastrointestinal (GI) tract, testes, mammary glands. ABCG2 is also expressed in stem cells (Cole et al., 1992; Jonker et al., 2005; Maliepaard et al., 2001; Mao, 2008; van Herwaarden et al., 2007; Van Herwaarden and Schinkel, 2006).

The physiological expression of ABCG2 at the blood-brain barrier (BBB) is thought to provide a fundamental protective role by severely limiting the exposure/ uptake of antineoplastic agents, mediating their active efflux into the blood vessel lumen (Cisternino et al., 2004). Mdr1a(-/-) mouse brains showed a carrier-mediated efflux of mitoxantrone (MX) and prazosin (PRA) with an upregulation in ABCG2 mRNA

expression in comparison to the wild-type mouse brain (Cisternino et al., 2004). More recently, *in vivo* knockdown of *Abcg2*(-/-) in mice showed an improvement in the brain penetration of temozolomide (used in the treatment of high-grade gliomas) by about 20% (de Gooijer et al., 2018). The expression of ABCG2 at the BBB has also been assessed for its protective mechanism against Alzheimer's disease. The accumulation of amyloid- β peptide (A β) in the brain is a hallmark for Alzheimer's disease. Human ABCG2 stably expressed in HEK293 cells demonstrated the cellular efflux of [3H] A β 1-40, this result was also corroborated in *in situ* brain perfusion technique of A β in *Abcg2* deficient mouse studies (Do et al., 2012).

The presence of ABCG2 in the placenta, specifically in syncytiotrophoblastic cells indicates that ABCG2 may provide a protective role against endogenous toxins and xenobiotics in the development of the foetus (Maliepaard et al., 2001). For example, glibenclamide, a sulphonylurea, used in the treatment of gestational diabetes, has shown to have limited tissue exposure within the placenta (Langer et al., 2000). An *in vitro* drug accumulation study of transporters expressed within the placenta (ABCB1, ABCC1, ABCC2, ABCC3, and ABCG2) revealed preferential efflux of [3H]-glibenclamide by ABCG2 in a drug selected cell line (MCF7-MX) (Gedeon et al., 2006), thereby protecting the foetus from neonatal hypoglycaemia induced by therapeutic drug treatment.

The expression of ABCG2 in the mammary gland seems to have a contrasting role to the protective one demonstrated in the placenta. ABCG2 has been reported to transport xenobiotics into the milk, which poses a serious health risk for breast-fed

infants (Anderson, 1991; Jonker et al., 2005). This also extends to the dairy industry, where polymorphisms in bovine ABCG2 reportedly have an effect on milk disposition and quality (Jonker et al., 2005; Otero et al., 2015; Sanchez et al., 2017).

ABCG2 is also located in the brush-border membrane of kidney proximal tubule cells, where it has been shown to mediate renal urate secretion (thought to be the main physiological function of ABCG2) (Huls et al., 2008; Woodward et al., 2009). A common single nucleotide polymorphism (SNP) rs2231142 (Q141K point mutation) (described later), demonstrated a 53% reduction in urate transport *in vitro* compared to the wild type protein (Woodward et al., 2009).

Furthermore, ABCG2 was observed to be localised at the apical membrane in the epithelium of both the small intestine and the colon, as well of the canalicular membranes of the liver. This localisation restricts the absorption of dietary toxins from the gastrointestinal tract and eliminates compounds from the liver, into the bile (Tucker et al., 2012). These sites are imperative to the absorption, distribution and elimination of endogenous and exogenous substances, and therefore limits the bioavailability of drugs with pharmacological importance (Ni et al., 2010b).

ABCG2 has also been found to be enriched in a population of primitive stem cells, likely to provide a protective role against xenobiotics. The overexpression of ABCG2 in colon cancer stem cells (CSCs)-like population from colon cancer tissues, is thought to be the main reason behind chemotherapy failure and tumour recurrence (Ni et al., 2010b; Xie et al., 2014).

An increase in the expression of ABCG2 has been linked to an increase in the efflux of numerous diverse pharmaceuticals used in the treatment of cancer, including mitoxantrone, methotrexate, topotecan derivatives, SN-38, and flavopiridol, leading to the conformation of multidrug resistance (Li et al., 2007). This has implications in the ADMETox (absorption, distribution, metabolism, elimination and toxicity) profile of many compounds with an intracellular target. For example, a recent study, showed that a novel camptothecin analogue, FL118, was able to bypass ABCG2-mediated drug resistance in colon and lung models, demonstrating an improved efficacy in comparison to the use of SN-38, a known ABCG2 substrate (see an extensive list of substrates and inhibitors later) (Westover et al., 2015).

1.2.5 Polymorphisms of ABCG2

The major SNP is a predictor of gout

ABCG2 is a highly polymorphic transporter that has over 80 single nucleotide polymorphisms (SNPs) residing within its gene coding region, many of which alter the substrate recognition/binding affinity and are of significant pharmacological and physiological relevance (Ni et al., 2010b). An example of a well characterised SNP within ABCG2 is the rs2231142 (or 421C>A) polymorphism, which brings about a glutamine to lysine mutation at position 141, Q141K. The mutation varies significantly with ethnicity, occurring in 35% of the Chinese population, whereas it is less recorded at less than 1% among African-Americans (Sissung et al., 2010). The mutation leads to instability residing within the nucleotide-binding domain, resulting in a folding and trafficking defect. The consequence of this mutation is

increased degradation, reduced function and temperature sensitivity, resulting in the reduced excretion of urate (Woodward et al., 2013, 2009). This increases urate plasma levels (hyperuricemia) leading to sodium urate crystal formation in and around the joints, the known cause of gout.

Many SNPs can influence pharmacokinetics

Single nucleotide polymorphisms with phenotypic differences in the expression or function of ABCG2 may have important roles in the disposition and efficacy of pharmacological agents (Heyes et al., *In Press*).

The Q141K (421C>A) SNP has been shown to result in the increased bioavailability of topotecan, also more than doubling the exposure to rosuvastatin, commonly used in the treatment of hypercholesterolemia (Sissung et al., 2010). The Q141K genetic variation has also been found to have a potential influence on the clinical efficacies, in terms of 5-year progression-free survival, in patients with advanced gastrointestinal stromal tumour undergoing imatinib therapy (Koo et al., 2015).

Several other polymorphisms have been associated with altered drug pharmacokinetics. Polymorphisms resulting in F431L (1291 T>C) and F489L (1465 T>C) mutations have shown an impaired porphyrin (pheophorbide A) transport, as well as decreased resistance to antineoplastic agents SN-38, mitoxantrone, and methotrexate. This may result in an increased risk of hypersensitivity to selective antineoplastic agents, especially for methotrexate, as transport was completely abolished (Deppe et al., 2014; Sjöstedt et al., 2017; Tamura et al., 2007b, 2007a, 2006; Yoshioka et al., 2007).

Despite the significant medical importance, the mechanism of transport as stated remains unknown. This has led to the understanding of the protein's mechanism being increasingly important. It is of such significance that the FDA provides guidelines for the investigation of potential interactions of pharmacological agents with ABCG2 at the drug development stage (Huang, 2009).

1.2.6 Site-directed mutagenesis

A great deal of time and effort has been put aside for *in vitro* site-directed mutagenesis studies of ABCG2, as these have previously provided crucial information about its function. Due to the vast range of transport substrates and wide tissue distribution implicating ABCG2 in MDR, the ability to understand this protein in a structural and functional sense is of great importance. Studies where strategic residues are mutated are key to an enhanced understanding within transporter studies, to ultimately produce a comprehensive map of the functionally important residues; i.e. the R482G/T mutation and its functional characteristics suggests the residue is important in substrate recognition, stabilisation, and/or transport activity (Cai et al., 2010; Chen et al., 2003; Clark et al., 2006; Ejendal et al., 2006; Honjo et al., 2001; Ozvegy-Laczka et al., 2005; Özvegy et al., 2002; Robey et al., 2003).

For this area of research mutations generally can be assigned to four crude but effective categories, established based on the characteristics displayed, as outlined below (Ni et al., 2010b).

The first category describes mutations which do not alter ABCG2 expression levels but affect overall substrate specificity and/or transport activity. Arginine 482 to glycine/threonine (TM3), a classic example, it is a gain-of-function mutation, and it results in the ability to transport daunorubicin, rhodamine 123, and lyso-tracker green. The mutant protein is also able to transport most native ABCG2 substrates with the exception of methotrexate, so R482G is also a loss of function mutant (Nakanishi et al., 2003; Ni et al., 2010b). P485A is another mutation which demonstrates no difference in expression characteristics but shows perturbed allocrite efflux. The mutation demonstrated a reduced efflux for BODIPY-prazosin by up to 70% (compared to wild type protein) but had no effect on the efflux of Hoechst 33342 or pheophorbide A (PhA). functional data regarding mitoxantrone (MX) is conflicting, with Haider et al., 2015 reporting a significant reduction in Ko143-inhibitable efflux of mitoxantrone, whereas Ni et al., 2011 concluded no effect on MX efflux (Haider et al., 2015; Ni et al., 2011).

The second category is reserved for mutations in which the protein biogenesis is disrupted, i.e. the protein demonstrates decreased stabilisation, is often expressed at a lower level and the subcellular distribution is impaired. Q141K, a common SNP (described above) is a classic example of a mutation resulting in disrupted protein biogenesis.

The third category is alterations in post-translational modifications. ABCG2 is glycosylated at a specific amino acid within the third extracellular loop (ECL3) between TMHs 5 and 6. Mutagenesis studies revealed glycosylation occurred at an

asparagine residue at position N596. Mutagenesis of asparagine to glutamine (N596Q) resulted in a glycosylation defective mutant protein. However, this mutation seemed to have no impact on localisation of the protein or transport activity, it showed a slight decrease in expression in comparison to wild type ABCG2 and therefore it was suggested that it could increase proteasomal-mediated protein degradation (Haider, 2011; Wakabayashi-Nakao, 2009). The formation of intra/intermolecular disulphide interactions has also been established as being significant in the stabilisation and oligomerisation of the protein transporter, Cys592, 608 and 603 (Henriksen et al., 2005; Kage et al., 2005; Wakabayashi et al., 2007).

The final category possesses mutations which neither have major effects on the expression or function of the protein, these residues are therefore suggested to be non-critical.

This method for categorising mutations is convenient. However, it is also crude as not all mutation characteristics are mutually exclusive. The isoleucine to alanine variant at position 573, proposed to reside within ECL3 (based on the HA epitope insertion determined topology model of ABCG2 (Wang et al., 2008)) resulted in a glycosylation defective mutant with reduced expression, maturation (post-translational glycosylation) and trafficking (and consequently reduced efflux capacity) which fits with a mutation of both disrupted biogenesis and altered post-translational modification. This was suggestive of a fundamental structural defect with reduced maturation being a symptom/ side effect of a potential destabilisation

of ECL3 and disruption of inter- and intramolecular disulphide bonds (Haider et al., 2015).

1.2.7 Oligomerisation of ABCG2

As a 'half-transporter', ABCG2 is required to form higher ordered structures, the minimum conformation being a dimer. Several studies have suggested that it homodimerises (Haider et al., 2011; Ni et al., 2010c), with several residues considered important in the formation of intermolecular interaction, namely residues in the extracellular loop between TM helices 5 and 6; Cys592, Cys603 and Cys608. Data are consistent with Cys592 and 608 forming an intramolecular disulphide 'bridge', assumed to stabilise the extracellular loop. It is also assumed that this stabilisation allows for accessibility of Cys603, which is thought to form intermolecular disulphide interactions, contributing to the oligomerisation of the transporter (Kage et al., 2005). However, substitution studies of cysteine residues including Cys603, reported no effect on the dimerisation of ABCG2 or any effect on functional activity, but that other cysteine residues may interact to form intramolecular disulphide bonds important for structural stability (Liu et al., 2008; Ni et al., 2010c). The molecular basis of dimerisation was also proposed to occur through a classic GxxxG motif, a motif that has been identified for transmembrane inter-helix association (Polgar et al., 2004; Russ and Engelman, 2000). Mutational analysis of G406 and G410 to lysine demonstrated a marked effect on the transport of mitoxantrone, pheophorbide A, and BODIPY-prazosin, but mutations retained susceptibility to cross-linking (disuccinimidyl suberate (DSS) or dithiobis

succinimidyl propionate (DPS)), demonstrating high molecular weight complex under non-reducing conditions, suggesting dimerisation (Polgar et al., 2004).

There is also data that provides evidence of higher oligomeric structures, including initial electron microscopy (EM) analysis (3D structure resolved at 18Å) suggesting the formation of an octameric structure, consisting of a tetramer of dimers (McDevitt et al., 2006; Mo and Zhang, 2012). Step-wise photo bleaching studies using TIRF (total internal reflection fluorescence) microscopy also demonstrated ABCG2s capacity to form higher-ordered oligomers supporting the formation of a tetrameric structure (i.e. a dimer of dimers) (Wong et al., 2016). Photon-counting histograms (PCH) using fluorescence correlation spectroscopy (FCS) confirmed the formation of both dimeric and tetrameric (up to 69% of the molecules analysed) in live cell analysis (Wong et al., 2016).

Recent structural studies of ABCG2 (discussed later) demonstrate the protein forming a dimeric complex, suggesting that oligomeric state seen in early EM may be a result of particle (protein) packing in EM grid preparations. Whereas the tetrameric evidence shown by Wong et al., 2016 may be indicative of physiological transient complexes.

1.2.8 Pharmacology of ABCG2

Repertoire of drugs

ABCG2 has a wide tissue distribution and the physiological role of ABCG2 hasn't been fully established but its role in disease related MDR has provided information on its transport capacity. ABCG2 possesses an extensive array of substrates and

inhibitors. The molecules are both clinically important and structurally diverse, including cytotoxic compounds (mitoxantrone, topotecan, flavopiridol and methotrexate), antibiotics, statins (rosuvastatin), immune-suppressants (including anti-HIV drugs), anti-diabetic agent (glibenclamide), anti-rheumatic agents (sulfasalazine), fluorescent dyes (e.g. Hoechst 33342), photosensitizers (pheophorbide A and protoporphyrin IX) and tyrosine kinase inhibitors (imatinib, gefitinib and nilotinib) (Table 1.1) (Mao and Unadkat, 2015; Van Herwaarden and Schinkel, 2006). At present exactly how ABCG2 transports a diverse range of structurally un-related molecules remains for the majority, unknown (although recent structural studies are helping to build upon this picture, discussed later).

ABCG2 drug substrates	
Drug	Reference
Anthracenes	
Mitoxantrone	(Litman et al., 2000; Robey et al., 2003)
Bisantrone	(Litman et al., 2000)
Aza-anthrapyrazole	(Hazlehurst et al., 1999; Rabindran et al., 2000)
Camptothecin derivatives	
Topotecan	(Litman et al., 2000)
SN-38	(Nakatomi et al., 2001)
Irinotecan	(Maliepaard et al., 2001; Rabindran et al., 1998)
Diflomotecan	(Sparreboom et al., 2004)
Polyglutamates	
Methotrexate	(Chen et al., 2003; Volk et al., 2002)
Methotrexate-Glu2	(Chen et al., 2003; Volk and Schneider, 2003)
Methotrexate-Glu3	
Nucleoside analogs	
AZT	(Wang et al., 2004, 2003)
AZT 5'-monophosphate	
Lamivudine (3TC)	
Other drugs	
Prazosin	(Litman et al., 2000)
Indolocarbazole	(Nakagawa et al., 2002)
Flavopiridol	(Robert W Robey et al., 2001b)
Canertinib (CI1033)	(Erlichman et al., 2001)
Imatinib mesylate (STI571)	(Burger et al., 2004)
Gefitinib (ZD1839)	(Elkind et al., 2005)
Nilotinib	(Brendel et al., 2007)
Glyburide	(Zhou et al., 2008)
Cimetidine	(Pavek, 2004)
Sulfasalazine	(Van Der Heijden et al., 2004)
Nitrofurantoin	(Merino et al., 2005)
Rosuvastatin	(Kitamura et al., 2008)
Pantoprazole	(Breedveld et al., 2004)

Substrates listed in this table are only for wild-type ABCG2

ABCG2 drug inhibitors		
Drug	IC50 (nM)	Reference
Tyrosine kinase inhibitors		
Gefitinib	300	(Ozvegy-Laczka et al., 2004)
Imatinib mesylate	170	(Houghton et al., 2004)
Erlotinib	ND	(Shi et al., 2007)
Nilotinib	ND	(Dohse et al., 2010; Hegedus et al., 2009)
Lapatinib	ND	(Dai et al., 2008)
HIV protease inhibitors		
Ritonavir	19,500	
Saquinavir	19,500	(Gupta et al., 2004)
Nelfinavir	12,500	
Lopinavir	7660	(Weiss et al., 2007)
HCV protease inhibitors		
Boceprevir	81,000	(Chu et al., 2013)
Telaprevir	30,000	(Fujita et al., 2013)
Calcium channel blockers		
Dipyridamole	6400	
Nicardipine	4800	
Nimodipine	13,700	(Zhang et al., 2005)
Nitrendipine	ND	
Antifungal azoles		
Ketoconazole	15,300	
Itraconazole	ND	(Gupta et al., 2007)
Fluconazole	ND	
Immunosuppressants		
Cyclosporin A	4300	
Tacrolimus	3600	(Gupta et al., 2007)
Sirolimus	1900	
Other drugs		
Novobiocin	50–100	(Shiozawa et al., 2004; Yang et al., 2003)
Tamoxifen	ND	(Sugimoto et al., 2003)
Reserpine	ND	(Zhou et al., 2001)
Omeprazole	10,000–50,000	
Pantoprazole	ND	(Breedveld et al., 2004)

ND not determined

Table 1.1 ABCG2 drug substrates and inhibitors. Table taken from (Mao and Unadkat, 2015).

Multiple drug binding sites

Studies in P-glycoprotein (Pgp, ABCB1), which shares many common substrates and inhibitors to ABCG2, have suggested that Pgp possess multiple drug binding sites, with at least four being identified in previous research (Martin et al., 2000).

Due to the commonality in transport substrates this led to a hypothesis that ABCG2 could act via a similar binding mechanism with multiple drug binding sites. This hypothesis was later investigated by Clark et al. who demonstrated the presence of at least two pharmacologically equivalent drug binding sites (one in each TMD of a dimeric conformation) by [³H]daunomycin displacement binding studies.

Doxorubicin, prazosin, and unlabelled daunomycin were able to fully displace bound [³H] daunomycin, whereas both mitoxantrone and Hoechst 33342 were only able to partially displace [³H] daunomycin, the partial displacement was suggestive of a level of cooperative binding/ allosteric interactions between the two sites (Clark et al., 2006).

The potential for multiple drug binding sites has most recently been explored structurally and will be discussed in context later.

Communication between NBDs and TMDs

Communication between the NBDs and the TMDs of ABCG2 is essential for allocrite transport. This has been demonstrated time and time again. For ABCG2, an example would be the substitution of the 'catalytic carboxylate' E211 (within the Walker-B of the NBD) to a non-acidic glutamine residue which completely abolished

ATPase activity. The mutation also resulted in the loss of methotrexate transport, thus demonstrating the need for ATP hydrolysis for allocrite efflux (Hou et al., 2009).

McDevitt et al. in 2008 further investigated the coupling of allocrite ($[^3\text{H}]$ daunomycin) binding in the ABCG2^{R482G} isoform to ATP hydrolysis. The investigation showed that the binding of $[^3\text{H}]$ daunomycin was displaced by ATP analogues (non-hydrolysable analogues ATP- γ -S and TNP-ATP) with less than 20% of initial binding of the radioligand observed. This indicated a transition from a high affinity allocrite binding state to a low affinity state upon ATP binding, this low affinity state was retained immediately after hydrolysis. Therefore, it was proposed that the 'power stroke' for translocation of allocrite was nucleotide binding, demonstrating an important communication between the NBD(s) and TMD(s) during allocrite transport cycle (McDevitt et al., 2006). Projection electron microscopy structures of ABCG2, although at low resolution, (5-7Å) also suggested conformational changes within the protein in the presence and absence of substrate (MX), supporting the 'ATP-switch' functional model (as described in section 1.1.4) and demonstrate the cross communication between the TMD(s) and NBD(s) in transport (Higgins and Linton, 2004; Linton and Higgins, 2007; Rosenberg et al., 2010).

Most recently Jackson et al., 2018 have shown Ko143 derivatives to knockdown estrogen-1-sulphate (E₁S) stimulated ATPase activity. One of these derivatives MZ29 was found to bind at a 2:1 ratio in the central cavity between the two-monomeric halves of ABCG2, again suggesting cross-talk between the inhibitor

bound within the TMD and the NBD, but exactly how this occurs is open to interpretation.

1.2.9 Structure of ABCG2

At the commencement of these current studies little was definitively known about the structure of ABCG2. Six transmembrane (TM) spanning helices were identified by topological prediction algorithms and confirmed by hemagglutinin (HA) epitope insertion experiments (Wang et al., 2008). It was the topological data presented by Wang et al. that was used in the development of a 'lateral slice' hypothesis for drug recognition/binding described in this current study.

There were low resolution cryo-EM structures and homology models which were refined using MsbA, Sav1866, and mouse Pgp structures (McDevitt et al., 2009, 2006; Ni et al., 2010b; Polgar et al., 2010; Rosenberg et al., 2010). However, the very low sequence similarity between ABCG2 and ABCB transporters (less than 15% identity in the TMDs) renders these models questionable and attempts to use such to interpret experimental data were beset by contradictions. For example, threonine residue at position 402 (in transmembrane helix 1, TMH1) close to the GXXXG motif (common in forming TMD:TMD interactions) (TXXXGXXXG), was proposed to be involved in dimerisation of ABCG2. Studies substituting T402 for alanine, leucine, or arginine report no particular effect on protein expression or plasma membrane localisation but demonstrated perturbed allocrite efflux (significant decrease in the efflux of mitoxantrone, Hoechst 33342, and BODIPY-prazosin and pheophorbide A). Homology modelling of ABCG2 based on the

bacterial transporter, Sav1866, suggested that T402 interacted with residues from TMH5 and/or 6 of the opposing monomer possibly forming inter-helical interactions promoting dimer formation (Polgar et al., 2010). Whilst another study using a homology model based on mouse Pgp concluded that T402 was outside of the putative drug binding site (situated at the TMD interface), and likely to form inter-helical interactions with TMH 2 or 3, proposing that it was possible that T402 formed part of the drug translocation pathway (Ni et al., 2010a). This demonstrates a need for better structural interpretation of ABCG2.

ABC transporter structural resolution revolution

During the last decade there have been significant advancements in membrane protein structural biology. This is owed to enhancements in membrane protein extraction and stabilisation as well as advances in cryo-EM imaging (Bai et al., 2015; Dörr et al., 2016; Gulati et al., 2014). Early cryo-EM images were recorded on photographic film and electron beam-induced sample movement resulted in image blurring which made particle determination difficult and often resulted in multiple structural intermediates within a single reconstruction. Three advances have yielded better reconstitutions: (i) the development of digital direct-electron detectors, which allow the recording of movies during electron-beam exposure, (ii) software improvements have allowed the correction of beam-induced movement, and (iii) powerful classification methods can determine multiple structures from a sample mixture. All of which now make cryo-EM a rival, in terms of resolution, for X-ray crystallography (Bai et al., 2015).

ABCG family structural revolution

This 'resolution revolution' has translated into an explosion of structures for the ABC transporter superfamily, resolved by either cryo-EM or X-ray crystallography utilising enhanced membrane extraction methods, with a representative structure for all but one human ABC subfamilies (ABCD). This explosion also resulted in the publication of a crystal structural of ABCG5/ABCG8 (ABCG5/G8), a close family member of ABCG2 (Lee et al., 2016).

ABCG5/G8 is a 6-TM spanning reverse half transporter that forms functional heterodimers to mediate the efflux of sterols (cholesterol and phytosterols) in the liver and intestine (Kerr et al., 2011). ABCG5/G8 was crystallised from protein co-expressed in *Pichia pastoris* and reconstituted into 1,2-Dimethyl-sn-glycero-3-phosphocholine (DMPC) bicelles to obtain a structural resolution of 3.9Å in a nucleotide-free state, 'open-inward' conformation (Lee et al., 2016).

Importantly, ABCG5/G8 is the first sterol transporter to be resolved at a near-atomic level and the first ABCG family member to be structurally characterised. ABCG5 and ABCG8 share around 25% sequence identity to ABCG2, and although low in the general scheme of homology modelling, László et al. were able to generate a homology model for ABCG2 demonstrating stability of the homology model in molecular dynamics simulations. It is also important to note that although ABCG5/G8 forms a heterodimeric structure they demonstrate a high degree of structural conservation, so this was not considered a problem in the generation of an ABCG2 homology model (László et al., 2016; Lee et al., 2016).

The ABCG proteins from these studies demonstrate TMH packing and interfacial contacts similar to that of type II importer fold (Locher et al., 2002; Rice et al., 2014). Early ABC exporter structures (Sav1866, MsbA, and mouse Pgp), from which ABCG2 was repeatedly modelled on, comprised of a type I structural fold, i.e. transmembrane domains that extend out of the membrane and contact the NBD via long intracellular looped regions that are thought to couple allocrite transport to nucleotide hydrolysis, called the coupling helix. Type I exporters also show domain swapping where TMHs from each half of the protein cross over to the opposing TMD, with an intracellular loop extending into the opposing NBD also thought to couple transport (Dawson and Locher, 2007; Zolnerciks et al., 2007).

The ABCG proteins resemble a structure most similar to type II bacterial importers, demonstrating a more compact TMD organisation, compared to type I exporters (Figure 1.5) (Hollenstein et al., 2007; Locher et al., 2002; Oldham et al., 2007b). It is therefore no surprise that the homology models of ABCG2 based upon the ABCB transporters were so inaccurate (as indicated at the beginning of section 1.2.9). Transmembrane domain swapping is also not present in this structure, rather the TMD appears to contact the NBD via an intracellular loop between TMHs 2 and 3 (ICL). This contact seems to be further stabilised through an amphipathic α -helix which precedes TMH1 (TMH1a), which is interfacial to the membrane bilayer, termed the connecting helix (László et al., 2016; Lee et al., 2016). The connecting helix demonstrates the potential to form multiple helical interactions, including interactions with ICL1 and an α -helix within the NBD containing Q141, which could

explain the protein destabilisation effect of the Q141K ABCG2 isoform (László et al., 2016; Lee et al., 2016).

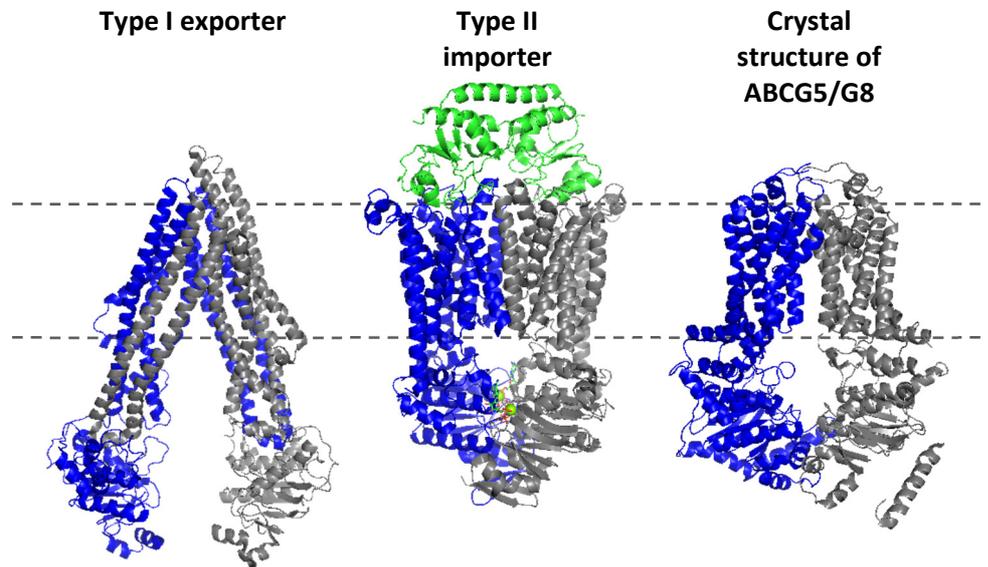


Figure 1.5 ABCG5/G8 structure resembles type II transporter fold. Type I exporter structural topology of human ABCB1 in the post hydrolytic open-inward conformation, demonstrating domain swapping and elongated TMD topology (PDB: 6GDI), Type II importer structural topology of ButCD-F in the nucleotide bound state with SBP (green), the protein possesses a more compact TMD organisation with no domain swapping (PDB: 4FI3). The ABCG5/G8 (PDB: 5DO7) structure adopts a structural topology more closely resembling a type II importer, with compact TMD organisation and no domain swapping. One half of each structure had been coloured blue for easy visualisation.

Although the NBDs of ABCG5/G8 were resolved in a nucleotide free state, they form contacts with each other at the extreme cytoplasmic end to form a closed conformation through a pair of NPxDF motifs (where x is any residue). These motifs are conserved amongst the ABCG sub-family and are essential for cholesterol efflux by ABCG1 (Lee et al., 2016). A final noteworthy feature of the ABCG secondary structure, is a series of short α -helices (TMH5a, 5b, and 6a) previously predicted to be part of continuous TM helices 5 and 6 (László et al., 2016; Lee et al., 2016) (see Figure 1.6 ABCG2 topology schematic based on HA-epitope insertion versus new structural information), which appear to form a structural cap at the extracellular

interface of the TMDs. Subsequent cryo-EM structures of ABCG2 were published by Taylor et al., 2017 and Jackson et al., 2018 confirming the ABCG family structural topology (discussed in detail in Chapter 6).

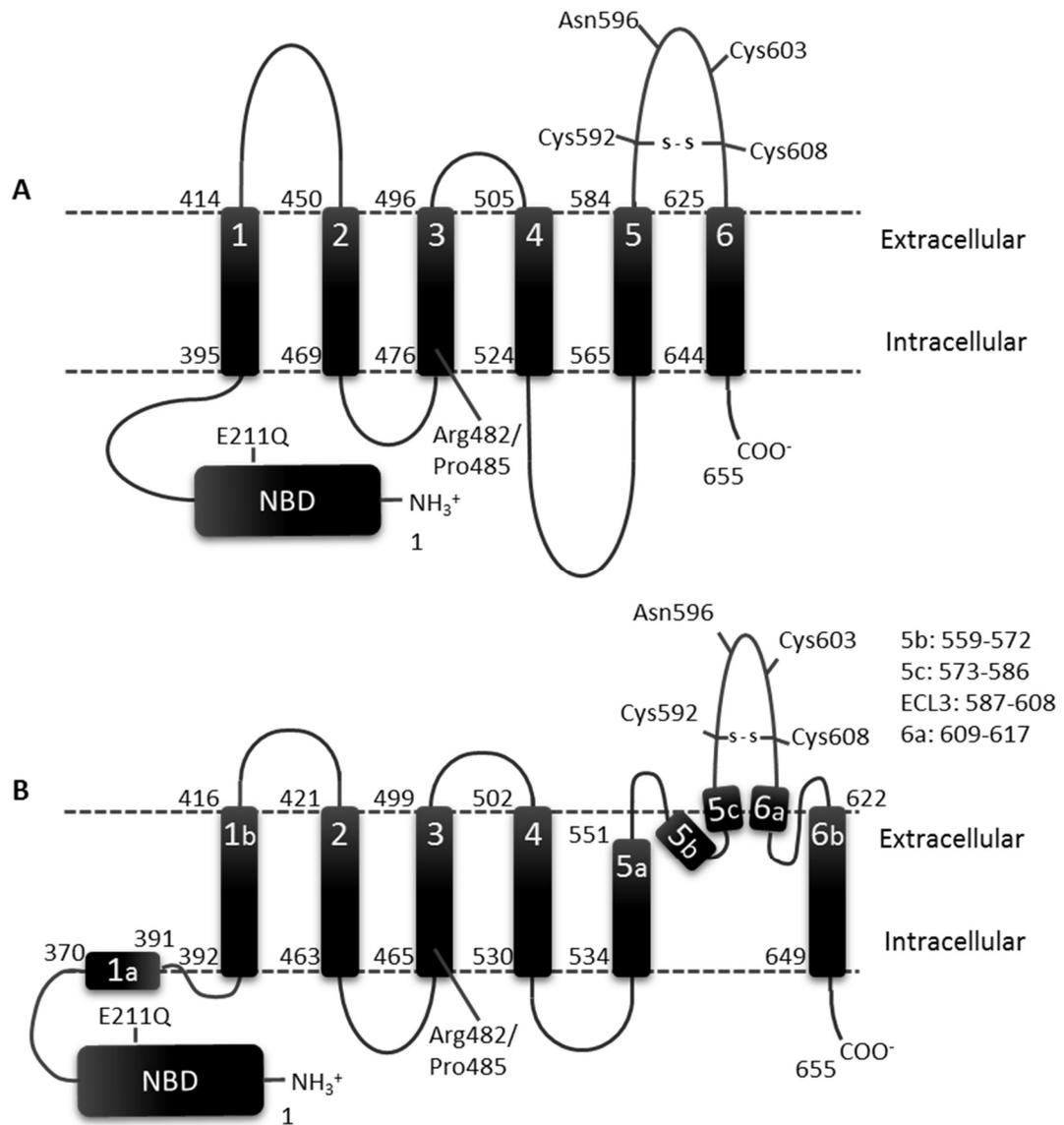


Figure 1.6 ABCG2 topology. A. ABCG2 topology based on HA-epitope insert (Wang et al., 2008). **B.** ABCG2 topology based on cryo-EM structure (Jackson et al., 2018). Residues E211 (catalytically inactive variant), R482, and P485 (residues thought to be involved in allocrite binding) are highlighted, as well as Asn596 (glycosylation site), and cysteine residues thought to form inter and intra molecular disulphide bonds.

This new structural information for ABCG2 was used in the interpretation of residues mutated in the current 'lateral slice' hypothesis and rational design of a new working model, termed the 'binding pocket mutations', described in Chapter 4.

1.3 Project aims

ABCB1 (P-glycoprotein: P-gp), ABCC1 (multidrug resistance protein: MRP1) and ABCG2 (breast cancer resistance protein: BCRP), contribute to the phenomenon of multi-drug resistance (MDR), possessing extensive substrate polyspecificity. The protein transporters export drugs of clinical importance including chemotherapeutic agents from cells, acquiring the characteristics of cellular resistance. Molecular explanations for MDR are therefore another aspect that is of high importance for the future of disease management. ABCG2 has lagged for many years in terms of its structural and functional understanding, in the shadow of its larger cousin ABCB1, leaving gaps in the fundamental knowledge of amino acid residues required for substrate polyspecificity and transport activity. So, a greater understanding of exactly how ABCG2 binds and transports a large number of structurally unrelated drug substrates is required.

Therefore, the overall aim of this project is to elucidate potential drug binding interactions of ABCG2.

The initial 'lateral slice' hypothesis resulted from consideration of the two residues (R482 and P485, described earlier (Haider et al., 2015; Honjo et al., 2001; Ni et al., 2011; Ozvegy-Laczka et al., 2005; Robey et al., 2003)) which have been shown to impact on allocrite specificity. R482 and P485 are approximately 1 helical turn

apart and reside towards the cytoplasmic interface of ABCG2, in the experimentally determined topology model as described by Wang et al., 2008. It was hypothesised that residues taken from a lateral transection of the TMD at equalling spatial positions to R482 and P485 may also form part of this transport substrate recognition site.

To test this hypothesis a total of 10 residues were selected for mutagenesis. It was hoped that downstream expression and functional characterisation of the mutants would then provide both structural and functional information of the selected residues and reveal their importance in drug substrate transport.

The fast-moving nature of the ABCG structural family then provided a framework for a further round of mutagenesis and functional analysis. The combined information on these two series of residues would hopefully aid in bridging the gap in our understanding of residues important for drug transport and build upon the current understanding of the transport mechanism of ABCG2, with the hope that this would lead to the generation of experimentally informed modulators of ABCG2 for the use in the treatment of human disease.

Chapter 2 Methods

2.1 Molecular Biology

2.1.1 Materials and reagents

All molecular reagents were purchased from either New England Biolabs (NEB) or Promega unless stated otherwise. The fluorescently labelled mammalian expression vector, pcDNATM-3.1/zeo(+)_sfGFP_ABCG2 was kindly provided by Dr Kelvin Wong.

2.1.2 Site-directed mutagenesis

Site-directed mutagenesis, based on the *in vitro* amplification of DNA by polymerase chain reaction (PCR) method (Sambrook and Russell, 2001), was set up in a final reaction volume of 50µl containing; template DNA (50ng), mutagenic primers (0.5mM) (Table 2.1), deoxynucleotides (dNTPs) (0.2mM), Pfu polymerase (1.5U), and preferred Pfu reaction buffer. PCR works by initial denaturation of the template DNA and then cycles through primer annealing and elongation steps, see Table 2.2 for details. The methylated template DNA was then digested using a site-specific restriction endonuclease; DpnI; a 37°C incubation of 1 hour with 20U of enzyme was sufficient to digest template DNA, the digest was stopped by heat inactivation at 80°C for 20mins.

Table 2.1 Mutagenic primers used in the generation of ABCG2 constructs. Amino acid changes are highlighted in yellow in the forward primer, and lower case bases are identical to the template.

Mutation	Primers 5' to 3'
T402A	cagatcattgtcGcagtcgtactgggac
L405A	gtcacagtcgtaGCgggactggttatag
S440A	ccaaccagtgtttcGCCagtgtttcagcc
S443A	cagcagtgttGcagccgtggaactc
S519A	gatggtggcttatGcTgccagttccatg
M523A	tcagccagttccGCAgcaactggccatag
L540A	gtagcaacacttGCCatgaccatctgt
M541A	gcaacacttctcGCCaccatgtg
I543A	cttctcatgaccGCCctgttttggtttatg
F545A	catgaccatctgtGCCgtgtttatgat
M548A	ctgttttggtttGCCatgatttttcagg
F571A	cttcagtacGCCagcattcca
L633A	gaatcacgtggccGCGgcttgatgattg
M636A	gccttggtgtGCGattgttattttcc
F640A	gtatgattgttattGCCctcacaattgcctac
I643A	gttattttcctcacaGCTgcctacctgaaattg

All reverse primers are the exact reverse complement.

Table 2.2 Polymerase chain reaction (PCR) cycling parameters for site-directed mutagenesis.

PCR step	Temperature (°C)	Time (mins)
Step 1 Denaturation	95	1
	95	1
Step 2 Annealing	55	1
Step 3 Extension	72	12
	72	10
Step 4 End	10	

The PCR program was cycled 18 times for the steps highlighted in yellow.

2.1.3 Transformation

Newly synthesised DNA was transformed into DH5 α competent cells. 100 μ l aliquots of DH5 α competent cells were thawed on ice. 5 μ l of the PCR amplicon was added to the competent cells. The competent cell mixture was then incubated on ice for 30 minutes, after which the cells were heat shocked for 1.5 minutes at 42°C and then incubated for a further 2 minutes on ice. The cell mixture was then supplemented with 900 μ l of Luria-Bertoni (LB) medium (1% (w/v) NaCl, 1% (w/v) tryptone and 0.05% (w/v) yeast extract) and incubated for 60 minutes at 37°C. Subsequently 10% and 90% of the transformation mixture was plated onto LB-agar (1.5% (w/v) agar) antibiotic selective plates (ampicillin 100 μ g/ml) and finally incubated overnight at 37°C.

2.1.4 Plasmid DNA purification

Single transformant colonies from LB-agar selective plates were inoculated into 5ml of LB media containing appropriate antibiotics. The bacterial cultures were grown overnight at 37°C, with orbital shaking (200rpm). For long-term storage, glycerol stocks of the resultant bacterial cultures were generated, by taking 500 μ l of the culture and adding 500 μ l of 30% (v/v) glycerol, the resultant glycerol stock was stored at -80°C. The remainder of the cultures were centrifuged at 4,000xg for 5 minutes and the supernatant discarded. The plasmid DNA was then purified from the bacterial pellet using a NucleoSpin[®] Plasmid Miniprep Kit (MACHEREY-NAGEL) according to the manufacturer's protocol.

The concentration and purity of the plasmid DNA were determined using a Nanodrop 2000® (Thermo Fisher Scientific™). The purity was confirmed by determination of the A_{260}/A_{280} ratio, where a ratio of between 1.70 and 1.90 was acceptable.

2.1.5 Restriction Digest

Restriction digest was used as a size based confirmative method of plasmid DNA identity. Restriction digests were set up in a reaction volume of 15µl and comprised of plasmid DNA (100-500ng) and restriction endonuclease (5-10U) with the preferred enzyme buffer. The reaction mixtures were incubated at 37°C for 90 minutes, then mixed with DNA loading dye (0.25% (w/v) bromophenol blue, 40% (w/v) sucrose) and resolved on a 1% (w/v) agarose gel containing 1µg/ml ethidium bromide (Thermo Fisher Scientific). Gels were prepared and resolved in TPE buffer (90mM Tris-phosphate, 1mM EDTA, pH8) and visualised under long wavelength ultraviolet (UV) light (>300nm).

2.1.6 DNA sequencing

Once restriction digest confirmed the size of the plasmid DNA the mutants were Sanger sequenced. Initially across the site of mutation and once incorporation of the mutation was verified, the entirety of ABCG2 was sequenced to ensure that the construct did not carry any further undesired mutations (sequencing primers shown in Table 2.3). The sequencing data was analysed against the predictive sequences using NCBI Blast, multiple sequence alignment tool. The sequence chromatogram was also assessed for trace quality using Chromas ([Technelysium Pty Ltd](#)).

Table 2.3 Sequencing primers with a schematic representation of their binding location.

Primers	Sequences (5' → 3')
SeqR1	TCGTGGTGCTCCATTTAT
SeqF0	GAGTGGCTTTCTACCTTGTC
SeqF2	GCAGGGACGAACAATC
SeqF482	AACTCTTTGTGGTAGA
GFP-rev	GCTGTTGTAGTTGTACTCCAG



2.2 Tissue culture

2.2.1 Materials and reagents

HEK293T (CRC-3216), a human cell line, derived from the human embryonic kidney 293 cell line was used for the expression of mammalian protein. The cell line expresses a mutant version of the simian vacuolating virus (SV40) large T antigen, allowing for transfections with high copy number of plasmids containing an SV40 origin of replication.

All tissue culture reagents were obtained from Sigma-Aldrich, except for linear transfection reagent polyethyleneimine (PEI, Polyscience Inc.) and Zeocin™ (Invitrogen). All plastic culture-ware was manufactured by Thermo Fisher Scientific unless stated otherwise.

2.2.2 Cell passage maintenance

HEK293T cells were maintained in T25(cm²) flasks at 37°C, 5% CO₂ in Dulbecco's Modified Eagle's Medium (DMEM, 4500mg/L-glucose, L-glutamine, sodium pyruvate and sodium bicarbonate) supplemented with 10% (v/v) foetal calf serum (non-heat-treated) (FCS; Sigma), and 100U/ml penicillin and 100µg/ml streptomycin. Once the cells reached 80-90% confluency, the media was removed, and the cells washed once with 2ml of sterile phosphate buffered saline (PBS). The cells were then detached from the plastic-ware by incubation at 37°C with 0.35ml of 1x trypsin/EDTA for 1-5 minutes. Cells were then resuspended in DMEM by repeated pipetting and pelleted at 500xg for 5 minutes to remove trypsin. The

pelleted cells were resuspended in 5ml of media before being re-seeded typically at a dilution of 1:10 of the original culture.

2.2.3 Transfection

Cells were seeded at 1.25×10^5 cells/ml (determined using a haemocytometer) into a 6-well plate 24 hours prior to transfection. Two hours prior to transfection the media was replaced with 5% (v/v) FCS supplemented medium. Cells were transfected using linear polyethyleneimine (PEI) (10mM working stock solution dissolved in PBS) (first described by Boussif et al., 1995) at a molar PEI nitrogen: DNA phosphorous ratio of 15:1. DNA was typically used at 2 μ g. Complexes of PEI/DNA were made by the addition of PEI (9 μ l) to DNA (2 μ g) and incubated for 15mins. 250 μ l of 5% FCS DMEM was then added to the preformed PEI/DNA complex, which was subsequently added drop wise to the cells. Twenty-four hours post transfection the media was replaced by 10% (v/v) FCS supplemented medium.

2.2.4 Zeocin selection of stable cell lines

Transiently transfected HEK293T cells were detached (trypsin/EDTA) and transferred to T25(cm²) flasks containing 5ml of DMEM supplemented with 10% (v/v) FCS. Following an overnight incubation (37°C, 5% CO₂) the cells received fresh media containing 200 μ g/ml ZeocinTM (Invitrogen). Cells were transferred to T75(cm²) flasks once confluency was reached and the Zeocin concentration maintained at 200 μ g/ml for ten to fifteen days, with media replacement every two to three days. Once healthy stably transfected cells were observed the concentration of Zeocin was reduced to 40 μ g/ml to maintain the selection. For long

term storage, stable cell lines were stored in 90% (v/v) FCS, 10% (v/v) DMSO in liquid nitrogen at 1×10^6 cells/ml.

2.3 SDS-PAGE gel electrophoresis and western blotting

2.3.1 Cell harvesting

A confluent T25(cm^2) flask yielded an adequate number of cells to be able to assess protein expression by SDS-PAGE gel electrophoresis and western blotting. Cells were washed once with 2ml of sterile phosphate buffered saline (PBS) and then incubated at 37°C with 0.35ml of 1x trypsin/EDTA for 1-5 minutes to detach the cells from the growing surface. Cells were washed a further time with ice-cold PBS via centrifugation and then resuspended in 200-500 μl of ice-cold PBS, 10% (v/v) glycerol containing EDTA-free protease inhibitor cocktail set III (Calbiochem), to prevent proteolysis. Cells were lysed by 2x10 second burst sonication at 50% output (Microsonics Instruments).

2.3.2 Protein concentration determination assay

A modified Lowry protein assay (conducted according to the manufacturer's protocol, Bio-Rad) was used to determine the protein concentration in the cell lysates, with a standard curve containing 0-10 μg bovine serum albumin (BSA) constructed in parallel. All samples were analysed in duplicate.

2.3.3 SDS-polyacrylamide gel electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Laemmli (1970). Equal quantities of protein lysate (15-

30µg) samples were first incubated with protein loading buffer (50mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 100mM 2-mercaptoethanol) at 37°C for 30-60mins, samples were then electrophoretically resolved on discontinuous 8-12% (w/v) polyacrylamide gels alongside molecular weight protein markers (SeeBlue® Plus2 Prestained Standard (Novex)), at 30mA/ gel until the protein dye-front fully resolved.

2.3.4 N-linked glycanase digestion

Cell lysates (15µg protein) were combined with 10x glycoprotein denaturing buffer (5% (w/v) SDS, 0.4M DTT), and the total reaction volume was made to 10µl with dH₂O. The samples were denatured by heating at 100°C for 10 minutes. Following denaturation, the reaction mixture was made to 20µl with the addition of 10x G7 reaction buffer (0.5M sodium phosphate (pH7.5 @ 25°C)), 10% NP40 and 1-2µl of Peptide: N-glycosidase F (PNGaseF) (BioLabs). The resultant reaction mixture was incubated at 37°C for 1 hour, prior to loading onto a 10% polyacrylamide gel.

2.3.5 Western blotting

Following SDS-PAGE electrophoresis, proteins were transferred (200mA for 2 hours, 4°C) onto a nitrocellulose membrane (Amersham) by electrophoresis, first described by Towbin et al., 1979. Membranes were then transiently stained with Ponceau S solution to confirm efficient transfer, washed for 2-3 minutes with PBS-Tween (PBS-T, 0.01% (v/v) Tween-20), and then incubated in blocking buffer (5% (w/v) skimmed milk powder in PBS-T) for 1 hour to prevent non-specific binding of antibody.

Membranes were then incubated with primary antibody, BXP-21 (Calbiochem), at 1:1000 (prepared in blocking buffer) for either 1 hour at room temperature or, preferentially overnight at 4°C. The nitrocellulose membranes were washed several times in PBS-T, 3-5 minutes each wash, to remove any unbound primary antibody before incubation with a secondary antibody, rabbit anti-mouse horseradish peroxidase (DAKO), at 1:2000 (in PBS-T) for 1 hour at room temperature. The membranes were then washed several further times in PBS-T (3-5 minutes per wash) to remove unbound secondary antibody. Following the final wash, the membranes were incubated with enhanced chemiluminescence reagent, ECL (Supersignal West Pico, Thermo Scientific) and developed using light sensitive film.

2.4 Confocal microscopy

2.4.1 Materials and reagents

35mm² glass bottomed dishes were purchased from MatTek Corp[®]. All other reagents were purchased from Sigma-Aldrich unless stated otherwise.

2.4.2 Poly-L-lysine coating

To ensure the efficient attachment of HEK293T cells to any untreated culture-ware, the culture dishes were coated prior to seeding with poly-L-lysine (0.1mg/ml) for at least 30 minutes. Excess poly-L-lysine was removed by twice washing with sterile PBS before cell seeding.

2.4.3 Cell seeding

Cells were seeded at 1.25×10^5 cells/ml, determined using a haemocytometer, in poly-L-lysine coated 35mm² glass bottomed dishes (MatTek Corp[®]), and incubated at 37°C, 5% CO₂ for 24 hours. Following incubation, cells were washed twice with pre-warmed (37°C) phenol-red free HBSS (Hank's Balanced Salt Solution).

2.4.4 Confocal Image Acquisition

Visualisation of cells was conducted on an LSM710 confocal laser scanning microscope (Zeiss, Jena, Germany) equipped with a Plan-Apochromat 63x/1.40 Oil Ph3 DIC M27 objective and argon laser. To be able to determine the localisation of the protein within the cell, an oil objective equipped with a phase contrast was used as opposed to the optimal x40 water objective which lacks the phase contrast. For detection of sfGFP-ABCG2, a 2% argon laser power was used with the pinhole set to 1AU (area unit) and the excitation set at 488nm and emission collected between 500-530nm (FITC channel). To be able to determine localisation of sfGFP-ABCG2 across a mixed-expressing population of cells, high resolution fluorescent channel images with the corresponding phase contrast (1024 x 1024 pixels, image depth of 12 bit) were acquired in at least three different viewing positions for any given culture dish, on three separate acquisition sessions.

2.5 Flow cytometry

2.5.1 Materials and reagents

All drugs, solvents and inhibitors were obtained from Sigma-Aldrich.

2.5.2 Fluorescence based drug accumulation assay

The flow cytometric based drug accumulation assay was based on the assay described by Robey et al. (2001a). Cells were seeded at 1×10^6 cells/ml in phenol-red free DMEM and incubated with either DMSO (solvent control 0.2% (v/v)) or mitoxantrone (10 μ M), pheophorbide A (10 μ M), or daunorubicin (10 μ M), in the presence or absence of the ABCG2 inhibitor, Ko143 (1 μ M, maximal ABCG2 specific concentration (Ambudkar et al., 2015)) at 37°C for 30 minutes (preincubation step) with occasional agitation. Cells were centrifuged to remove excess drug at 350xg, 4°C for 5 minutes prior to a second incubation at 37°C for 60 minutes, with either phenol-red free DMEM only or phenol-red free DMEM plus Ko143 (for samples originally incubated with drug and inhibitor). Cells were centrifuged a further time (350xg, 5 minutes, 4°C) and resuspended in phenol-red free DMEM prior to analysis by flow cytometry.

A preincubation of 30mins has previously been demonstrated to be sufficient to reach steady state kinetics (i.e. equilibrium of drug uptake vs. efflux, represented by consistent levels of MX fluorescence in cells) (Wong, 2015).

2.5.3 Flow cytometry acquisition

Data was collected using either an FC500 flow cytometer or a MoFlo Astrios flow cytometer (Beckman Coulter), depending on the fluorescence parameters of drug substrate used in the assay and the detection capabilities of the instrument (see Table 2.4 for more details). GFP fluorescence representative of protein expression was collected simultaneously to drug fluorescence and was determined using an excitation wavelength of 488nm and an emission wavelength of 526nm.

Table 2.4 Drug fluorescent properties.

	Excitation	Emission	Flow cytometer channel
Mitoxantrone	635nm	670nm	FL4: 633/75 (FC500) 488-710/45 (Astrios)
Pheophorbide A	400nm	710nm	355-692/75 (Astrios)
Daunorubicin	490nm	630nm	488-710/45 (Astrios)

Flow cytometer channel settings were applied based on highest fluorescent signal output. A minimum of 10,000 events (cells) collected for each condition after gating.

2.5.4 Flow cytometry data analysis

Data were analysed using Kaluza analysis version 1.5 (Beckman Coulter). Data was collected for each mutant variant along with wild type and HEK (untransfected) cell lines for three difference conditions: DMSO (vehicle control), drug only, and drug plus inhibitor (Ko143). Both GFP fluorescence (FL1 (FC500) or 488-513/26 (Astrios)) and drug fluorescence (see parameters in Table 2.4) were collected. Dot plots of forward scatter (FS, particle size) versus side scatter (SS, particle granularity) allowed for visualisation of particle size and granularity within a sample, allowing for gating to exclude debris from the cellular population (Figure 2.1 A). A subsequent dot plot of GFP fluorescence (x axis) versus FS (y axis) (FC500) or FS area (x) versus height (y) (Astrios) was plotted and gated for cellular population in Figure 2.1 A. to detect disproportions between cell size and signal, gating was applied to discriminate outliers' representative of cellular aggregates. GFP fluorescence histograms (representative of protein expression (sfGFP-ABCG2) gated for the monodispersed cellular population were used to identify expression characteristics. Due to the nature of a heterogeneous stable cell line, some GFP (protein expression) profiles demonstrated a split in their distribution which were

denoted low and high (high expression population comparable to wild type expression) (e.g. S519A Figure 2.1 B). Dot plots of GFP fluorescence versus drug fluorescence (DMSO, drug only, and drug plus Ko143) (gated for monodispersity) were plotted, drug fluorescence is proportional to drug accumulation within cells. The low GFP fluorescence population demonstrates a consistently high level of drug fluorescence across drug only and drug plus Ko143 (ABCG2 specific inhibitor) (Figure 2.1 C/D, fuchsia). Whereas, the drug fluorescence in the high GFP fluorescence population for the drug only sample is reduced (Figure 2.1 C, teal) compared to the drug plus Ko143 sample (Figure 2.1 C, grey), this difference in drug fluorescence is representative of Ko143 (ABCG2 specific) inhibitable efflux of drug substrate and therefore a measure of ABCG2 drug efflux. As the low GFP fluorescence (protein expression) population showed no function, the high GFP fluorescence population, comparable to wild type GFP fluorescence, was used in subsequent analysis. Overlay histograms of drug fluorescence (gated for high GFP fluorescence) for DMSO, drug and drug plus Ko143 were plotted and the median fluorescence extracted. The median drug fluorescence value for the drug only and drug plus Ko143 samples were baseline corrected for DMSO 'drug' (background) fluorescence (e.g. drug only – DMSO). The relative efflux, the ratio of median drug fluorescence defined as: $(\text{drug plus Ko143} - \text{drug only}) / \text{drug only}$ was calculated as a measure of efflux (Figure 2.1 E and Figure 3.15 D). As the data shows some variation in the median GFP fluorescence (protein expression) of the mutant variant compared to the wild type the ratio between wild type expression and mutant

expression was calculated (i.e. WT/mutant) and the value obtained for normalised relative efflux (i.e. relative efflux * ratio (expression)) (Figure 3.15 E).

The expression gating and normalisation ensured that data represented the function of the mutant protein and not their expression level (implications of excluding protein expression (GFP fluorescence) analysis discussed in Chapter 3).

The values obtained for WT expression normalised relative efflux were analysed using Prism 7 (GraphPad), a one-way ANOVA was conducted with a Dunnett's multiple comparisons post test against the wild type protein to assess if drug efflux was significantly perturbed in any mutant.

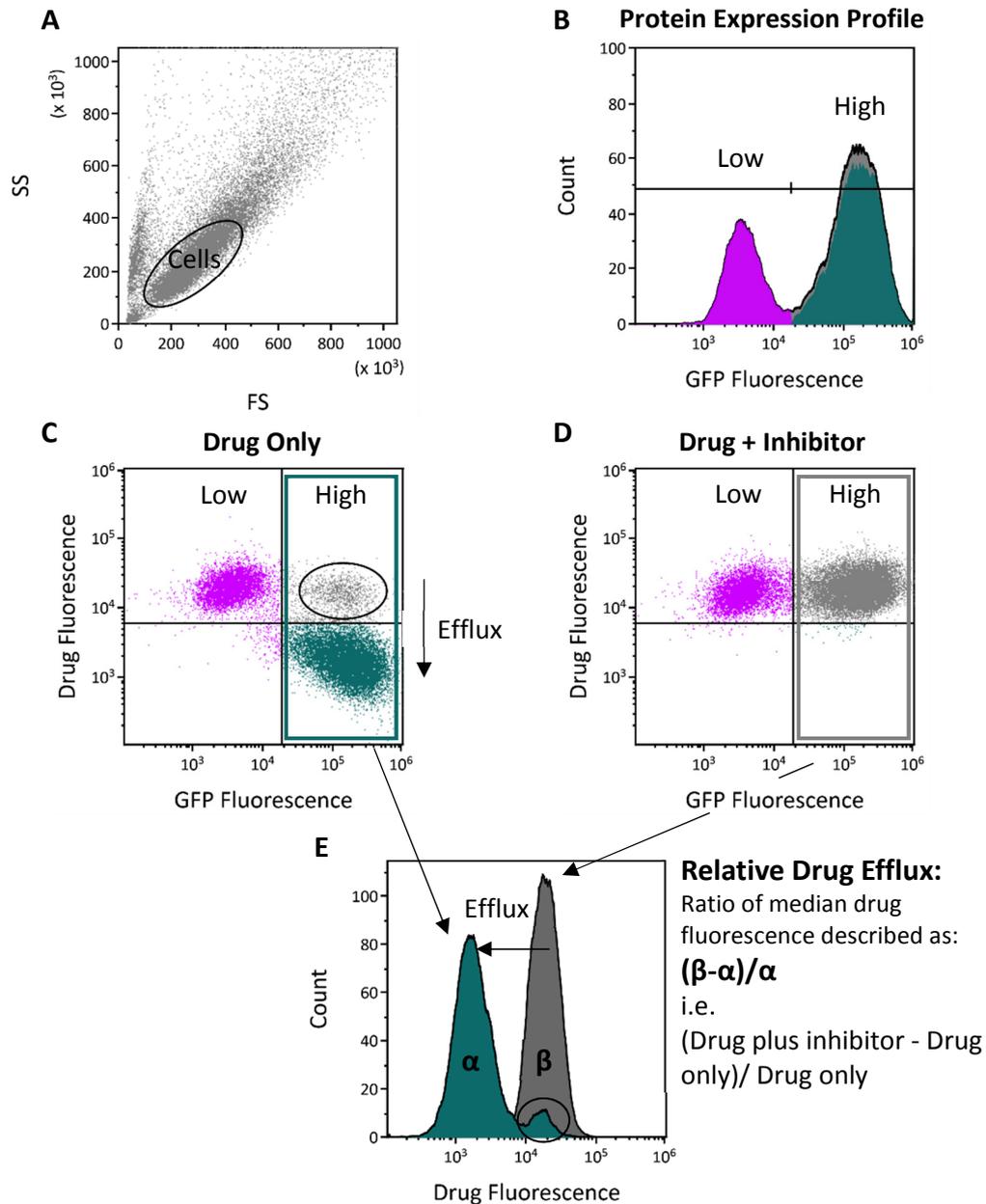


Figure 2.1 Flow cytometric data analysis of an ABCG2 variant. **A.** Plot of forward scatter vs. side scatter, representing particle size vs. granularity. Ellipse shows gating to discriminate intact cells from cellular debris. **B.** GFP fluorescence histogram representing GFP-tagged ABCG2 protein expression. Mutant protein S519A shows a phenotypic split in protein expression denoted low (fuchsia) and high (teal, comparable to WT). **C/D.** Dot plots of GFP fluorescence vs. drug fluorescence. C demonstrates reduced drug fluorescence (cellular accumulation) for the high expressing population (teal, lower right quadrant), this increased upon the addition of Ko143 (G2 specific inhibitor) (D, grey, upper right quadrant). This is representative of ABCG2 specific drug efflux. The lower expressing population in C/D shows no efflux (fuchsia, upper left quadrant) so was excluded from further analysis. **E.** Overlay of drug fluorescence (gated for high expression) for drug only (teal, α) and drug plus inhibitor (Ko143) (grey, β). Relative drug efflux is a ratio of the median drug fluorescence defined as: $(\beta - \alpha) / \alpha$. Cells circled (C/E) shows a negligible proportion of high expressing cells that appear to be functionally impaired. Data is representative of one of three repeat experiments for an ABCG2 variant S519A.

2.6 Molecular visualisation and drug docking

All molecular docking studies were performed using a model for ABCG2 based upon the structure of the homologous ABCG5/G8 (Lee et al., 2016), kindly provided by Dr. Thomas Stockner (Khunweeraphong et al., 2017). To enable the docking, coordinates, parameters for mitoxantrone, pheophorbide A, and daunorubicin were obtained from ChemSpider and both the protein and drug were converted to PDBQT (Vina executable) files, via AutoDockTools (v1.5.6). A grid box centred on the approximate geometric midpoint of the TMD of an ABCG2 monomer (residues 517/518) with dimensions 36x20x20 (Å) (Figure 2.2) was assigned at this stage as the search region for AutoDock Vina (Trott and Olson, 2009) molecular docking software. The exhaustiveness of the search was manually set to 128 (repeats of docking calculations, increases the chance of finding the global minimum of the docked drug, 128 used by László et al., 2016), all other parameters being the default, to find the most energetically favourable pose. Binding energy outputs for each drug were between 5.5 and 7 kcal/mol which is within the range seen by László et al., 2016. Visual inspection of the PDBQT output files was accomplished using PyMOL© (Schrodinger, LLC).

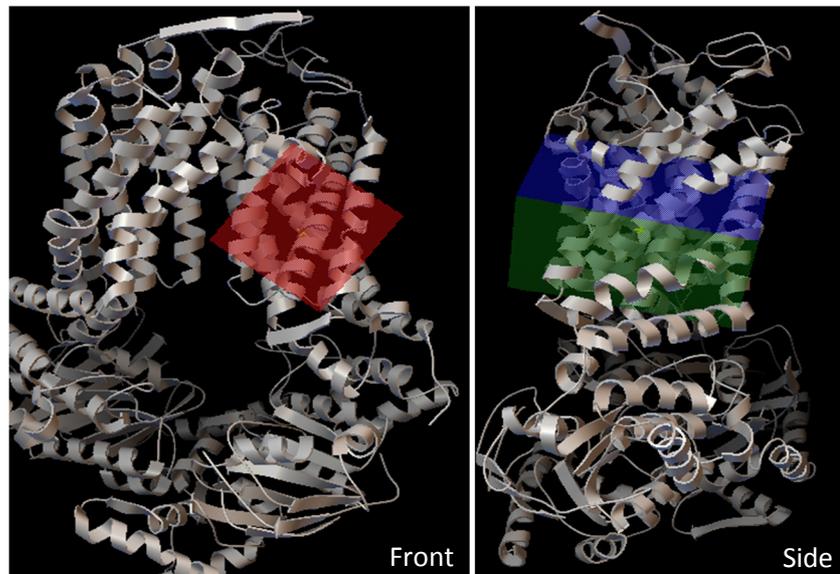


Figure 2.2 ABCG2 homology model with grid box assigned. ABCG2 grid box was assigned using python molecular viewer, the grid box was centred on the estimated geometric midpoint of the lateral hypothesis residue and enclosed the majority of the TMD. The figure is representative of both front and a 90° rotation, side view.

2.7 Membrane based assays

2.7.1 Sodium butyrate analysis

Confocal analysis

Cells were seeded at $0.5-1 \times 10^5$ cells/ml (determined using a haemocytometer), into a poly-L-lysine coated (section 2.4.2) black-walled clear flat-bottomed 96-well plate (Greiner) in supplemented (10% (v/v) FCS and 100U/ml penicillin and 100 μ g/ml streptomycin) DMEM, and incubated at 37°C, 5% CO₂ for 24 hours. Following the 24-hour incubation cells were washed once with sterile PBS and the media replaced with phenol-red free DMEM supplemented with increasing concentrations of sodium butyrate (transcriptional activator) (0, 1, 2, 3, 5, and 10mM). Cells were visualised at 0 hours and 24 hours post sodium butyrate addition on an ImageXpress (IX) Ultra confocal plate reader (Molecular Devices), equipped with a plan-apochromat 40x objective, with an excitation wavelength of 488nm and an

emission bandwidth filter of 520/50nm (GFP Fluorescence), image acquisition parameters (i.e. gain) were set based on wild type expression at 0 hours post incubation with sodium butyrate and kept consistent across all data collection.

Flow cytometry analysis

Cells were seeded at 1.25×10^5 cells/ml (determined using a haemocytometer), into a 6-well plate (Greiner) in DMEM, and incubated at 37°C, 5% CO₂ for 24 hours. Following incubation cells were washed once with sterile PBS and the media replaced with DMEM supplemented with increasing concentrations of sodium butyrate (0, 1, 2.5, 5, and 10mM). following a further 24-hour incubation (37°C, 5% CO₂) cells were washed with sterile PBS and then detached from the plastic-ware by incubation at 37°C with 0.35ml of 1x trypsin/EDTA for 1-5 minutes. Cells were then resuspended in phenol-red free DMEM by repeated pipetting and pelleted at 500xg for 5 minutes to remove trypsin. Cells were finally resuspended in 0.5-1ml iced-cold phenol red free DMEM and analysed for GFP fluorescence (fluorescence channel: FL1) using an FC500 flow cytometer (Beckman Coulter). Cells were gated as described in section 2.5.4 and the median GFP fluorescence extracted for each sodium butyrate concentration. A percentage increase in protein expression between the lowest and highest concentration of sodium butyrate was calculated as follows:

$$\frac{(MGF_{L10mM} - MGF_{L0mM})}{MGF_{L0mM}} \times 100$$

MGF_L: Median GFP Fluorescence

2.7.2 Tissue culture

For larger-scale isolation of membranes approx. 8.5×10^6 cells were seeded into 45cm^2 culture plates. For ABCG2 expressing cell lines the growth medium (DMEM, 10% FCS) was supplemented with $40\mu\text{g/ml}$ ZeocinTM. The culture plates were incubated at 37°C , 5% CO_2 for 36 hours before the addition of 10mM sodium butyrate (transcriptional activator). The cells were harvested 24 hours later by detachment of the cells via scraping into sterile PBS. The cell suspension was washed by centrifugation (350xg , 5 minutes, 4°C) with sterile PBS and the cell pellet either stored at -80°C or resuspended 5x the cell pellet weight (g) (approximately 1g from one 45cm^2 plate) in membrane isolation buffer I (10mM Tris, 250mM sucrose, 0.2mM CaCl_2 , pH7.4) with the addition of protease inhibitor (EDTA-free protease inhibitor cocktail set III, Calbiochem).

2.7.3 Membrane preparation

Cells were disrupted by nitrogen cavitation (2x 10 minutes at 1000Psi) in a Parr Instruments 25ml capacity pressure cell. Cellular debris was pelleted by centrifugation at 1500xg for 15 minutes at 4°C . The resultant supernatant was subjected to ultracentrifugation at $100,000\text{xg}$ for 1 hour at 4°C to isolate a total membrane fraction. The membrane pellet (glassy in nature) was resuspended at 100mg (wet weight) per millilitre in modified membrane isolation buffer IV (50mM Tris, 250mM sucrose, pH8) and syringe homogenised by at least 20 passes through a 25-gauge needle. The membranes were used immediately or stored in single use aliquots at -80°C .

2.7.4 Membrane and protein stability analysis

Chemical and thermal stability of ABCG2 was assessed using NanoTemper Technology's differential scanning fluorimetry instrument, Prometheus NT.48. Determination of the most stabilising buffer condition was monitored by measuring the shift of intrinsic tryptophan fluorescence at emission wavelengths of 330nm and 350nm upon protein thermal un-folding. Membranes overexpressing sfGFP-ABCG2 were subjected to a melting curve of 20-95°C in 1°C/ minute intervals. The integrated fluorescence at 350/330nm was plotted against temperature to obtain the melting temperature (T_m) of the protein for each buffer condition.

2.7.5 Vesicular transport assay

A vesicular transport assay (VTA) was used to determine ATP dependent ABCG2 specific drug uptake into membrane vesicles. Membrane vesicles (20µg membrane protein) were combined with either DMSO (solvent control, 0.01% v/v), or 100 µM lucifer yellow in the presence or absence of the ABCG2 inhibitor, Ko143 (1µM). Each test condition was supplemented with 5mM of MgSO₄ only, or with the addition (5mM) of either AMP or ATP, the reaction volume was made to 10µl with TSN buffer (50mM Tris, 250mM Sucrose, 150mM NaCl, pH8) prior to a 10 minute incubation in a 37°C water bath. PCR tubes were used for efficient heat transfer. Following the incubation, samples were quickly transferred to ultracentrifuge tubes containing 190µl of ice-cold TSN buffer and the free drug and nucleotide separated from membrane vesicles by ultracentrifugation at 60,000xg for 5 minutes at 4°C. The resultant membrane pellet was resuspended in 200µl of TSN buffer and transferred to a 96-well black-out plate (Greiner) prior to analysis by fluorescence

spectroscopy. Lucifer yellow excitation wavelength was set at 430nm and emission collected at 538nm using a SpectraMax® M2 microplate reader (Molecular Devices).

Data was analysed using Prism 7 (GraphPad), where relative drug accumulation within vesicles was defined as the ratio between ATP (stimulated activity) and AMP (control) for each experimental condition. The solvent only sample (DMSO) represented background fluorescence.

2.7.6 ATPase assay

ATPase activity was measured by the release of free orthophosphate (Pi) in a colorimetric assay, originally described by Chifflet et al., 1988.

Membranes (20µg), were combined with 2X assay buffer (100mM KCl, 5mM EGTA, 2.6mM ouabain, 4mM DTT), 5mM MgSO₄ and either DMSO (solvent control, 0.01% v/v) or 100µM lucifer yellow in the presence or absence of 1µM Ko143. The final reaction volume was made up to 10µl with TSN buffer (allowing for the additional volume of AMP or ATP). Each test condition was supplemented with either AMP or ATP, added at regular time intervals (typically 30 seconds), and incubated in a 37°C water bath for 20-30 minutes. A 5% SDS solution (10µl) was used to stop the reaction. Each reaction was then transferred to an individual well of a 96-well plate already containing 50µl/ well of dH₂O, followed by 40µl of phosphate detection reagent (1% (w/v) ammonium molybdate, 0.014% (w/v) antimony potassium tartrate, 2.5N sulphuric acid) and 20µl of 1% (w/v) ascorbic acid. The plate was then incubated for 7-12 minutes at room temperature to allow the colour to develop and absorption read at 880nm using a SpectraMax® M2 microplate reader (Molecular

Devices). Longer incubation resulted in reduced sensitivity of the phosphate-detection reagent. Relative Pi release was defined as the absorbance at 880nm ($A_{880\text{nm}}$) for experimental condition in the presence of ATP minus the background absorbency in the AMP control (e.g. $A_{880\text{nm G2 LY ATP}} - A_{880\text{nm G2 LY AMP}}$).

Chapter 3 Generation, expression and functional characterisation of ABCG2 'lateral slice' mutants

3.1 Development of the 'lateral slice' hypothesis

ABCG2 (also known as BCRP) is an efflux ABC transporter that is recognised to export a wide range of substances. ABCG2 is widely distributed in humans, predominantly residing within tissues which have protective roles, such as the blood brain barrier (Allen et al., 2002; Zhang et al., 2003). ABCG2 is known for its role in drug resistance, particularly to chemotherapeutic agents (Mao and Unadkat, 2015). Despite this significant role in tissue exposure to drugs and xenobiotics, exactly how ABCG2 binds and transports a large number of structurally unrelated drug substrates remains unknown.

A significant body of evidence has been accumulated by pharmacological or mutagenesis studies of ABCG2 with the aim of identifying drug binding sites (Cai et al., 2010; Clark et al., 2006; Haider et al., 2015; Polgar et al., 2010). Arginine 482 to glycine/threonine (TMH3) is a classic example of a mutation that does not alter protein expression levels but has a profound effect on overall substrate specificity and/or transport activity. R482G/T shows a gain in the ability to efflux substrates such as daunorubicin, rhodamine 123, and lyso-tracker green which are not transported by the wild type protein (Honjo et al., 2001; Robey et al., 2003). The mutant protein is also able to transport most native ABCG2 substrates, with the exception of methotrexate (Chen et al., 2003). The R482G/T mutation and its

behaviour suggests the residue's importance in substrate recognition and/or transport activity (Clark et al., 2006).

A later study investigated a series of proline residues conserved across ABCG2 homologues in eukaryotes. Proline residues often form flexible hinge regions, particularly in α -helices, so are thought to have a fundamental role in dynamic conformational changes. Alone amongst a series of single proline to alanine mutations, P485A demonstrated reduced efflux for BODIPY-prazosin by up to 70% but had no effect on the efflux of mitoxantrone or Hoechst 33342 (Haider et al., 2015; Ni et al., 2011). Due to the spatial positioning of residue P485 (one turn of α -helix C-terminal to Arg482) it was suggested that P485 could introduce structural flexibility in a drug binding pocket (Ni et al., 2011).

Previous research on other ABC MDR transporters has argued that drug recognition requires involvement from multiple helices (Deeley and Cole, 2006; Loo and Clarke, 2000). In the ABCG2 topology model as described by Wang et al. 2008, T402 in TMH1 is in an approximate equal topological position with residue R482 (TMH3) towards the cytoplasmic interface of the transmembrane domains; when threonine 402 was mutated to alanine it also demonstrated differential drug efflux patterns (Ni et al., 2010a; Polgar et al., 2010). Therefore, this enabled me to hypothesise that R482 and P485 are a pair of residues that contribute to a drug substrate binding pocket, and that residues in topologically equivalent positions in transmembrane helix 1, 2, 4, 5, and 6 may also be incorporated into this binding pocket. Therefore, in this study a pair of residues from TM 1,2,4,5 and 6 were identified and for

feasibility of down-stream analysis, a total of 10 single amino acid mutations were designed (Figure 3.1), including the previously published T402A mutation.

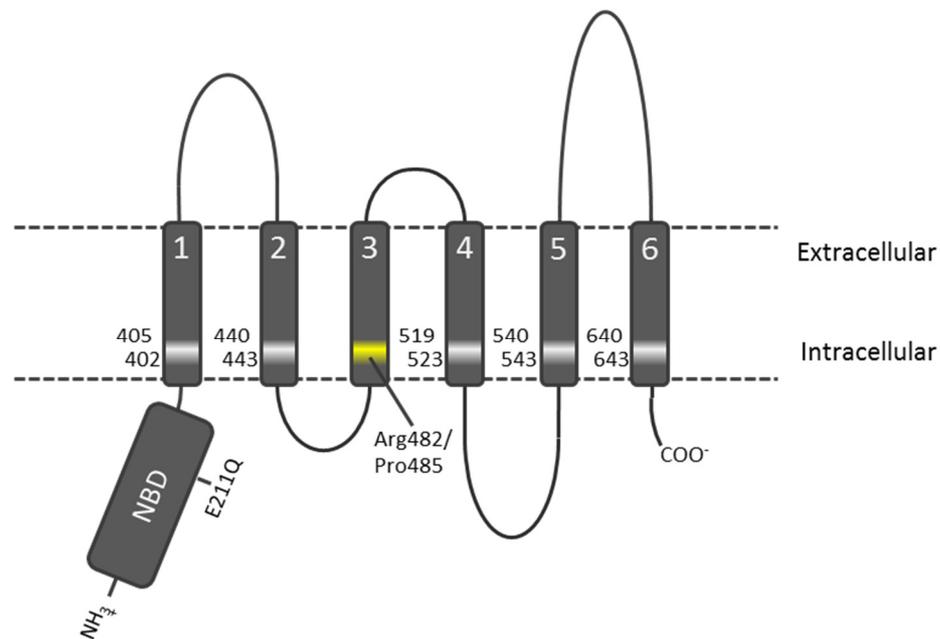


Figure 3.1 The lateral slice hypothesis mutations. The previously characterised residues arginine 482 and proline 485, highlighted in yellow, are located approximately one helical turn apart, towards the intracellular interface. The 10 residues selected for investigation now were based on the equivalent topological arrangement to R482 and P485, from the remaining transmembrane helices (highlighted in white). Schematic based upon the topology model of Wang et al 2008 as this was the only available experimental topology at the start of this study.

This chapter will proceed to investigate the effect of single amino acid mutants on the protein's expression, localisation and capacity to efflux drug substrate. Several *in vitro* techniques were available to investigate transporter variant function, including assay systems employing isolated membranes and intact cells. Membrane-based assay systems include ATPase assays and vesicular transport assays. Baculovirus-infected insect cells (Sf9, *Spodoptera frugiperda*) have often been used for membrane-based assays as they express higher yields of protein expression than mammalian cells (HEK293T) (Murphy et al., 2004). The downfall of a heterologous expression system such as insect cells, is that protein post

translational modification may not recapitulate the native modifications (Murphy et al., 2004). ABCG2 has a glycosylation site in the third extracellular loop which is not glycosylated in insect cell expression (Clark et al., 2006); however, studies have shown that glycosylation is not required for localisation to the plasma membrane or function of the protein (Diop and Hrycyna, 2005). Moreover, insect cells also have reduced plasma membrane cholesterol content; cholesterol has long been associated with potential regulatory effects of ABCG2 and therefore protein isolated from insect cells demonstrate a requirement for cholesterol supplementation (Pál et al., 2007; Telbisz et al., 2014, 2007).

Homologous expression of ABCG2 in a mammalian cell line (HEK293T), has demonstrated protein to be expressed at full length, to be localised to the plasma membrane and to function as an allocrite (transport substrate) exporter (Robey et al. 2003; Haider et al. 2015). The drawback of this system is that stable protein expression requires a time-consuming and costly selection process and the yields of the protein are still reduced compared to a viral delivery method (Figure 3.2). Even so, homologous expression is favourable and therefore a mammalian expression system was preferable to other available expression systems and used within this study (Figure 3.2).

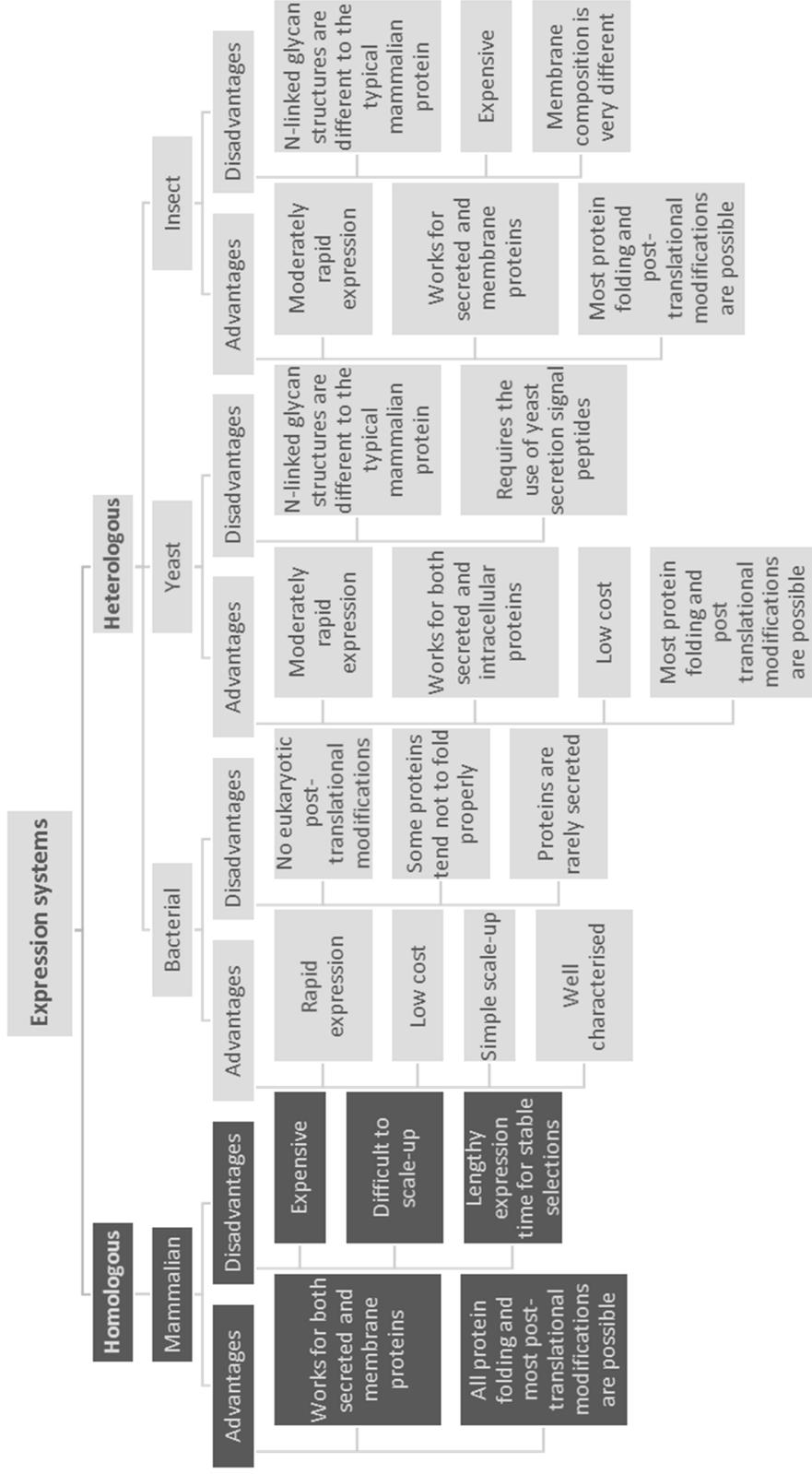
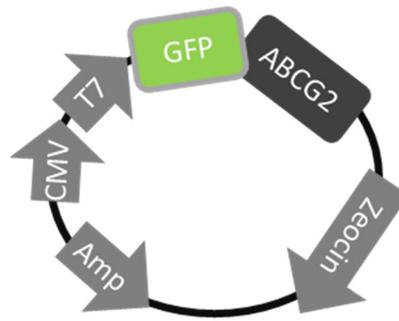


Figure 3.2 An overview of systems for the expression of recombinant proteins. Heterologous expression systems for ABCG2 include bacteria, yeast and insect cells. All three of these systems have the advantage of rapid to moderate expression of protein, with the main disadvantage being impaired post translation modifications. Homologous expression is within mammalian cell lines, with the protein being expressed and folded with the correct translational modifications present. However, the disadvantages include increased cost due to the length and difficulty in scaling-up expression. Adapted from www.sigmaaldrich.com/technical-documents/articles/biology/protein-expression-systems.

3.2 Construct information

The mammalian expression plasmid, pcDNATM-3.1/zeo(+) was used in the generation of the ABCG2 mutant constructs (Figure 3.3). The plasmid contains; a highly efficient cytomegalovirus (CMV) promoter to enhance the expression of the gene of interest, multiple cloning sites, and an antibiotic resistance cassette (ampicillin) for bacterial selection. The pcDNATM-3.1/zeo(+) plasmid also contains a selectable marker for the cytotoxic agent, Zeocin[®], which enables the isolation of stable mammalian cell lines.

The plasmid was engineered with a super folder green fluorescent protein (sfGFP) tag, which fuses N-terminally to the protein, to allow for easy visualisation. GFP has been shown to dimerise at high local concentrations; ABCG2 also forms higher ordered structures, predominantly dimeric conformations. As the GFP-ABCG2 fusion protein is expressed constitutively the GFP might be driven to form dimers, potentially stabilising ABCG2 and masking mutations that could demonstrate a structural effect. Therefore, the GFP used contains a mutation at alanine 206 to lysine that prevents GFP dimerisation (Von Stetten et al., 2012). The tagging of ABCG2 with N-terminal sfGFP has previously been shown not to affect localisation or function (Haider et al., 2015; Wong et al., 2016).



pcDNA3.1/zeo_sfGFP_ABCG2

Figure 3.3 Plasmid map of pcDNATM-3.1/zeo(+)_sfGFP_ABCG2. The plasmid contains a strong promoter for enhanced gene expression, multiple cloning sites, and a bacterial resistance gene for genetic engineering. The plasmid also contains a GFP fusion tag (GFP-ABCG2 linker sequence: 5' TCCGGACTCAGATCTCGAGCTCAAGCTTCGAATTCT 3') for visualisation of the protein and selectable marker, Zeocin[®] for stable expression of the protein in a mammalian cell line.

3.3 Construct generation

The 10 variants were generated as single amino acid substitutions to alanine.

Alanine was used because the side chain beyond the beta carbon position is eliminated, but it retains a high propensity to form α -helices, and thus is unlikely to cause perturbation to overall protein folding (Lefevre et al., 1997). The mutant constructs were generated by site-directed mutagenesis based on the *in vitro* amplification of DNA by PCR method (Sambrook & Russell 2001; summarised in Figure 3.4) using the pcDNATM-3.1/zeo(+)_sfGFP_ABCG2 vector, kindly provided by Dr Kelvin Wong (University of Nottingham).

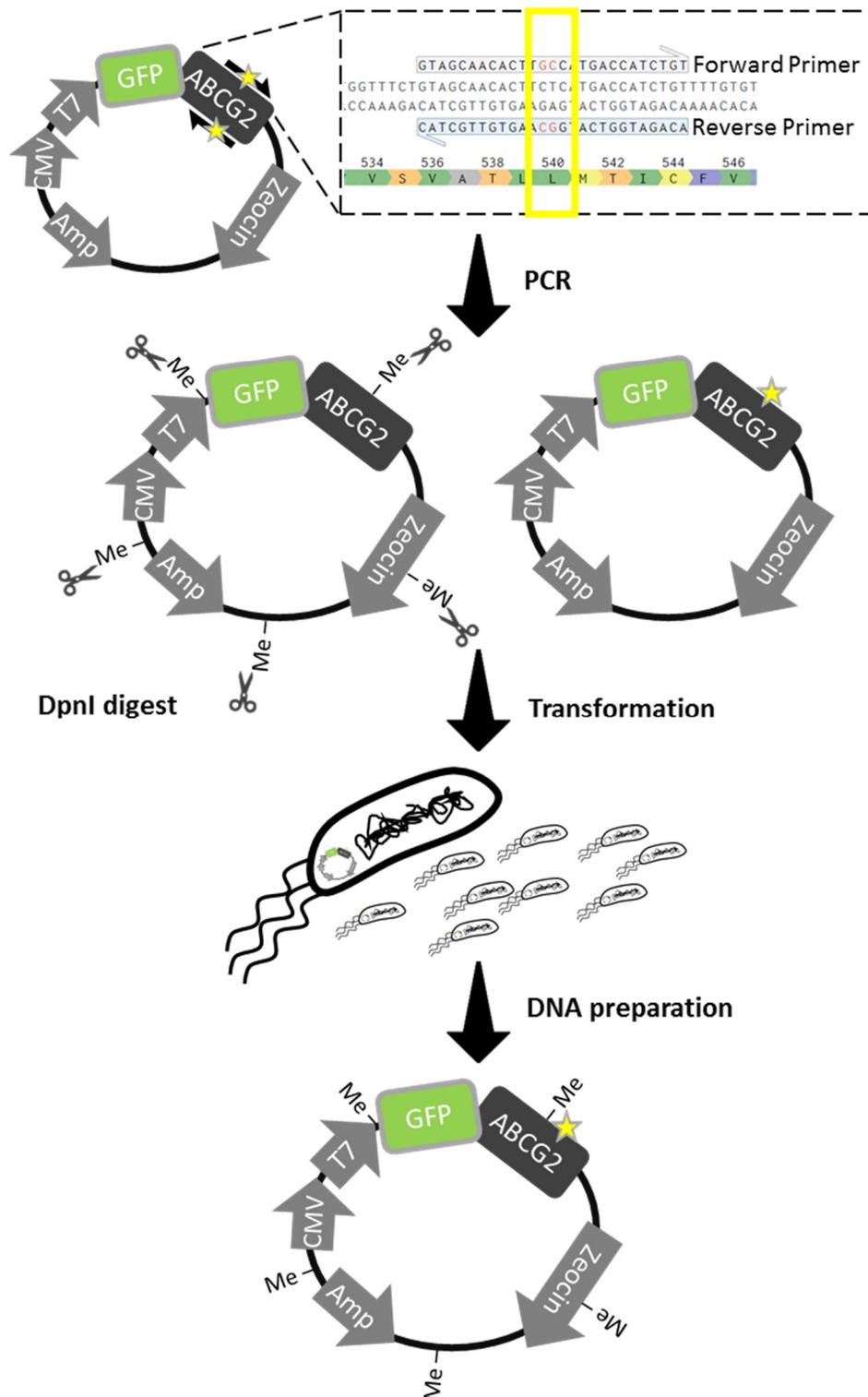


Figure 3.4 Schematic representation of the generation of mutant constructs. Primers containing base mutations anneal to the template DNA via their complementary regions and PCR elongation incorporates the mutation. The hemi-methylated template DNA is digested by DpnI restriction endonuclease. The newly synthesised DNA is then transformed into DH5α (*E. coli*) competent cells and the plasmid DNA with the incorporated mutation isolated.

Initially the newly generated constructs were digested with a restriction endonuclease, PvuII, in order to confirm the plasmid size. PvuII cuts in a site-specific manner (Figure 3.5 A.) resulting in DNA banding at 4486, 1749 and 1421bp for pcDNATM-3.1/zeo(+)_sfGFP_ABCG2 constructs. The I643A mutation itself incorporated an extra PvuII restriction site resulting in banding at 4486, 1749, 1089 and 332bp, the digested plasmids demonstrated the correct banding pattern as can be seen in Figure 3.5 B.

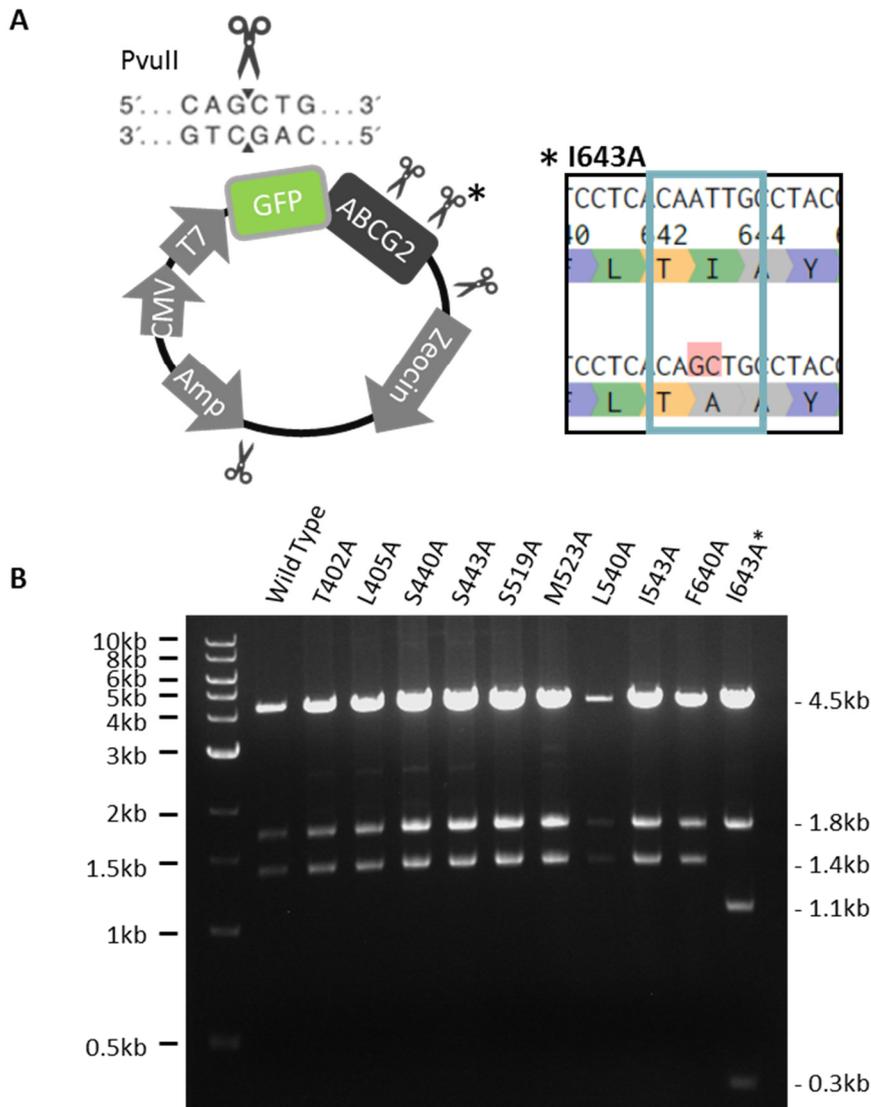


Figure 3.5 A. PvuII restriction endonuclease digest. PvuII cuts DNA between the G and the C of the DNA sequence CAGCTG to create a blunt ended cut. pcDNA™-3.1/zeo(+)_sfGFP_ABCG2 contains three PvuII recognition sequences two in the plasmid backbone, and one in the ABCG2 gene sequence, indicated by the scissor icon. The incorporation of an I643A mutation inserts a fourth cut site as indicated with an asterisk and shown in the right-hand panel. **B. Agarose gel electrophoresis analysis of mutant constructs digested with PvuII restriction endonuclease.** All constructs demonstrate the expected banding patterning, wild type through to F640A banding at 4486, 1749, and 1421bp, I643A additional PvuII site with banding at 4486, 1749, 1089, and 332bp.

To confirm the incorporation of the specific mutation and no further mutagenesis, the plasmids were sent for Sanger sequencing across the entire length of the sfGFP_ABCG2 gene (depicted in Figure 3.6). DNA sequence alignments for a single mutant are shown in Appendix 1 to demonstrate the data fully.

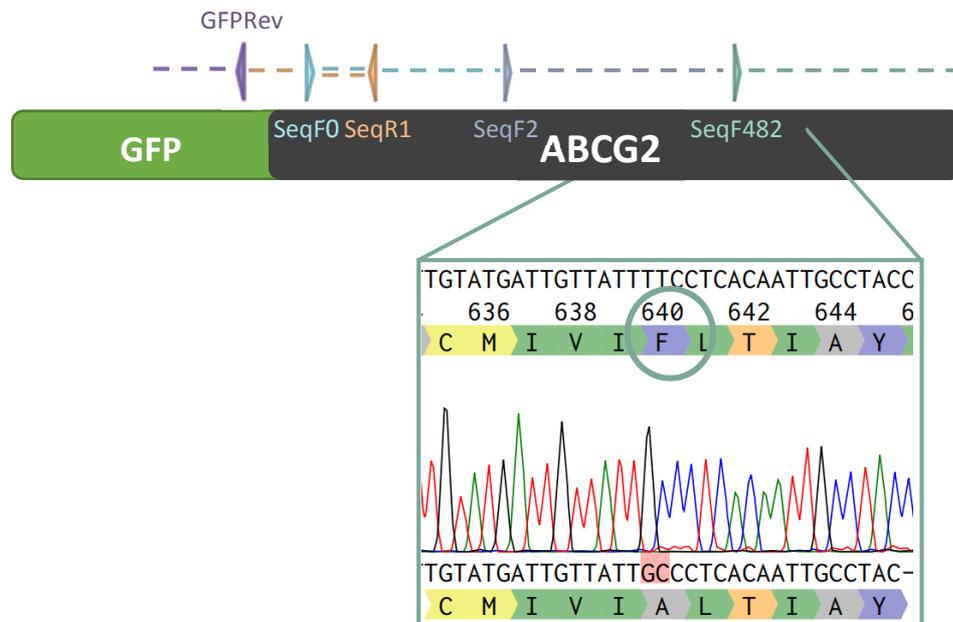


Figure 3.6 DNA mutagenesis confirmed by Sanger sequencing. Schematic representation of the sequencing primers spanning the entirety of ABCG2 and proportion of sfGFP. An example, SeqF482 sequencing data obtained for the F640A mutation, TT to GC base change is highlighted.

3.4 Protein expression

HEK293T cells, a mammalian cell line derived from human embryonic kidney cells were used in this study for all localisation and functional investigations. The 'T' of HEK293T cells is indicative of the insertion of a temperature sensitive simian vacuolating virus 40 (SV40) large T-antigen gene, which at 37°C results in larger amounts of replication competent T-antigen. The presence of an SV40 origin of replication in the mammalian expression plasmid (pcDNATM-3.1/zeo(+)) hijacks the HEK293T SV40 T-antigen, which acts as a transcriptional activator, resulting in constitutive expression of recombinant protein (sfGFP_ABCG2) (DuBridges et al., 1987).

To determine the expression characteristics of the 10 ABCG2 variants, HEK293T cells were transiently transfected with the sfGFP tagged constructs alongside the wild type construct and any additional controls (E211Q, Walker B mutant). Forty-eight hours post transfection the cell pellets were harvested, lysed in an ice-cold buffer, then analysed for total protein and ABCG2 specific expression (section 2.2.3 and 2.3.1-2.3.2). Protein (15-20µg of whole cell lysate) was analysed on discrete 8-10% (w/v) polyacrylamide gels. Total protein expression was visualised by staining with InstantBlueTM and served as a measure of equal protein loading as seen in Figure 3.7 below. ABCG2 protein expression was detected by immunoblotting with a monoclonal antibody, BXP21, which recognises an epitope within the nucleotide binding domain of ABCG2 (Allen et al., 1999). Immunoblotting detected an immunoreactive band at approximately 100kDa, consistent with sfGFP (27kDa) tagged ABCG2 (72kDa), for all mutants. The majority of the mutants also

demonstrated a secondary immunoreactive band at a slightly reduced molecular weight, approximately 88kDa (Figure 3.7). This secondary banding could be indicative of proteolysis or reduced maturation (incomplete glycosylation). To investigate the nature of this lower molecular weight band, samples were enzymatically treated with N-glycosidase F (PNGase F) which cleaves between the asparagine and the inner most GlcNAc. The treatment confirmed the secondary immunoreactive band as non-glycosylated protein (Figure 3.8) which could be indicative of immature protein in the trafficking pathway.

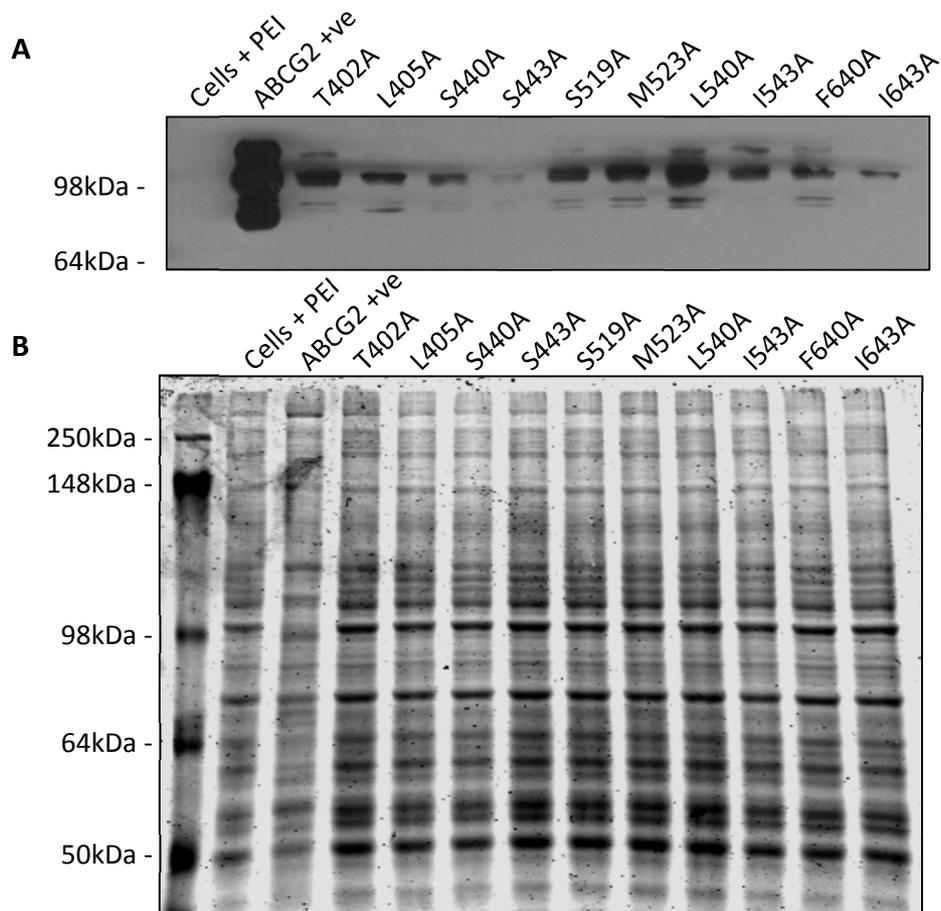


Figure 3.7 Immunoblot analysis of ABCG2 mutant variants. 20µg of cell lysate were prepared 48 hours post transfection and immunoblotted against BXP-21 antibodies. **A.** All mutants demonstrate an immunoreactive band consistent with ABCG2^{WT} (positive control, +ve) at approximately 100kDa. **B.** InstantBlue™ stained SDS-PAGE gel of cell lysates confirm equal total protein loading.

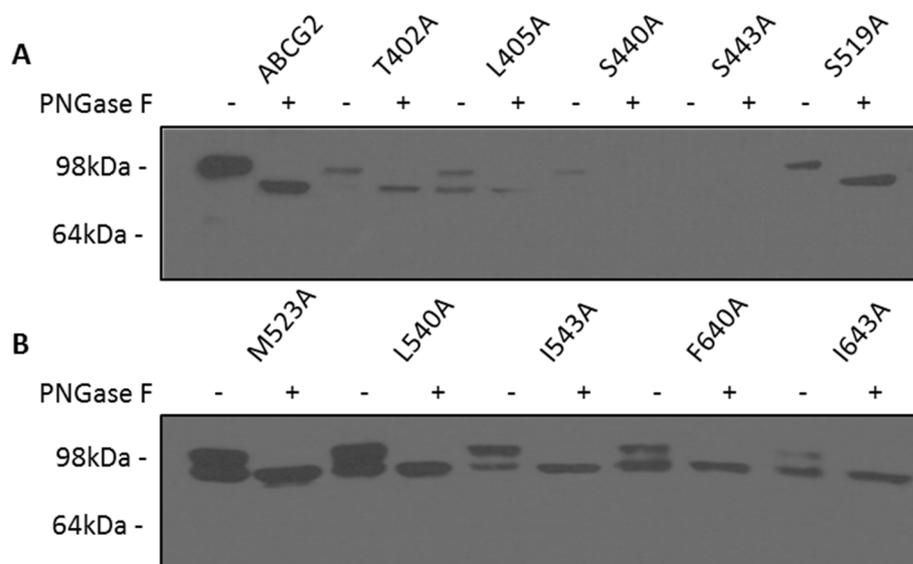


Figure 3.8 Immunoblot analysis of ABCG2 mutant expression and glycosylation. 15µg of cell lysate were digested with PNGase F and immunoblotted against BXP-21 antibodies. **A.** ABCG2^{WT} demonstrates a single immunoreactive band at approx. 100kDa without PNGase F digestion which is reduced to around 88kDa upon treatment. **A/B.** Secondary banding in majority of the ABCG2 mutants at 88kDa are consistent with non-glycosylated protein.

On reflection, all mutant constructs are able to express full length ABCG2. However, they do so at varying levels of expression, possibly due to varying efficiencies of transfection. Therefore, subsequent analysis of mutant constructs was performed in stably selected cell lines, established as described in section 2.2.4, to minimise this effect.

3.5 Localisation

Wild type ABCG2 localises to the plasma membrane in order to efflux drug substrate from the cell. Therefore, confocal microscopy was used to determine whether the replacement of amino acids with alanine had resulted in perturbation of protein localisation. Such an observation would suggest an effect on protein folding, as the latter is intrinsically linked to the release of membrane proteins from the endoplasmic reticulum (ER) (Ellgaard et al., 2016).

Confocal analysis revealed that all ABCG2 mutant variants were trafficked to the plasma membrane (PM).

Mutant variants T402A, L405A, and L540A demonstrated expression at the PM consistent with the wild type protein, whereas S440A and S519A demonstrated a slight increase in PM expression.

S443A, S519A and L543A demonstrated mixed expression: high PM (compared to WT expression) mixed with low cytosolic fluorescence, this may be representative of free GFP or protein within the trafficking pathway.

M523A also demonstrates mixed plasma membrane expression, with expression comparable to high expression in approximately 20% of cells, the remaining 80% of the cells demonstrate a decrease expression level (compared to WT), this could be suggestive of increased protein turn-over, or transfection efficiency before stable cell selection (i.e. copy number of plasmid DNA within each individual cell, a mixed expression stable cell line).

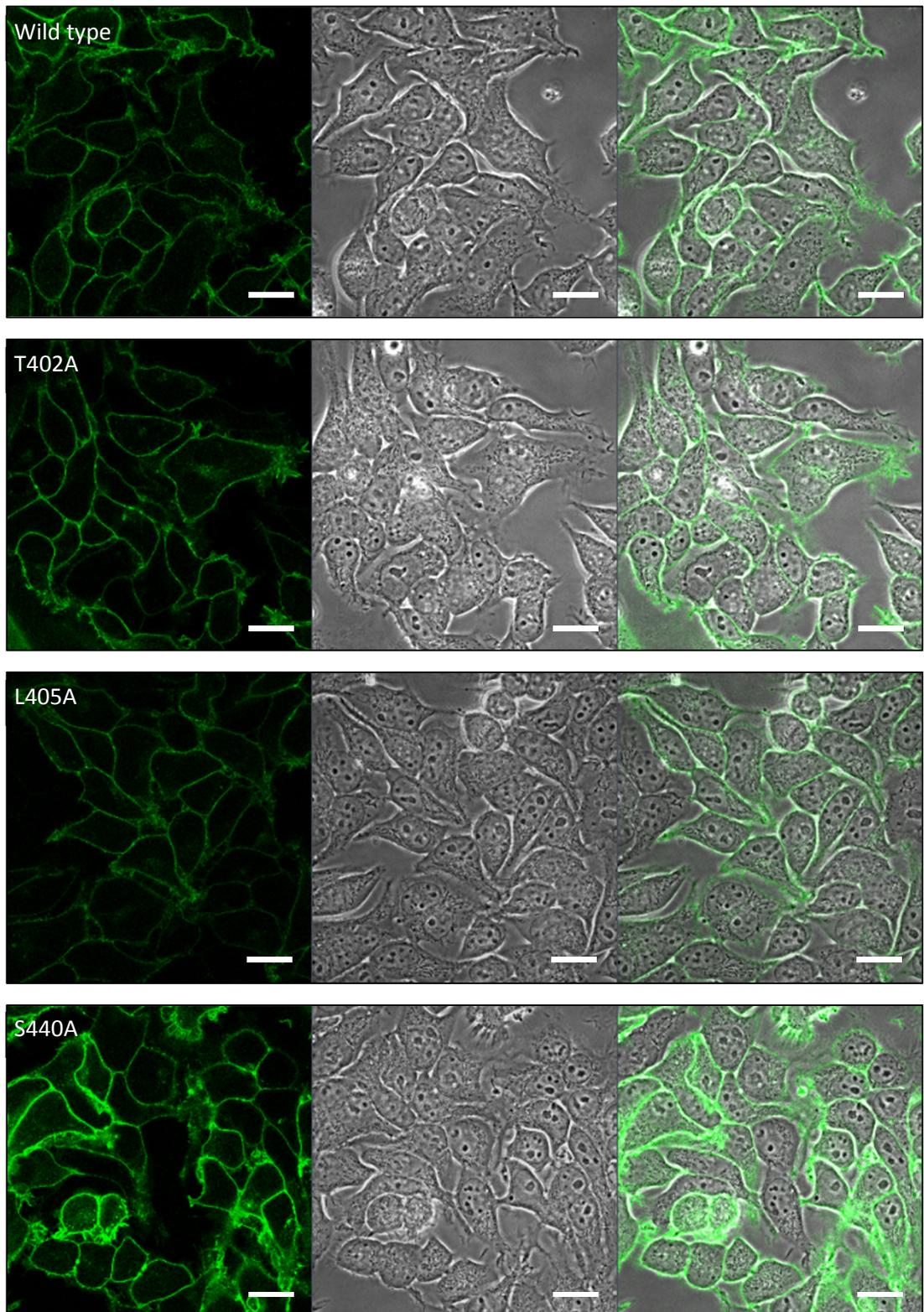
F640A demonstrated low PM expression, only in approx. 40% of cells, the remaining cells have no apparent expression but were able to confer resistance to the cytotoxic selection agent. This is perplexing and would suggest either untagged protein expression (i.e. GFP not transcribed with the protein) or a gain in the ability of the F640A mutation to efflux zeocin (equilibrium between expression and cytotoxicity reached over increasing cell passages).

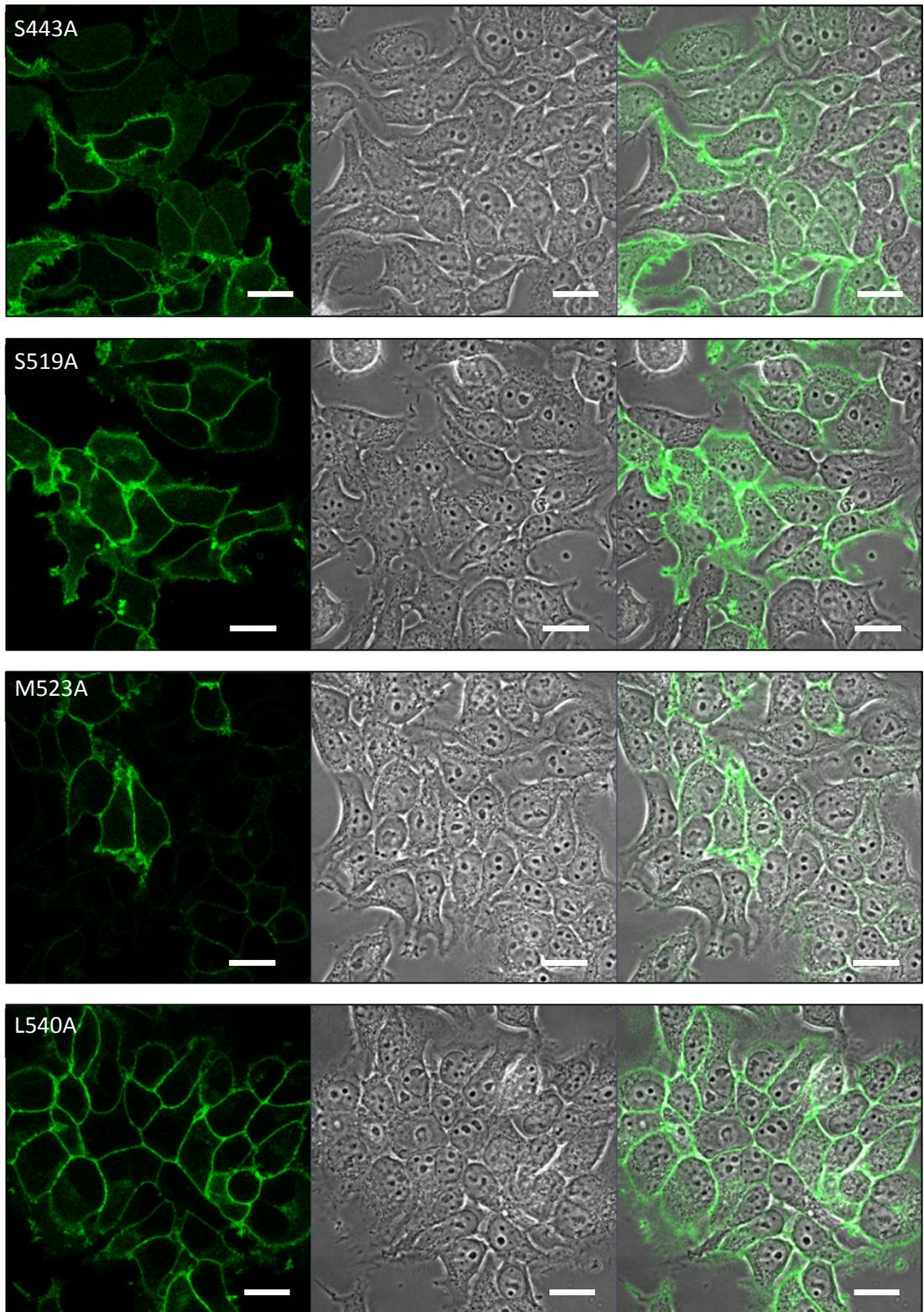
I643A shows a slightly lower level of PM expression compared to wild type.

E211Q shows PM localisation but looks slightly punctated in the dataset shown in figure 3.9, this could be representative of increase protein internalisation suggestive of a stability effect of the mutation (Ellgaard et al., 2016; Kelm et al., 2004).

Retrospective analysis would include:

- Incorporation of a house keeping enzyme (such as GAPDH) to confirm equal protein loading (/transfer to nitrocellulose membrane).
- Western blot analysis of GFP species in the stable cell line, to determine if cytosolic GFP fluorescence (in confocal analysis) is attributed to free GFP, GFP-G2 in the trafficking pathway, or GFP-G2 breakdown products.
- Densitometry analysis of anti-GFP and anti-G2 blots normalised for a housekeeping enzyme could give a percentage of G2 that is GFP tagged, this could confirm the phenotype visualised for F640A in confocal analysis.
- If western blot analysis of GFP species is suggestive of protein breakdown, pulse-chase or cycloheximide-chase (inhibits protein synthesis by blocking translational elongation) assays could be used to assess protein turnover rate, evaluating effect of mutation on protein stability (Kao et al., 2015, 2014; Tatematsu et al., 2008).





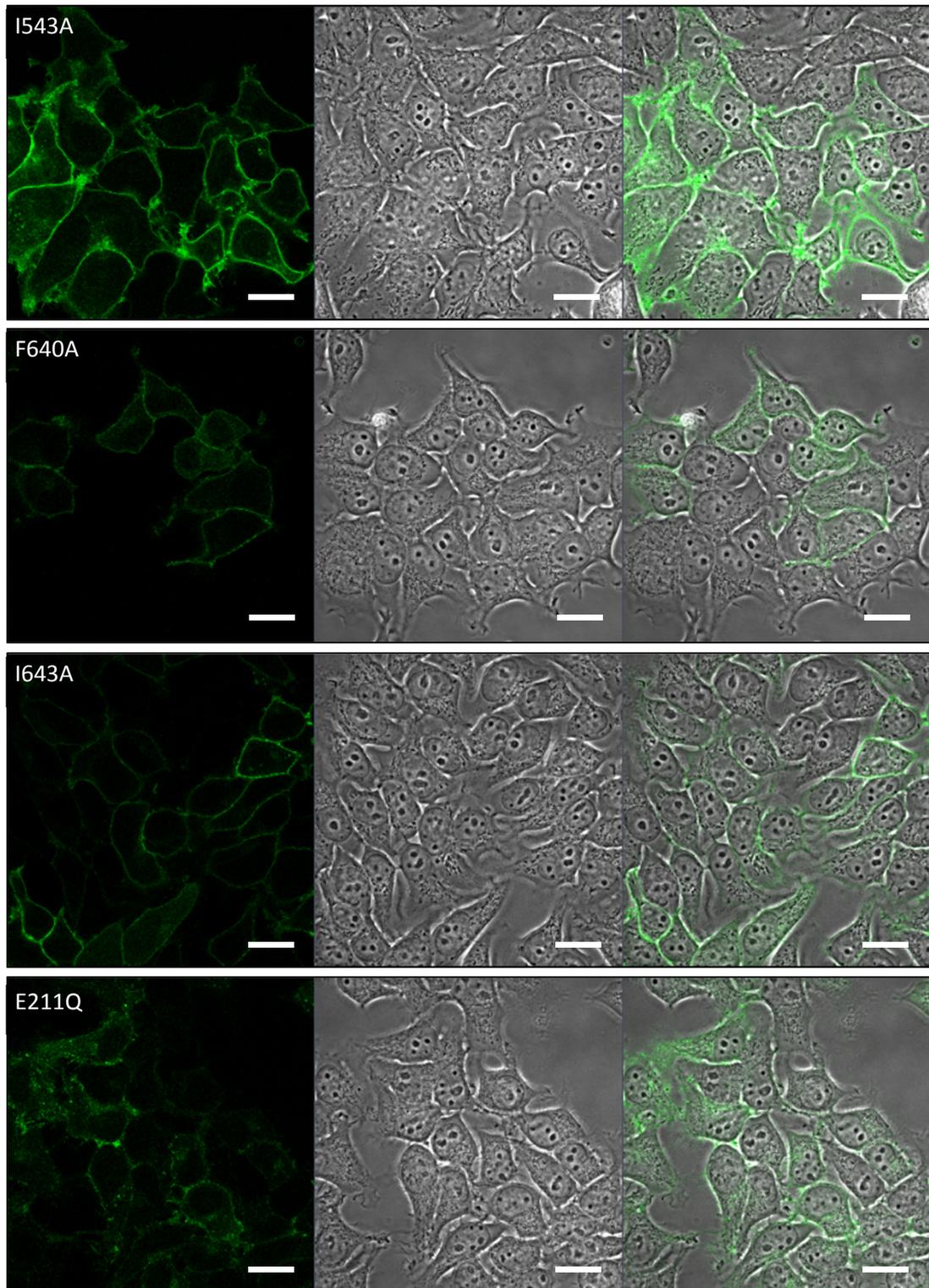


Figure 3.9 Confocal analysis of ABCG2 mutants. Stable ABCG2 mutant cell lines were visualised for protein localisation by confocal microscopy. All mutants demonstrate the ability to traffic protein to the plasma membrane. Visualisation of cells was conducted on an LSM710 confocal laser scanning microscope (Zeiss, Jena, Germany) equipped with a Plan-Apochromat 63x/1.40 Oil Ph3 DIC M27 objective and argon laser. For detection of sfGFP-ABCG2, a 2% argon laser power was used with excitation set at 488nm and emission collected between 500-530nm (FITC channel). Data are representative of one of three independent repeats. Presented dataset was acquired with identical acquisition settings to allow comparison of expression levels. Scale bar is representative of 20µm.

Western blot analysis of ABCG2 mutant variants transiently expressed in HEK293T cells was able to show overall protein expression. Confocal microscopy of ABCG2 mutant variants stably expressed in HEK293T cells showed protein localisation, confocal microscopy also showed a variation in the protein expression phenotype from cell to cell. Therefore, it is important to consider a technique for functional characterisation of the ABCG2 mutant variants, that can account for varying protein expression phenotypes.

3.6 Functional characterisation

For analysis of the mutant ABCG2 function, a robust, reproducible, and sensitive method for analysing function was required. Drug accumulation assays in intact cells usually monitor the accumulation of an isotope-labelled or fluorescent drug. To monitor protein-specific drug transport, a specific inhibitor is added and the difference in drug accumulation is attributed to the efflux capacity of the protein (Robert W Robey et al., 2001a). For ABCG2 there is only one available radiolabelled drug substrate ([³H]-daunomycin), whereas there are several fluorescent substrates (Strouse et al., 2013), and so a fluorescence based transport assay was adopted. Previous investigations have used 96-well plate-based accumulation assays; however, there are drawbacks for this technique as the protein expression level cannot be simultaneously monitored on a per cell basis (Hoechst 33342 live-cell nuclear stain used in the analysis on a cell by cell basis is a substrate of ABCG2 and therefore not suitable) and cell confluency must be equal to or greater than 95% to achieve reproducible results (Haider et al. 2015). Flow cytometry was therefore

selected for analysis of ABCG2 function (Hegedús et al., 2009; Robert W Robey et al., 2001a).

Flow cytometry is a highly sensitive laser-based technique in which a liquid stream containing particles passes through a focused beam of light. The signals generated are converted by an electrical network to values that are proportional to the light intensity (Figure 3.10) (Longobardi Givan, 2001). In this way, flow cytometry is able to analyse protein expression (GFP fluorescence) and drug (substrate) accumulation (fluorescence) simultaneously on a cell by cell basis.

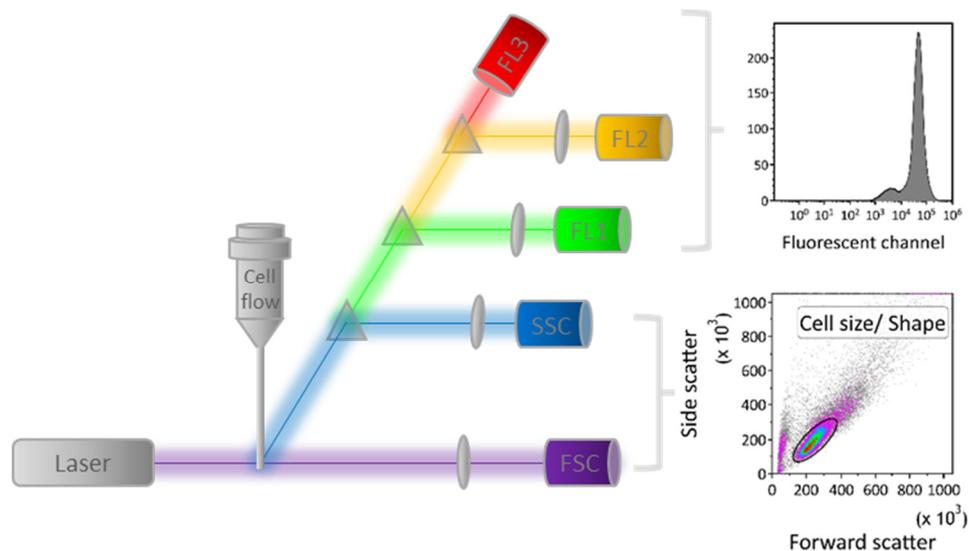


Figure 3.10 Schematic representation of flow cytometry. Cell suspension enters the flow cell where the cells are aligned into a single stream, the cells pass through a focused laser-beam, the results signal provides information on their cell size/shape based on their forward and side scatter and the fluorescent intensities of each cell defined by the fluorescent probe used in any given experiment.

3.7 Flow cytometric analysis of ABCG2 mutant expression

Flow cytometric data analysis is built upon the principle of gating. Gates or regions are assigned to cells with common characteristics, to investigate and quantify populations of interest (Figure 3.11). Forward scatter versus side scatter gating of density plots can be used to distinguish the cellular population from debris that

usually lie in the bottom left of the dot plot. A subsequent dot plot of GFP fluorescence (x axis) versus FS (y axis) (FC500) or FS area (x) versus height (y) (Astrios) enables gating to distinguish monodispersed cells from cellular aggregates, this gating reduces the chance of false positives by removing high fluorescence outliers (associated with cellular aggregates). The final type of gating is associated to single channel histograms (the fluorescent channel), gating can be applied to distinguish protein (GFP) expressing cells versus non-expressing cells, or due to the nature of mixed cell lines, variation in the protein expression level away from the wild type profile. This gating can then be applied to drug fluorescence to determine function (Lugli et al., 2010; Radcliff and Jaroszeski, 1998) (Described in 2.5, Figure 2.1).

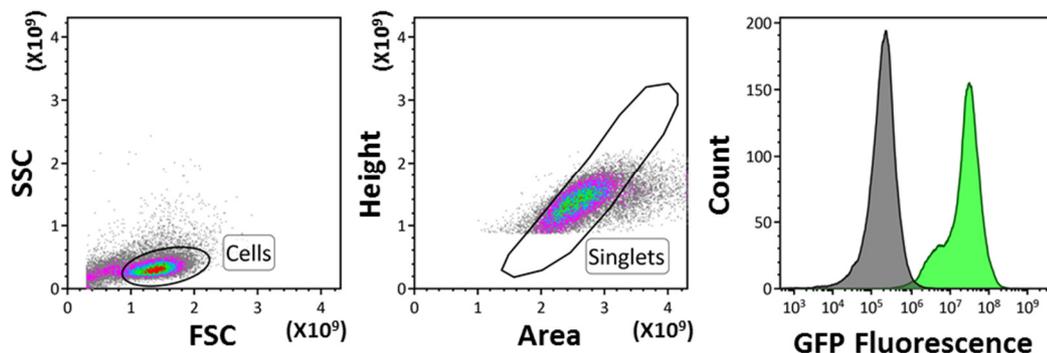


Figure 3.11 Flow cytometric gating parameters. Left to right: Forward scatter (FSC) versus side scatter (SSC) plot demonstrating gating of cells (inside ellipse) from cellular debris. Forward scatter height versus area plot (Astrios) gating singlets (monodispersed cells, inside oblong) from cellular aggregates. Single channel histogram for the fluorescence of wild type protein (sfGFP, green) including the low-level autofluorescence (grey) induced by untransfected HEK293T cells.

Western blotting and confocal microscopy had previously suggested variation in the expression level and phenotype of the ABCG2 mutants (Figure 3.7 (transient expression) and Figure 3.9 (stable expression)). Flow cytometric analysis was used to quantitatively assess the overall protein expression level across ABCG2 mutant

cell lines (Figure 3.12). The median GFP fluorescence was extracted for cells gated for cell size/shape and dispersity across ABCG2 wild type and all mutant variants. Relative expression was calculated for all mutants compared to the wild type protein, defined as: the ratio between median GFP fluorescence (MGF_L) for the mutant over the MGF_L for wild type protein. A ratio of greater than one suggests an increase in overall protein expression, whereas a ratio of less than one suggests a reduction in the overall level of protein expression, with wild type set to 1. T402A, S440A, and L540A show a significant increase in protein expression compared to $ABCG2^{WT}$ ($P \leq 0.05$), whereas E211Q, S443A, S519A, I543A, and F640A show a significant decrease in protein expression ($P \leq 0.05$, compared to WT).

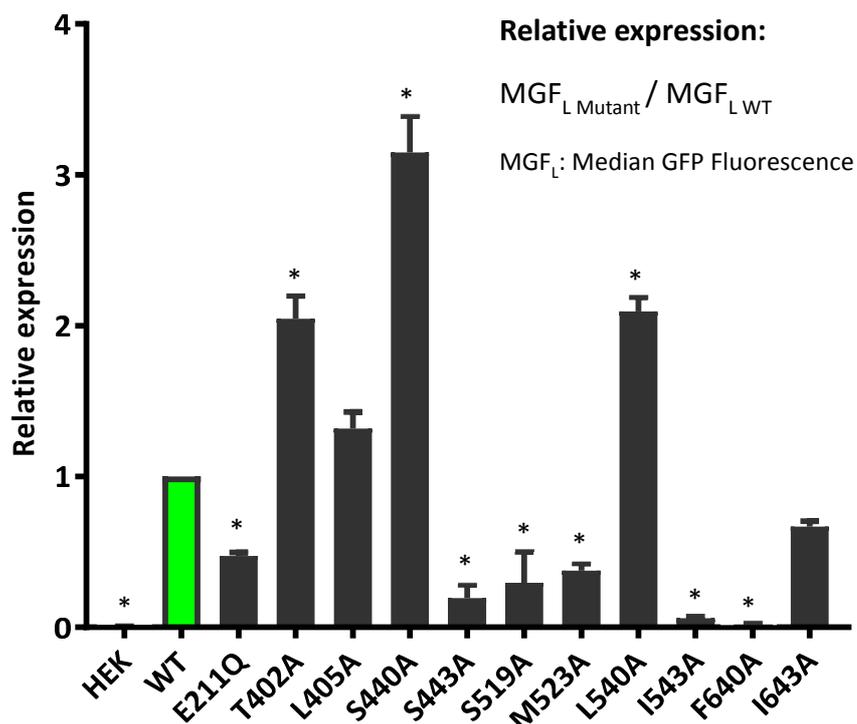


Figure 3.12 Relative expression of ABCG2 mutant variants stably expressed in HEK293T cell. The median GFP fluorescence (GFP-ABCG2 expression) for each ABCG2 mutant variant was obtained by flow cytometry (FC500, fluorescence channel: FL1). The relative expression defined as the ratio between the median GFP fluorescence (MGF_L) of the mutant and the MGF_L of wild type was determined. A ratio of >1 suggests an increase in the overall expression of protein, a ratio of <1 suggests a decrease in protein expression. Data were subjected to a one-way ANOVA with a Dunnett's post test against WT expression, asterisk denotes a P -value ≤ 0.05 . Data are representative of $N \geq 3$ independent repeats.

Extracting the median for protein expression across the entirety of a cellular population provides basic quantitative expression analysis. However, it does not take into account the 'mixed' aspect of the cellular population. A closer look at the protein expression histograms for all ABCG2 mutants demonstrate a phenotypic split common of a mixed cell population, Figure 3.13 below demonstrates this for the wild type population, shown by a left skewing of the data representing low expression.

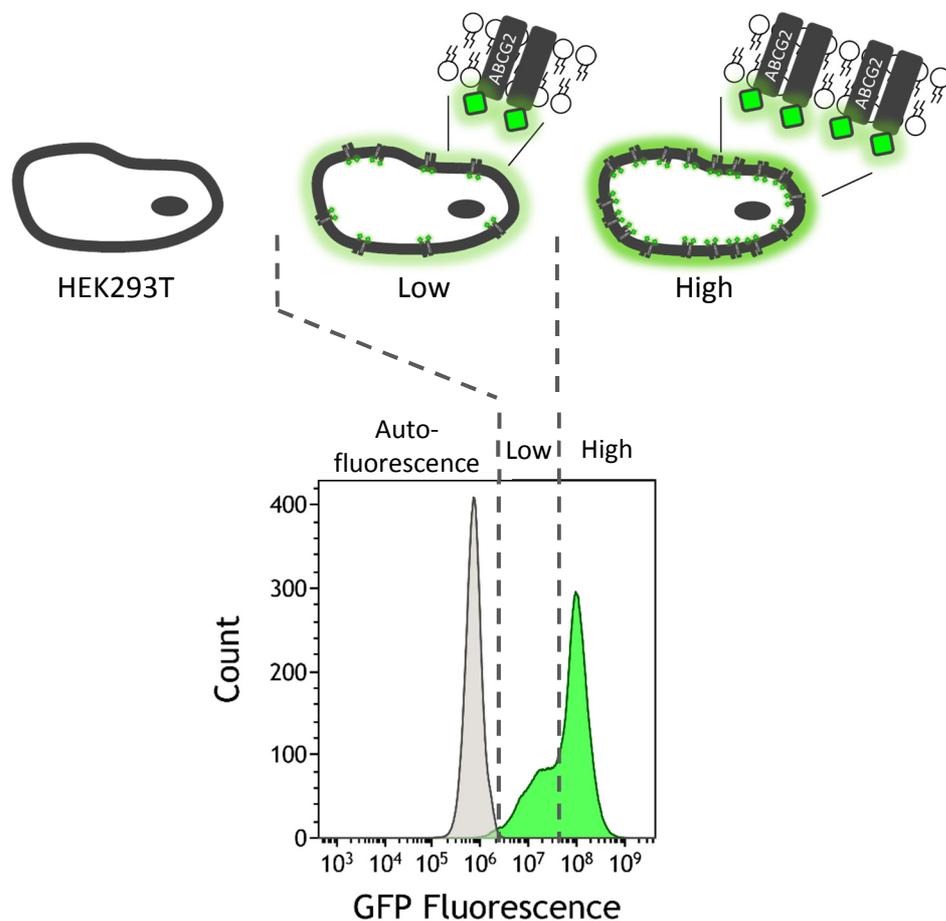


Figure 3.13 Protein expression phenotype across total population. Single fluorescence channel histogram for wild type protein expression (sfGFP) demonstrates a shouldering region which could be applicable to cells expressing low levels of protein due to the inherent nature of a mixed cell line. The grey peak is representative of untransfected HEK293T cells and therefore provides a level of background fluorescence.

This variation in expression of protein can have consequences in the interpretation of functional data, shown in Figure 3.14 and Figure 3.15. To put this into context, there are three main theoretical phenotypes applied to drug efflux of a wild type substrate: Firstly, an impaired phenotype where a mutant's ability to efflux a drug substrate is reduced. Secondly, a wild type phenotype where an ABCG2 variant is comparable in drug efflux to the wild type protein. Thirdly, an enhanced transport phenotype, resulting from the ability of an ABCG2 variant to increase the rate in which it is able to efflux a native drug substrate. A separate phenotype i.e. a gain in the ability to export a non-native drug substrate would be the equivalent of the R482G/T ABCG2 isoform (Honjo et al., 2001).

When protein expression varies the functional phenotype can be misinterpreted; for example, a mutant expressing low levels of ABCG2 with enhanced drug efflux would appear to have wild type drug efflux capability when compared to wild type efflux, if differences in protein expression level were not accounted for. This is described as a false negative data interpretation, as can be seen in Figure 3.14. Equally, a cell highly expressing protein could have reduced efflux capacity but due to the increased level of protein results in a masking of this phenotype as total efflux is the same compared to WT, resulting in a false positive (Figure 3.14).

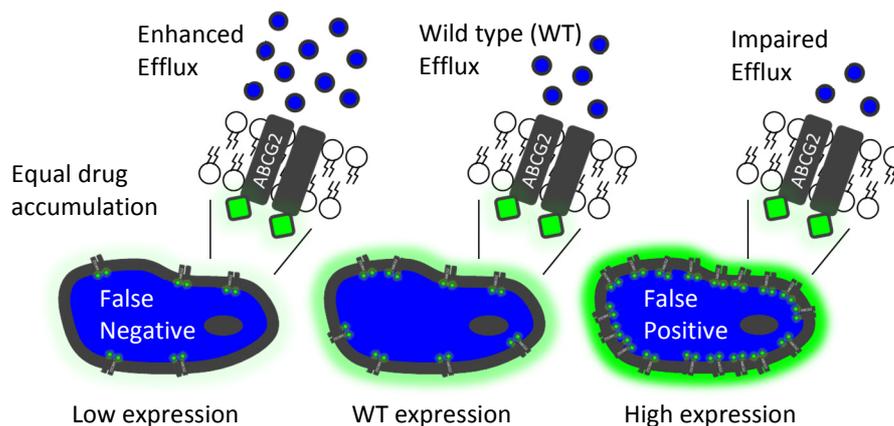


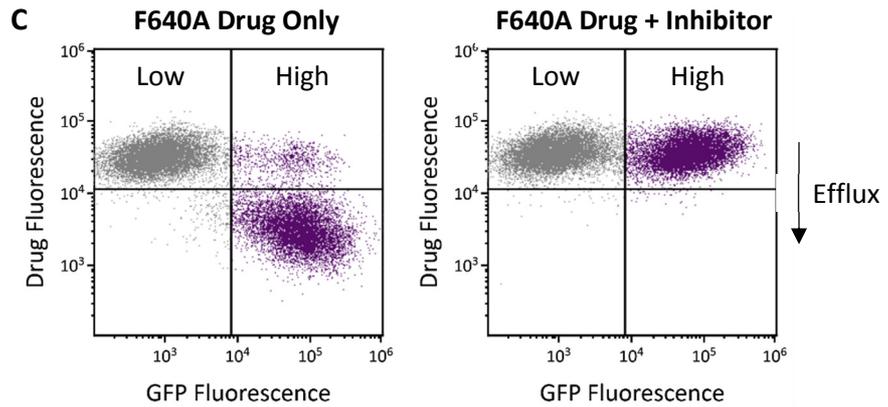
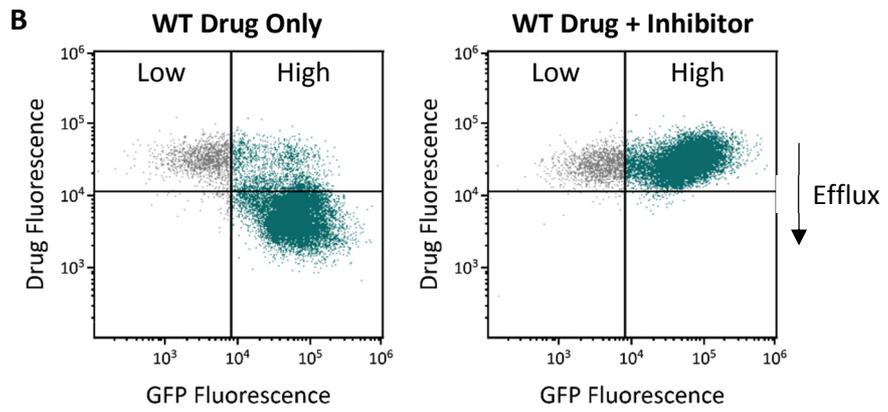
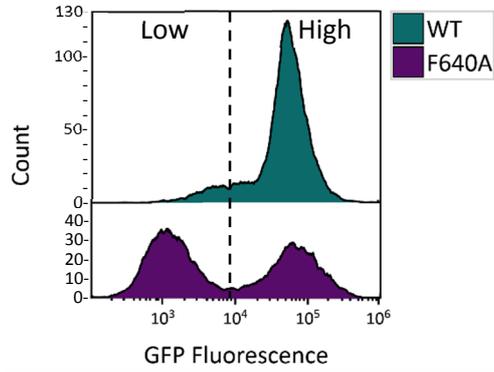
Figure 3.14 Implication of expression upon functional data interpretation. Schematic representation of low, normal, and high-level protein expression and the ratio of efflux for an impaired, normal (wild type), and enhanced drug efflux. An example of false negative and false positive interpretation has been highlighted.

In order to demonstrate how relative expression levels were accounted for in functional data processing, it is pertinent to look at the WT protein and a mutant example, in this instance F640A (Figure 3.15). The WT ABCG2 cell line demonstrated a consistent expression level (teal, Figure 3.15 A.), whereas only about half of the F640A cells demonstrate a WT-level of expression (Figure 3.15 A, purple high). The remaining cells showed a much lower protein expression level (Figure 3.15 A, purple low). [GFP fluorescence histograms for all remaining ABCG2 mutants can be seen in Appendix 2].

Figure 3.15 demonstrates typical data obtained from flow. GFP fluorescence versus drug fluorescence dot plots coloured for expression gating (expression gating represented by dashed line in Figure 3.15 A, low expression coloured grey, high expression coloured teal for WT and purple for F640A) demonstrates that the high GFP/protein expressing population correlates to functional protein, indicated by the decrease in drug fluorescence (y-axis) for the drug only sample compared to the drug plus Ko143 sample. In order to calculate drug efflux, the median drug

fluorescence value is extracted for drug \pm Ko143 (defined as $(\beta-\alpha)/\alpha$, where α is the median drug fluorescence (MDF_L) for the drug only sample and β is the MDF_L for drug plus inhibitor (Ko143)). Relative drug efflux values across the whole (GFP fluorescence) population (ignoring the expression phenotype) erroneously suggests an impaired drug efflux for F640A (0.7 vs. WT 4.5) (Figure 3.15, D). When the relative efflux is corrected for the 'high' expression gate (Figure 3.15, A), F640A is demonstrated to have a higher inhibited drug transport (10.2 vs. WT 4.5) (Figure 3.15, D). The mutant protein expression level (median GFP fluorescence) within the gated parameters still varies compared to the WT (Figure 3.15, E), therefore a normalisation of drug efflux was performed for the GFP fluorescence in the gated high expressing cells, this resulted in F640A having an expression normalised efflux of 8.4, i.e. an enhanced efflux.

A Protein Expression Profile



D

		Whole GFP Fluorescence Population		High GFP Fluorescence Population		
		Drug Fluorescence (median)	Relative Efflux $((\beta-\alpha)/\alpha)$	Drug Fluorescence (median)	Relative Efflux $((\beta-\alpha)/\alpha)$	
WT	α	Drug	5486		5180	
	β	Drug + Ko143	28272	4.2	28474	4.5
F640A	α	Drug	21719		3268	
	β	Drug + Ko143	36345	0.7	36577	10.2

E

	GFP Fluorescence: High Population (median)	Ratio (WT/mutant)	Relative Efflux (Table D)	Expression Normalised Efflux (Relative Efflux * Ratio)	Efflux Phenotype
WT	55690		4.5		
F640A	67759	0.82	10.2	8.4	Enhanced

Figure 3.15 Expression normalised drug efflux of WT and F640A variants. **A.** GFP-tagged ABCG2 protein expression for WT (teal) and mutant variant F640A (purple). Dotted line represents expression level gating. F640A's expression split is roughly 1:1 across the low and high populations. **B/C.** Dot plots of GFP (expression) fluorescence versus drug accumulation. **B.** WT, Left plot: drug only sample shows the majority of cells in the lower right quadrant (teal) representative of high protein expression and reduced drug accumulation. Right plot: drug plus inhibitor sample shows the majority of cells in the upper right quadrant (teal) representative of high protein expression and high drug accumulation. The shift in drug fluorescence across the two samples is representative of ABCG2 specific drug efflux. **C.** F640A high expressing population (purple) shows the same efflux pattern as WT upon addition of inhibitor, whereas there is no effect on the low expressing population (grey). **D.** Relative efflux values across the whole population (ignoring the expression phenotype) erroneously suggests an impaired drug efflux phenotype for F640A (0.7 vs. WT 4.2). When the relative efflux is corrected for the expression gate in panel A, F640A is demonstrated to have a higher inhibited drug transport (10.2 vs. WT 4.5). **E.** Normalisation of drug efflux for the GFP fluorescence in the gated high expressing cells results in F640A having an expression normalised efflux of 8.4, i.e. an enhanced efflux. Data are representative of one experiment of three independent repeats.

3.8 Flow cytometric functional analysis of ABCG2 mutants

Flow cytometric based drug accumulation assays were performed with three drug substrates, mitoxantrone (MX), Pheophorbide A (PhA), both well described drug substrates of ABCG2, and the final drug, daunorubicin (DNR). Daunorubicin is a non-native drug substrate for wild type ABCG2 but is exported by the R482G/T variants, therefore was used to access any gain in efflux specificity (Robey et al., 2005, 2003).

All three drug substrates were used at a concentration of 10 μ M within the drug accumulation assay. This concentration was used as previous studies for cytotoxicity assays with increasing drug concentration demonstrated an IC₅₀ (i.e. concentration at which 50% of cells die) for each drug substrate within the nM range (MX 107 \pm 64nM (Robey et al., 2003), PhA 90 \pm 10nM (Robey et al., 2005), DNR 40 \pm 1nM (Robey et al., 2003)), therefore a concentration of 10 μ M should be in excess (by several orders of magnitude), ensuring protein- (ABCG2) drug substrate saturation (implying V_{max} is reached).

A pilot assay for function used at a fixed concentration of MX (10 μ M) with varying concentration of the ABCG2 specific inhibitor, Ko143 (Figure 3.16). A normalised value for % inhibition of MX efflux was determined for each concentration of Ko143 as follows:

1. Relative inhibition (RI): $(MDF_{L\text{ MX plus Ko143}} - MDF_{L\text{ MX only}}) / MDF_{L\text{ MX only}}$
2. normalised % inhibition: $(RI_{\text{max[Ko143]}} / RI_{\text{x[Ko143]}}) * 100$

where, MDF_L is the median drug fluorescence and $RI_{\text{x[Ko143]}}$ is the relative inhibition of efflux at any given concentration.

The $\log_{10}IC_{50}$ of Ko143 obtained from curve fitting of the data (Figure 3.16) was -7.342 (95% confidence interval (CI) from -7.453 to -7.217). This indicates that 45nM of Ko143 is sufficient to cause 50% inhibition of sfGFP-ABCG2_{WT} protein (as demonstrated by 50% inhibition of MX efflux). Ko143 was used at a concentration of 1 μ M in subsequent assays, at this concentration Ko143 is in excess by several orders of magnitude and was determined as the maximal ABCG2 specific concentration in studies by Ambudkar et al. (2015).

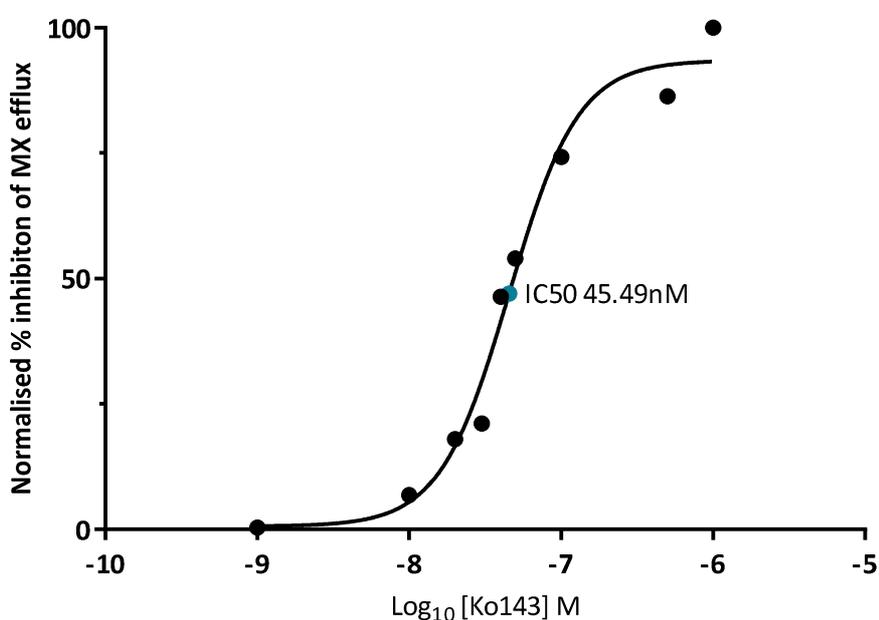


Figure 3.16 Inhibition of mitoxantrone (MX) efflux in the presence of various concentrations of Ko143. HEK293T cells stably expressing ABCG2_{WT} protein were incubated with 10 μ M MX only or in the presence of various concentration of the ABCG2 specific inhibitor, Ko143. The median mitoxantrone fluorescence was measured using an FC500 flow cytometer. Data is representative of the relative inhibition $((MDF_{L\text{ MX plus Ko143}} - MDF_{L\text{ MX only}}) / MDF_{L\text{ MX only}})$ percentage normalised for the maximum Ko143 concentration $((RI_{\max[\text{Ko143}]} / RI_{x[\text{Ko143}]}) * 100)$ (MDF_L is the median drug fluorescence and $RI_{x[\text{Ko143}]}$ is the relative inhibition of efflux at any given concentration), with fitted curve (using Sigmoidal, 4PL, X is log(concentration) equation in GraphPad Prism 7.04).

Functional data analysis for the entire series of mutants with all three drugs was then performed as described above (sections 2.5 and 3.7 (Figure 3.15)).

Representative overlay histograms of vehicle (DMSO), drug, and drug plus inhibitor associated fluorescence are shown in Figure 3.17 for three mutants, and the data for the remaining mutants is in Appendix 3. For each mutant there are drug specific fluorescence histograms in the absence or presence of the ABCG2 specific inhibitor Ko143. A leftward shift of the fluorescence in the absence of inhibitor provides a measure of mutant-specific drug efflux. For the wild type protein this leftward shift was observed with both MX and PhA but not with DNR, confirming that the WT protein can transport 2 of the 3 test substrates. For L405A, the leftward shift is reduced for MX and almost indistinguishable from the inhibited sample for PhA, indicating greater accumulation (less efflux) of the drug substrate and thereby demonstrating a perturbed efflux of drug substrate. The mutant F640A demonstrates the complete opposite to L405A showing a greater leftward shift in fluorescence for the drug substrate only sample, this indicates an enhanced efflux for drug substrate. F640A also shows the ability to export a non-native drug substrate, showing a leftward shift for DNR. A potential limitation of this drug accumulation assay is that the mechanism of Ko143 (i.e. competitive vs. non-competitive action) is unknown. If the binding mechanism is competitive then mutating residues in a putative drug binding site may also affect the binding of inhibitor. If the binding of Ko143 had been altered a leftward shift in the fluorescence for drug plus Ko143 would be anticipated, this phenotype was not observed for any data set (example of this can be seen in Figure 3.17). Although

there is a possibility that any effect on Ko143 binding could be masked due as it is present in molar excess in the accumulation assay. Follow up studies would aim to investigate the direct substrate and inhibitor (fluorescently labelled) binding via techniques such as microscale thermophoresis or fluorescence correlation microscopy (discussed further in Chapter 6).

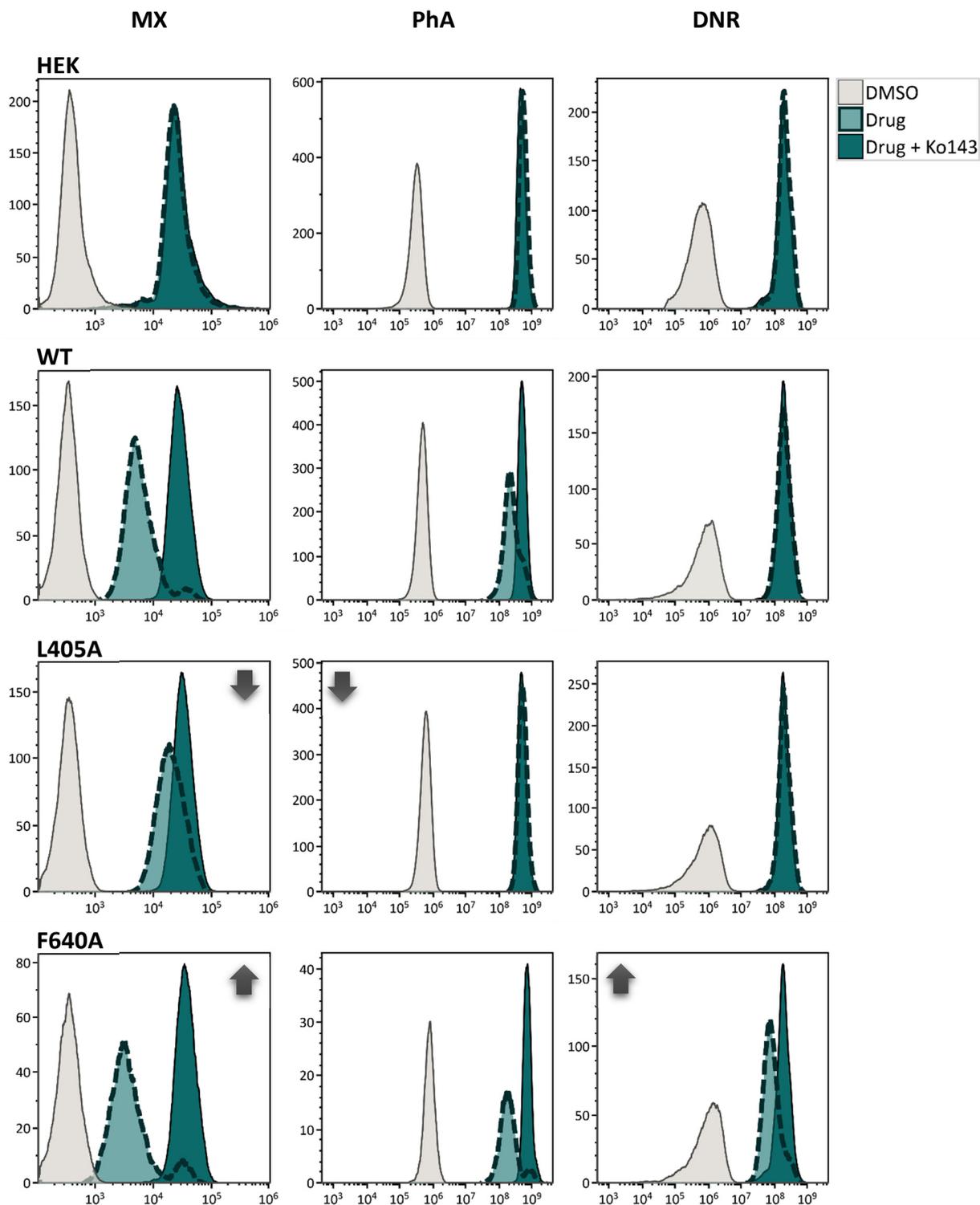


Figure 3.17 ABCG2 mutants stably expressed in HEK293T cells demonstrate perturbed drug efflux. Data is representative of Kaluza© single fluorescence channel overlay histograms for a typical dataset of $N \geq 3$ independent experiments for mitoxantrone (MX), pheophorbide A (PhA), and daunorubicin (DNR) (gated for high expression). A leftward shift in drug fluorescence in the absence of Ko143 is representative of drug efflux (see key, top left). Statistical significance (arrows, $P \leq 0.05$) were determined using a one-way ANOVA analysis of expression normalised relative efflux values (Table 3.1) followed by a Dunnett's multiple comparisons test against wild type drug efflux.

The flow cytometric based drug accumulation assay and the data processing described thus far was used to assess the functional effect across all mutations.

A decrease in the ability to efflux drug substrate was seen for several of the mutants, namely T402A, L405A, S440A, L540A, and I543A for mitoxantrone (Figure 3.18 A., Table 3.1). Four of these residues, T402A, L405A, S440A and I543A, also showed a reduced efflux of Ko-143 inhibitable efflux of pheophorbide A (Figure 3.18 B., Table 3.1), indicating a potential common interaction/recognition among these residues (László et al., 2016). Conversely, L540A demonstrates a decrease for mitoxantrone but not for pheophorbide, suggesting a potential drug specific interaction. As the L540A mutant demonstrates a non-significant but slight reduction in pheophorbide A efflux and given its spatial proximity to I543A, it is tempting to speculate that the residues (T402A, L405A, S440A, L540A, and I543A) could form part of a drug binding cavity with distinct drug binding regions (Clark et al., 2006). Interestingly, F640A located in transmembrane helix 6 (TMH6) demonstrates a significant ($P \leq 0.001$) increase in the efflux of mitoxantrone, it also demonstrates a marginal (not significant) increase in pheophorbide A efflux (Figure 3.18 B., Table 3.1). F640A alongside M523A (TMH4) gained the ability to efflux a non-native drug substrate daunorubicin (Figure 3.18 C., Table 3.1), exclusively documented as a drug substrate for the R482G/T ABCG2 isoform (Honjo et al., 2001). The data suggests that the replacement of the bulky aromatic group of phenylalanine for the small aliphatic residue (alanine) could result in multiple outcomes: altered drug binding by direct binding or access to the active site, an improved ability to alter conformations of the transmembrane domain during the

transport cycle, or an improved TMD communication with the nucleotide binding domain as the bulky aromatic phenylalanine is replaced by the small aliphatic residue.

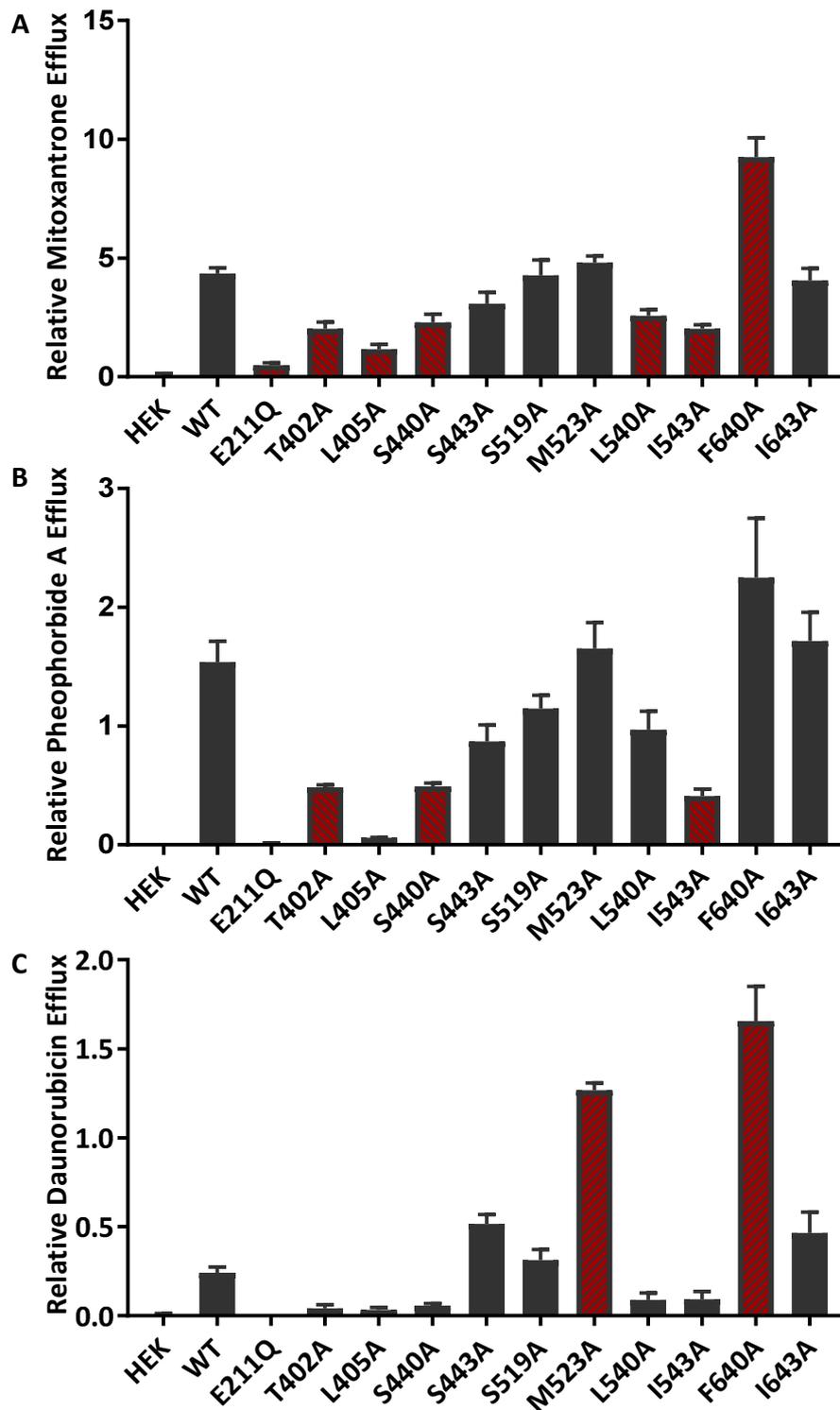


Figure 3.18 Function of ABCG2 wild type and mutant variants stably expressed in HEK293T cells. Data is representative of the mean \pm SEM for normalised relative drug efflux for MX (A), PhA (B), and DNR (C) from at least three independent experiments. Statistical significance (red hatching, $P \leq 0.05$) was determined using a one-way ANOVA analysis followed by a Dunnett's multiple comparisons test against wild type drug efflux.

	Mitoxantrone		Pheophorbide A		Daunorubicin	
	Mean	SEM	Mean	SEM	Mean	SEM
HEK	0.06***	0.08	-0.02***	0.01	-0.01	0.02
WT	4.35	0.24	1.54	0.17	0.24	0.03
E211Q	0.49***	0.10	-0.02***	0.03	-0.01	0.02
T402A	2.02**	0.29	0.48**	0.02	0.04	0.02
L405A	1.16***	0.21	0.06***	0.00	0.03	0.01
S440A	2.29**	0.35	0.49**	0.03	0.06	0.01
S443A	3.08	0.48	0.87	0.14	0.52	0.05
S519A	4.27	0.65	1.15	0.11	0.31	0.06
M523A	4.80	0.29	1.65	0.22	1.27***	0.04
L540A	2.57*	0.25	0.97	0.16	0.09	0.04
I543A	2.03**	0.17	0.41**	0.06	0.09	0.04
F640A	9.26***	0.81	2.25	0.50	1.65***	0.20
I643A	4.06	0.51	1.71	0.24	0.47	0.12

Table 3.1 Function of ABCG2 wild type and mutant variants stably expressed in HEK293T cells. Data is representative of the mean \pm SEM for normalised relative drug efflux for mitoxantrone (MX), pheophorbide A (PhA), and daunorubicin (DNR) from at least three independent experiments. Statistical significance (* $P \leq 0.05$ ** $P \leq 0.01$ *** $P \leq 0.001$) was determined using a one-way ANOVA analysis followed by a Dunnett's multiple comparisons test against wild type drug efflux.

3.9 Summary

This chapter set out to investigate 10 residues within ABCG2 as part of a 'lateral slice' hypothesis of drug binding.

- Single point mutations were generated for the 'lateral slice' residues via site-directed mutagenesis in an sfGFP-ABCG2 fusion construct, suitable for a mammalian-based expression system and downstream fluorescence analysis.
- All ABCG2 mutants were confirmed to express full length protein by western blot analysis with BXP-21 antibodies, although a variation in protein expression levels was observed. A secondary band visible at a reduced molecular weight on the western blot was attributed to a reduced glycosylation level, indicative of protein in the trafficking pathway.
- Fluorescence based confocal analysis of the mutants demonstrate predominant localisation to the plasma membrane, confirming that the mutations resulted in no major structural effect.
- A fluorescence-based flow cytometric drug accumulation assay was used to detect the functional phenotype of the ABCG2 mutant isoforms. Relative drug efflux gave a measure of drug fluorescence (mitoxantrone, pheophorbide A, or daunorubicin) in the presence of an ABCG2 specific inhibitor Ko143, in comparison to the absence of inhibitor. Due to the variation in ABCG2 mutant protein expression level, GFP fluorescence data

(sfGFP-ABCG2 expression) was collected simultaneously and mutant protein efflux data normalised for expression (mutant compared to wild type).

- One mutant, L540A demonstrated reduced efflux for mitoxantrone compared to the wild type protein.
- Mutants T402A, L405A, S440A, and I543A demonstrated reduced allocrite efflux for both mitoxantrone and pheophorbide A. The results obtained for T402A was consistent with data published previously (Polgar et al., 2010).
- The mutant F640A was the only mutation to show an enhanced efflux capability, with a significant increase in the efflux of mitoxantrone.
- F640A, along with M523A was also able to show a gain in efflux for daunorubicin, a non-native drug substrate reported to be transported by the R482G and T ABCG2 variants (Honjo et al., 2001).
- The combined data suggests a significant role for the 'lateral slice' residues in allocrite recognition and/or transport, but exactly how remains unknown without a high-resolution structure for interpretation.
- At this point in the current investigations a crystal structure of the ABCG5/G8 sterol transporter was published, shortly after an ABCG2 homology model based on the ABCG5/G8 structure was published. These structures provided novel insights into the domain organisation and transmembrane helical arrangement of G sub-family proteins.
- Moving forwards, I will use the homology model as a framework for which current data can be interpreted to develop a picture of drug transport.

Chapter 4 Using the structural revolution to enhance the understanding of ABCG2 drug binding

4.1 The structural revolution

In the introduction I described how little was known about the structure of ABCG2 at the start of this project. At the time, only low resolution negative stain cryo-electron microscopy provided 3D structural data of ABCG2, allowing for visualisation of an overall shape and oligomeric state (McDevitt et al., 2006; Rosenberg et al., 2010) and six transmembrane (TM) spanning helices were identified by topological prediction algorithms, confirmed by hemagglutinin (HA) epitope insertion experiments (Wang et al., 2008). It was the topological data presented by Wang et al. that was used in the selection of residues for investigation in the 'lateral slice' hypothesis. This postulated that residues located in topologically equivalent positions to R482 and P485 in the cytoplasmic halves of other transmembrane helices would be involved in substrate recognition and/or binding.

Chapter 3 described data showing that several mutations in the lateral slice had an impact on the efflux of allocrite. Fortunately, as the data collection was coming to an end for these residues, new structural data for the ABCG family was published (László et al., 2016; Lee et al., 2016). The X-ray structure of ABCG5/G8 described a compact overall structure with no crossing-over of helices from the TMDs to the NBDs (Lee et al., 2016). This Type II exporter fold was in contrast to the Type I exporter exemplified by Sav1866 (Jackson et al., 2018; Taylor et al., 2017).

Homology models of ABCG2 based upon this structure predicted the presence of multiple drug binding sites (László et al., 2016).

This new fold enabled a structural interpretation of the existing lateral slice mutations and also enabled a series of new mutations to be designed, to test predicted binding sites in the homology model. The use of mutagenesis to validate ABC transporter folds has been shown previously, when mutations were made in ABCB1/Pgp to test the type I exporter domain swapping (TMH cross-over) shown by the bacterial homolog, Sav1866 (Zolnerciks et al., 2007).

This chapter describes the construction and analysis of a series of mutations in a predicted drug binding site. Integration of data from both series of mutants (and from a fellow PhD student investigating a different region of the TMD) will enable a later chapter to describe a working model for multidrug binding by ABCG2.

4.2 Structural mapping

The lateral slice mutants M523A and F640A demonstrated, to differing extents, a perturbation of drug efflux; both mutations demonstrated a gain in the ability to efflux a non-native drug substrate daunorubicin, by approximately 5 and 7-fold respectively compared to wild type. When these two particularly interesting mutations were mapped on to the recently published homology model of ABCG2 (László et al., 2016) they were shown to be in close spatial proximity (around 3.5Å) (Figure 4.1). Furthermore, they appear to be located to a pocket at the lipid interface of the transmembrane domains, this pocket being lined by residues from transmembrane helices (TMHs) 4, 5, and 6. This region in the ABCG5/G8 crystal

structure showed electron density consistent with cholesterol and was proposed as a potential sterol binding site (Lee et al., 2016). The hypothesis of this chapter is that this pocket of residues in ABCG2 may form interactions with drug substrates, acting as a drug recognition site for hydrophobic molecules accessing the protein via the lipid interface. To test this hypothesis, 6 new single mutations to alanine (torsionally neutral amino acid) were proposed from residues that appeared to line the pocket, namely M541, F545, M548, F571, L633, and M636 (Figure 4.1/ Table 4.1).

	TMH4			TMH5					TMH6		
	S519	M523	L540	M541	I543	F545	M548	F571	L633	M636	F640
S519									4.0	4.5	
M523				4.5			4.0	4.9			3.5
L540					4.3						
M541		4.5									
I543			4.3								
F545											
M548		4.0						3.7			4.4
F571		4.9					3.7				4.2
L633	4.0									4.6	
M636	4.5								4.6		4.2
F640		3.5					4.4	4.2		4.2	

Table 4.1 Table of distances (below 5Å) between the residues in the lipid exposed pocket. Highlighted yellow are residues from the 'lateral slice' hypothesis.

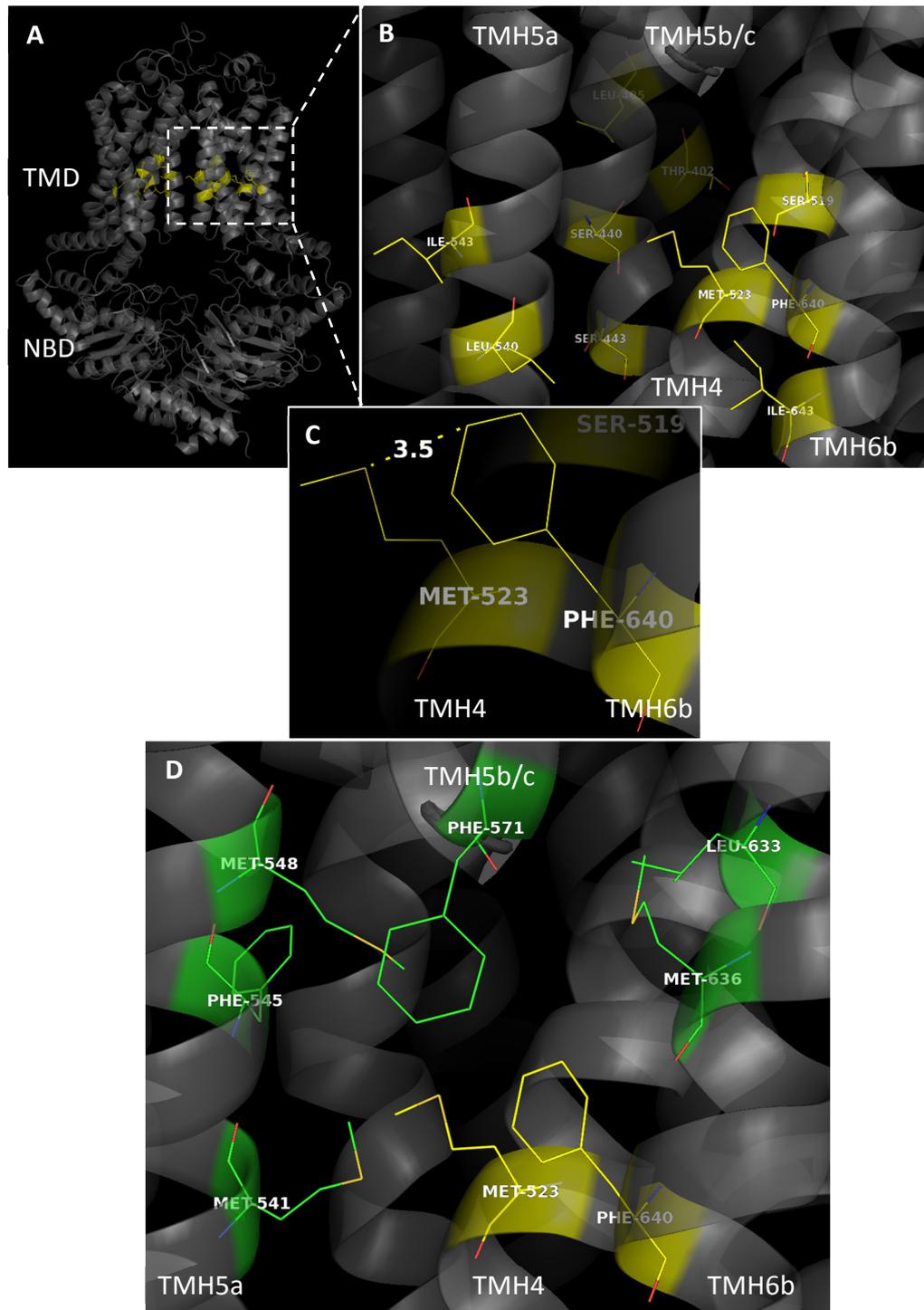


Figure 4.1 Structural mapping of lateral slice residues and proposal of a drug binding pocket. **A.** whole ABCG2 model structure (László et al., 2016) with functionally significant lateral slice residues highlighted in yellow. **B.** A zoomed in region demonstrating the positioning of several lateral slice residues to a lipid exposed pocket. **C.** enlarged image of residues M523 and F640 (gain in function when mutated to alanine) highlighted in ‘pocket’ demonstrating a spatial proximity of 3.5Å **D.** residues highlighted (green) that appear to line the pocket and are hypothesised to form part of a drug binding region. Figure was produced using PyMOL© (Schrodinger, LLC), with the coordinates of ABCG2 obtained from <http://abcg.hegelab.org>.

4.2 Construct generation

The mutants were generated by site-directed mutagenesis based on the *in vitro* amplification of DNA by PCR (as described in Figure 3.4 and methodological details in section 2.1) in the pcDNATM-3.1/zeo(+)_sfGFP_ABCG2 vector.

The newly generated constructs were digested with PvuII, to confirm the correct plasmid size (Figure 4.2 A.). To confirm the incorporation of the specific mutation and that no further mutagenesis had occurred, the plasmid was sent for Sanger sequencing across the entirety of ABCG2 (example in Figure 4.2 B., example of full sequence in Appendix 1). The sequencing also confirmed the incorporation of the sfGFP tag N-terminal to the ABCG2 gene sequence.

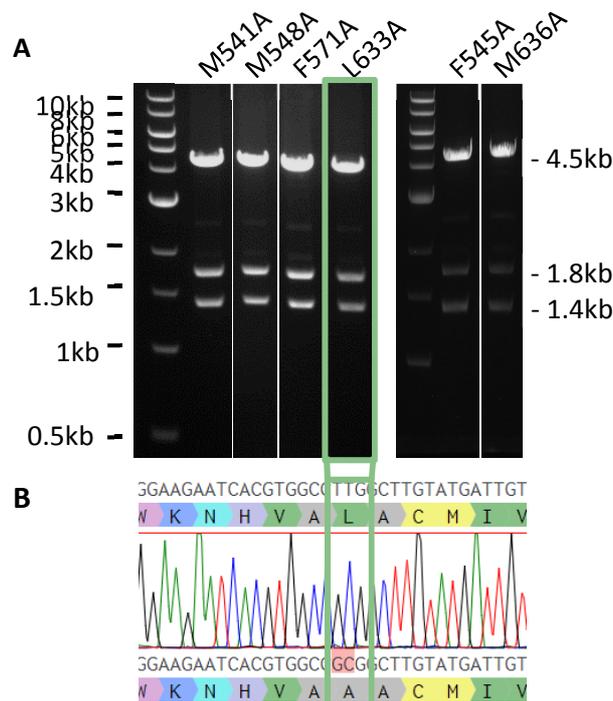


Figure 4.2 Confirmation of ABCG2 “binding pocket” mutant construct generation. A. Electrophoretic gel of DNA fragments digested with PvuII restriction endonuclease confirms correct plasmid size by a distinct banding pattern. Breaks in the gel result from the removal of lanes for other “putative” constructs for this set of mutants. **B.** Example chromatographic sequencing data for mutant L633A demonstrates a leucine (TTG) to alanine (GCG) mutation with high quality trace sequence.

4.3 Construct expression and localisation

Stable cell lines (HEK293T) were generated for each of the new mutations as described in section 2.2.4. For M548, the stable cell line was very difficult to maintain, with ABCG2 expression declining within 10 passages. Time constraints precluded an investigation into the reason for this. For all experiments with this mutant, stably transfected cells were recovered from liquid nitrogen stocks and used within 5-8 passages. The expression characteristics of the 'binding pocket' mutants were confirmed by western blot of cell lysates (10µg). M541A, F545A, F571A, and M636A all demonstrated an immunoreactive band at around 100kDa, consistent with sfGFP tagged wild type ABCG2 (Figure 4.3 B). M548A and L633A also demonstrated an immunoreactive band at 100kDa. However, the expression was greatly reduced, which could be indicative of reduced protein stability and/or a potential structural perturbation. L633A also demonstrated moderate expression of a second immunoreactive band (approx. 90kDa) (Figure 4.3 B), this banding seems to be of an intermediate weight and may be representative of partial glycosylation (as was seen for a previous mutant, I573A (Haider et al. 2015)).

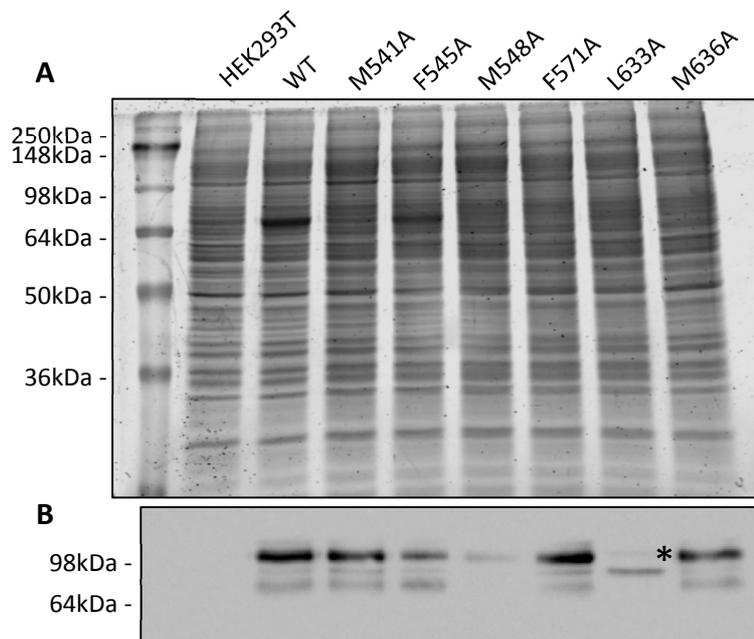
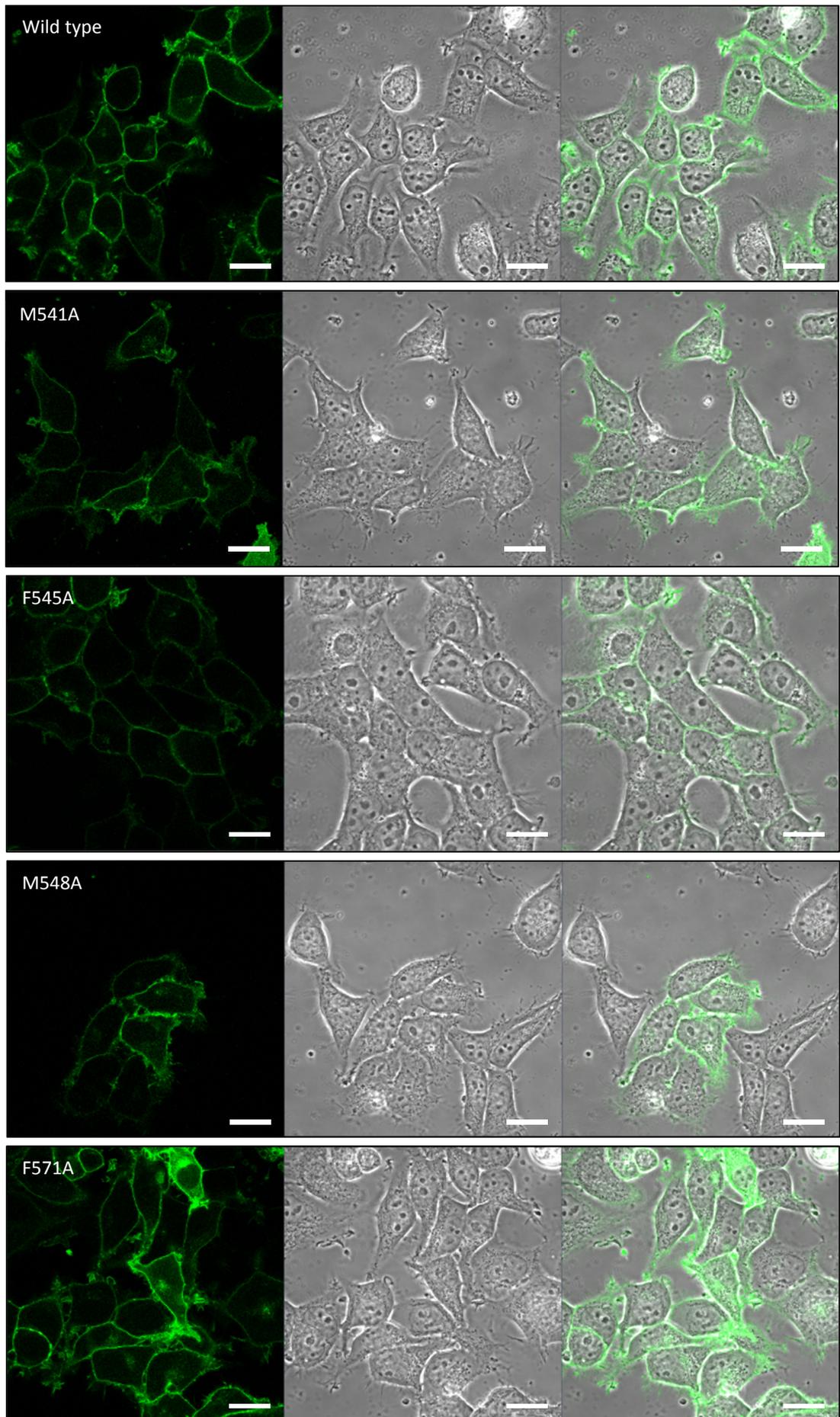


Figure 4.3 Expression of ABCG2_{WT} and mutant variants stably expressed in HEK293T cells. Cell lysates (20µg) were prepared from stably transfected HEK293T cells expressing mutant ABCG2. **A.** 10% SDS-PAGE acrylamide gel instant blue stained for total protein expression demonstrated equal protein loading. **B.** western blotting against ABCG2 shows an immunoreactive band at 100kDa, consistent with sfGFP-tagged ABCG2. M548A has a reduced expression over the wild type protein, whereas, L633A demonstrated an immunoreactive band at 90kDa (indicated with an asterisk), consistent with partial glycosylation (Haider et al., 2015).

As two of the ‘binding pocket’ mutations (M548A and L633A) showed a variation in the expression characteristics by western blot (Figure 4.3), it was necessary to validate that the mutations could localise correctly to the plasma membrane. Confocal microscopy was used to assess the localisation of mutant protein by exciting the sfGFP fluorophore fused N-terminal to ABCG2. Figure 4.4 below, demonstrates that all mutant protein was able to localise to the plasma membrane, comparable to wild type, with the exception of L633A, that was intracellularly retained. Upon closer inspection L633A demonstrated a reticulated patterning in its localisation which may suggest, combined with the maturation status of the protein, that it affects protein structure and/or folding (Figure 4.4, (Haider et al., 2015)).



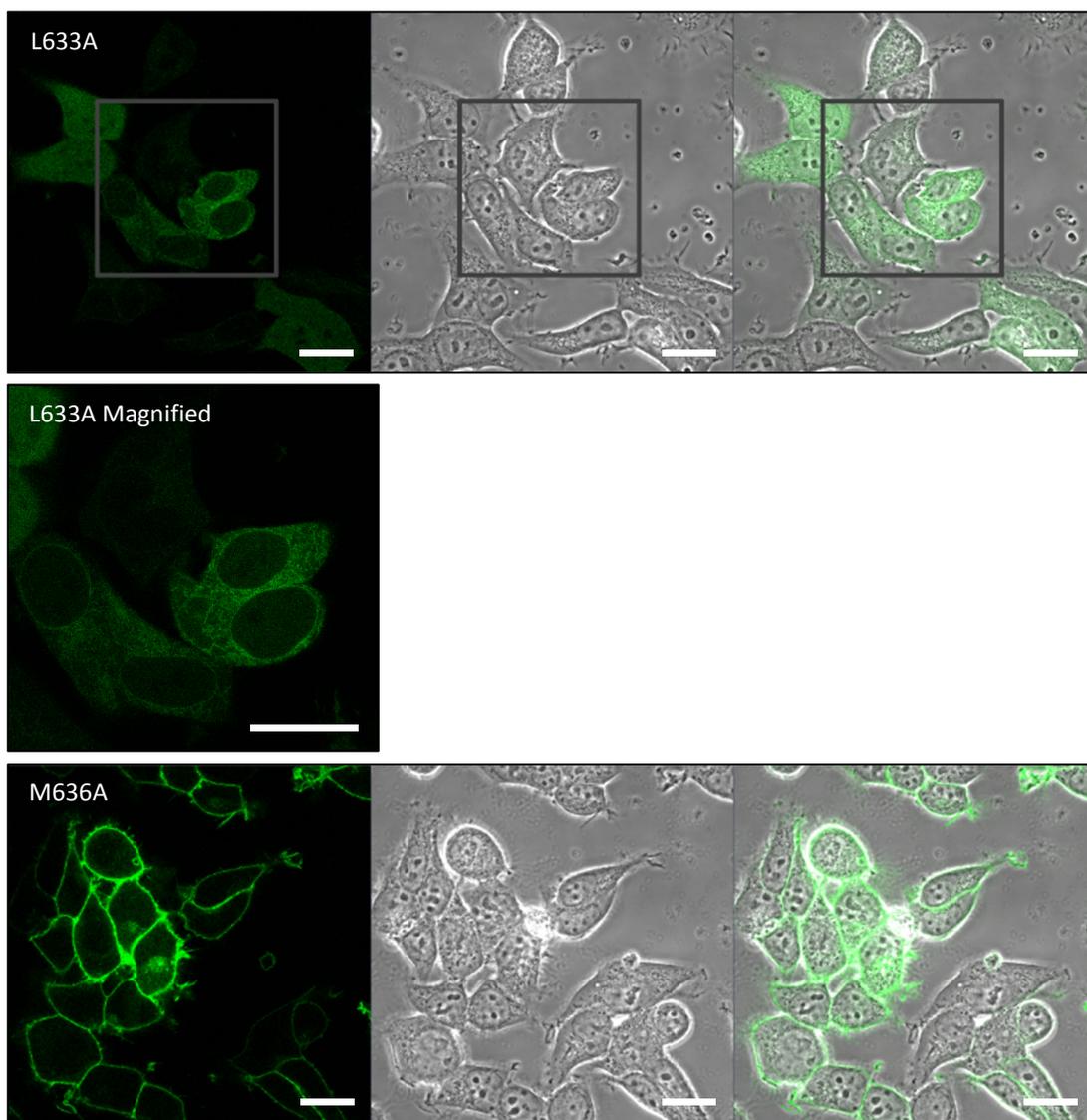


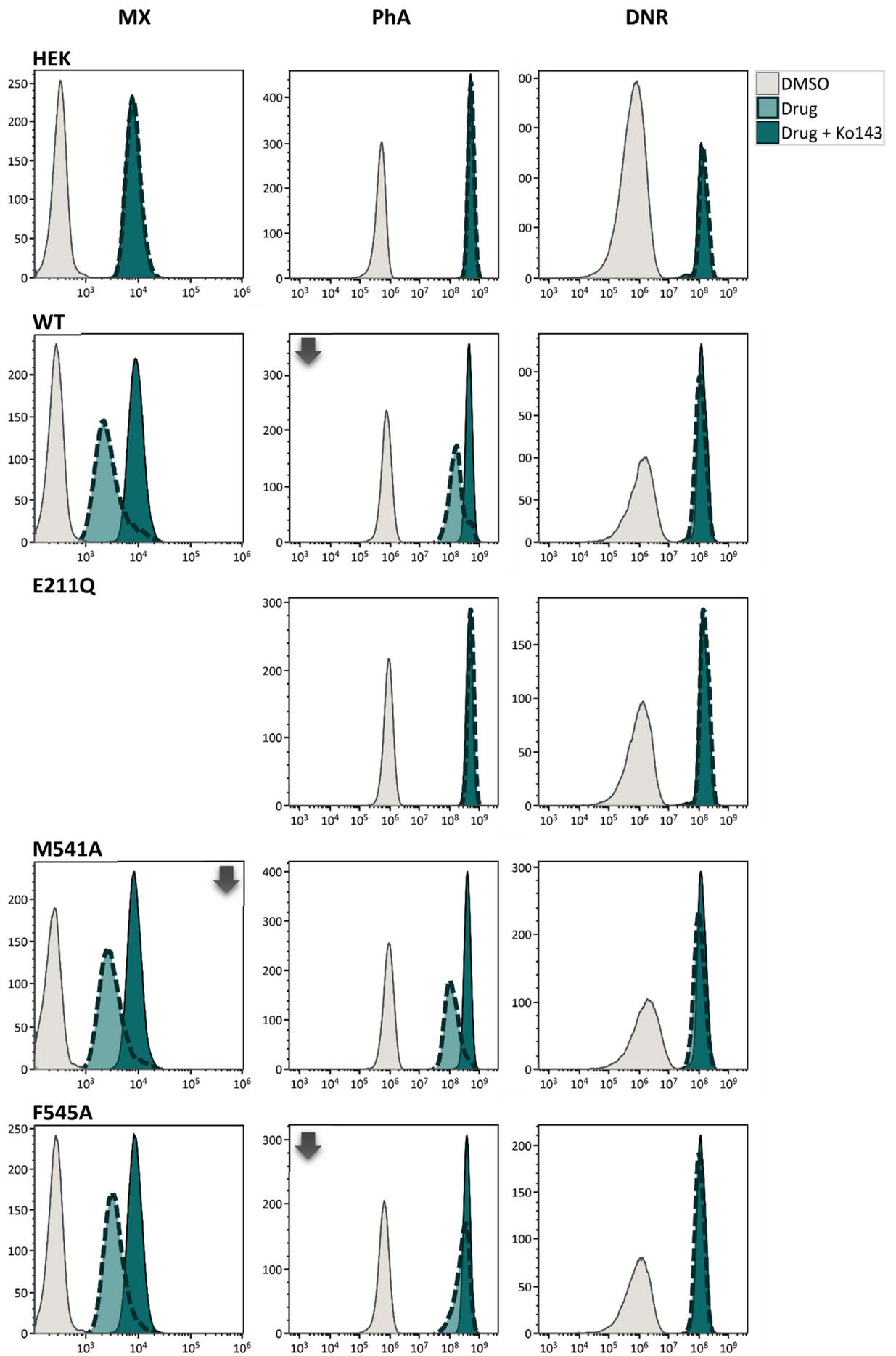
Figure 4.4 Cell surface localisation of ABCG2 ‘binding pocket’ variants stably expressed in HEK293T cells. Stably transfected HEK293T cells expressing ABCG2 mutant variants were visualised for protein localisation by confocal microscopy. All mutants were able to demonstrate plasma membrane localisation bar L633A which demonstrated intracellularly retained protein. On close inspection (zoomed image) the protein appears to have a striated patterning, suggestive of a trafficking issue. Visualisation of cells was conducted on an LSM710 confocal laser scanning microscope (Zeiss, Jena, Germany) quipped with a Plan-Apochromat 63x/1.40 Oil Ph3 DIC M27 objective and argon laser. For detection of sfGFP-ABCG2, a 2% argon laser power was used with excitation set at 488nm and emission collected between 500-530nm (FITC channel). Data are representative of one of three independent repeats. Presented dataset was acquired with identical acquisition settings to allow comparison of expression levels. Scale bar is representative of 20µm.

4.4 Functional characterisation

Having ascertained that the mutants retained plasma membrane targeting, with the notable exception of L633A, it was important to assess the functional impact of the single point mutations on the ability to efflux drug substrate. To be able to obtain an unbiased measure of drug efflux flow cytometry was used. Flow cytometry allowed for the collection of multiparametric data (i.e. GFP fluorescence (protein expression) and drug fluorescence) to allow for expression-based gating and normalisation of drug efflux (as discussed previously, Chapter 3).

As L633A was not trafficked to the plasma membrane it was excluded from functional analysis. All other ABCG2 mutant variants showed drug export function that was statistically different (one-way ANOVA with a Dunnett's multiple comparisons to WT efflux, $P \leq 0.05$) from the ABCG2_{WT} protein, which was investigated in parallel to these new mutants. M541A differed from WT only in its ability to export mitoxantrone, which was reduced by about half (1.58 ± 0.33 compared to 3.51 ± 0.60). F571A also differed from WT only in its ability to export MX (2.31 ± 0.3 compared to 3.51 ± 0.60). F545A retained WT levels of MX transport but reduced pheophorbide A transport by over a half (0.91 ± 0.36 compared to 2.39 ± 0.91). M636A demonstrates a reduction in MX export (2.30 ± 0.3 compared to WT 3.51 ± 0.60) and also shows reduced PhA transport ($P \leq 0.05$ based on mean difference in relative efflux, 1.24 ± 0.25 compared to 2.39 ± 0.91). M548A differs from WT across all tested compounds; export of mitoxantrone is increased by about a third (5.08 ± 0.44 compared to 3.51 ± 0.60), in contrast the transport of PhA is reduced by about a half (0.99 ± 0.35 compared to 2.39 ± 0.91). Of particular

interest, M548A (like M523A and F640A in Chapter 3) gained the ability to export daunorubicin, a non-native drug substrate (1.60 ± 0.11 compared to (WT) 0.33 ± 0.04), such data combined would suggest a pivotal role for M548A in drug recognition Figure 4.5 and Table 4.2.



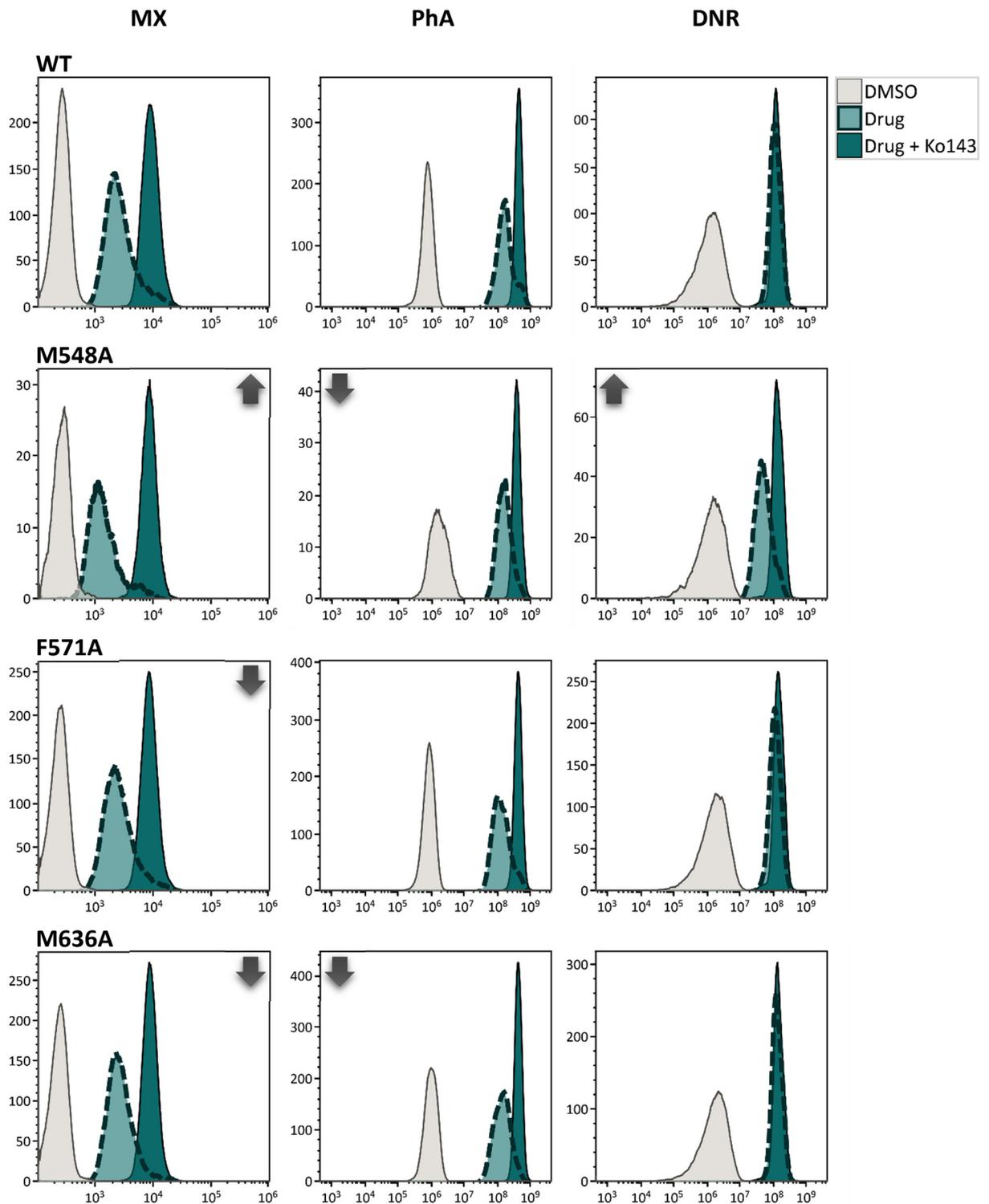


Figure 4.5 Function of ABCG2 wild type and ‘binding pocket’ mutant variants stably expressed in HEK293T cells. Data is representative of Kaluza© single fluorescence channel overlay histograms for a typical dataset of $N \geq 3$ independent experiments for mitoxantrone (MX), pheophorbide A (PhA), and daunorubicin (DNR) (gated for high expression). Statistical significance (arrows, $P \leq 0.05$) were determined using a one-way ANOVA analysis of expression normalised relative efflux values (Table 4.2) followed by a Dunnett’s multiple comparisons test against wild type drug efflux.

	Mitoxantrone		Pheophorbide A		Daunorubicin	
	Mean	SEM	Mean	SEM	Mean	SEM
HEK	0.08***	0.07	0.00***	0.09	0.00***	0.00
WT	3.51	0.60	2.39	0.91	0.33	0.04
E211Q					0.15	0.05
M541A	1.58***	0.33	1.51	0.17	0.21	0.02
F545A	2.81	0.67	0.91**	0.36	0.39	0.06
M548A	5.08***	0.44	0.99**	0.35	1.60***	0.11
F571A	2.31*	0.30	1.85	0.15	0.27	0.02
L633A	0.59***	0.07	0.04***	0.15	0.05	0.05
M636A	2.30*	0.49	1.24*	0.25	0.18	0.03

Table 4.2 Function of ABCG2 wild type and 'binding pocket' mutant variants stably expressed in HEK293T cells.

Data is representative of the mean \pm SEM for normalised relative drug efflux for mitoxantrone (MX), pheophorbide A (PhA), and daunorubicin (DNR) from at least three independent experiments. Statistical significance (* $P \leq 0.05$ ** $P \leq 0.01$ *** $P \leq 0.001$) was determined using a one-way ANOVA analysis followed by a Dunnett's multiple comparisons test against wild type drug efflux.

4.5 Summary

This chapter aimed to use a new homology model for ABCG2, based on the crystal structure of ABCG5/G8, to interpret the functional effect of the 'lateral slice' residues.

- Structural mapping of the lateral slice located residues T402, L405 (from TMH1), S440, and S443 (from TMH2) at the entrance to a side chain lined cavity, sandwiched between predominantly TMHs 3 and 4 (involvement of TMH3 in drug binding and translocation currently being investigated by Parth Kapoor, a PhD student from Dr Ian Kerr's research group).
- The remaining lateral slice residues resided within a lipid exposed cavity. Due to the functional perturbation of these residues upon mutation to alanine, it was hypothesised that this cavity could be a multidrug binding cavity ('binding pocket') for allocrite translocation via the lipid interface.
- To test this secondary hypothesis a further 6 residues lining this 'binding pocket' were mutated and characterised as previously described.
- One of the 'binding pocket' mutants L633A demonstrated immature (reduced glycosylation) protein expression, which, when analysed by fluorescence based confocal microscopy, showed intracellular reticular retention of protein. The results were suggestive of a structural and/or folding defect.
- All other mutants were able to express full length protein (confirmed by western blot) and were also predominantly targeted to the plasma membrane.

- Mutants M541A and F571A demonstrated a reduced efflux capacity for mitoxantrone only, whilst F545A demonstrated a reduced efflux for pheophorbide A only.
- M636A demonstrated a reduced efflux for both mitoxantrone and pheophorbide A.
- Mutant M548A showed particularly interesting efflux parameters, demonstrating an enhanced efflux for mitoxantrone and a gain in efflux for daunorubicin, like the lateral slice mutant F640A. However, M548A also resulted in a reduced efflux for pheophorbide A, suggestive of a complex role for M548 in drug recognition.
- The results suggest that some residues may coordinate the specific binding of some allocrites but not others, thus enhancing the possibility of the lipid exposed cavity being a multidrug binding site. In order to interpret the data any further (i.e. assess direct drug binding) the translocation cycle needs to be broken down into its individual catalytic steps.
- In the next chapter I will try and elucidate these steps using a membrane-based ATPase assay. I will also use the structural model to perform molecular drug docking to assess the feasibility of drug binding to the lipid exposed 'drug binding pocket' and build a picture of how the residues could be interacting with allocrite.

Chapter 5 Enhancing the understanding of ABCG2-mediated allocrite recognition and translocation

5.1 'Gain in function' mutation still requires the energy from ATP hydrolysis to drive the efflux of drug substrates

Some ABCG2 mutants demonstrated enhanced allocrite transport. For example, F640A demonstrated an increased efflux of mitoxantrone (approx. 2-fold difference compared to wild type) and a 7-fold gain in efflux capacity for the non-native drug substrate, daunorubicin (Chapter 3, Figure 3.18). To aid in the interpretation of the mechanistic basis of the effects observed, it is imperative to confirm that the tight coupling of ATP hydrolysis and allocrite transport is preserved. Indeed, for some ABC transporters (e.g. MsbA and LmrA) it has been shown that allocrite transport can occur in the absence of the nucleotide binding cassettes, instead functioning via a proton coupled mechanism (Singh et al., 2016; Venter et al., 2003).

To confirm ATPase dependent transport, the plasmid containing the catalytically inactive Walker B mutation, E211Q was introduced into ABCG2_{F640A} through sub-cloning, via a double restriction endonuclease digest with Bsu36I and PpuMI (Figure 5.1 A/B.). E211 is essential for coordinating the hydrolytic attack of the γ -phosphate of ATP. Therefore mutagenesis to a glutamine produces a catalytically inactive mutant, as shown by Hou et al., 2009 and Moody et al., 2002 (described in Chapter 1).

The correct plasmid size and insertion of E211Q containing gene fragment was confirmed by restriction digest with EcoNI and PflmI. The E211Q mutation contains an insertion of the restriction endonuclease EcoNI which when combined with PflmI cuts at distinct sites to give a banding pattern of approx. 6.5kb and 1kb (Figure 5.1 C.).

In order to study the expression and functional characteristics of the ABCG2_{E211Q/F640A} double mutant the plasmid was transfected into HEK293T cells and selected for stable expression using the cytotoxic agent ZeocinTM. Expression of the double mutant was confirmed by western blotting as a single immunoreactive band at approximately 100kDa, consistent with sfGFP tagged wild type ABCG2 and the corresponding single mutations, albeit at a higher expression level (Figure 5.1 D.). ABCG2_{E211Q/F640A} mutant protein was also shown to be located to the plasma membrane (data not shown).

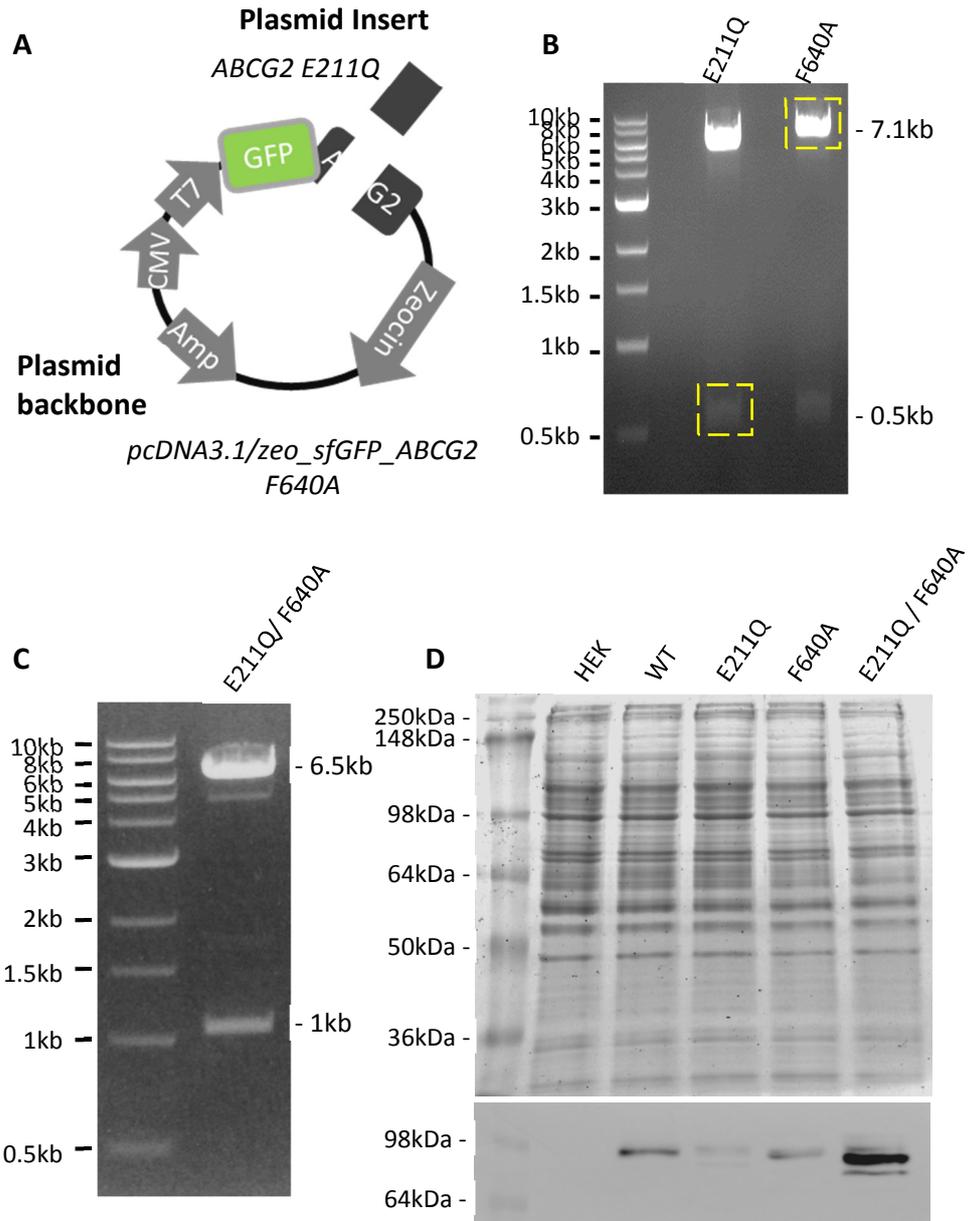


Figure 5.1 Generation and expression of E211Q F640A double mutant. **A.** Schematic representation of restriction endonuclease digest, with insertion of E211Q containing gene fragment into the plasmid backbone containing the F640A mutation. **B.** Restriction digest products were separated on a 1% (w/v) agarose gel; E211Q present in a 0.5kb fragment and F640A backbone were extracted and ligated prior to transformation and DNA extraction. **C.** Confirmation of double mutant construct by restriction endonuclease digest with EcoNI and PflMI which produces a diagnostic banding pattern of approx. 6.5kb and 1kb. **D.** Protein expression analysed by gel electrophoresis and western blotting demonstrates equal loading of total protein (upper part) and a single immunoreactive band for E211Q/F640A double mutant and its control counterparts at 100kDa (lower part), consistent with sfGFP-tagged ABCG2.

Drug export and protein expression were measured by flow cytometry as previously described (section 2.5 and Chapter 3). ABCG2^{F640A} demonstrated a slightly increased mitoxantrone efflux (17 vs. 13 for wild type (Figure 5.2)), which is broadly consistent with previously acquired data (Figure 3.17 and Figure 3.18). The catalytically inactive single mutant control, E211Q, demonstrated no Ko143 inhibitable efflux of mitoxantrone, as expected. When the two mutations were combined ABCG2 behaved like the E211Q mutation, demonstrating a complete absence of mitoxantrone efflux. This data is consistent with F640A still requiring ATP hydrolysis to facilitate drug efflux.

As the F640A mutation previously demonstrated a gain in the ability to efflux a non-native drug substrate, daunorubicin, this was also functionally investigated for the E211Q/F640A double mutant. The F640A single mutant demonstrated a gain in efflux (7-fold change in comparison to wild type) consistent with previous data. The double mutant produced results consistent with mitoxantrone allocrite efflux, the combined mutation rendered ABCG2 functionally inactive.

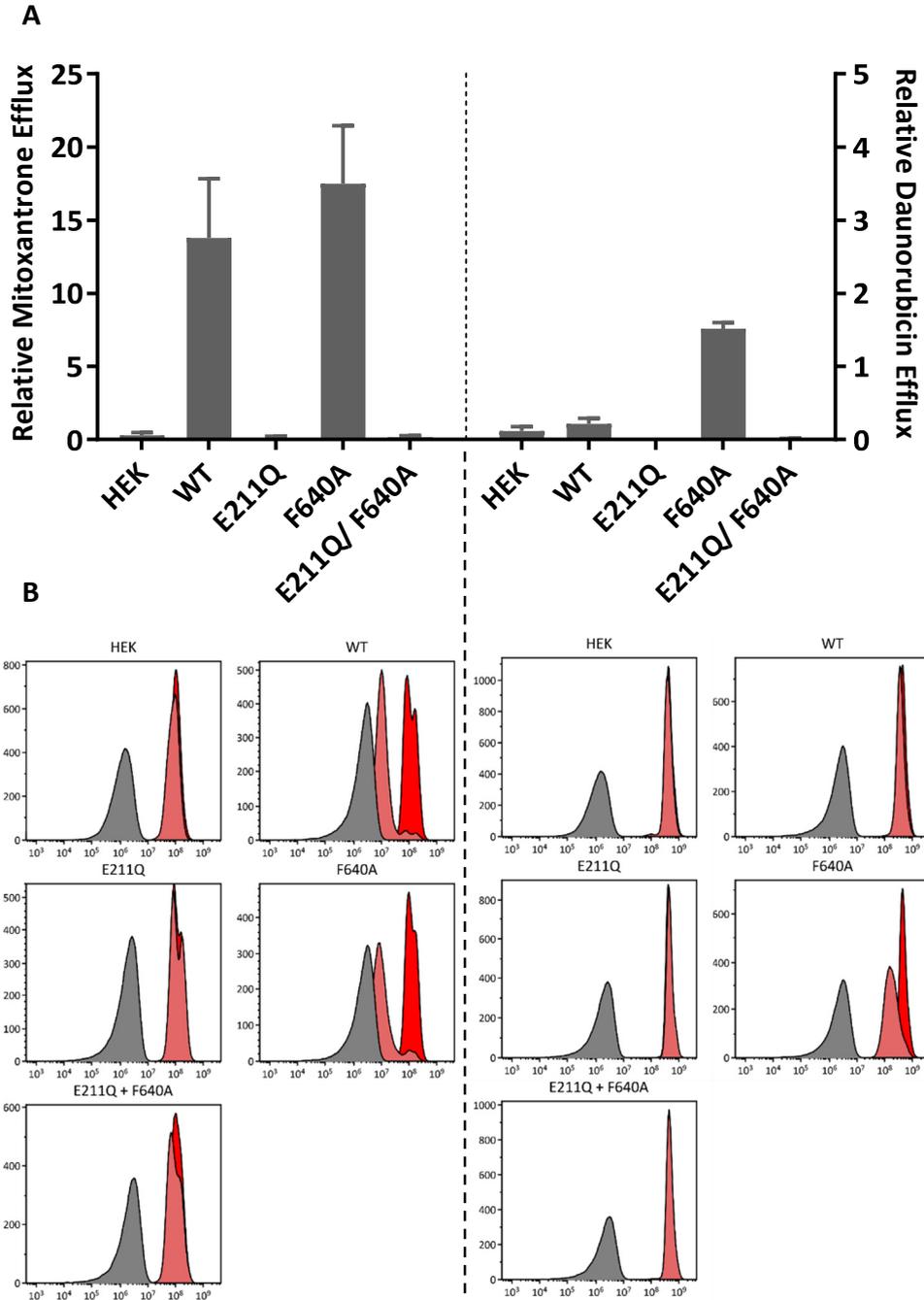


Figure 5.2 E211Q/F640A double mutant still requires the energy from ATP hydrolysis to drive efflux. A. Flow cytometric drug accumulation assays demonstrate that both wild type and F640A are able to efflux mitoxantrone, whereas only F640A can efflux daunorubicin ($P \leq 0.001$, when compared to wild type), consistent with prior data. The E211Q/F640A double mutant behaves functionally like the E211Q mutation, demonstrating that F640A required the energy from ATP hydrolysis to efflux drug substrate. **B.** Flow cytometric overlay histograms of data presented in A., a leftward shift in fluorescence in the drug only sample (light red) compared to the inhibited (dark red) sample demonstrates the level of drug efflux. Data are representative of N=3 independent repeats.

In this type of experiment flow cytometry can report that ATP hydrolysis is required for allocrite efflux, but the catalytic cycle is made up of many steps (as described in Chapter 1). Understanding exactly how allocrite binding and/or translocation is coupled to each step within the catalytic cycle will require *in vitro* investigations of protein, either in isolated membranes or purified and reconstituted into proteoliposomes. Preliminary *in vitro* investigations with isolated membranes (data not shown) indicated that for any such investigations the protein expression level would have to be improved, especially for low expressing ABCG2 mutants. Therefore, the next section presents investigations conducted to increase the ABCG2 protein expression level.

5.2 Sodium butyrate increases ABCG2 expression of *in vitro* investigation

Sodium butyrate is a histone deacetylase (HDAC) inhibitor which has been reported to induce gene expression in cultured mammalian cells (Davie, 2003). Inhibition of HDACs only affects 2% of mammalian genes (Davie, 2003), therefore to test its effectiveness on ABCG2 gene expression dose response assays were conducted. sfGFP-tagged wild type ABCG2 and the low expressing mutant M523A, stably expressed in HEK293T cells, were seeded into a 96-well plate and subjected to increasing concentrations of sodium butyrate (0-10mM). Cells were analysed (fluorescence based confocal microscopy) immediately after sodium butyrate addition and again 24 hours later. This demonstrated a concentration dependent increase in sfGFP-ABCG2 expression, with no apparent protein localisation or cytotoxic effect (Figure 5.3 A/B).

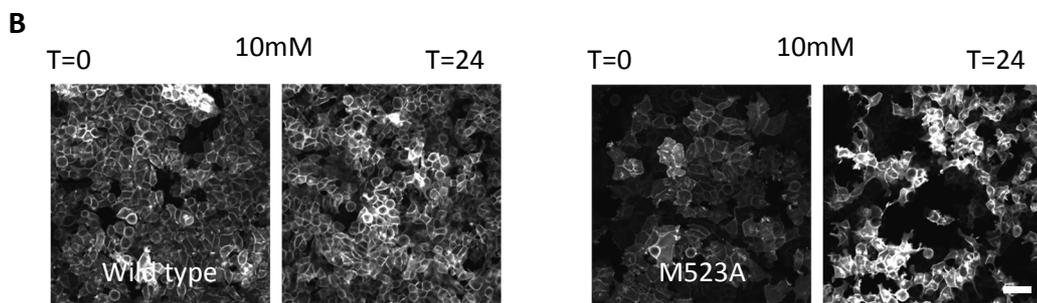
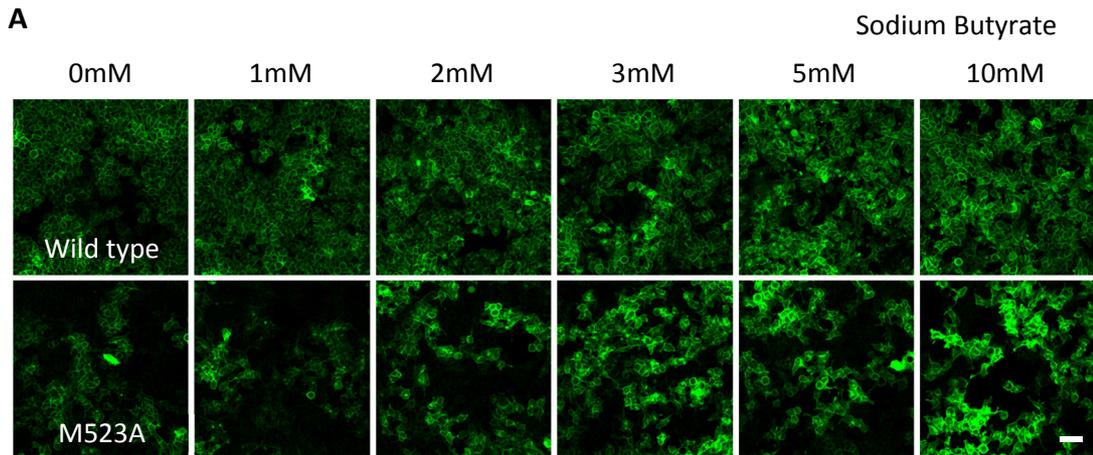


Figure 5.3 Sodium butyrate increases expression of ABCG2 in stably transfected HEK293T cells. HEK293T cells stably transfected with sfGFP-tagged ABCG2 variants were incubated with the indicated concentrations of sodium butyrate, and analysed by confocal microscopy at T=0 and T=24 hours. A concentration dependent increase in sfGFP-fluorescence under identical image capture conditions was observed, consistent with elevated ABCG2 expression. Images are representative of data collected at T=24. B. At T=0 and T=24 comparison to demonstrate no obvious signs of cytotoxicity. ImageXpress (IX) Ultra confocal plate reader (Molecular Devices), using a plan-apochromat 40× objective, with excitation wavelength of 488 nm and emission bandpass filter of 525/50 nm. Data are representative of N=3 independent repeats. Scale bar represents 50µm.

The interpretation of this data was qualitative due to the lack of an appropriate live-cell nuclear stain; Hoechst 33342 being unsuitable as it is an allosteric inhibitor of ABCG2. However, flow cytometry was used to gain quantitative data for several low expressing ABCG2 mutant cell lines. Flow cytometry demonstrated a rightward shift in sfGFP fluorescence, associated to protein expression, in a sodium butyrate concentration dependent manner (Figure 5.4 A). Sodium butyrate at 10mM enhanced ABCG2 protein expression by up to 80% (in most cases) in comparison to

untreated cells (Figure 5.4 B.). It was subsequently decided that cells would be incubated with 10mM sodium butyrate 24 hours prior to cell harvest and membrane isolation to increase protein yield and subsequently increase the signal to noise ratio in downstream experiments.

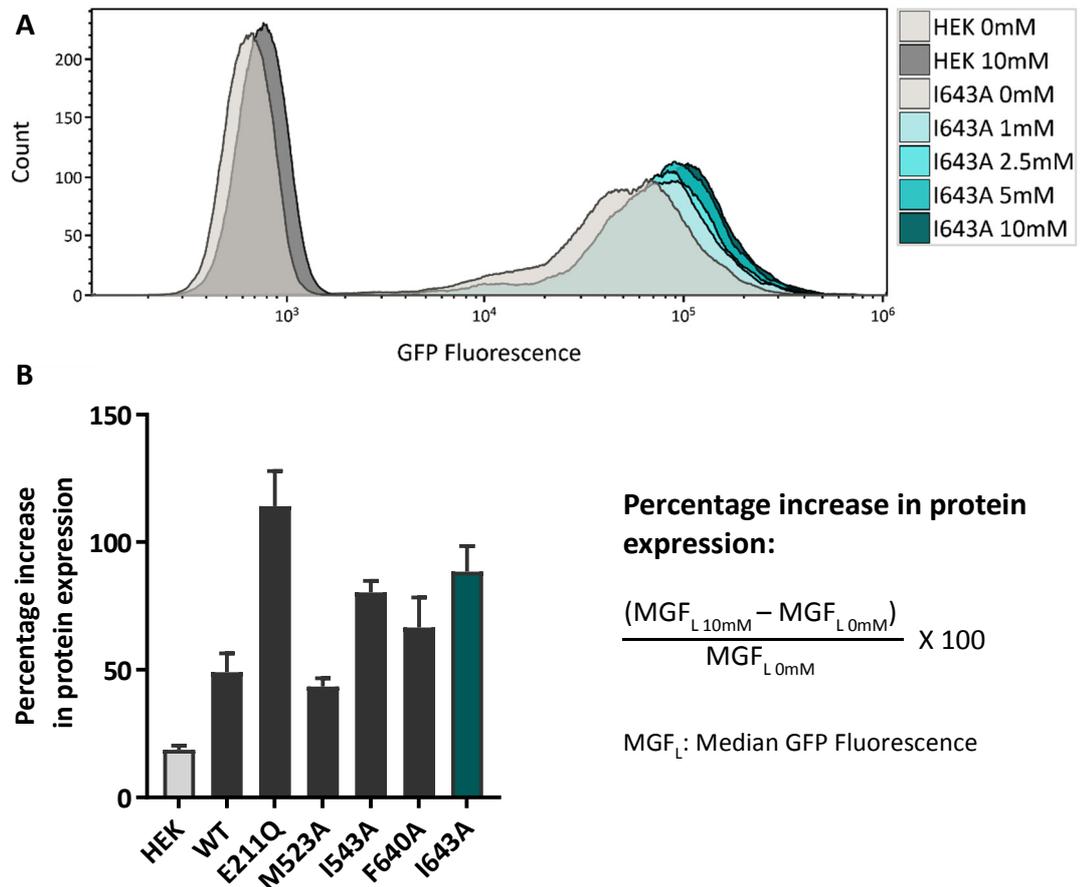


Figure 5.4 Sodium butyrate increases expression of ABCG2 in stably transfected HEK293T cells. HEK293T cells stably transfected with sfGFP-tagged ABCG2 variants were incubated with increasing concentrations of sodium butyrate. **A.** Example of flow cytometric data for I643A with increasing concentration of sodium butyrate being represented by darker shades of teal, the rightward shift in fluorescence is representative of a concentration dependent increase in protein expression (sfGFP-ABCG2 fluorescence). **B.** Graphical representation of percentage increase in protein expression levels from 0mM versus 10mM sodium butyrate. Data demonstrates at least 50% increase in protein expression. Data representative of N≥3 independent experiments.

5.3 Preservation of ABCG2 function in isolated membranes

Increased cellular expression of ABCG2 'opens the door' for *in vitro* analysis of the protein. Such analysis requires the optimisation of buffer conditions to preserve ABCG2 function. Optimisation was assessed with a vesicular transport assay (VTA) (Methods, section 2.7.5 and overview in Figure 5.5). Due to the hydrophobic nature of many ABCG2 drug substrates, non-specific vesicular accumulation can occur resulting in a decreased signal to noise ratio. For this reason, a water soluble ABCG2 drug substrate, Lucifer yellow, was preferentially used in the assay (Hegedűs et al., 2009). Lucifer yellow is a fluorescent drug substrate with an excitation/ emission of 430nm and 538nm respectively. The fluorescence spectra of Lucifer yellow significantly overlaps with the fluorescence spectra of the GFP tag in the construct used in the current investigations to date. A construct of ABCG2 (His-SNAP-ABCG2) that was developed for protein purification in parallel studies (Horseý, 2018), was used as an alternative.

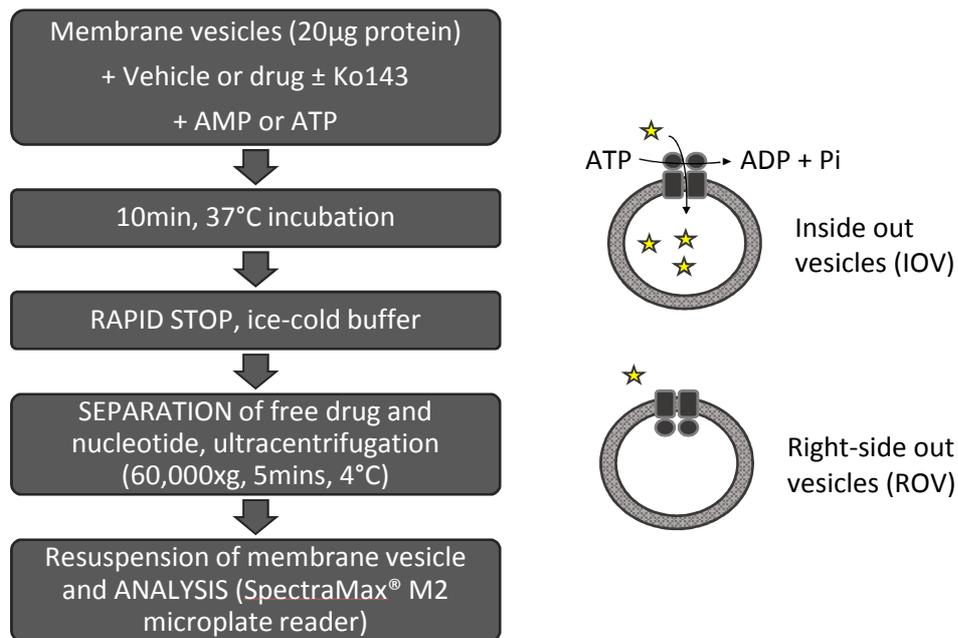


Figure 5.5 Vesicular transport assay overview. Vesicles of an inside-out configuration will accumulate allocrite in an ATP-dependent manner (as seen on the right-hand side of the figure) expected to be inhibited by Ko143. Right-side out vesicles would not transport drug (lucifer yellow) as the NBDs could not access the ATP in sealed vesicles.

The initial attempts of the VTA assay in a 10mM Tris, 250mM sucrose pH7.4 buffer composition demonstrated membrane destabilisation resulting in a visible precipitation (Figure 5.6 A.). Precipitation was further identified as being Mg^{2+} and ATP dependent. Trials were conducted to monitor the effect of increasing concentration of magnesium ($MgCl_2$ and $MgSO_4$) on membrane precipitation. This was analysed by both visual inspection of the size and consistency of the precipitant after pelleting by brief centrifugation and by western blotting, whereby both the supernatant and pellet were analysed for ABCG2 content. This resulted in a threshold concentration of 5mM magnesium, above which significant membrane (i.e protein) precipitation occurred (Figure 5.6 B). Therefore, a maximum concentration of 5mM $MgSO_4$ was carried forward in further assays, as it showed

the least amount of membrane precipitation for a concentration of magnesium that would support ATP hydrolysis.

Membrane precipitation upon ATP addition was investigated in the same manner and was attributed to ATP acidity. ATP was made fresh for every experiment, and due to the low volume needed, the ATP solution itself could not be pH corrected. To combat this, ATP was dissolved in a solution with a higher buffering capacity (50mM Tris) and was used at a maximal concentration of 5mM (matching the concentration of magnesium) throughout subsequent assays. Figure 5.6 demonstrates the precipitation before (C.) and after (D.) buffer optimisation. Parallel to these investigations the membranes/protein were subjected to thermostability assays (Prometheus, NanoTemper), presented in Figure 5.6. The outcome of these investigations resulted in a final buffer for *in vitro* examination of ABCG2 function consisting of; 50mM Tris, pH8, 250mM sucrose (Figure 5.6 E, circled) with 5mM Mg.ATP added for nucleotide dependent effects. The addition of NaCl (150mM) had little effect on protein stability (Figure 5.6 E, circled, data points are representative of 50mM Tris pH8, 250mM sucrose \pm 150mM NaCl) but demonstrated a marginal effect on function (Figure 5.7 A.) so was subsequently included. In VTA experiments, in this optimised buffer, ABCG2 function was confirmed by ATP dependent, Ko143 inhibited accumulation of Lucifer yellow into membrane vesicles (Figure 5.7 B.).

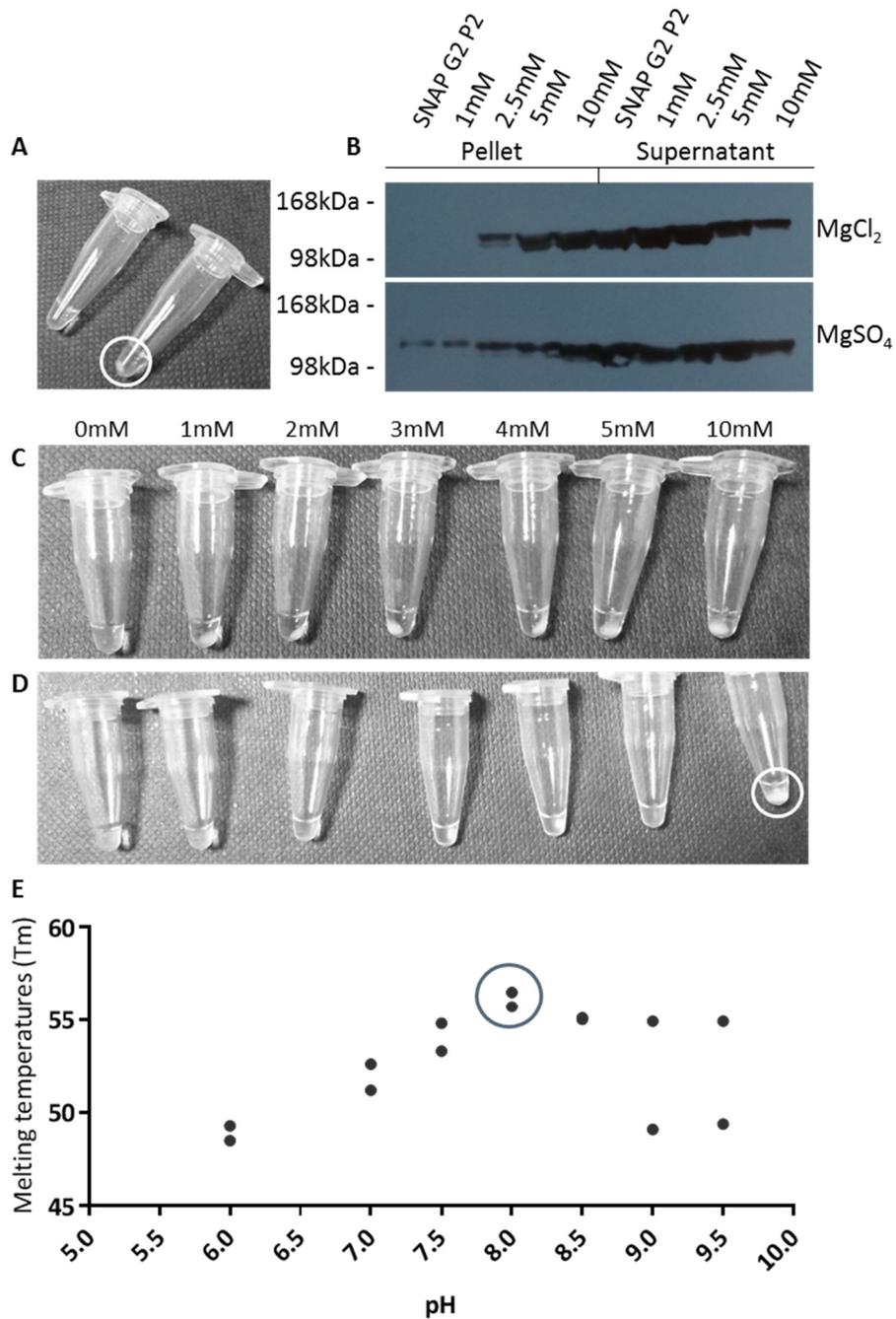


Figure 5.6 Optimisation of buffering capacity and stabilisation of ABCG2 in isolated membranes. **A.** Visualisation of membrane precipitation (circled) with original buffer conditions. **B.** Western blot analysis of pelleted precipitant and resultant supernatant (20µg) demonstrated that precipitation was attributed to increasing magnesium concentrations. **C/D.** Visualisation of precipitation before (C) and after (D) buffer optimisation for increasing concentrations of Mg²⁺ and ATP. Precipitation was only visualised at 10mM (Mg²⁺ and ATP) (circled) in the optimised buffer conditions (50mM Tris, 250mM sucrose, 150mM NaCl, pH 8: TSN). **E.** Prometheus data investigated the effect of different buffer compositions (represented by each data point) on protein stability, buffers of pH 8 (circled) demonstrate the highest protein stabilising effect.

Relative Lucifer Yellow Accumulation:

Ratio of lucifer yellow fluorescence defined as: $F_{LY\ ATP} / F_{LY\ AMP}$

F_{LY} : Lucifer yellow fluorescence

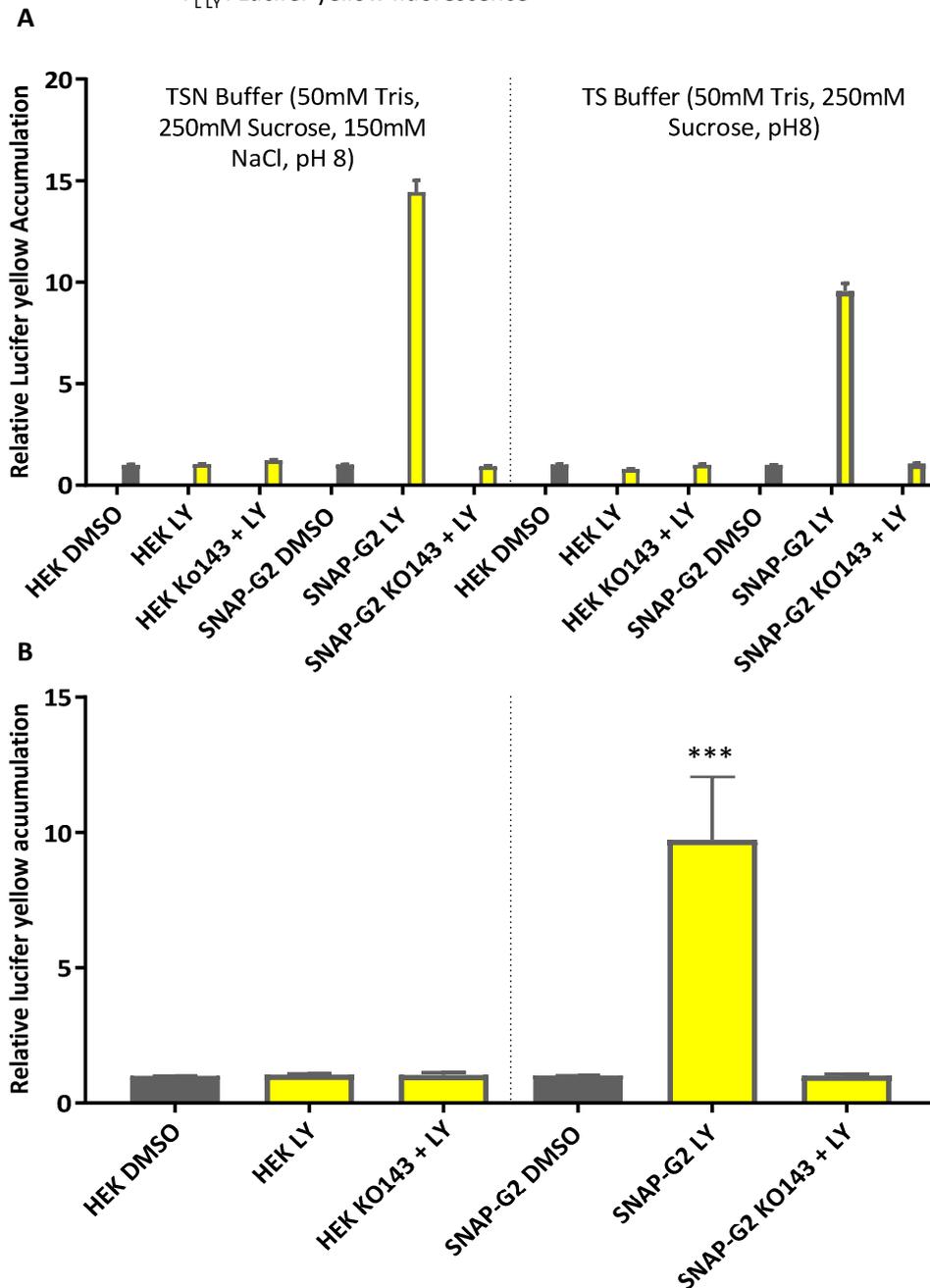


Figure 5.7 ABCG2 function is conserved in isolated membrane vesicles. A. vesicular transport assay with the optimised buffer conditions from protein stability assays demonstrates an ABCG2 dependent fold change (10-15) in lucifer yellow (LY) accumulation. Data is representative of N=1 independent repeats, SEM is representative of multiple read points throughout a well of a 96-well plate. **B.** ABCG2 specific ATP-dependent Ko143 inhibitable accumulation of Lucifer yellow. VTA is representative of at least three independent repeats with optimised TSN buffer. Data were analysed using a one-way ANOVA with a Sidak's multiple comparisons test between the vehicle control (DMSO) and test conditions (LY and Ko143 plus LY) for each cell line. The asterisks denote a P-value of 0.001 for SNAP-tagged ABCG2 in the presence of LY compared to DMSO.

5.4 Function of ABCG2 in isolated membrane validated by ATPase assays

The vesicular transport assay (VTA) validates protein stability and functionality in a 'cell-free' environment. However, it provides no more information about the individual steps of the drug transport cycle than the in-cell flow cytometry-based assay. Moving forwards, the conditions highlighted by the VTA were applied to an ATPase assay. The assay used was a modified version of that originally presented by Ambudkar, 1998, whereby the release of inorganic phosphate (Pi) upon ATP hydrolysis is monitored using a colourimetric technique, first described by Chifflet *et al.*, 1988 (see methods section 2.7.6 for further details). However, many subsequent ABCG2 investigations have struggled to recapitulate this assay (e.g. Clark *et al.*, 2006; McDevitt *et al.*, 2008) and so implementation of an ATPase assay to investigate these mutants was not considered a trivial undertaking.

As with the flow cytometry and vesicular transport assays, this assay monitors an ABCG2 specific response by the addition of an ABCG2 inhibitor, Ko143. HEK293T membranes (i.e. not expressing ABCG2) were used to determine the background activity of endogenous ATPase proteins. The experiments presented investigated the parameters of the assay, and therefore the amount of Pi released was not quantified by a standard curve, instead raw absorbance values indicated ATP hydrolysis. HEK293T membranes in the presence of DMSO, Lucifer yellow and Lucifer yellow + Ko143 displayed a low level of Pi release. ABCG2 expressing membranes showed a much higher level of Pi release in the presence of either DMSO or Lucifer yellow, which was completely inhibited by Ko143 (Figure 5.8,

denoted by asterisk). The vehicle control also showed a high level of Pi release which suggested that the ABCG2 activity in membranes is already stimulated, potentially by an endogenous substrate or uncoupled from ligand transport. To confirm that the Pi release was ABCG2 dependent a further control was to employ the catalytically inactive mutant E211Q, which showed a Pi release profile the same as HEK293T membranes (Figure 5.8).

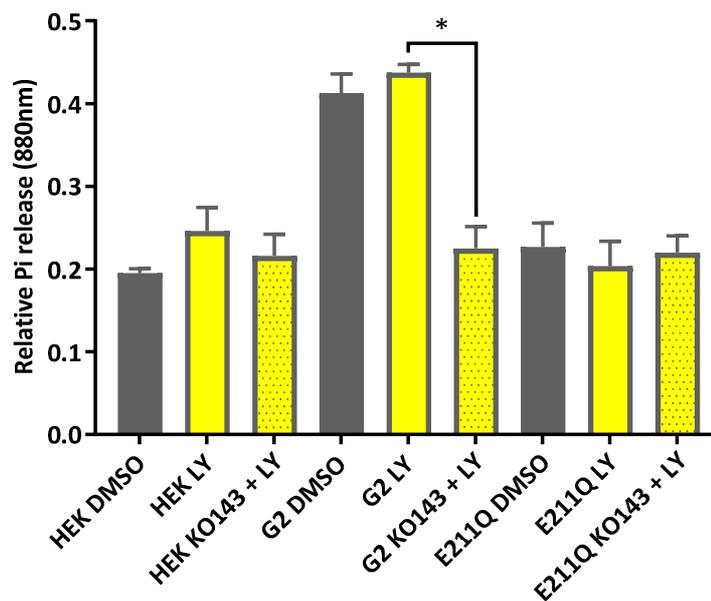


Figure 5.8 Functional assessment of sfGFP-ABCG2 expressed in HEK293T membranes. Membranes (20µg protein) were incubated with vehicle control (DMSO), or lucifer yellow only (100µM) or LY plus Ko143 (1µM), in the presence of 5mM Mg.ATP or Mg.ATP. Data are representative of relative Pi release define as the absorbance at 880nm (A_{880nm}) for samples incubated with ATP minus AMP (baseline correction) for $N \geq 3$ independent experiments. The results show a low-level Pi release for both the HEK and E211Q control membranes. The DMSO and LY ABCG2 samples demonstrate a high level of Pi release which is reduced to a control level when inhibited with Ko143 (denoted by asterisk), demonstrating ABCG2 specific Pi release (i.e. protein ATPase activity).

The similar ATPase activity for DMSO and Lucifer yellow treated membranes was not expected, retrospectively, analysis of ABCG2 expressing membranes in the absence of DMSO could confirm if DMSO is acting as a substrate. Typical ABC transporters show a 'basal' ATPase activity, that is stimulated by allocrites. Previous research for ABCG2 reported that several drug substrates have a more profound

effect on ATPase activity (Robey et al., 2003), including the alpha-1 adrenergic receptor antagonist; prazosin (PRA). The fluorescent derivative, BODIPY-prazosin, was also established to be a transport substrate of ABCG2 by flow cytometry (Figure 5.9).

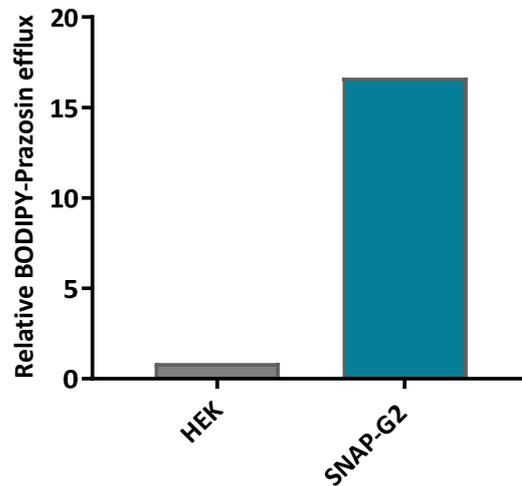


Figure 5.9 BODIPY-Prazosin is an allocrite of ABCG2. Results represent the fold change in intracellular (BODIPY-prazosin) fluorescence in the absence compared to the presence of Ko143 obtained by a flow cytometric drug accumulation assay. HEK293T control samples demonstrates no relative efflux, in comparison SNAP-tagged ABCG2_{WT} demonstrates a fold change in drug accumulation/efflux of approx. 15. Data are representative of N=1.

As for Lucifer yellow, incorporating prazosin into the ATPase assay shows an ABCG2 specific ATPase activity, but the difference between the DMSO vehicle control and allocrite is insignificant (Figure 5.10).

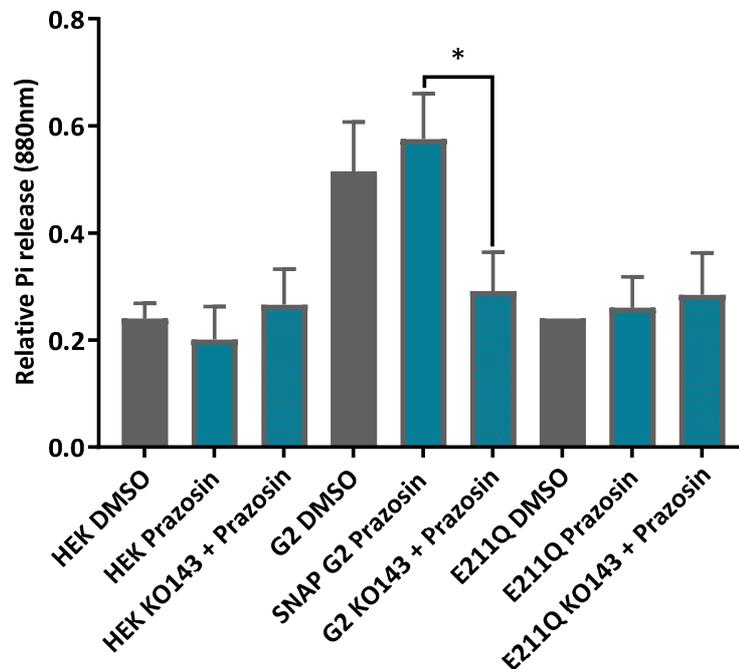


Figure 5.10 Functional assessment of sfGFP-ABCG2 expressed in HEK293T membranes. Membranes (20µg protein) were incubated with vehicle control (DMSO), or prazosin only (10µM) or prazosin plus Ko143 (1µM), in the presence of 5mM Mg.ATP or Mg.ATP. Data are representative of relative Pi release define as the absorbance at 880nm (A_{880nm}) for samples incubated with ATP minus AMP (baseline correction) for $N \geq 3$ independent experiments. The results show a low-level Pi release for both the HEK and E211Q (sfGFP-tagged) control membranes. The DMSO and prazosin ABCG2 (SNAP-tagged) samples demonstrate a high level of Pi release which is reduced to a control level when inhibited with Ko143 (denoted by asterisk), demonstrating ABCG2 specific Pi release (i.e. protein ATPase activity).

Throughout the assays conducted thus far Mg.ATP has been used at highest tolerated level, it was postulated whether at this level ATPase activity was 'artificially' stimulated and therefore masked any potential allocrite stimulated response. To assess this the concentration of Mg.ATP was reduced. The results showed that at any given Mg.ATP concentration the vehicle control and the allocrite samples produced a similar response, demonstrating an increase in Pi

release in a concentration dependent manner (Figure 5.11). The results also clearly show that the use of 5mM ATP produced the best signal to noise ratio (Figure 5.11).

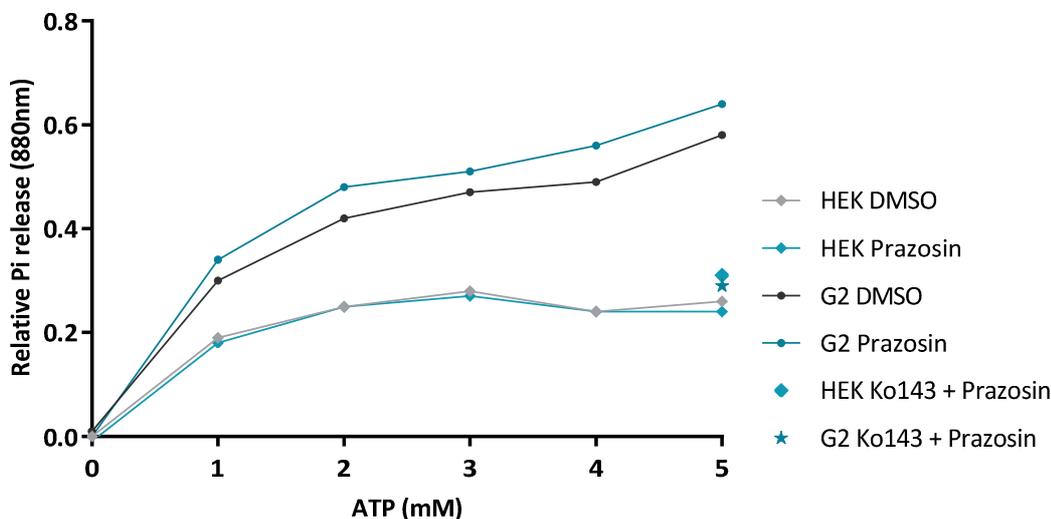


Figure 5.11 ABCG2 demonstrates an ATP concentration dependent increase in Pi release. Membranes (20 μ g) (see key) were subjected to increasing concentrations of nucleotide. HEK membranes with both DMSO (vehicle control) and prazosin show a slight increase in Pi release up to 3mM. G2 expressing membranes demonstrate a steady increase in Pi release, showing that 5mM ATP produces the best signal to noise ratio. Ko143 inhibited sample at a fixed ATP concentration of 5mM are consistent with previous data demonstrating full ABCG2 inhibition (comparable to HEK only membrane). Data are representative of relative Pi release ($A_{880nm} [ATP] - A_{880nm} [AMP]$) for a single experiment.

As all previous attempts at demonstrating an allocrite stimulated response were unsuccessful, a dose response assay was conducted with prazosin to see if this was an allocrite concentration related outcome (i.e. allocrite concentration too low). In such assays the percentage of solvent was kept consistent across all reaction conditions. The results show a dose dependent response in ATP turnover with increasing concentration of prazosin, with an EC_{50} of approximately 0.5 μ M (Figure 5.12). Notwithstanding this, the vehicle only sample still displayed a high level of Pi release.

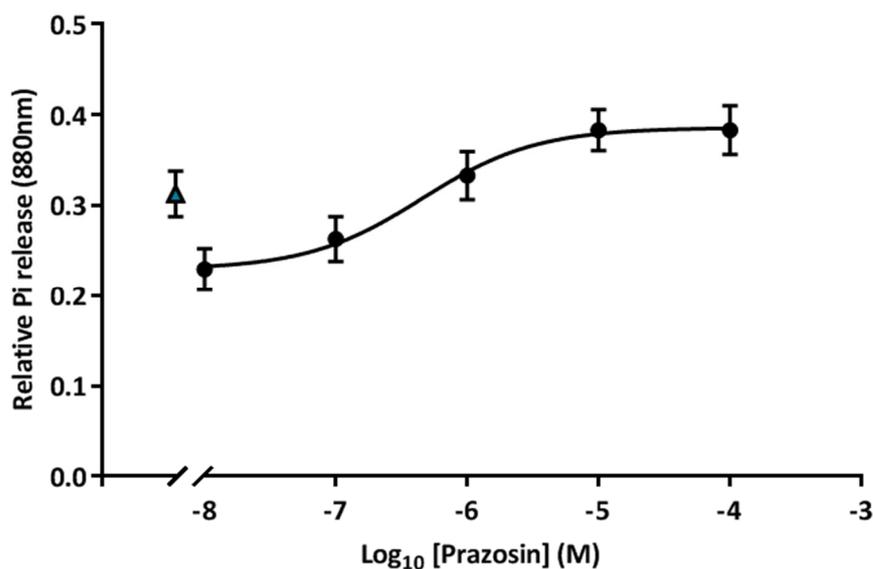


Figure 5.12 Dose response of ABCG2_{WT} membranes with increasing concentrations of prazosin. HEK293T membranes expressing ABCG2_{WT} protein were subjected to increasing concentrations of prazosin in the presence of 5mM ATP (stimulated response). The relative Pi release was calculated as the ratio between ABCG2_{WT} membranes and HEK control membranes for each concentration. DMSO vehicle control (blue triangular data point on split x-axis) demonstrates a higher Pi release than some of the ‘stimulated’ conditions limiting the interpretation of the data. Error bars are representative of the SEM of technical repeats from one experiment. Curve fitted using a log(agonist) vs. response (three parameters) equation in GraphPad Prism 7.04.

As the assay was producing a drug dose response (excluding solvent control from analysis) it was questioned how an ABCG2 mutant may behave. Membranes for the mutant F640A (sfGFP-tagged), a gain in efflux mutation, were investigated for their ATPase activity both with fixed concentration drug in the presence or absence of Ko143 or dose response with prazosin. ABCG2-F640A expressing membranes demonstrated reduced ATPase activity (Figure 5.13), however wild type protein expression levels were approx. 3.2x higher than F640A (western blot Figure 5.13 B.). When corrected for the variation in expression level F640A’s relative ATPase activity is much higher compared to wild type (Figure 5.13 B.), which is consistent with other forms of functional data. When analysed in a dose response manner, the

F640A demonstrated a traditional dose response curve (Figure 5.13) which when analysed gave an approximate EC_{50} of $1.5\mu\text{M}$ slightly reduced compared to ABCG2_{WT} $2\mu\text{M}$. This could suggest that the F640A mutation may increase the drug binding affinity for prazosin, however, the data are only representative of one experiment and therefore no conclusive results could be determined.

The research described in this section provides a foundation for the analysis of ABCG2 function by an ATPase assay, but sadly due to the time constraints of the PhD the ATPase investigations had to come to a close. There is an anticipation that these studies will be carried forward and help interpret mutagenic data on the mechanistic properties of ABCG2, with thoughts and ideas explored in the discussion.

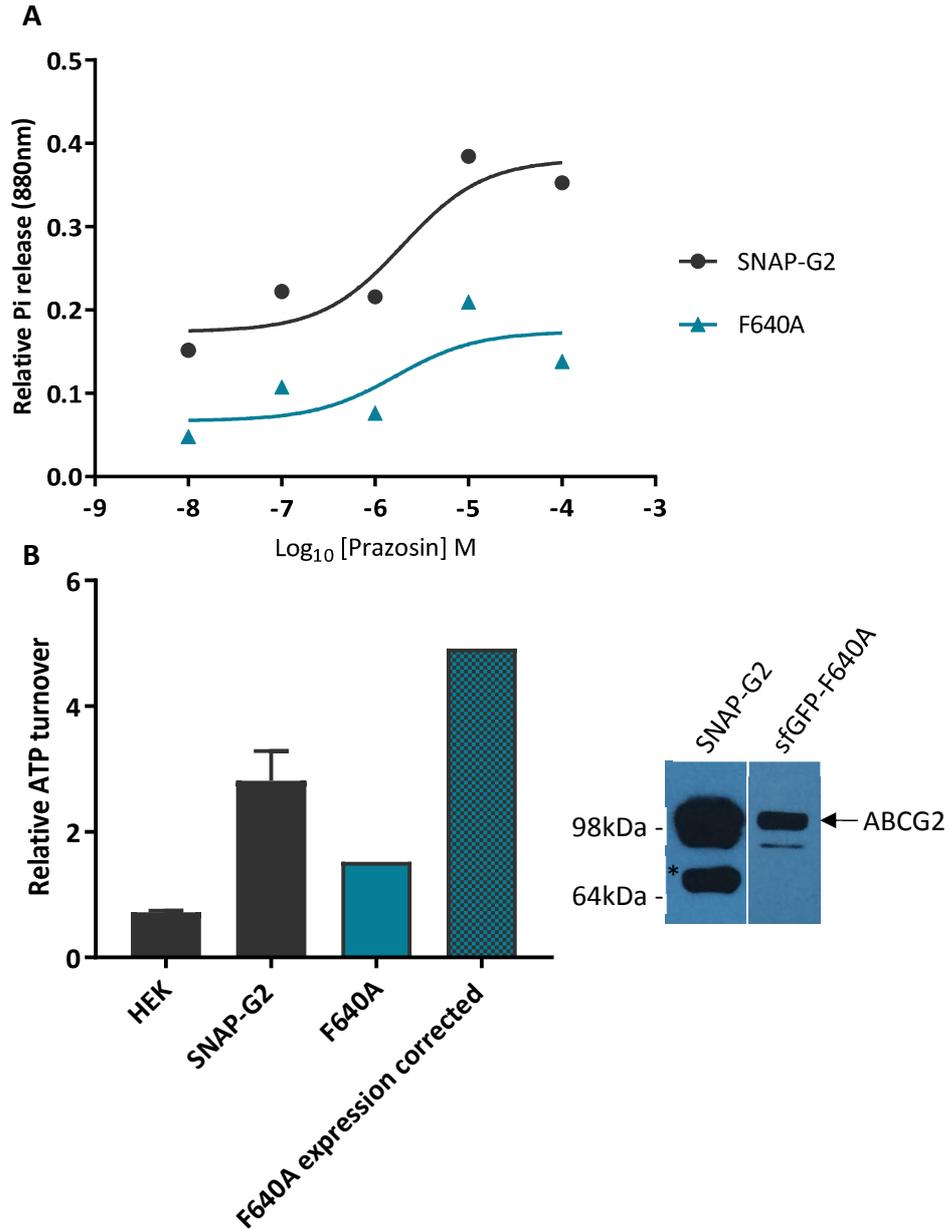


Figure 5.13 F640A demonstrates reduced ATPase activity. **A.** HEK293T membranes (20 μ g) expressing ABCG2_{WT} or ABCG2_{F640A} protein were subjected to increasing concentrations of prazosin in the presence of 5mM ATP (stimulated response). The relative Pi release was calculated as the ratio between the ABCG2 expressing membranes and HEK control membranes for each concentration. A curve was fitted to the data using a log(agonist) vs. response (three parameters) equation (GraphPad Prism 7.04). **B.** Relative ATP turnover (ratio, drug + inhibitor/ drug, for baseline corrected (ATP-AMP) values) demonstrates reduced ABCG2 specific ATP turnover for F640A in comparison to wild type, however when corrected for expression (approx. 3.2x lower expression by densitometry of band indicated ABCG2 on western blot, band marked with an asterisk could be the result of unlabelled ABCG2) results in an increase in overall relative ATP turnover.

5.5 Structural interpretation of ABCG2 mutants

In this final results section, I want to combine the experimental efforts presented across the project with the recent structural data and perform molecular drug docking, using it as a tool to build a picture of drug binding in ABCG2.

To be able to do this Vina (molecular docking software) was used to dock known ABCG2 drug substrates (mitoxantrone, pheophorbide A, and an R482G substrate; daunorubicin) into the protein. Docking was performed using an ABCG2 homology model based on the ABCG5/G8 structure as it was crystallised in a nucleotide-free state, which likely corresponds to a high affinity allocrite binding state (Khunweeraphong et al., 2017; Lee et al., 2016). In contrast, the cyro-EM structures of ABCG2 were all solved in an antibody inhibited state (Jackson et al., 2018; Taylor et al., 2017) with unknown consequence for the TMD conformation and, by extension, for allocrite accessibility within the TMD.

A grid box was assigned to the ABCG2 structural model, enclosing the majority of one half of the TMD (the half that was modelled against ABCG5 as it contains more similarities to ABCG2), centred at the estimated geometric midpoint of the lateral slice hypothesis residues (Figure 5.14).

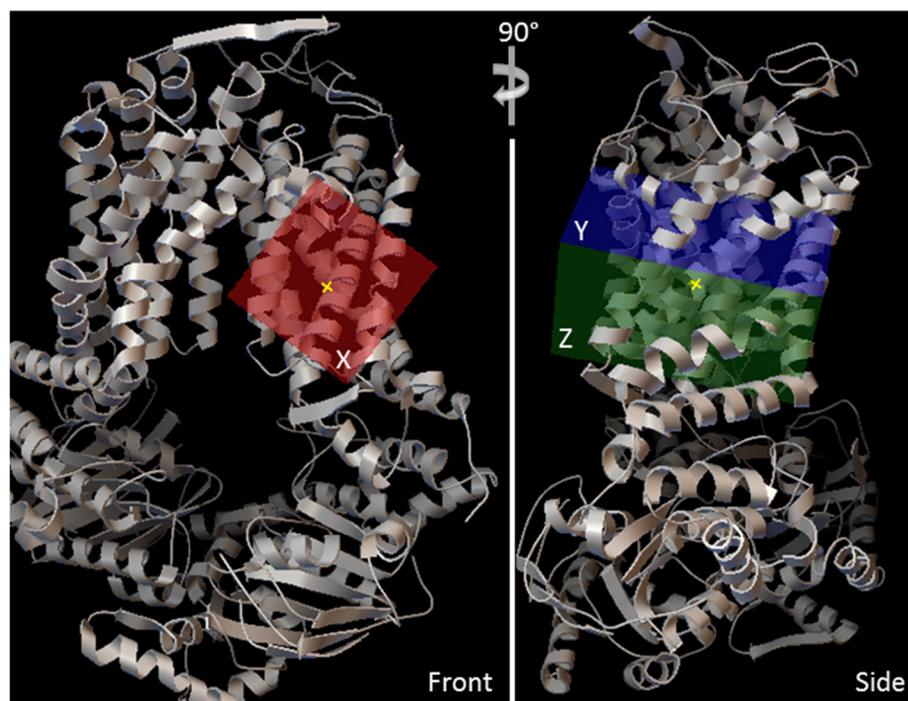


Figure 5.14 ABCG2 homology model with grid box assigned. ABCG2 grid box was assigned using Python molecular viewer; the grid box was centred on the estimated geometric midpoint (denoted by yellow cross) of the lateral hypothesis residues and enclosed the majority of the TMD. The figure is representative of both front view and a 90° side view rotation.

Docking of mitoxantrone, pheophorbide A, and daunorubicin was performed with high exhaustiveness (time and searches to find a global minimum) and revealed two putative drug binding sites. One, exclusive to mitoxantrone (MX), is an ‘internal’ site interfacial to TMH 1b, 2, 3 (major contributions), and 4 of the same TMD. This is consistent with the hypothesis of TMH 3 being integral to substrate selectivity and transport (under current investigation by Parth Kapoor, a PhD student in Dr Ian Kerr’s research group). This region includes residues T402, L405 (TMH1), S440, and S443 (TMH2) from the lateral slice selection, which when mutated demonstrate a perturbation to drug efflux, with other potential interaction (based on proximity, $\leq 5\text{\AA}$) being described in Figure 5.15.

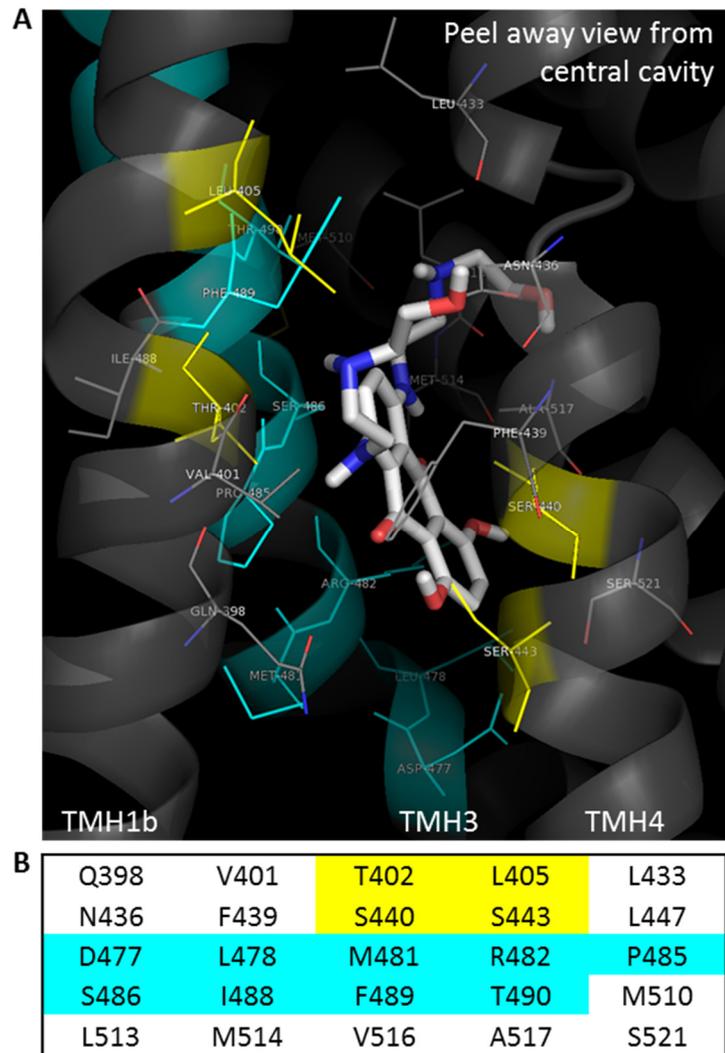


Figure 5.15 The internal binding site exclusively acts as a docking site for mitoxantrone.

A. PyMol image of ABCG2 homology model with mitoxantrone docked to a site predominantly lined by residues from TMH1b, 2 and 3, including residues T402, L405, S440, and S443, highlighted yellow, investigated as part of the lateral slice hypothesis. And TMH3 residues, including R482 and P485, highlighted in cyan, investigated as part of another functional theory, all demonstrating a functional effect when mutated to alanine. **B.** Tabulation of residues at equal to or less than 5Å to the docked mitoxantrone (residues are highlighted to match the colours denoted in the PyMol image).

The second site is located at the lipid interface between TMH 5a, b, and c, and 6a. This secondary site is able to bind all three drug substrates tested (mitoxantrone, pheophorbide A, and daunorubicin) and contains all other residues described in the lateral slice hypothesis and the structurally informed 'binding pocket' mutations (Figure 5.16 and Figure 5.17). For mitoxantrone, the docking orientation within the binding pocket in multiple poses was very similar in the two distinct sites. For pheophorbide A and daunorubicin, although the docking site was consistent, the orientation of the molecules showed variation. Figure 5.17 shows the top binding pose (in terms of binding energy, around 5.5kcal/mol) for each drug (with other poses of similar binding energies are shown in Appendix 4. This provides a strong structural explanation for the experimental data in Chapters 3 and 4.

Both putative binding sites demonstrate partial similarities with other ABCG2 drug docking studies (including T402, L405, S440, and I543) and could indicate that the two sites are distinct drug access sites, or that the two sites could form part of a translocation pathway (see discussion for further details) (Ferreira et al., 2017; László et al., 2016).

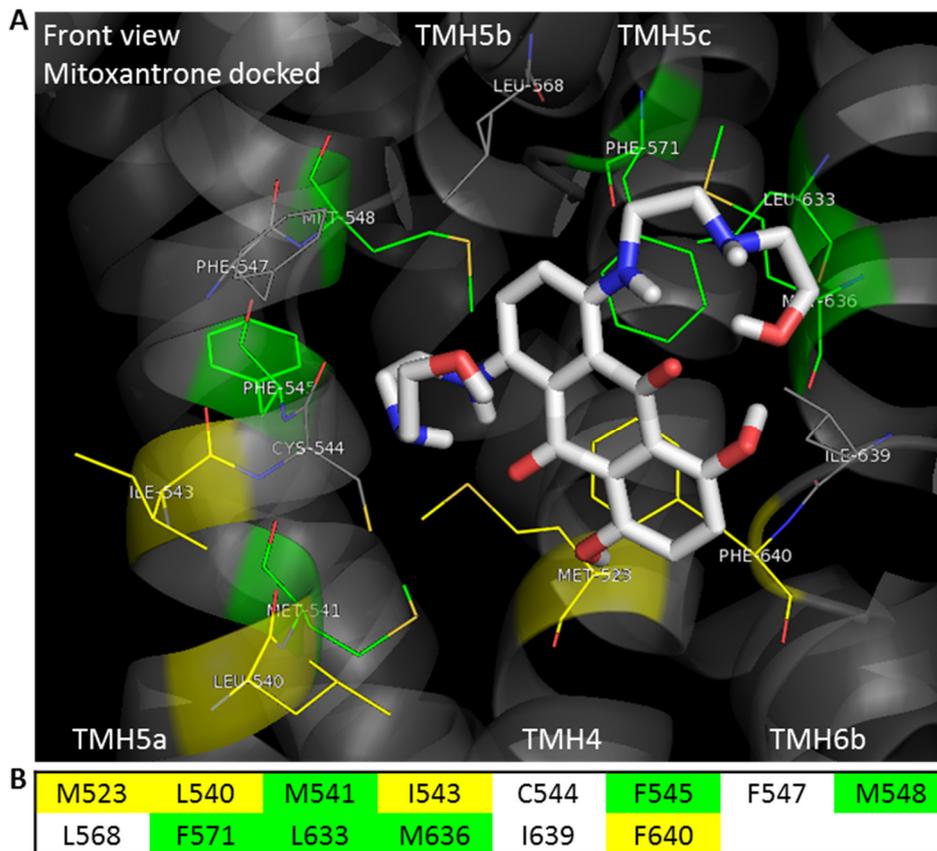
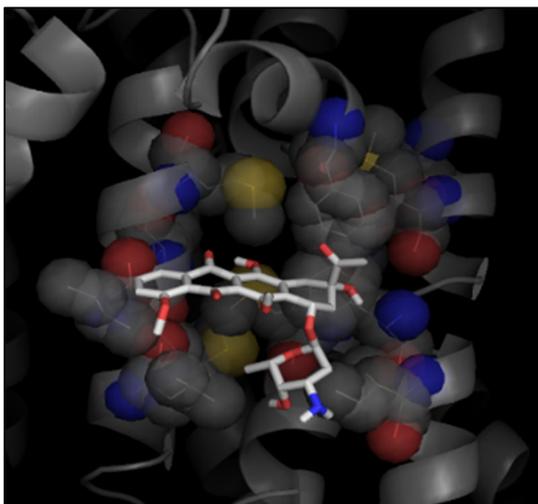


Figure 5.16 Multidrug binding site at the lipid interface of ABCG2. **A.** PyMol image of ABCG2 homology model with mitoxantrone docked to a site lined by residues from TMH5 and 6. MX is shown in white sticks with potential interactions from residues shown in line form (yellow; lateral slice residues, green; binding pocket residues, and grey; other potential residues). **B.** Summary of residues at equal to or less than 5Å to the docked substrate (residues are highlighted to match the colours denoted in the PyMol image).

Daunorubicin



Pheophorbide A

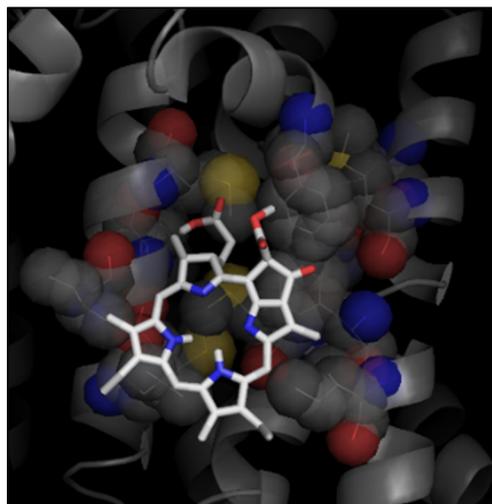


Figure 5.17 Daunorubicin and pheophorbide A dock to a lipid exposed surface site. The figure depicts both daunorubicin and pheophorbide A docked to the ABCG2 homology model (Khunweeraphong et al., 2017) in multiple binding poses (depicted by A-D), with potential interacting residues ($\leq 5\text{\AA}$) summarised in Figure 5.16 shown in spheres.

5.6 Summary

This chapter set out to extend the interpretations of ABCG2 mutant variants that show a functional perturbation.

- Chapters 3 and 4 showed that a flow cytometric based drug accumulation assay was a reliable and robust measure of overall drug efflux but provided no detailing of the individual steps of the catalytic cycle. To be able to investigate these steps, 'cell-free', membrane-based assays such as an ATPase assay had to be employed.
- A histone deacetylase (HDAC) inhibitor, sodium butyrate, was used to try and increase the protein expression level, with an aim of increasing the signal to noise ratio in downstream assays. 10mM of sodium butyrate supplemented to the growth media 24 hour prior to cell harvesting demonstrated an up to 80% increase in protein expression. Therefore, it was implemented for membrane preparations.
- Buffer optimisation via protein unfolding assays and vesicular transport assays demonstrated that the optimal conditions for the protein in isolated membrane preparations were a 50mM Tris, 250mM sucrose, 150mM NaCl solution at pH8, and that MgSO₄ and ATP at a maximal concentration of 5mM could be used.
- The optimal reaction conditions were carried forward into an ATPase assay where ABCG2 specific inorganic phosphate (Pi) release was seen, in a drug and nucleotide 'dose' dependent manner. The incorporation of an ABCG2 mutant, F640A, which demonstrated a gain in the ability to efflux

daunorubicin, raised questions with regards to ABCG2 expression levels versus Pi release. The ATPase assay still requires optimisation (see discussion) but has promise of understanding the underlying mechanism of drug transport (nucleotide linked).

- The second part of this chapter covered the structural interpretation of residues and hypothesised drug binding and/or interaction sites which was achieved by drug docking. Two putative drug binding sites were revealed, one exposed to the lipid interface, lined by residues from TMH5a, b, c and 6b, containing lateral slice and binding pocket residues. A second site, exclusive to mitoxantrone, was observed with major contributions from TMH3 with other contributions from TMH1b, 2, and 4. The sites could represent distinct drug recognition sites or form part of a translocation pathway, which will be deliberated over in the discussion.

Chapter 6 Discussion

At the beginning of this project I proposed a hypothesis of a 'lateral slice' through the transmembrane domain: that residue pairs in TMHs 1, 2, 4, 5, and 6 in equal spatial arrangement to R482 and P485 in TMH 3 formed part of a binding pocket for multidrug recognition and transport. A total of 10 residues were mutated to alanine as it is torsionally neutral but still has the propensity to form alpha helical structures, so is unlikely to affect protein folding. Expression and localisation characterisation of these residues demonstrated characteristics similar to the wild type protein.

When functionally assessed almost all mutations demonstrated a functional perturbation, showing a reduction in the ability to efflux either mitoxantrone and/or pheophorbide A. Two particular mutations: M523A and F640A, demonstrated wild type like efflux of PhA and an enhanced efflux of MX. Also observed was a gain in their ability to efflux a non-native drug substrate daunorubicin, similar to the R482G mutation (Honjo et al., 2001; Robey et al., 2003). These data supported the original 'lateral slice' hypothesis, implicating the importance of the residues in drug transport.

At this time there was an explosion of structural research for ABCG2, starting with the publication of a structural model of ABCG5/G8, a close protein family member to ABCG2. This structure revealed a novel fold for G family proteins and it wasn't long before the first homology model of ABCG2 based on the ABCG5/G8 structure was produced (László et al., 2016; Lee et al., 2016). It was at this point that I was

able to take my data for the 'lateral slice' residues and interpret it using the structural model as a framework.

Structural interpretation revealed that the 'lateral slice' residues were split between two spatially distinct regions of the protein. One, a region sandwiched between TMHs 1b, 2, 3, and 4, which contains residues T402, L405, S440, and S443. Indeed, further investigation of this site has continued with a series of TMH3 mutations made by a subsequent PhD student (Parth Kapoor) within the Kerr research group. A second region, containing the remaining 'lateral slice' residues, is located at the lipid interface with contributions from TMHs 4, 5, and 6, with residues M523 and F640 residing within approximately 3.5Å of each other.

This led to the development of a secondary hypothesis, that the lipid exposed region could be a putative multi-drug binding site and that residues lining this region would affect drug recognition and transport.

The secondary hypothesis was validated, as the ABCG2 mutant variants: M541A, F545A, M548A, F571A, and M636A demonstrated a perturbation in efflux of one or more of the drugs tested, confirming their importance in drug binding and or translocation. One mutation from this investigation, L633A resulted in a maturation defective mutation which demonstrated reticular retention, indicative of a folding defect.

The project progressed to extend the experimental investigation to better understand the involvement of residues in the mechanism of transport. This was conducted in two ways: one, by investigating assays that can break down the steps

in the transport cycle i.e. by probing the ATP hydrolysis of the protein (ATPase assay). Two, investigating the ability of drug substrate to dock in the ABCG2 homology model, to gain an insight as to whether the proposed drug binding pocket(s) have the propensity to bind drug substrate.

This discussion commences by addressing a number of finer points involving specific interactions or residue pairs of interest and the associated downstream investigations. The discussion will then proceed to present an overarching hypothesis for drug transport using the results presented in this study, whilst drawing links to external research on ABCG2. Following this, directions for further research are provided.

6.1 Leucine to alanine mutation at position 633 may cause destabilisation of the re-entry helix

Of all the mutants in this study, only one, L633A, resulted in a folding defect and structural interpretation of this mutant is warranted. L633A is located at the N-terminal end of TMH6b. Immunoblot analysis of L633A demonstrated reduced protein expression of a species with an intermediary molecular weight of approximately 90kDa. A previously investigated isoleucine to alanine substitution at position 573 demonstrated similar expression characteristics, being most comparable to the glycosylation defective mutant, N596Q (Diop and Hrycyna, 2005; Haider et al., 2015). Peptide: N-glycosidase F (PNGaseF) digestion of the glycoprotein between the asparagine and the innermost GLcNAC confirmed the difference in molecular weight to be consistent with core glycosylation (Haider et al., 2015). Visual analysis of L633A by confocal microscopy is again consistent with

I573A. Both mutant variants demonstrated protein predominantly localised within a reticular compartment and co-localisation of I573A with anti-calnexin antibodies confirmed endoplasmic reticulum retention. As the glycosylation defective mutant (N596Q) is able to traffic to the plasma membrane and is functionally active, this mutation is suggestive of a folding defect in early protein processing (Diop and Hrycyna, 2005). Structural analysis was conducted to assess potential interactions due to the similarities in the characteristics of L633A to I573A. It was revealed that I573A is localised to the middle of the ‘V’ within the re-entry helix, between THM5b and 5c, on average 3.5Å apart (depending on the model examined) from L633A in TMH6b (Figure 6.1). At this distance the residues are likely to form hydrophobic interactions.

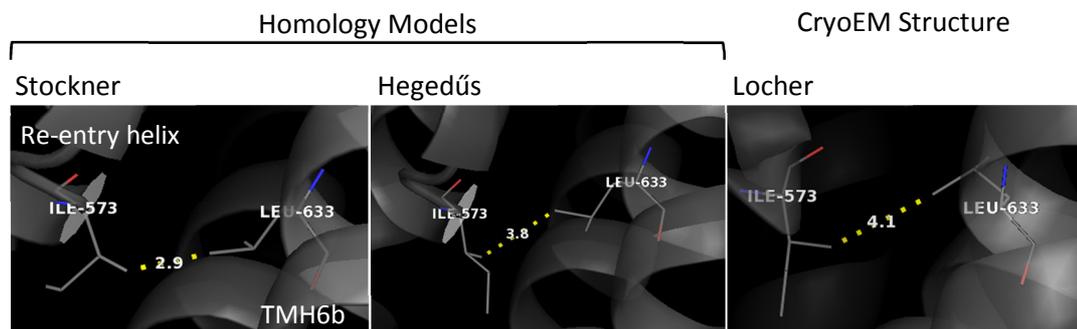


Figure 6.1 Leucine at position 633 may form inter-helical interactions with isoleucine at position 573 to stabilise the re-entry helix. Pymol molecular viewing of residues L633 and I573 in the ABCG2 homology models (based on ABCG5G8) as presented by Stockner and Hegedűs research groups, and cryoEM structure resolved by Locher’s research group show on average a 3.5Å separation of residue sidechains, presumed to form direct hydrophobic interactions due to similarities in expression characteristics (Jackson et al., 2018; Khunweeraphong et al., 2017; László et al., 2016).

Indeed, structural studies have postulated the ability for cholesterol to bind within the vicinity of this pair of residues (Ferreira et al., 2017; Jackson et al., 2018).

Cholesterol is a known regulator of protein function (Pál et al., 2007; Telbisz et al., 2014, 2007) and this raised the question whether mutagenesis caused

destabilisation of the protein through the loss of an interaction with cholesterol? However, investigations of cholesterol's role in ABCG2 function shows no evidence of providing a fundamental stabilisation role for protein expression or localisation (Pál et al., 2007; Telbisz et al., 2007). The data presented here on L633, and previous data on I573, suggest a significant role of the re-entry helix in the structure and function of ABCG2.

Testing this hypothesis would require investigations into other residues that could potentially contribute to re-entry helix stabilising interactions. Investigations could question whether substitution to more bulky hydrophobic residues (such as tryptophan etc) could maintain the proposed hydrophobic interaction.

Substitutions of the residues to cysteine, could also assess the potential to form cross-links and assess the effect on the restriction of the protein's conformational changes and therefore the re-entry helix's impact on allocrite transport.

6.2 Methionine: aromatic interactions may be central to drug recognition and/or protein stabilisation

ABCG2 mutant variants, M523A and F640A, from the lateral slice hypothesis both demonstrated expression and localisation characteristics comparable to the wild type protein. When functionally analysed both mutants conferred the ability to efflux a non-native drug substrate and resided within the lipid exposed potential binding cavity as described in chapter 4 and 5. M548A, described as part of the secondary hypothesis implicating residues lining the pocket to be involved in drug recognition and transport, also demonstrated the ability to efflux a non-native drug substrate as well as a reduced efflux for the drug substrate pheophorbide A. These

residues are of particular interest as the only other mutant to show a gain in efflux is R482G, a well-known variant implicated in drug recognition (Clark et al., 2006; Ejendal et al., 2006; Ozvegy-Laczka et al., 2005). However, it is found to reside within a structurally distinct region to the aforementioned residues and therefore may have a different role (László et al., 2016).

So, the question here is how are the residues producing the observed effect?

Out of all amino acids, Methionine is probably one of the least well understood. It has the propensity to form hydrophobic interactions, but does it have the propensity to form further interactions? There is evidence that the sulphur of the methionine can interact well beyond typical hydrophobic interaction distances with aromatic amino acids (tyrosine, phenylalanine, and tryptophan), extending distances to 5-6Å. Longer-range methionine: aromatic interactions have been implicated in providing structural stability and propensity to coordinate ligands, in tumour necrosis factor (TNF) receptors (TRAIL-DR5 and LT α -TNFR1) (Valley et al., 2012). There is evidence for thousands of these interactions throughout proteins which haven't been fully characterised. Due to this, the residues within the surface site were evaluated for potential methionine: aromatic (Met: aromatic) interactions. The interactions are thought to occur between Met and aromatic groups with preference of a 30-60% rotation around the aromatic ring (Valley et al., 2012). M523 and F640 fit these parameters with the interaction also being maintained in the ABCG5 sub-structure (Figure 6.2).

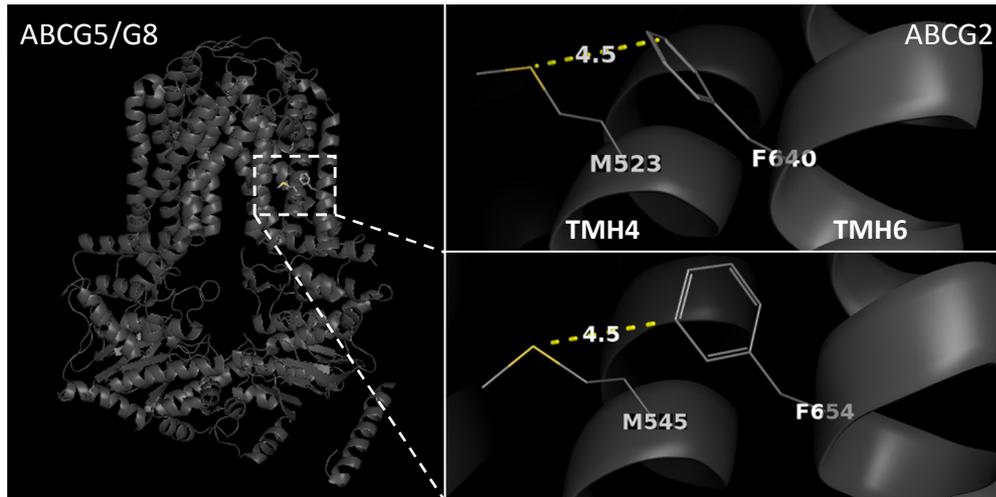


Figure 6.2 Methionine: phenylalanine interaction between TMH4 and TMH6 is conserved. Figure depicts a methionine: phenylalanine interaction within ABCG2 which when either residue is mutated leads to functional perturbation, this interaction is maintained in the ABCG5 crystal structure (PDB: 5DO7), suggesting a potential importance for this interaction across G-protein family. Figure created using PyMol (Scripps Research Institute).

Upon evaluation, the surface site, proposed to be an allocrite binding site, contains a plethora of these potential interactions with several methionine and phenylalanine residues being investigated by mutagenesis (Figure 6.3 A.: including M523, M541, M548, M636 and F545, F571, F640). Interestingly, on an ABCG2-wide search there appears to be three distinct regions (Figure 6.3) for these potential interactions that could be pivotal for allocrite association or propagation of conformation changes with the potential to form interactions at a greater distance, aiding in the resetting of the protein upon drug translocation?

Future investigations into Met: aromatic interactions would include coevolution analysis, to determine if these residue interactions are conserved across species.

Molecular dynamic simulations could also assess the propensity of these regions to propagate signal transduction via conformational changes required for allocrite efflux.

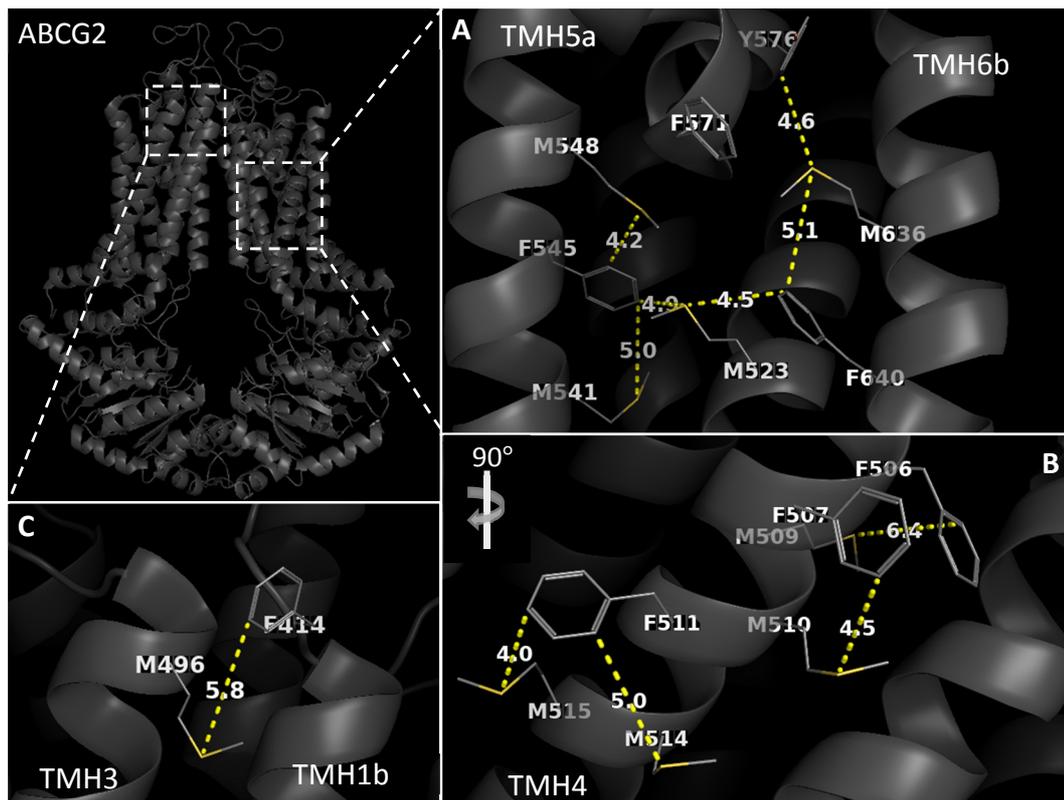


Figure 6.3 Methionine: aromatic interactions may be crucial for either ligand: and/or protein: protein interactions/ stabilisation. ABCG2 structural analysis reveals several potential inter and or intra-helical methionine: aromatic interactions. **A.** ‘surface site’ proposed to have the propensity to bind both cholesterol and allocrite, demonstrates a high proportion of met: aromatic residues within proximity ($\leq 6\text{\AA}$) to form interactions. **B.** A second potential site (extracellular end of TMH4) for met: aromatic residues, intra-helical interactions could stabilise protein conformations. **C.** Potential interhelical structural stabilisation interaction between residues of TMH3 and TMH1b. figure were generated in PyMol (The Scrips Research Institute) using PDB: 6ETI.

6.3 Hypothesis of an ABCG2 mediated allocrite transport cycle/mechanism

I would now like to propose a mechanism for allocrite transport based upon the current experimental data in conjunction with recently published structural data for ABCG2. Additionally, a summary of the work that would be required to test the transport mechanism is presented.

At this point it is important to note that; both the potentially stabilising interaction partnership of I573 and L633, and the Met:Phe (aromatic) interaction, intertwine

with an overarching drug transport mechanism hypothesis that may aid in discussing the initial stages of drug binding and the propagation of movement.

6.3.1 Current investigations and context

To put the current investigations into context, little was known mechanistically about ABCG2 allocrite transport. Experimental data implicated a hand full of residues in allocrite transport, however, structural interpretation of the protein was limited due to the resolution of the structural data.

The past few years have reaped the rewards of a decade's worth of improvements, in both cryo-EM detection systems and computing algorithms, combined with an increased understanding of the role the native lipid environment exerts on the protein, leading to enhanced protein extraction methods (Bai et al., 2015; Dörr et al., 2016; Gulati et al., 2014).

6.3.2 Recent structural advances

In 2016, a crystal structure for the heterodimeric sterol transporter ABCG5/G8 was resolved at 3.9Å (Lee et al., 2016). This was the first ABCG family protein to be resolved at a near-atomic level, providing insights into the transmembrane helical domain organisation and interfacial contacts. In the same year, a homology model for ABCG2 was published, based on the ABCG5/G8 crystal structure (László et al., 2016). Since then, a further four structural papers have been published to date, including the first reasonably high resolution cryoEM structure of ABCG2 (Ferreira et al., 2017; Jackson et al., 2018; Khunweeraphong et al., 2017; Taylor et al., 2017).

The structures revealed a new protein fold for ABCG proteins. The transmembrane domains are shorter, and almost entirely contained within the lipid bilayer. There is also no evidence of domain swapping (i.e. the crossing over of transmembrane helices from one TMD monomer to the other) implying structural topology similar to that of a type II importer (Rice et al., 2014).

Four further notable features of the TMH organisation are;

- 1) The 'elbow/ connecting helix', the topologically predicted transmembrane helix 1 is split into a and b; TMH1a lies interfacial to the lipid bilayer and forms part of a tri-helix bundle 'connecting' the TMD to the NBD (Khunweeraphong et al., 2017). TMH1b forms part of the standard TMD architecture (Khunweeraphong et al., 2017; Lee et al., 2016).
- 2) The 'coupling' helix. ABC exporters usually contain two intracellular loops per TMD that extend into the NBD, which are thought to propagate signal transduction between the TMD and NBD to facilitate allocrite efflux (Dawson and Locher, 2007; Kim and Chen, 2018). ABCG2 only contains one ICL that is long enough to extend into the NBD (László et al., 2016).
- 3) A region of 'split' helices towards the extracellular interface of the protein which forms a 'V-like' structure, termed the 're-entry' or 'capping' helix. This is comprised of residues from both TMH5 and 6, split into 5a, **b**, and **c**, and **6a** and **b** (letters denoted in bold are regions within the re-entry helix).
- 4) The protein, resolved in an apo-state resembles the classic 'open-inward' configuration shown by other ABC exporters (Sav1866, Pgp) (Dawson and

Locher, 2007; Jackson et al., 2018; Taylor et al., 2017). However, unlike these structures the NBDs are in 'constant contact' via an ABCG specific NPxDF motif (where x is any residue).

6.3.3 Structural evidence of multiple allocrite binding sites

Molecular dynamic simulations and allocrite docking by László et al. demonstrate the possibility of at least four distinct allocrite binding sites (Figure 6.4), consistent with early allocrite binding studies suggesting multiple binding sites (Clark et al., 2006).

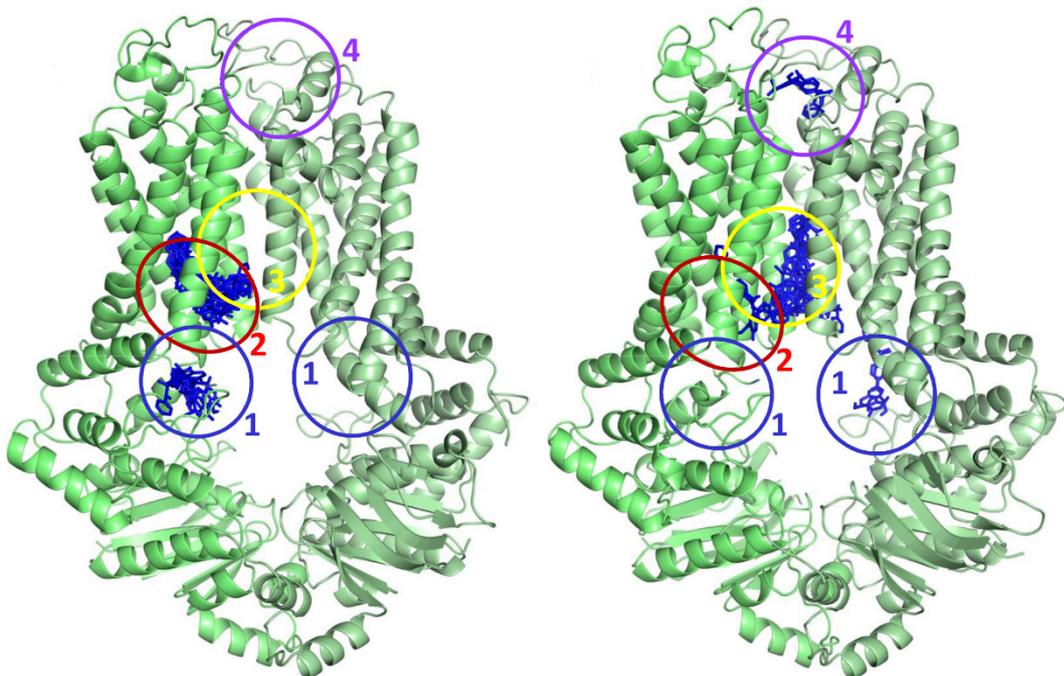


Figure 6.4 *In silico* docking of allocrites in an ABCG2 homology model reveals multiple potential allocrite interaction sites. Figure shows two ABCG2 conformations in which there are 4 distinct regions where sulfasalazine can dock in multiple poses. Figure taken from László et al., 2016.

Cryo-EM structural data of ABCG2, resolved to a near-atomic resolution (overall resolution between 3.1 and 3.8Å) in the presence of FAB complexes bound to the extracellular looped region, initially presented a structure in which cholesterol

could be assigned to a region of electron density within a central cavity (a region interfacial to the two TMDs, similar to site 3 presented by László et al., 2016). Later structural publications from the same research group, presented a Ko143 (specific ABCG2 inhibitor) derivative (MZ29) bound in a 2:1 ratio (inhibitor molecule to protein dimer) or a tariquidar derivative (MB136) bound in a 1:1 ratio within the central cavity between TMH2 of one monomer and TMH5a of the opposing monomer (Figure 6.5).

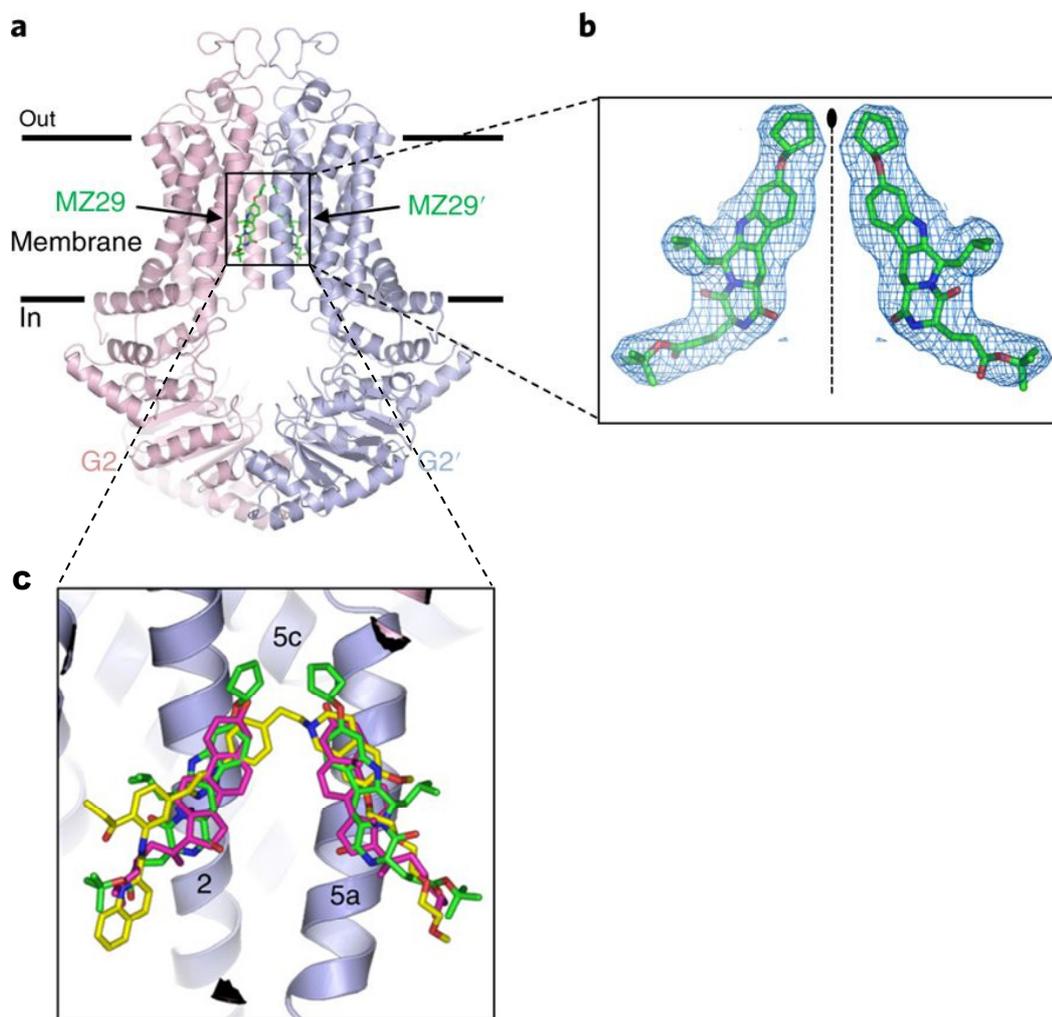


Figure 6.5 CryoEM structure of ABCG2 with molecules bound to the central cavity. A. ABCG2 full length structure with Ko143 derivative: MZ29 bound to a central cavity. **B.** Enhanced region demonstrating electron density to which the MZ29 was refined, demonstrating a 2:1 binding ratio and 2-fold symmetry. **C.** Magnified region showing MZ29 (green), a tariquidar derivative: MB136 (yellow), bound in a 1:1 ratio, and cholesterol docked (Taylor et al., 2017) (pink). Figure adapted from (Jackson et al., 2018).

6.3.4 Cox proposed drug binding sites

The work presented in chapters 4 and 5 utilised the ABCG2 structural models as a framework for interpreting experimental data and designing new experiments. The three allocrites that were used in the scope of the functional studies in chapter 3 and 4; pheophorbide A, daunorubicin, and mitoxantrone were subsequently used in the molecular allocrite docking studies in chapter 5.

As the cryoEM structures were obtained in the presence of inhibitory antibodies and/or small molecule inhibitors, with unknown consequence for the proteins conformation, the homology model (based on the uninhibited apo-state ABCG5/G8 crystal structure) provided by Thomas Stockner was preferentially used for molecular allocrite docking (as seen in chapter 5). This resulted in the discovery of two potential allocrite binding sites contained within a single TMD.

A lipid exposed site orientated just below the 're-entry' helix, was able to bind all three allocrites tested within this project, encompassing several of the residues investigated as part of the 'lateral slice' hypothesis. Whereas, a second site, buried between TMHs 3 (TMH3 residues in ongoing studies by Parth Kapoor in Dr Ian Kerr's research group) and 4 with interactions from TMHs 1b and 2 demonstrated the ability to only bind mitoxantrone in this project. A number of residues investigated here demonstrated partial overlap with the aforementioned sites in the previous section (6.3.3), as summarised in Table 6.1 below.

Hegedűs research group	Site 2	T402	L405	S440	S443		
	Site 3	S440	S443	M541	F545	L540	I543
Locher research group	Cavity 1	L405	S440	I543			

Table 6.1 Cox proposed drug binding residues that overlap with other ABCG2 structural publications. Residues highlighted in yellow are residues within a ‘lipid exposed’ surface site (I543 is in close proximity to this site). All other residues reside within a ‘buried’ site (Cox et al., 2018; Jackson et al., 2018; László et al., 2016).

6.3.5 Cox hypothesis of the ABCG2 multi-drug transport mechanism

I propose that these two allocrite binding sites are potential drug recognition sites, the first is a lipid exposed surface pocket and the second is a slit like cavity accessible via the cytoplasmic interface. How the allocrites are translocated from here is open to interpretation, and is an area of hot debate, as discussed below.

Cryo-EM structures resolved by Kaspar Locher’s research group propose a transport mechanism in which allocrites bind within the central cavity and, is then translocated to a secondary ‘upper’ cavity via a ‘leucine plug’ (L555), for release. As they were able to show electron density consistent with cholesterol (a debatable transport substrate), which is an analogue of a “*bona fide*” substrate of ABCG2 (estrogen-1-sulphate: E₁S) they conclude that the central cavity is a multi-drug binding site for ABCG2 (Jackson et al., 2018; Taylor et al., 2017).

Whilst this model for transport is appealing it hinges on the central point that allocrites associate in a lower central cavity that is occluded from an upper ‘smaller’ cavity (designated for allocrite release) by the formation of a leucine plug (L555). Leucine at position 555 has been experimentally investigated as the residue forms part of a putative steroid-binding element (SBE). A leucine to alanine mutation at

position 555 did not demonstrate any structural (expression maintained) or functional perturbation (vanadate-sensitive ATPase activity reduced, however drug transport retained in intact insect cells), inconsistent with the results expected for a residue with a key role in structure and or transport (Telbisz et al., 2007).

I propose an alternative transport mechanism where the recognition of allocrites occurs *before* translocation to the central cavity and that this initial recognition propagates a signal transduction (via conformational changes) from the TMD to the NBD.

At a more detailed level my hypothesis is a 4-stage cycle (as shown in Figure 6.6), consisting of drug association to an “access site”, signal propagation resulting in the binding of ATP simultaneously resulting in the dimerisation of the NBD and translocation of drug from the “access site” via the central cavity to an extracellular release site, ATP hydrolysis then provides the energy to force the dissociation of the allocrite and subsequently reset the protein to a basal state, ADP dissociates starts the cycle again.

This hypothesis stems from the fact that common drug substrates for ABCG2 are hydrophobic in nature. In an open-inward conformation the central cavity is exposed to the intracellular compartment, i.e. an aqueous environment. It is unfeasible to suggest that hydrophobic allocrites would associate to the protein via this route. It is more likely that that hydrophobic allocrites would interact with the protein via the lipid interface, i.e. via a lipid exposed transmembrane binding site such as the two presented within this project.

The aforementioned methionine: aromatic (phenylalanine) interactions within the surface site could be altered upon allocrite binding, and thus result in a signal propagation (via conformational changes and steric interactions) to the NBD via ICL1.

ICL1 has been recently investigated along these lines and suggested as an allosteric conduit (“polar relay”) between the NBDs and TMDs, to act as a molecular spring coupling the domains (Khunweeraphong et al., 2017). I propose that it is this initial drug binding that alters the affinity for ATP binding to a higher state. Where the association of ATP then results in the formation of a closed NBD dimer consistent with the mechanism of NBD dimerisation presented by Smith et al., 2002.

It is this mechanism of dimer formation that I speculate loads the ‘molecular spring’ of ICL1 propagating allosteric movement of TMHs to alter the affinity for allocrites at the surface site, as indicated by McDevitt et al., 2008, who demonstrated a reduced affinity for allocrite binding upon the addition of ATP. This subsequently could promote translocation through the central cavity to an external release site, in a multi-stage process. The energy derived from the hydrolysis of ATP would drive the dissociation of the allocrite via NBD dimer dissociation and a resetting of the protein to an open-inward conformation, aided by the release of the ‘molecular spring’. I propose this would occur in an asymmetric fashion, consistent with the sequential hydrolysis of ATP proposed within the constant contact model for efflux (see Figure 6.6 for a schematic overview).

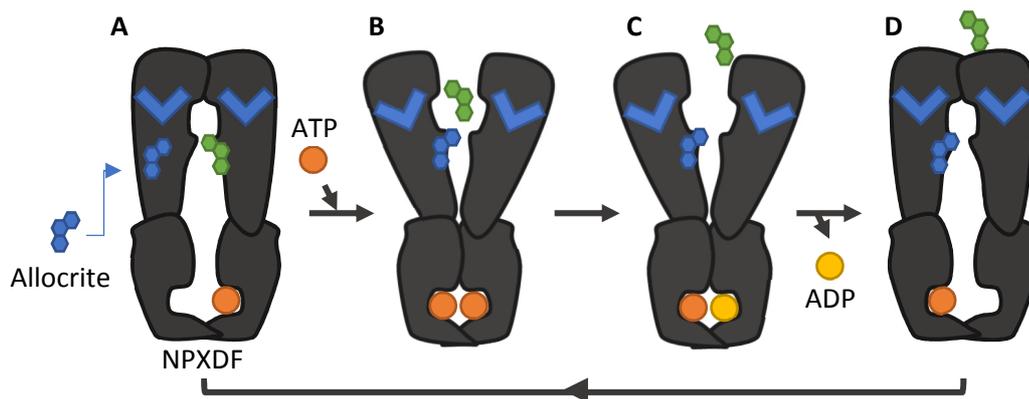


Figure 6.6 Schematic model of ABCG2 mediated allocrite efflux. The model presented is based on the alternate access 'constant contact' model for drug efflux. **A.** Allocrite, (green hexagons) associated from the previous transport cycle is associated to a secondary site, with ATP (orange circle) associated to the corresponding NBD. Binding of a second allocrite (blue hexagons) in the opposing TMD triggers **(B.)** the binding of a second ATP, causing NBD dimer formation and translocation of the first allocrite to a release site, and the second allocrite to a secondary site. **C.** The hydrolysis of ATP in the 'first' NBD causes the full displacement of the first allocrite and **(D.)** resetting of the protein to an 'inward-open' configuration. The cycle re-initiates with allocrite and ATP bound in the opposing TMD/NBD to the previous cycle. The re-entry helix, situated above the proposed lipid exposed allocrite binding site, is denoted by a blue V-shape, and may have an involvement in the translocation pathway.

There are still lots of questions and the only way to gain the answers will be through further investigations, including; residues mutagenesis using the new structures as a frame work for the design of new experiments and data interpretation. The development of techniques to monitor direct drug binding interactions are also crucial.

6.4 Directions for further research

6.4.1 Mechanism of ABCG2 regulation by cholesterol

Cholesterol has been shown to be fundamental for functional ABCG2, exactly how cholesterol regulates ABCG2 activity is unknown. It has been recently shown to have the propensity (via molecular dynamic simulations) to bind in the vicinity of the putative surface exposed drug binding site; so, one question is whether

cholesterol could act as a structural 'place holder' for allocrite binding? i.e. in other words, that cholesterol occupies the allocrite binding site until it is displaced prior to drug transport.

Cholesterol association could be investigated by photon counting histogram (PCH) analysis with a fluorescent cholesterol analogue (Chen et al., 1999; Sezgin et al., 2016). The comparable analysis between an allocrite free versus an allocrite associated protein could ascertain whether allocrite association results in the dissociation of cholesterol analogue. Investigations could also ascertain residues involved in direct binding of cholesterol via similar methods.

6.4.2 Role of re-entry helix in structural stability and or transport

As described in section 6.1 residues within the re-entry helix have been implicated in protein stability. Both cholesterol and allocrites again have been shown to bind within the vicinity of the re-entry helix (i.e. in the surface site situated below the re-entry helix). It would be interesting to see if any of the residues within the re-entry helix show a high-degree of co-evolution and whether substitutions would have equal effects for intracellular trafficking and folding defects.

6.4.3 Role of methionine: aromatic interactions in allocrite efflux

Met: aromatic residue interactions may contribute to structural stability and or allocrite binding and or cholesterol regulation, as described in section 6.2. One region with many potential interactions is the 'surface site'. As mentioned previously, this is situated within a region with potential stabilising, regulation, and allocrite binding importance. Further investigation of these potential interactions

would require similar approaches to those investigating the re-entry helix: co-evolution analysis could indicate potential interacting residue pairs and substitution of these residues would look for functional perturbation. It would also be interesting if a substitution to another aromatic residue could maintain these interactions or whether phenylalanine, shown in most of these potential interactions, is required.

6.4.4 Investigations into residues forming the translocation pathway/ the transport hypothesis

Combined experimental and structural studies are gradually aiding in developing a better understanding of an allocrite efflux mechanism as described in section 6.3. However, at the moment this picture remains incomplete. The understanding of the mechanism is of great importance in being able to comprehend the protein in the context of multidrug resistance and the involvement in disease.

Molecular dynamic simulations of ABCG2 could be used to help build a picture of the residues potentially involved in drug transport, and subsequent substitution of these residues would be used to investigate their functional importance. A major current limitation for interpretation of a transport mechanism is the existing structures are resolved either in the presence of an inhibitory antibody which locks the protein into a conformation of uncertain significance, and/or the presence of synthetic Ko143 derivatives. Ideally, structural information for an uninhibited structure, in the presence of nucleotide, or allocrite would better facilitate the interpretation of a transport mechanism. Whilst structural biology will provide this in due course, other direct techniques remain valuable: drug binding studies by

microscale thermophoresis (MST) and/ or fluorescence correlation spectroscopy (FCS) of ABCG2 mutants, would enable researchers to separate allosteric interactions from direct allocrite interactions and gain a better picture of transport. These techniques would require the optimisation of protein expression and purification, which has progressed in parallel to this project in a thesis by a colleague (Horsey, 2018).

6.5 Summary

This project set out to investigate the molecular basis of allocrite polyspecificity in ABCG2. In undertaking this project, a number of complementary hypotheses were proposed and subsequently investigated. This confirmed that several of the investigated residues have significant roles in the transport mechanism of ABCG2, as presented in Cox et al., 2018. There is significant scope for further investigation of the hypotheses developed during this project. The premise of which has formed part of a successful grant application funded by the BBSRC (ref: BB/S001611/1, “Channelling a path for substrates through a multidrug transporter”) within the Kerr research group at the University of Nottingham.

References

- Abolhoda, A., Wilson, A.E., Ross, H., Danenberg, P. V., Burt, M., Scotto, K.W., 1999. Rapid activation of MDR1 gene expression in human metastatic sarcoma after in vivo exposure to doxorubicin. *Clin. Cancer Res.* 5, 3352–3356.
- Allen, J.D., Brinkhuis, R.F., Wijnholds, J., Schinkel, A.H., 1999. The mouse Bcrp1/Mxr/Abcp gene: Amplification and overexpression in cell lines selected for resistance to topotecan, mitoxantrone, or doxorubicin. *Cancer Res.* 59, 4237–4241.
- Allen, J.D., van Loevezijn, A., Lakhai, J.M., van der Valk, M., van Tellingen, O., Reid, G., Schellens, J.H.M., Koomen, G.-J., Schinkel, A.H., 2002. Potent and specific inhibition of the breast cancer resistance protein multidrug transporter in vitro and in mouse intestine by a novel analogue of fumitremorgin C. *Mol. Cancer Ther.* 1, 417–25.
- Allikmets, R., Schriml, L.M., Hutchinson, A., Romano-Spica, V., Dean, M., 1998. A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res.* 58, 5337–9.
- Ambudkar, S. V., 1998. ATPase Activity in Crude Membranes of Human MDR1-Transfected Mammalian Cells. *Methods Enzymol.* 292, 504–514.
- Ambudkar, S. V., Gottesman, M.M., Zoghbi, S.S., Lu, S., Weidner, L.D., Pike, V.W., Innis, R.B., Mulder, J., Shukla, S., Hall, M.D., 2015. The Inhibitor Ko143 Is Not Specific for ABCG2. *J. Pharmacol. Exp. Ther.* 354, 384–393.
- Anderson, P.O., 1991. Drug use during breast-feeding. *Clin. Pharm.* 10, 594–624.
- Andress, E.J., Nicolaou, M., McGeoghan, F., Linton, K.J., 2017. ABCB4 missense mutations D243A, K435T, G535D, I490T, R545C, and S978P significantly impair the lipid floppase and likely predispose to secondary pathologies in the human population. *Cell. Mol. Life Sci.* 74, 2513–2524.
- Andress, E.J., Nicolaou, M., Romero, M.R., Naik, S., Dixon, P.H., Williamson, C., Linton, K.J., 2014. Molecular mechanistic explanation for the spectrum of cholestatic disease caused by the S320F variant of ABCB4. *Hepatology* 59, 1921–1931.
- Auricchio, A., Trapani, I., Allikmets, R., 2015. Gene Therapy of ABCA4-Associated Diseases. *Cold Spring Harb. Perspect. Med.* 5, a017301–a017301.
- Awortwe, C., Kaehler, M., Rosenkranz, B., Cascorbi, I., Bruckmueller, H., 2018. MicroRNA-655-3p regulates Echinacea purpurea mediated activation of ABCG2. *Xenobiotica* 48, 1050–1058.
- Babenko, A.P., Bryan, J., 2003. SUR Domains that Associate with and Gate KATP Pores Define a Novel Gatekeeper. *J. Biol. Chem.* 278, 41577–41580.
- Bai, X., McMullan, G., Scheres, S.H.W., 2015. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* 40, 49–57.

- Bailey-Dell, K.J., Hassel, B., Doyle, L.A., Ross, D.D., 2001. Promoter characterization and genomic organization of the human breast cancer resistance protein (ATP-binding cassette transporter G2) gene. *Biochim. Biophys. Acta* 1520, 234–41.
- Bakos, É., Evers, R., Szakács, G., Tusnády, G.E., Welker, E., Szabó, K., De Haas, M., Van Deemter, L., Borst, P., Váradi, A., Sarkadi, B., 1999. Functional multidrug resistance protein (MRP1) lacking the N-terminal transmembrane domain. *J. Biol. Chem.* 273, 32167–32175.
- Berntsson, R.P.-A., Smits, S.H.J., Schmitt, L., Slotboom, D.J., Poolman, B., 2010. A structural classification of substrate-binding proteins. *FEBS Lett.*
- Boël, G., Smith, P.C., Ning, W., Englander, M.T., Chen, B., Hashem, Y., Testa, A.J., Fischer, J.J., Wieden, H.-J., Frank, J., Gonzalez, R.L., Hunt, J.F., 2014. The ABC-F protein ETTA gates ribosome entry into the translation elongation cycle. *Nat. Struct. Mol. Biol.* 21, 143–51.
- Boussif, O., Lezoualc'h, F., Zanta, M.A., Mergny, M.D., Scherman, D., Demeneix, B., Behr, J.P., 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. U. S. A.* 92, 7297–301.
- Breedveld, P., Zelcer, N., Pluim, D., Sönmezer, Ö., Tibben, M.M., Beijnen, J.H., Schinkel, A.H., Van Tellingen, O., Borst, P., Schellens, J.H.M., 2004. Mechanism of the pharmacokinetic interaction between methotrexate and benzimidazoles: Potential role for breast cancer resistance protein in clinical drug-drug interactions. *Cancer Res.* 64, 5804–5811.
- Brendel, C., Scharenberg, C., Dohse, M., Robey, R.W., Bates, S.E., Shukla, S., Ambudkar, S. V., Wang, Y., Wennemuth, G., Burchert, A., Boudriot, U., Neubauer, A., 2007. Imatinib mesylate and nilotinib (AMN107) exhibit high-affinity interaction with ABCG2 on primitive hematopoietic stem cells. *Leukemia* 21, 1267–1275.
- Burger, H., Van Tol, H., Boersma, A.W.M., Brok, M., Wiemer, E.A.C., Stoter, G., Nooter, K., 2004. Imatinib mesylate (STI571) is a substrate for the breast cancer resistance protein (BCRP)/ABCG2 drug pump. *Blood* 104, 2940–2942.
- Cai, X., Bikadi, Z., Ni, Z., Lee, E.-W., Wang, H., Rosenberg, M.F., Mao, Q., 2010. Role of basic residues within or near the predicted transmembrane helix 2 of the human breast cancer resistance protein in drug transport. *J. Pharmacol. Exp. Ther.* 333, 670–81.
- Carrier, I., Urbatsch, I.L., Senior, A.E., Gros, P., 2007. Mutational analysis of conserved aromatic residues in the A-loop of the ABC transporter ABCB1A (mouse Mdr3). *FEBS Lett.* 581, 301–308.
- Cascorbi, I., 2011. P-glycoprotein: Tissue Distribution, Substrates, and Functional Consequences of Genetic Variations, in: *Handbook of Experimental Pharmacology*. pp. 261–283.
- Chen, S., Oldham, M.L., Davidson, A.L., Chen, J., 2013. Carbon catabolite repression of the maltose transporter revealed by X-ray crystallography. *Nature* 499, 364–368.

- Chen, Y., Müller, J.D., So, P.T.C., Gratton, E., 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophys. J.* 77, 553–567.
- Chen, Z.-S., Robey, R.W., Belinsky, M.G., Shchaveleva, I., Ren, X.-Q., Sugimoto, Y., Ross, D.D., Bates, S.E., Kruh, G.D., 2003. Transport of methotrexate, methotrexate polyglutamates, and 17 β -estradiol 17-(β -D-glucuronide) by ABCG2: effects of acquired mutations at R482 on methotrexate transport. *Cancer Res.* 63, 4048–54.
- Cheng, S.H., Rich, D.P., Marshall, J., Gregory, R.J., Welsh, M.J., Smith, A.E., 1991. Phosphorylation of the R domain by cAMP-dependent protein kinase regulates the CFTR chloride channel. *Cell* 66, 1027–1036.
- Chifflet, S., Torriglia, A., Chiesa, R., Tolosa, S., 1988. A method for the determination of inorganic phosphate in the presence of labile organic phosphate and high concentrations of protein: Application to lens ATPases. *Anal. Biochem.* 168, 1–4.
- Chu, X., Cai, X., Cui, D., Tang, C., Ghosal, A., Chan, G., Green, M.D., Kuo, Y., Liang, Y., Maciolek, C.M., Palamanda, J., Evers, R., Prueksaritanont, T., 2013. In Vitro Assessment of Drug-Drug Interaction Potential of Boceprevir Associated with Drug Metabolizing Enzymes and Transporters. *Drug Metab. Dispos.* 41, 668–681.
- Cisternino, S., Mercier, C., Bourasset, F., Roux, F., Scherrmann, J.-M., 2004. Expression, Up-Regulation, and Transport Activity of the Multidrug-Resistance Protein Abcg2 at the Mouse Blood-Brain Barrier. *Cancer Res.* 64, 3296–3301.
- Clark, R., Kerr, I.D., Callaghan, R., 2006. Multiple drugbinding sites on the R482G isoform of the ABCG2 transporter. *Br. J. Pharmacol.* 149, 506–15.
- Cole, S.P.C., Bhardwaj, G., Gerlach, J.H., Mackie, J.E., Grant, C.E., Almquist, K.C., Stewart, A.J., Kurz, E.U., Duncan, A.M.V., Deeley, R.G., 1992. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* (80-). 258, 1650–1654.
- Cordon-Cardo, C., O'Brien, J.P., Casals, D., Rittman-Grauer, L., Biedler, J.L., Melamed, M.R., Bertino, J.R., 1989. Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood-brain barrier sites. *Proc. Natl. Acad. Sci. U. S. A.* 86, 695–8.
- Cox, M.H., Kapoor, P., Briggs, D.A., Kerr, I.D., 2018. Residues contributing to drug transport by ABCG2 are localised to multiple drug-binding pockets. *Biochem. J.* BCJ20170923.
- Dai, C.L., Tiwari, A.K., Wu, C.P., Su, X.D., Wang, S.R., Liu, D.G., Ashby, C.R., Huang, Y., Robey, R.W., Liang, Y.J., Chen, L.M., Shi, C.J., Ambudkar, S. V., Chen, Z.S., Fu, L.W., 2008. Lapatinib (Tykerb, GW572016) reverses multidrug resistance in cancer cells by inhibiting the activity of ATP-binding cassette subfamily B member 1 and G member 2. *Cancer Res.* 68, 7905–7914.
- Davie, J.R., 2003. Inhibition of histone deacetylase activity by butyrate. *J. Nutr.* 133, 2485S–2493S.
- Dawson, R.J.P., Hollenstein, K., Locher, K.P., 2007. Uptake or extrusion: Crystal structures of full ABC transporters suggest a common mechanism. *Mol. Microbiol.*

- Dawson, R.J.P., Locher, K.P., 2007. Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett.* 581, 935–938.
- Dawson, R.J.P., Locher, K.P., 2006. Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–5.
- de Gooijer, M.C., de Vries, N.A., Buckle, T., Buil, L.C.M., Beijnen, J.H., Boogerd, W., van Tellingen, O., 2018. Improved Brain Penetration and Antitumor Efficacy of Temozolomide by Inhibition of ABCB1 and ABCG2. *Neoplasia (United States)* 20, 710–720.
- de Vree, J.M., Jacquemin, E., Sturm, E., Cresteil, D., Bosma, P.J., Aten, J., Deleuze, J.F., Desrochers, M., Burdelski, M., Bernard, O., Oude Elferink, R.P., Hadchouel, M., 1998. Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis. *Proc. Natl. Acad. Sci. U. S. A.* 95, 282–7.
- Deeks, E.D., 2016. Lumacaftor/Ivacaftor: A Review in Cystic Fibrosis. *Drugs* 76, 1191–1201.
- Deeley, R.G., Cole, S.P.C., 2006. Substrate recognition and transport by multidrug resistance protein 1 (ABCC1). *FEBS Lett.*
- Deppe, S., Ripperger, A., Weiss, J., Ergün, S., Benndorf, R.A., 2014. Impact of genetic variability in the ABCG2 gene on ABCG2 expression, function, and interaction with AT1 receptor antagonist telmisartan. *Biochem. Biophys. Res. Commun.* 443, 1211–1217.
- Diederichs, K., Diez, J., Greller, G., Müller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W., Welte, W., 2000. Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*. *EMBO J.* 19, 5951–61.
- Diop, N.K., Hrycyna, C.A., 2005. N-Linked Glycosylation of the Human ABC Transporter ABCG2 on Asparagine 596 Is Not Essential for Expression, Transport Activity, or Trafficking to the Plasma Membrane[†]. *Biochemistry* 44, 5420–5429.
- Dixon, P.H., Weerasekera, N., Linton, K.J., Donaldson, O., Chambers, J., Egginton, E., Weaver, J., Nelson-Piercy, C., de Swiet, M., Warnes, G., Elias, E., Higgins, C.F., Johnston, D.G., McCarthy, M.I., Williamson, C., 2000. Heterozygous MDR3 missense mutation associated with intrahepatic cholestasis of pregnancy: evidence for a defect in protein trafficking. *Hum. Mol. Genet.* 9, 1209–17.
- Do, T.M., Noel-Hudson, M.S., Ribes, S., Besengez, C., Smirnova, M., Cisternino, S., Buyse, M., Calon, F., Chimini, G., Chacun, H., Scherrmann, J.M., Farinotti, R., Bourasset, F., 2012. ABCG2-and ABCG4-mediated efflux of amyloid- β peptide 1-40 at the mouse blood-brain barrier. *J. Alzheimer's Dis.* 30, 155–166.
- Dohse, M., Scharenberg, C., Shukla, S., Robey, R.W., Volkmann, T., Deeken, J.F., Brendel, C., Ambudkar, S. V., Neubauer, A., Bates, S.E., 2010. Comparison of ATP-binding cassette transporter interactions with the tyrosine kinase inhibitors imatinib, nilotinib, and dasatinib. *Drug Metab. Dispos.* 38, 1371–1380.

- Dörr, J.M., Scheidelaar, S., Koorengel, M.C., Dominguez, J.J., Schäfer, M., van Walree, C.A., Killian, J.A., 2016. The styrene–maleic acid copolymer: a versatile tool in membrane research. *Eur. Biophys. J.*
- Doyle, L.A., Yang, W., Abruzzo, L. V, Krogmann, T., Gao, Y., Rishi, A.K., Ross, D.D., 1998. A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc. Natl. Acad. Sci.* 95, 15665–15670.
- DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.M., Miller, J.H., Calos, M.P., 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell. Biol.* 7, 379–87.
- Ebert, B., Seidel, A., Lampen, A., 2005. Identification of BCRP as transporter of benzo[a]pyrene conjugates metabolically formed in Caco-2 cells and its induction by Ah-receptor agonists. *Carcinogenesis* 26, 1754–1763.
- Ee, P.L.R., Kamalakaran, S., Tonetti, D., He, X., Ross, D.D., Beck, W.T., 2004. Identification of a Novel Estrogen Response Element in the Breast Cancer Resistance Protein (ABCG2) Gene. *Cancer Res.* 64, 1247–1251.
- Ejendal, K.F.K., Diop, N.K., Schweiger, L.C., Hrycyna, C.A., 2006. The nature of amino acid 482 of human ABCG2 affects substrate transport and ATP hydrolysis but not substrate binding. *Protein Sci.* 15, 1597–607.
- Elcock, A.H., McCammon, J.A., 2001. Identification of protein oligomerization states by analysis of interface conservation. *Proc. Natl. Acad. Sci.* 98, 2990–2994.
- Elkind, N.B., Szentpétery, Z., Apáti, Á., Özvegy-Laczka, C., Várady, G., Ujhelly, O., Szabó, K., Homolya, L., Váradi, A., Buday, L., Kéri, G., Német, K., Sarkadi, B., 2005. Multidrug transporter ABCG2 prevents tumor cell death induced by the epidermal growth factor receptor inhibitor Iressa (ZD1839, Gefitinib). *Cancer Res.* 65, 1770–1777.
- Ellgaard, L., McCaul, N., Chatsisvili, A., Braakman, I., 2016. Co- and Post-Translational Protein Folding in the ER. *Traffic* 17, 615–638.
- Erlichman, C., Boerner, S.A., Hallgren, C.G., Spieker, R., Wang, X.Y., James, C.D., Scheffer, G.L., Maliepaard, M., Ross, D.D., Bible, K.C., Kaufmann, S.H., 2001. The HER tyrosine kinase inhibitor CI1033 enhances cytotoxicity of 7-ethyl-10-hydroxycamptothecin and topotecan by inhibiting breast cancer resistance protein-mediated drug efflux. *Cancer Res.* 61, 739–748.
- Ferreira, R.J., Bonito, C.A., Cordeiro, M.N.D.S., Ferreira, M.-J.U., dos Santos, D.J.V.A., 2017. Structure-function relationships in ABCG2: insights from molecular dynamics simulations and molecular docking studies. *Sci. Rep.* 7, 15534.
- Filipits, M., Suchomel, R.W., Dekan, G., Haider, K., Valdimarsson, G., Depisch, D., Pirker, R., 1996. MRP and MDR1 gene expression in primary breast carcinomas. *Clin. Cancer Res.* 2, 1231–7.
- Fredriksson, R., Nordström, K.J.V., Stephansson, O., Hägglund, M.G.A., Schiöth, H.B., 2008. The solute carrier (SLC) complement of the human genome: Phylogenetic classification reveals four major families. *FEBS Lett.* 582, 3811–3816.

- Fujita, Y., Noguchi, K., Suzuki, T., Katayama, K., Sugimoto, Y., 2013. Biochemical interaction of anti-HCV telaprevir with the ABC transporters P-glycoprotein and breast cancer resistance protein. *BMC Res. Notes* 6, 445.
- Fukuda, Y., Aguilar-Bryan, L., Vaxillaire, M., Dechaume, A., Wang, Y., Dean, M., Moitra, K., Bryan, J., Schuetz, J.D., 2011. Conserved intramolecular disulfide bond is critical to trafficking and fate of ATP-binding cassette (ABC) transporters ABCB6 and sulfonylurea receptor 1 (SUR1)/ABCC8. *J. Biol. Chem.* 286, 8481–92.
- Galimberti, S., Marchetti, A., Buttitta, F., Carnicelli, V., Pellegrini, S., Bevilacqua, G., Petrini, M., 1998. Multidrug resistance related genes and p53 expression in human non small cell lung cancer. *Anticancer Res.* 18, 2973–6.
- Gedeon, C., Behravan, J., Koren, G., Piquette-Miller, M., 2006. Transport of Glyburide by Placental ABC Transporters: Implications in Fetal Drug Exposure. *Placenta* 27, 1096–1102.
- George, A.M., Jones, P.M., 2012. Perspectives on the structure-function of ABC transporters: the Switch and Constant Contact models. *Prog. Biophys. Mol. Biol.* 109, 95–107.
- Glavinas, H., Krajcsi, P., Cserepes, J., Sarkadi, B., 2004. The role of ABC transporters in drug resistance, metabolism and toxicity. *Curr. Drug Deliv.* 1, 27–42.
- Gros, P., Croop, J., Roninson, I., Varshavsky, A., Housman, D.E., 1986. Isolation and characterization of DNA sequences amplified in multidrug-resistant hamster cells. *Proc. Natl. Acad. Sci. U. S. A.* 83, 337–41.
- Gulati, S., Jamshad, M., Knowles, T.J., Morrison, K.A., Downing, R., Cant, N., Collins, R., Koenderink, J.B., Ford, R.C., Overduin, M., Kerr, I.D., Dafforn, T.R., Rothnie, A.J., 2014. Detergent-free purification of ABC (ATP-binding-cassette) transporters. *Biochem. J.* 461, 269–278.
- Gupta, A., Unadkat, J.D., Mao, Q., 2007. Interactions of azole antifungal agents with the human breast cancer resistance protein (BCRP). *J. Pharm. Sci.* 96, 3226–3235.
- Gupta, A., Zhang, Y., Unadkat, J.D., Mao, Q., 2004. HIV protease inhibitors are inhibitors but not substrates of the human breast cancer resistance protein (BCRP/ABCG2). *J. Pharmacol. Exp. Ther.* 310, 334–41.
- Haider, A.J., 2011. Function and Dimerisation of the Human Multidrug Resistance Pump: ABCG2. University of Nottingham.
- Haider, A.J., Briggs, D., Self, T.J., Chilvers, H.L., Holliday, N.D., Kerr, I.D., 2011. Dimerization of ABCG2 Analysed by Bimolecular Fluorescence Complementation. *PLoS One* 6, e25818.
- Haider, A.J., Cox, M.H., Jones, N., Goode, A.J., Bridge, K.S., Wong, K., Briggs, D., Kerr, I.D., 2015. Identification of residues in ABCG2 affecting protein trafficking and drug transport, using co-evolutionary analysis of ABCG sequences. *Biosci. Rep.* 35.

- Hazlehurst, L.A., Foley, N.E., Gleason-Guzman, M.C., Hacker, M.P., Cress, A.E., Greenberger, L.W., De Jong, M.C., Dalton, W.S., 1999. Multiple mechanisms confer drug resistance to mitoxantrone in the human 8226 myeloma cell line. *Cancer Res.* 59, 1021–1028.
- Hegedus, C., Özvegy-Laczka, C., Apáti, A.Á., Magócsi, M., Német, K., Orfi, L., Kéri, G., Katona, M., Takáts, Z., Váradi, A., Szakács, G., Sarkadi, B., 2009. Interaction of nilotinib, dasatinib and bosutinib with ABCB1 and ABCG2: Implications for altered anti-cancer effects and pharmacological properties. *Br. J. Pharmacol.* 158, 1153–1164.
- Hegedűs, C., Szakács, G., Homolya, L., Orbán, T.I., Telbisz, Á., Jani, M., Sarkadi, B., 2009. Ins and outs of the ABCG2 multidrug transporter: An update on in vitro functional assays. *Adv. Drug Deliv. Rev.* 61, 47–56.
- Henriksen, U., Fog, J.U., Litman, T., Gether, U., 2005. Identification of intra- and intermolecular disulfide bridges in the multidrug resistance transporter ABCG2. *J. Biol. Chem.* 280, 36926–36934.
- Heyes, N., Kapoor, P., Kerr, I.D., n.d. Polymorphisms of the multidrug pump ABCG2: a systematic review of their effect on protein expression, function and drug pharmacokinetics. Press.
- Higgins, C.F., 1992. ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* 8, 67–113.
- Higgins, C.F., Hiles, I.D., Salmond, G.P., Gill, D.R., Downie, J.A., Evans, I.J., Holland, I.B., Gray, L., Buckel, S.D., Bell, A.W., 1986. A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature* 323, 448–50.
- Higgins, C.F., Linton, K.J., 2004. The ATP switch model for ABC transporters. *Nat. Struct. Mol. Biol.* 11, 918–926.
- Hopfner, D.R., Almquist, K.C., Leslie, E.M., Gerlach, J.H., Grant, C.E., Deeley, R.G., Cole, S.P., 1997. Membrane topology of the multidrug resistance protein (MRP). A study of glycosylation-site mutants reveals an extracytosolic NH₂ terminus. *J. Biol. Chem.* 272, 23623–30.
- Holland, I.B., Blight, M.A., 1999. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J. Mol. Biol.* 293, 381–399.
- Hollenstein, K., Frei, D.C., Locher, K.P., 2007. Structure of an ABC transporter in complex with its binding protein. *Nature* 446, 213–216.
- Honjo, Y., Hrycyna, C.A., Yan, Q.W., Medina-Pérez, W.Y., Robey, R.W., van de Laar, A., Litman, T., Dean, M., Bates, S.E., 2001. Acquired mutations in the MXR/BCRP/ABCP gene alter substrate specificity in MXR/BCRP/ABCP-overexpressing cells. *Cancer Res.* 61, 6635–9.
- Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., Tainer, J.A., 2000. Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* 101, 789–800.

- Horse, A.J., 2018. Analysis of ABCG2 pharmacology and oligomerisation by solution based fluorescence correlation microscopy. University of Nottingham.
- Hou, Y., Li, C.-Z., Palaniyandi, K., Magtibay, P.M., Homolya, L., Sarkadi, B., Chang, X., 2009. Effects of putative catalytic base mutation E211Q on ABCG2-mediated methotrexate transport. *Biochemistry* 48, 9122–31.
- Houghton, P.J., Germain, G.S., Harwood, F.C., Schuetz, J.D., Stewart, C.F., Buchdunger, E., Traxler, P., 2004. Imatinib Mesylate Is a Potent Inhibitor of the ABCG2 (BCRP) Transporter and Reverses Resistance to Topotecan and SN-38 in Vitro. *Cancer Res.* 64, 2333–2337.
- Hrycyna, C.A., Ramachandra, M., Germann, U.A., Cheng, P.W., Pastan, I., Gottesman, M.M., 1999. Both ATP sites of human P-glycoprotein are essential but not symmetric. *Biochemistry* 38, 13887–13899.
- Huang, S.-M., 2009. Clinical Drug Interaction Studies-Study Design, Data Analysis, and Clinical Implications Guidance for Industry DRAFT GUIDANCE.
- Huls, M., Brown, C.D.A., Windass, A.S., Sayer, R., van den Heuvel, J.J.M.W., Heemskerk, S., Russel, F.G.M., Masereeuw, R., 2008. The breast cancer resistance protein transporter ABCG2 is expressed in the human kidney proximal tubule apical membrane. *Kidney Int.* 73, 220–225.
- Hung, L.W., Wang, I.X., Nikaido, K., Liu, P.Q., Ames, G.F.L., Kim, S.H., 1998. Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* 396, 703–707.
- Igarashi, Y., Aoki, K.F., Mamitsuka, H., Kuma, K.I., Kanehisa, M., 2004. The evolutionary repertoires of the eukaryotic-type ABC transporters in terms of the phylogeny of ATP-binding domains in eukaryotes and prokaryotes. *Mol. Biol. Evol.* 21, 2149–2160.
- Jackson, S.M., Manolaridis, I., Kowal, J., Zechner, M., Taylor, N.M.I., Bause, M., Bauer, S., Bartholomaeus, R., Bernhardt, G., Koenig, B., Buschauer, A., Stahlberg, H., Altmann, K.-H., Locher, K.P., 2018. Structural basis of small-molecule inhibition of human multidrug transporter ABCG2. *Nat. Struct. Mol. Biol.* 25, 333–340.
- Jacquemin, E., De Vree, J.M., Cresteil, D., Sokal, E.M., Sturm, E., Dumont, M., Scheffer, G.L., Paul, M., Burdelski, M., Bosma, P.J., Bernard, O., Hadchouel, M., Elferink, R.P., 2001. The wide spectrum of multidrug resistance 3 deficiency: from neonatal cholestasis to cirrhosis of adulthood. *Gastroenterology* 120, 1448–58.
- Jones, P.M., George, A.M., 1999. Subunit interactions in ABC transporters: towards a functional architecture. *FEMS Microbiol. Lett.* 179, 187–202.
- Jonker, J.W., Merino, G., Musters, S., Van Herwaarden, A.E., Bolscher, E., Wagenaar, E., Mesman, E., Dale, T.C., Schinkel, A.H., 2005. The breast cancer resistance protein BCRP (ABCG2) concentrates drugs and carcinogenic xenotoxins into milk. *Nat. Med.* 11, 127–129.
- Juliano, R.L., Ling, V., 1976. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *BBA - Biomembr.* 455, 152–162.

- Kage, K., Fujita, T., Sugimoto, Y., 2005. Role of Cys-603 in dimer/oligomer formation of the breast cancer resistance protein BCRP/ABCG2. *Cancer Sci.* 96, 866–72.
- Kao, S.-H., Wang, W.-L., Chen, C.-Y., Chang, Y.-L., Wu, Y.-Y., Wang, Y.-T., Wang, S.-P., Nesvizhskii, A.I., Chen, Y.-J., Hong, T.-M., Yang, P.-C., 2015. Analysis of Protein Stability by the Cycloheximide Chase Assay. *Bio-protocol* 5.
- Kao, S.-H., Wang, W.-L., Chen, C.-Y., Chang, Y.-L., Wu, Y.-Y., Wang, Y.-T., Wang, S.-P., Nesvizhskii, A.I., Chen, Y.-J., Hong, T.-M., Yang, P.-C., 2014. GSK3 β controls epithelial–mesenchymal transition and tumor metastasis by CHIP-mediated degradation of Slug. *Oncogene* 33, 3172–3182.
- Kelm, K.B., Hoyer, G., Huang, J.C., Michaelis, S., 2004. The Internalization of Yeast Ste6p Follows an Ordered Series of Events Involving Phosphorylation, Ubiquitination, Recognition and Endocytosis. *Traffic* 5, 165–180.
- Kerr, I.D., 2002. Structure and association of ATP-binding cassette transporter nucleotide-binding domains. *Biochim. Biophys. Acta - Biomembr.*
- Kerr, I.D., Haider, A.J., Gelissen, I.C., 2011. The ABCG family of membrane-associated transporters: you don't have to be big to be mighty. *Br. J. Pharmacol.* 164, 1767–79.
- Khunweeraphong, N., Stockner, T., Kuchler, K., 2017. The structure of the human ABC transporter ABCG2 reveals a novel mechanism for drug extrusion. *Sci. Rep.* 7.
- Kim, I.W., Peng, X.H., Sauna, Z.E., FitzGerald, P.C., Xia, D., Müller, M., Nandigama, K., Ambudkar, S. V., 2006. The conserved tyrosine residues 401 and 1044 in ATP sites of human P-glycoprotein are critical for ATP binding and hydrolysis: Evidence for a conserved subdomain, the A-loop in the ATP-binding cassette. *Biochemistry* 45, 7605–7616.
- Kim, Y., Chen, J., 2018. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science (80-.)*. 359, 915–919.
- Kitamura, S., Maeda, K., Wang, Y., Sugiyama, Y., 2008. Involvement of multiple transporters in the hepatobiliary transport of rosuvastatin. *Drug Metab. Dispos.* 36, 2014–2023.
- Kjellström, U., 2015. Reduced macular function in ABCA4 carriers. *Mol. Vis.* 21, 767–82.
- Koo, D.-H., Ryu, M.-H., Ryoo, B.-Y., Beck, M.Y., Na, Y.-S., Shin, J.-G., Lee, S.S., Kim, E.-Y., Kang, Y.-K., 2015. Association of ABCG2 polymorphism with clinical efficacy of imatinib in patients with gastrointestinal stromal tumor. *Cancer Chemother. Pharmacol.* 75, 173–82.
- Krishna, R., Mayer, L.D., 2001. Modulation of P-glycoprotein (PGP) mediated multidrug resistance (MDR) using chemosensitizers: recent advances in the design of selective MDR modulators. *Curr Med Chem Anticancer Agents* 1, 163–174.
- Krishnamurthy, P., Ross, D.D., Nakanishi, T., Bailey-Dell, K., Zhou, S., Mercer, K.E., Sarkadi, B., Sorrentino, B.P., Schuetz, J.D., 2004. The stem cell marker Bcrp/ABCG2 enhances hypoxic cell survival through interactions with heme. *J. Biol. Chem.* 279, 24218–24225.

- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- Lamers, M.H., Perrakis, A., Enzlin, J.H., Winterwerp, H.H.K., de Wind, N., Sixma, T.K., 2000. The crystal structure of DNA mismatch repair protein MutS binding to a G-T mismatch. *Nature* 407, 711–717.
- Langer, O., Conway, D.L., Berkus, M.D., Xenakis, E.M.-J., Gonzales, O., 2000. A Comparison of Glyburide and Insulin in Women with Gestational Diabetes Mellitus. *N. Engl. J. Med.* 343, 1134–1138.
- László, L., Sarkadi, B., Hegedűs, T., Sarkadi, B., Ozvegy-Laczka, C., Ozvegy-Laczka, C., 2016. Jump into a New Fold—A Homology Based Model for the ABCG2/BCRP Multidrug Transporter. *PLoS One* 11, e0164426.
- Lawson, J., O'Mara, M.L., Kerr, I.D., 2008. Structure-based interpretation of the mutagenesis database for the nucleotide binding domains of P-glycoprotein. *Biochim. Biophys. Acta - Biomembr.*
- Lee, J.-Y., Kinch, L.N., Borek, D.M., Wang, J., Wang, J., Urbatsch, I.L., Xie, X.-S., Grishin, N. V., Cohen, J.C., Otwinowski, Z., Hobbs, H.H., Rosenbaum, D.M., 2016. Crystal structure of the human sterol transporter ABCG5/ABCG8. *Nature* 533, 561–564.
- Lefevre, F., Rémy, M.-H., Masson, J.-M., 1997. Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function. *Nucleic Acids Res.* 25, 447–448.
- Leier, I., Jedlitschky, G., Buchholz, U., Cole, S.P.C., Deeley, R.G., Keppler, D., 1994. The MRP gene encodes an ATP-dependent export pump for leukotriene C4 and structurally related conjugates. *J. Biol. Chem.* 269, 27807–27810.
- Leonard, G.D., Fojo, T., Bates, S.E., 2003. The role of ABC transporters in clinical practice. *Oncologist* 8, 411–24.
- Leslie, E.M., Deeley, R.G., Cole, S.P.C., 2005. Multidrug resistance proteins: role of P-glycoprotein, MRP1, MRP2, and BCRP (ABCG2) in tissue defense. *Toxicol. Appl. Pharmacol.* 204, 216–237.
- Li, Y.-F., Polgar, O., Okada, M., Esser, L., Bates, S.E., Xia, D., 2007. Towards understanding the mechanism of action of the multidrug resistance-linked half-ABC transporter ABCG2: a molecular modeling study. *J. Mol. Graph. Model.* 25, 837–51.
- Linton, K.J., 2007. Structure and function of ABC transporters. *Physiology (Bethesda)*. 22, 122–130.
- Linton, K.J., Higgins, C.F., 2007. Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Arch.* 453, 555–67.
- Linton, K.J., Higgins, C.F., 1998. The Escherichia coli ATP-binding cassette (ABC) proteins. *Mol. Microbiol.*

- Litman, T., Brangi, M., Hudson, E., Fetsch, P., Abati, A., Ross, D.D., Miyake, K., Resau, J.H., Bates, S.E., 2000. The multidrug-resistant phenotype associated with overexpression of the new ABC half-transporter, MXR (ABCG2). *J. Cell Sci.* 113 (Pt 11), 2011–21.
- Liu, Y., Yang, Y., Qi, J., Peng, H., Zhang, J.-T., 2008. Effect of cysteine mutagenesis on the function and disulfide bond formation of human ABCG2. *J. Pharmacol. Exp. Ther.* 326, 33–40.
- Locher, K.P., Lee, A.T., Rees, D.C., 2002. The *E. coli* BtuCD structure: A framework for ABC transporter architecture and mechanism. *Science* (80-). 296, 1091–1098.
- Longobardi Givan, A., 2001. *Flow cytometry first principles*, Second Edi. ed. Wiley and Sons.
- Loo, T.W., Clarke, D.M., 2000. Identification of residues within the drug-binding domain of the human multidrug resistance P-glycoprotein by cysteine-scanning mutagenesis and reaction with dibromobimane. *J. Biol. Chem.* 275, 39272–39278.
- Lugli, E., Roederer, M., Cossarizza, A., 2010. Data analysis in flow cytometry: The future just started. *Cytom. Part A* 77A, 705–713.
- Maki, N., Hafkemeyer, P., Dey, S., 2003. Allosteric modulation of human P-glycoprotein. Inhibition of transport by preventing substrate translocation and dissociation. *J. Biol. Chem.* 278, 18132–18139.
- Maliepaard, M., Scheffer, G.L., Faneyte, I.F., van Gastelen, M.A., Pijnenborg, A.C., Schinkel, A.H., van De Vijver, M.J., Scheper, R.J., Schellens, J.H., 2001. Subcellular localization and distribution of the breast cancer resistance protein transporter in normal human tissues. *Cancer Res.* 61, 3458–64.
- Maliepaard, M., Van Gastelen, M.A., Tohgo, A., Hausheer, F.H., Van Waardenburg, R.C.A.M., De Jong, L.A., Pluim, D., Beijnen, J.H., Schellens, J.H.M., 2001. Circumvention of breast cancer resistance protein (BCRP)-mediated resistance to camptothecins in vitro using non-substrate drugs or the BCRP inhibitor GF120918. *Clin. Cancer Res.* 7, 935–941.
- Mao, Q., 2008. BCRP/ABCG2 in the placenta: Expression, function and regulation. *Pharm. Res.*
- Mao, Q., Unadkat, J.D., 2015. Role of the Breast Cancer Resistance Protein (BCRP/ABCG2) in Drug Transport—an Update. *AAPS J.* 17, 65–82.
- Marie, J.P., Zittoun, R., Sikic, B.I., Marie, J.P., 1991. MULTIDRUG RESISTANCE (MDR1) GENE-EXPRESSION IN ADULT ACUTE LEUKEMIAS - CORRELATIONS WITH TREATMENT OUTCOME AND INVITRO DRUG SENSITIVITY\rDrug resistance in hematologic malignancies. *Blood* 78, 88s–90s.
- Martin, C., Berridge, G., Higgins, C.F., Mistry, P., Charlton, P., Callaghan, R., 2000. Communication between multiple drug binding sites on P-glycoprotein. *Mol. Pharmacol.* 58, 624–32.
- McDevitt, C.A., Collins, R., Kerr, I.D., Callaghan, R., 2009. Purification and structural analyses of ABCG2. *Adv. Drug Deliv. Rev.*

- McDevitt, C.A., Collins, R.F., Conway, M., Modok, S., Storm, J., Kerr, I.D., Ford, R.C., Callaghan, R., 2006. Purification and 3D structural analysis of oligomeric human multidrug transporter ABCG2. *Structure* 14, 1623–32.
- McDevitt, C.A., Crowley, E., Hobbs, G., Starr, K.J., Kerr, I.D., Callaghan, R., 2008. Is ATP binding responsible for initiating drug translocation by the multidrug transporter ABCG2? *FEBS J.* 275, 4354–4362.
- Merino, G., Jonker, J.W., Wagenaar, E., van Herwaarden, A.E., Schinkel, A.H., 2005. The breast cancer resistance protein (BCRP/ABCG2) affects pharmacokinetics, hepatobiliary excretion, and milk secretion of the antibiotic nitrofurantoin. *Mol. Pharmacol.* 67, 1758–1764.
- Miyake, K., Mickley, L., Litman, T., Zhan, Z., Robey, R., Cristensen, B., Brangi, M., Greenberger, L., Dean, M., Fojo, T., Bates, S.E., 1999. Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: Demonstration of homology to ABC transport genes. *Cancer Res.* 59, 8–13.
- Mlejnek, P., Kosztyu, P., Dolezel, P., Bates, S.E., Ruzickova, E., 2017. Reversal of ABCB1 mediated efflux by imatinib and nilotinib in cells expressing various transporter levels. *Chem. Biol. Interact.* 273, 171–179.
- Mo, W., Zhang, J.-T., 2012. Human ABCG2: structure, function, and its role in multidrug resistance. *Int. J. Biochem. Mol. Biol.* 3, 1–27.
- Moody, J.E., Millen, L., Binns, D., Hunt, J.F., Thomas, P.J., 2002. Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters. *J. Biol. Chem.* 277, 21111–4.
- Murphy, C.I., Piwnicka-Worms, H., Grünwald, S., Romanow, W.G., Francis, N., Fan, H.-Y., 2004. Overview of the Baculovirus Expression System. *Curr. Protoc. Mol. Biol.* 65, 16.9.1-16.9.10.
- Nakagawa, R., Hara, Y., Arakawa, H., Nishimura, S., Komatani, H., 2002. ABCG2 confers resistance to indolocarbazole compounds by ATP-dependent transport. *Biochem. Biophys. Res. Commun.* 299, 669–675.
- Nakanishi, T., Bailey-Dell, K.J., Hassel, B.A., Shiozawa, K., Sullivan, D.M., Turner, J., Ross, D.D., 2006. Novel 5' untranslated region variants of BCRP mRNA are differentially expressed in drug-selected cancer cells and in normal human tissues: Implications for drug resistance, tissue-specific expression, and alternative promoter usage. *Cancer Res.* 66, 5007–5011.
- Nakanishi, T., Doyle, L.A., Hassel, B., Wei, Y., Bauer, K.S., Wu, S., Pumplin, D.W., Fang, H.-B., Ross, D.D., 2003. Functional characterization of human breast cancer resistance protein (BCRP, ABCG2) expressed in the oocytes of *Xenopus laevis*. *Mol. Pharmacol.* 64, 1452–62.

- Nakatomi, K., Yoshikawa, M., Oka, M., Ikegami, Y., Hayasaka, S., Sano, K., Shiozawa, K., Kawabata, S., Soda, H., Ishikawa, T., Tanabe, S., Kohno, S., 2001. Transport of 7-Ethyl-10-hydroxycamptothecin (SN-38) by Breast Cancer Resistance Protein ABCG2 in Human Lung Cancer Cells. *Biochem. Biophys. Res. Commun.* 288, 827–832.
- Ni, Z., Bikadi, Z., Cai, X., Rosenberg, M.F., Mao, Q., 2010a. Transmembrane helices 1 and 6 of the human breast cancer resistance protein (BCRP/ABCG2): identification of polar residues important for drug transport. *Am. J. Physiol. Cell Physiol.* 299, C1100-9.
- Ni, Z., Bikadi, Z., Rosenberg, M.F., Mao, Q., 2010b. Structure and function of the human breast cancer resistance protein (BCRP/ABCG2). *Curr. Drug Metab.* 11, 603–17.
- Ni, Z., Bikadi, Z., Shuster, D.L., Zhao, C., Rosenberg, M.F., Mao, Q., 2011. Identification of Proline Residues in or near the Transmembrane Helices of the Human Breast Cancer Resistance Protein (BCRP/ABCG2) That Are Important for Transport Activity and Substrate Specificity. *Biochemistry* 50, 8057–8066.
- Ni, Z., Mark, M.E., Cai, X., Mao, Q., 2010c. Fluorescence resonance energy transfer (FRET) analysis demonstrates dimer/oligomer formation of the human breast cancer resistance protein (BCRP/ABCG2) in intact cells. *Int. J. Biochem. Mol. Biol.* 1, 1–11.
- Nigam, S.K., 2014. What do drug transporters really do? *Nat. Rev. Drug Discov.*
- Oda, Y., Dockhorn-Dworniczak, B., Jürgens, H., Roessner, A., 1997. Expression of multidrug resistance-associated protein gene in Ewing's sarcoma and malignant peripheral neuroectodermal tumor of bone. *J. Cancer Res. Clin. Oncol.* 123, 237–239.
- Oldham, M.L., Khare, D., Quioco, F.A., Davidson, A.L., Chen, J., 2007a. Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450, 515–521.
- Oldham, M.L., Khare, D., Quioco, F.A., Davidson, A.L., Chen, J., 2007b. Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450, 515–521.
- Otero, J.A., Miguel, V., González-Lobato, L., García-Villalba, R., Espín, J.C., Prieto, J.G., Merino, G., Álvarez, A.I., 2015. Effect of bovine ABCG2 polymorphism Y581S SNP on secretion into milk of enterolactone, riboflavin and uric acid. *Animal* 10, 238–247.
- Ozvegy-Laczka, C., Hegedus, T., Várady, G., Ujhelly, O., Schuetz, J.D., Váradi, A., Kéri, G., Orfi, L., Németh, K., Sarkadi, B., 2004. High-affinity interaction of tyrosine kinase inhibitors with the ABCG2 multidrug transporter. *Mol. Pharmacol.* 65, 1485–1495.
- Ozvegy-Laczka, C., Köblös, G., Sarkadi, B., Váradi, A., 2005. Single amino acid (482) variants of the ABCG2 multidrug transporter: major differences in transport capacity and substrate recognition. *Biochim. Biophys. Acta* 1668, 53–63.
- Özvegy, C., Váradi, A., Sarkadi, B., 2002. Characterization of Drug Transport, ATP Hydrolysis, and Nucleotide Trapping by the Human ABCG2 Multidrug Transporter. *J. Biol. Chem.* 277, 47980–47990.

- Pál, a, Méhn, D., Molnár, E., Gedey, S., Mészáros, P., Nagy, T., Glavinas, H., Janáky, T., von Richter, O., Báthori, G., Szente, L., Krajcsi, P., 2007. Cholesterol potentiates ABCG2 activity in a heterologous expression system: improved in vitro model to study function of human ABCG2. *J. Pharmacol. Exp. Ther.* 321, 1085–1094.
- Pan, Y.-Z., Morris, M.E., Yu, A.-M., 2009. MicroRNA-328 negatively regulates the expression of breast cancer resistance protein (BCRP/ABCG2) in human cancer cells. *Mol. Pharmacol.* 75, 1374–9.
- Pavek, P., 2004. Human Breast Cancer Resistance Protein: Interactions with Steroid Drugs, Hormones, the Dietary Carcinogen 2-Amino-1-methyl-6-phenylimidazo(4,5-b)pyridine, and Transport of Cimetidine. *J. Pharmacol. Exp. Ther.* 312, 144–152.
- Polgar, O., Ierano, C., Tamaki, A., Stanley, B., Ward, Y., Xia, D., Tarasova, N., Robey, R.W., Bates, S.E., 2010. Mutational analysis of threonine 402 adjacent to the GXXXG dimerization motif in transmembrane segment 1 of ABCG2. *Biochemistry* 49, 2235–45.
- Polgar, O., Robey, R.W., Morisaki, K., Dean, M., Michejda, C., Sauna, Z.E., Ambudkar, S. V., Tarasova, N., Bates, S.E., 2004. Mutational analysis of ABCG2: Role of the GXXXG motif. *Biochemistry* 43, 9448–9456.
- Rabeh, W.M., Bossard, F., Xu, H., Okiyoneda, T., Bagdany, M., Mulvihill, C.M., Du, K., di Bernardo, S., Liu, Y., Konermann, L., Roldan, A., Lukacs, G.L., 2012. Correction of both NBD1 energetics and domain interface is required to restore Δ F508 CFTR folding and function. *Cell* 148, 150–63.
- Rabindran, S.K., He, H., Singh, M., Brown, E., Collins, K.I., Annable, T., Greenberger, L.M., 1998. Reversal of a novel multidrug resistance mechanism in human colon carcinoma cells by fumitremorgin C. *Cancer Res.* 58, 5850–5858.
- Rabindran, S.K., Ross, D.D., Doyle, L.A., Yang, W., Greenberger, L.M., 2000. Fumitremorgin C reverses multidrug resistance in cells transfected with the breast cancer resistance protein. *Cancer Res.* 60, 47–50.
- Radcliff, G., Jaroszeski, M.J., 1998. Basics of flow cytometry. *Methods Mol. Biol.* 91, 1–24.
- Rice, A.J., Park, A., Pinkett, H.W., 2014. Diversity in ABC transporters: Type I, II and III importers. *Crit. Rev. Biochem. Mol. Biol.* 49, 426–437.
- Robey, R.W., Honjo, Y., Morisaki, K., Nadjem, T.A., Runge, S., Risbood, M., Poruchybsky, M.S., Bates, S.E., 2003. Mutations at amino-acid 482 in the ABCG2 gene affect substrate and antagonist specificity. *Br. J. Cancer* 89, 1971–1978.
- Robey, R.W., Honjo, Y., van de Laar, A., Miyake, K., Regis, J.T., Litman, T., Bates, S.E., 2001a. A functional assay for detection of the mitoxantrone resistance protein, MXR (ABCG2). *Biochim. Biophys. Acta - Biomembr.* 1512, 171–182.
- Robey, R.W., Medina-Peèrez, W.Y., Nishiyama, K., Lahusen, T., Miyake, K., Litman, T., Senderowicz, A.M., Ross, D.D., Bates, S.E., 2001b. Overexpression of the ATP-binding cassette half-transporter, ABCG2 (MXR/BCRP/ABCP1), in flavopiridol-resistant human breast cancer cells. *Clin. Cancer Res.* 7, 145–152.

- Robey, R.W., Pluchino, K.M., Hall, M.D., Fojo, A.T., Bates, S.E., Gottesman, M.M., 2018. Revisiting the role of ABC transporters in multidrug-resistant cancer. *Nat. Rev. Cancer.*
- Robey, R.W., Steadman, K., Polgar, O., Bates, S.E., 2005. ABCG2-mediated transport of photosensitizers: Potential impact on photodynamic therapy. *Cancer Biol. Ther.* 4, 187–194.
- Roninson, I.B., Chin, J.E., Choi, K.G., Gros, P., Housman, D.E., Fojo, A., Shen, D.W., Gottesman, M.M., Pastan, I., 1986. Isolation of human *mdr* DNA sequences amplified in multidrug-resistant KB carcinoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 83, 4538–42.
- Rosenberg, M.F., Bikadi, Z., Chan, J., Liu, X., Ni, Z., Cai, X., Ford, R.C., Mao, Q., 2010. The Human Breast Cancer Resistance Protein (BCRP/ABCG2) Shows Conformational Changes with Mitoxantrone. *Structure* 18, 1688–1689.
- Rosenberg, M.F., Kamis, A.B., Callaghan, R., Higgins, C.F., Ford, R.C., 2003. Three-dimensional structures of the mammalian multidrug resistance P-glycoprotein demonstrate major conformational changes in the transmembrane domains upon nucleotide binding. *J. Biol. Chem.* 278, 8294–9.
- Rosenberg, M.F., Velarde, G., Ford, R.C., Martin, C., Berridge, G., Kerr, I.D., Callaghan, R., Schmidlin, A., Wooding, C., Linton, K.J., Higgins, C.F., 2001. Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *EMBO J.* 20, 5615–25.
- Rosmorduc, O., Poupon, R., 2007. Low phospholipid associated cholelithiasis: association with mutation in the *MDR3/ABCB4* gene. *Orphanet J. Rare Dis.* 2, 29.
- Russ, W.P., Engelman, D.M., 2000. The GxxxG motif: A framework for transmembrane helix-helix association. *J. Mol. Biol.* 296, 911–919.
- Russel, F.G.M., 2010. Transporters: Importance in drug absorption, distribution, and removal, in: *Enzyme- and Transporter-Based Drug-Drug Interactions: Progress and Future Challenges*. Springer New York, New York, NY, pp. 27–49.
- Sambrook, J., Russell, D.W., 2001. *Molecular Cloning, A laboratory Manual*, in: Volume 2. Cold Spring Harbor Laboratory Press, New York.
- Sanchez, M.P., Govignon-Gion, A., Croiseau, P., Fritz, S., Hozé, C., Miranda, G., Martin, P., Barbat-Leterrier, A., Letaïef, R., Rocha, D., Brochard, M., Boussaha, M., Boichard, D., 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genet. Sel. Evol.* 49, 68.
- Sarkadi, B., Homolya, L., Szakács, G., Váradi, A., 2006. Human Multidrug Resistance ABCB and ABCG Transporters: Participation in a Chemoimmunity Defense System. *Physiol. Rev.* 86, 1179–1236.
- Schinkel, 1999. P-Glycoprotein, a gatekeeper in the blood-brain barrier. *Adv. Drug Deliv. Rev.* 36, 179–194.

- Senior, A.E., al-Shawi, M.K., Urbatsch, I.L., 1995. The catalytic cycle of P-glycoprotein. *FEBS Lett.* 377, 285–289.
- Sezgin, E., Can, F.B., Schneider, F., Clausen, M.P., Galiani, S., Stanly, T.A., Waithe, D., Colaco, A., Honigmann, A., Wüstner, D., Platt, F., Eggeling, C., 2016. A comparative study on fluorescent cholesterol analogs as versatile cellular reporters. *J. Lipid Res.* 57, 299–309.
- Sharom, F.J., 2008. ABC multidrug transporters: structure, function and role in chemoresistance. *Pharmacogenomics* 9, 105–127.
- Sheppard, D.N., Welsh, M.J., 1999. Structure and function of the CFTR chloride channel. *Physiol. Rev.* 79, S23–S45.
- Shi, Z., Peng, X.X., Kim, I.W., Shukla, S., Si, Q.S., Robey, R.W., Bates, S.E., Shen, T., Ashby, C.R., Fu, L.W., Ambudkar, S. V., Chen, Z.S., 2007. Erlotinib (Tarceva, OSI-774) antagonizes ATP-binding cassette subfamily B member 1 and ATP-binding cassette subfamily G member 2-mediated drug resistance. *Cancer Res.* 67, 11012–11020.
- Shiozawa, K., Okai, M., Soda, H., Yoshikawa, M., Ikegami, Y., Tsurutani, J., Nakatomi, K., Nakamura, Y., Doi, S., Kitazaki, T., Mizuta, Y., Murase, K., Yoshida, H., Ross, D.D., Kohno, S., 2004. Reversal of breast cancer resistance protein (BCRP/ABCG2)-mediated drug resistance by novobiocin, a coumermycin antibiotic. *Int. J. Cancer* 108, 146–151.
- Sikic, B.I., Fisher, G.A., Lum, B.L., Halsey, J., Beketic-Oreskovic, L., Chen, G., 1997. Modulation and prevention of multidrug resistance by inhibitors of P-glycoprotein. *Cancer Chemother. Pharmacol.* 40 Suppl, S13-9.
- Singh, H., Velamakanni, S., Deery, M.J., Howard, J., Wei, S.L., Van Veen, H.W., 2016. ATP-dependent substrate transport by the ABC transporter MsbA is proton-coupled. *Nat. Commun.* 7.
- Sissung, T.M., Baum, C.E., Kirkland, C.T., Gao, R., Gardner, E.R., Figg, W.D., 2010. Pharmacogenetics of membrane transporters: an update on current approaches. *Mol. Biotechnol.* 44, 152–67.
- Sjöstedt, N., van den Heuvel, J.J.M.W., Koenderink, J.B., Kidron, H., 2017. Transmembrane Domain Single-Nucleotide Polymorphisms Impair Expression and Transport Activity of ABC Transporter ABCG2. *Pharm. Res.* 34, 1626–1636.
- Smit, J.J., Schinkel, A.H., Mol, C.A., Majoor, D., Mooi, W.J., Jongsma, A.P., Lincke, C.R., Borst, P., 1994. Tissue distribution of the human MDR3 P-glycoprotein. *Lab. Invest.* 71, 638–49.
- Smith, A.J., Timmermans-Hereijgers, J.L., Roelofsen, B., Wirtz, K.W., van Blitterswijk, W.J., Smit, J.J., Schinkel, A.H., Borst, P., 1994. The human MDR3 P-glycoprotein promotes translocation of phosphatidylcholine through the plasma membrane of fibroblasts from transgenic mice. *FEBS Lett.* 354, 263–6.
- Smith, P.C., Karpowich, N., Millen, L., Moody, J.E., Rosen, J., Thomas, P.J., Hunt, J.F., 2002. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol. Cell* 10, 139–49.

- Sparreboom, A., Gelderblom, H., Marsh, S., Ahluwalia, R., Obach, R., Principe, P., Twelves, C., Verweij, J., McLeod, H.L., 2004. Diflomotecan pharmacokinetics in relation to ABCG2 421C>A genotype. *Clin. Pharmacol. Ther.* 76, 38–44.
- Strouse, J.J., Ivnitski-Steele, I., Waller, A., Young, S.M., Perez, D., Evangelisti, A.M., Ursu, O., Bologa, C.G., Carter, M.B., Salas, V.M., Tegos, G., Larson, R.S., Oprea, T.I., Edwards, B.S., Sklar, L.A., 2013. Fluorescent substrates for flow cytometric evaluation of efflux inhibition in ABCB1, ABCC1, and ABCG2 transporters. *Anal. Biochem.*
- Su, W., Kumar, V., Ding, Y., Ero, R., Serra, A., Lee, B.S.T., Wong, A.S.W., Shi, J., Sze, S.K., Yang, L., Gao, Y.-G., 2018. Ribosome protection by antibiotic resistance ATP-binding cassette protein. *Proc. Natl. Acad. Sci. U. S. A.* 115, 5157–5162.
- Sugimoto, Y., Tsukahara, S., Imai, Y., Sugimoto, Y., Ueda, K., Tsuruo, T., 2003. Reversal of breast cancer resistance protein-mediated drug resistance by estrogen antagonists and agonists. *Mol. Cancer Ther.* 2, 105–112.
- Tamura, A., Onishi, Y., An, R., Koshiba, S., Wakabayashi, K., Hoshijima, K., Priebe, W., Yoshida, T., Kometani, S., Matsubara, T., Mikuriya, K., Ishikawa, T., 2007a. In vitro evaluation of photosensitivity risk related to genetic polymorphisms of human ABC transporter ABCG2 and inhibition by drugs. *Drug Metab. Pharmacokinet.* 22, 428–40.
- Tamura, A., Wakabayashi, K., Onishi, Y., Takeda, M., Ikegami, Y., Sawada, S., Tsuji, M., Matsuda, Y., Ishikawa, T., 2007b. Re-evaluation and functional classification of non-synonymous single nucleotide polymorphisms of the human ATP-binding cassette transporter ABCG2. *Cancer Sci.* 98, 231–239.
- Tamura, A., Watanabe, M., Saito, H., Nakagawa, H., Kamachi, T., Okura, I., Ishikawa, T., 2006. Functional validation of the genetic polymorphisms of human ATP-binding cassette (ABC) transporter ABCG2: identification of alleles that are defective in porphyrin transport. *Mol. Pharmacol.* 70, 287–296.
- Tanna, P., Strauss, R.W., Fujinami, K., Michaelides, M., 2017. Stargardt disease: clinical features, molecular genetics, animal models and therapeutic options. *Br. J. Ophthalmol.* 101, 25–30.
- Tatematsu, K., Yoshimoto, N., Okajima, T., Tanizawa, K., Kuroda, S., 2008. Identification of ubiquitin ligase activity of RBCK1 and its inhibition by splice variant RBCK2 and protein kinase C β . *J. Biol. Chem.* 283, 11575–85.
- Taylor, N.M.I., Manolaridis, I., Jackson, S.M., Kowal, J., Stahlberg, H., Locher, K.P., 2017. Structure of the human multidrug transporter ABCG2. *Nature* 546, 504.
- Telbisz, Á., Hegedüs, C., Váradi, A., Sarkadi, B., Özvegy-Laczka, C., 2014. Regulation of the function of the human ABCG2 multidrug transporter by cholesterol and bile acids: effects of mutations in potential substrate and steroid binding sites. *Drug Metab. Dispos.* 42, 575–85.
- Telbisz, Á., Müller, M., Özvegy-Laczka, C., Homolya, L., Szente, L., Váradi, A., Sarkadi, B., 2007. Membrane cholesterol selectively modulates the activity of the human ABCG2 multidrug transporter. *Biochim. Biophys. Acta - Biomembr.* 1768, 2698–2713.

- ter Beek, J., Guskov, A., Slotboom, D.J., 2014. Structural diversity of ABC transporters. *J. Gen. Physiol.* 143, 419–35.
- Tews, D.S., Nissen, A., Külgen, C., Gaumann, A.K.A., 2000. Drug Resistance-associated Factors in Primary and Secondary Glioblastomas and their Precursor Tumors. *J. Neurooncol.* 50, 227–237.
- To, K.K.W., Robey, R.W., Knutsen, T., Zhan, Z., Ried, T., Bates, S.E., 2009. Escape from hsa-miR-519c enables drug-resistant cells to maintain high expression of ABCG2. *Mol. Cancer Ther.* 8, 2959–2968.
- Tomblin, G., Bartholomew, L., Gimi, K., Tyndall, G.A., Senior, A.E., 2004a. Synergy between Conserved ABC Signature Ser Residues in P-glycoprotein Catalysis. *J. Biol. Chem.* 279, 5363–5373.
- Tomblin, G., Bartholomew, L.A., Tyndall, G.A., Gimi, K., Urbatsch, I.L., Senior, A.E., 2004b. Properties of P-glycoprotein with mutations in the “catalytic carboxylate” glutamate residues. *J. Biol. Chem.* 279, 46518–46526.
- Towbin, H., Staehelin, T., Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U. S. A.* 76, 4350–4.
- Trock, B.J., Leonessa, F., Clarke, R., 1997. Multidrug Resistance in Breast Cancer: a Meta-analysis of MDR1/gp170 Expression and Its Possible Functional Significance. *JNCI J. Natl. Cancer Inst.* 89, 917–931.
- Trott, O., Olson, A.J., 2009. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, NA-NA.
- Tucker, T.G.H.A., Milne, A.M., Fournel-Gigleux, S., Fenner, K.S., Coughtrie, M.W.H., 2012. Absolute immunoquantification of the expression of ABC transporters P-glycoprotein, breast cancer resistance protein and multidrug resistance-associated protein 2 in human liver and duodenum. *Biochem. Pharmacol.* 83, 279–285.
- Turner, J.G., Gump, J.L., Zhang, C., Cook, J.M., Marchion, D., Hazlehurst, L., Munster, P., Schell, M.J., Dalton, W.S., Sullivan, D.M., 2006. ABCG2 expression, function, and promoter methylation in human multiple myeloma. *Blood* 108, 3881–3889.
- Ueda, K., Cornwell, M.M., Gottesman, M.M., Pastan, I., Roninson, I.B., Ling, V., Riordan, J.R., 1986. The mdrl gene, responsible for multidrug-resistance, codes for P-glycoprotein. *Biochem. Biophys. Res. Commun.* 141, 956–962.
- Urbatsch, I.L., Gimi, K., Wilke-Mounts, S., Lerner-Marmarosh, N., Rousseau, M.E., Gros, P., Senior, A.E., 2001. Cysteines 431 and 1074 Are Responsible for Inhibitory Disulfide Cross-linking between the Two Nucleotide-binding Sites in Human P-glycoprotein. *J. Biol. Chem.* 276, 26980–26987.
- Urbatsch, I.L., Gimi, K., Wilke-Mounts, S., Senior, A.E., 2000. Investigation of the role of glutamine-471 and glutamine-1114 in the two catalytic sites of P-glycoprotein. *Biochemistry* 39, 11921–11927.

- Urbatsch, I.L., Gimi, K., Wilke-Mounts, S., Seniors, A.E., 2000. Conserved Walker A Ser residues in the catalytic sites of P-glycoprotein are critical for catalysis and involved primarily at the transition state step. *J. Biol. Chem.* 275, 25031–25038.
- Valley, C.C., Cembran, A., Perlmutter, J.D., Lewis, A.K., Labello, N.P., Gao, J., Sachs, J.N., 2012. The methionine-aromatic motif plays a unique role in stabilizing protein structure. *J. Biol. Chem.* 287, 34979–34991.
- Van der Blik, A.M., Baas, F., Ten Houte de Lange, T., Kooiman, P.M., Van der Velde-Koerts, T., Borst, P., 1987. The human mdr3 gene encodes a novel P-glycoprotein homologue and gives rise to alternatively spliced mRNAs in liver. *EMBO J.* 6, 3325–31.
- Van Der Heijden, J., De Jong, M.C., Dijkmans, B.A.C., Lems, W.F., Oerlemans, R., Kathmann, I., Schalkwijk, C.G., Scheffer, G.L., Scheper, R.J., Jansen, G., 2004. Development of sulfasalazine resistance in human T cells induces expression of the multidrug resistance transporter ABCG2 (BCRP) and augmented production of TNF α . *Ann. Rheum. Dis.* 63, 138–143.
- Van Herwaarden, A.E., Schinkel, A.H., 2006. The function of breast cancer resistance protein in epithelial barriers, stem cells and milk secretion of drugs and xenotoxins. *Trends Pharmacol. Sci.*
- van Herwaarden, A.E., Wagenaar, E., Merino, G., Jonker, J.W., Rosing, H., Beijnen, J.H., Schinkel, A.H., 2007. Multidrug Transporter ABCG2/Breast Cancer Resistance Protein Secretes Riboflavin (Vitamin B2) into Milk. *Mol. Cell. Biol.*
- Venter, H., Shilling, R.A., Velamakanni, S., Balakrishnan, L., van Veen, H.W., 2003. An ABC transporter with a secondary-active multidrug translocator domain. *Nature* 426, 866–870.
- Volk, E.L., Farley, K.M., Wu, Y., Li, F., Robey, R.W., Schneider, E., 2002. Overexpression of wild-type breast cancer resistance protein mediates methotrexate resistance. *Cancer Res.* 62, 5035–40.
- Volk, E.L., Schneider, E., 2003. Wild-type breast cancer resistance protein (BCRP/ABCG2) is a methotrexate polyglutamate transporter. *Cancer Res.* 63, 5538–43.
- Von Stetten, D., Noirclerc-Savoye, M., Goedhart, J., Gadella, T.W.J., Royant, A., 2012. Structure of a fluorescent protein from *Aequorea victoria* bearing the obligate-monomer mutation A206K. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*
- Wakabayashi-Nakao, K., 2009. ABCG2 ubiquitination and proteasomal degradation. *Adv. Drug Deliv. Rev.* 66–72.
- Wakabayashi, K., Nakagawa, H., Tamura, A., Koshiba, S., Hoshijima, K., Komada, M., Ishikawa, T., 2007. Intramolecular disulfide bond is a critical check point determining degradative fates of ATP-binding cassette (ABC) transporter ABCG2 protein. *J. Biol. Chem.* 282, 27841–27846.
- Walker, J.E., Saraste, M., Runswick, M., Gay, N.J., 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945–951.

- Wang, H., Lee, E.W., Cai, X., Ni, Z., Zhou, L., Mao, Q., 2008. Membrane topology of the human breast cancer resistance protein (BCRP/ ABCG2) determined by epitope insertion and immunofluorescence. *Biochemistry* 47, 13778–13787.
- Wang, J., Grishin, N., Kinch, L., Cohen, J.C., Hobbs, H.H., Xie, X.S., 2011. Sequences in the nonconsensus nucleotide-binding domain of ABCG5/ABCG8 required for sterol transport. *J. Biol. Chem.* 286, 7308–7314.
- Wang, X., Furukawa, T., Nitanda, T., Okamoto, M., Sugimoto, Y., Akiyama, S.-I., Baba, M., 2003. Breast cancer resistance protein (BCRP/ABCG2) induces cellular resistance to HIV-1 nucleoside reverse transcriptase inhibitors. *Mol. Pharmacol.* 63, 65–72.
- Wang, X., Nitanda, T., Shi, M., Okamoto, M., Furukawa, T., Sugimoto, Y., Akiyama, S.I., Baba, M., 2004. Induction of cellular resistance to nucleoside reverse transcriptase inhibitors by the wild-type breast cancer resistance protein. *Biochem. Pharmacol.* 68, 1363–1370.
- Weiss, J., Rose, J., Storch, C.H., Ketabi-Kiyanvash, N., Sauer, A., Haefeli, W.E., Efferth, T., 2007. Modulation of human BCRP (ABCG2) activity by anti-HIV drugs. *J. Antimicrob. Chemother.* 59, 238–245.
- Westover, D., Ling, X., Lam, H., Welch, J., Jin, C., Gongora, C., Del Rio, M., Wani, M., Li, F., 2015. FL118, a novel camptothecin derivative, is insensitive to ABCG2 expression and shows improved efficacy in comparison with irinotecan in colon and lung cancer models with ABCG2-induced resistance. *Mol. Cancer* 14, 92.
- Wong, K., 2015. Dynamics and oligomerisation of ABCG2 investigated using various fluorescence techniques. University of Nottingham.
- Wong, K., Briddon, S.J., Holliday, N.D., Kerr, I.D., 2016. Plasma membrane dynamics and tetrameric organisation of ABCG2 transporters in mammalian cells revealed by single particle imaging techniques. *Biochim. Biophys. Acta* 1863, 19–29.
- Wong, K., Ma, J., Rothnie, A., Biggin, P.C., Kerr, I.D., 2014. Towards understanding promiscuity in multidrug efflux pumps. *Trends Biochem. Sci.* 39, 8–16.
- Woo, J.S., Zeltina, A., Goetz, B.A., Locher, K.P., 2012. X-ray structure of the *Yersinia pestis* heme transporter HmuUV. *Nat. Struct. Mol. Biol.* 19, 1310–1315.
- Woodward, O.M., Köttgen, A., Coresh, J., Boerwinkle, E., Guggino, W.B., Köttgen, M., 2009. Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10338–10342.
- Woodward, O.M., Tukaye, D.N., Cui, J., Greenwell, P., Constantoulakis, L.M., Parker, B.S., Rao, A., Köttgen, M., Maloney, P.C., Guggino, W.B., 2013. Gout-causing Q141K mutation in ABCG2 leads to instability of the nucleotide-binding domain and can be corrected with small molecules. *Proc. Natl. Acad. Sci. U. S. A.* 110, 5223–8.
- Wright, J., Muench, S.P., Goldman, A., Baker, A., 2018. Substrate polyspecificity and conformational relevance in ABC transporters: new insights from structural studies. *Biochem. Soc. Trans.* 46, 1475–1484.

- Xie, Y., Xu, K., Linn, D.E., Yang, X., Guo, Z., Shimelis, H., Nakanishi, T., Ross, D.D., Chen, H., Fazli, L., Gleave, M.E., Qiu, Y., 2008. The 44-kDa Pim-1 kinase phosphorylates BCRP/ABCG2 and thereby promotes its multimerization and drug-resistant activity in human prostate cancer cells. *J. Biol. Chem.* 283, 3349–3356.
- Xie, Z.-Y., Lv, K., Xiong, Y., Guo, W.-H., 2014. ABCG2-mediated multidrug resistance and tumor-initiating capacity of side population cells from colon cancer. *Oncol. Res. Treat.* 37, 666–8, 670–2.
- Xu, J., Liu, Y., Yang, Y., Bates, S., Zhang, J.-T., 2004. Characterization of oligomeric human half-ABC transporter ATP-binding cassette G2. *J. Biol. Chem.* 279, 19781–19789.
- Yang, C.H., Chen, Y.C., Kuo, M.L., 2003. Novobiocin sensitizes BCRP/MXR/ABCP overexpressing topotecan-resistant human breast carcinoma cells to topotecan and mitoxantrone. *Anticancer Res.* 23, 2519–2523.
- Yoshioka, S., Katayama, K., Okawa, C., Takahashi, S., Tsukahara, S., Mitsuhashi, J., Sugimoto, Y., 2007. The identification of two germ-line mutations in the human breast cancer resistance protein gene that result in the expression of a low/non-functional protein. *Pharm. Res.* 24, 1108–1117.
- Yuan, Y.R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P.J., Hunt, J.F., 2001. The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. *J. Biol. Chem.* 276, 32313–32321.
- Zhang, W., Ding, W., Chen, Y., Feng, M., Ouyang, Y., Yu, Y., He, Z., 2011. Up-regulation of breast cancer resistance protein plays a role in HER2-mediated chemoresistance through PI3K/Akt and nuclear factor-kappa B signaling pathways in MCF7 breast cancer cells. *Acta Biochim. Biophys. Sin. (Shanghai)*. 43, 647–653.
- Zhang, W., Mojsilovic-Petrovic, J., Andrade, M.F., Zhang, H., Ball, M., Stanimirovic, D.B., 2003. The expression and functional characterization of ABCG2 in brain endothelial cells and vessels. *FASEB J.* 17, 2085–7.
- Zhang, Y., Gupta, A., Wang, H., Zhou, L., Vethanayagam, R.R., Unadkat, J.D., Mao, Q., 2005. BCRP transports dipyridamole and is inhibited by calcium channel blockers. *Pharm. Res.* 22, 2023–2034.
- Zhang, Z., Chen, J., 2016. Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* 167, 1586–1597.e9.
- Zhou, L., Naraharisetti, S.B., Wang, H., Unadkat, J.D., Hebert, M.F., Mao, Q., 2008. The breast cancer resistance protein (Bcrp1/Abcg2) limits fetal distribution of glyburide in the pregnant mouse: an Obstetric-Fetal Pharmacology Research Unit Network and University of Washington Specialized Center of Research Study. *Mol. Pharmacol.* 73, 949–959.

Zhou, S., Schuetz, J.D., Bunting, K.D., Colapietro, A.M., Sampath, J., Morris, J.J., Lagutina, I., Grosveld, G.C., Osawa, M., Nakauchi, H., Sorrentino, B.P., 2001. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat. Med.* 7, 1028–1034.

Zolnerciks, J.K., Wooding, C., Linton, K.J., 2007. Evidence for a Sav1866-like architecture for the human multidrug transporter P-glycoprotein. *FASEB J.* 21, 3937–3948.

Appendices

Appendix 1 Example sequence alignment for ABCG2_{F640A} mutant variant. Example sequencing data for the F640A ABCG2 mutant (Query) aligned against predicted sequence (Sbjct). Sequence alignment was performed using NCBI nucleotide Blast, F640A mutation is highlighted yellow in the F482 primer sequence trace.

F640A R1

Score	Expect	Identities	Gaps	Strand
1720 bits(931)	0.0	961/982(98%)	6/982(0%)	Plus/Minus
Query 19	CTTTCCTTGCAGCTAAGACATCTAATAACGAAGATTGCGCTCCACCTGTGGGTCCCAGGA	78		
Sbjct 1948	CTTTCCTTGCAGCTAAGACATCTAATAACGAAGATTGCGCTCCACCTGTGGGTCCCAGGA	1889		
Query 79	TGGCGTTGAGACCAGGTTTCATGATCCCATTGATATTCGATAATATTTCTTTCTCAACTG	138		
Sbjct 1888	TGGCGTTGAGACCAGGTTTCATGATCCCATTGATATTCGATAATATTTCTTTCTCAACTG	1829		
Query 139	GTTTTCGACAAGGTAGAAAGCCACTCTTCAGTTTTACTCGATAGCAGATGTTATGAAAAC	198		
Sbjct 1828	GTTTTCGACAAGGTAGAAAGCCACTCTTCAGTTTTACTCGATAGCAGATGTTATGAAAAC	1769		
Query 199	TTAACACAGCTCCCTTCAGTAAATGCCTTCAGGTCATTGGAAGCTGTCGCGGGGAAGCCAT	258		
Sbjct 1768	TTAACACAGCTCCCTTCAGTAAATGCCTTCAGGTCATTGGAAGCTGTCGCGGGGAAGCCAT	1709		
Query 259	TGGTGTTCCTTGTGACACTGGGATAAAAACTTCGACATTACTGGAAGACATAGAATTCG	318		
Sbjct 1708	TGGTGTTCCTTGTGACACTGGGATAAAAACTTCGACATTACTGGAAGACATAGAATTCG	1649		
Query 319	AAGCTTGAGCTCGAGATCTGAGTCCGGACTTGTACAGCTCGTCCATGCCGAGAGTGATCC	378		
Sbjct 1648	AAGCTTGAGCTCGAGATCTGAGTCCGGACTTGTACAGCTCGTCCATGCCGAGAGTGATCC	1589		
Query 379	CGGCGGCGGTACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGC	438		
Sbjct 1588	CGGCGGCGGTACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGC	1529		
Query 439	TCAGTTTGGACTGGGTGCTCAGGTAGTGGTTGTGGGCAGCAGCACGGGGCCGTCGCCGA	498		
Sbjct 1528	TCAGTTTGGACTGGGTGCTCAGGTAGTGGTTGTGGGCAGCAGCACGGGGCCGTCGCCGA	1469		
Query 499	TGGGGGTGTTCTGCTGGTAGTGGTTCGGCAGCTGCACGCTGCCGTCCTCGATGTTGTGGC	558		
Sbjct 1468	TGGGGGTGTTCTGCTGGTAGTGGTTCGGCAGCTGCACGCTGCCGTCCTCGATGTTGTGGC	1409		
Query 559	GGATCTTGAAGTTGGCCTTGATGCCGTTCTTCTGCTTGTGACGCGGTGATATAGACGTTGT	618		
Sbjct 1408	GGATCTTGAAGTTGGCCTTGATGCCGTTCTTCTGCTTGTGACGCGGTGATATAGACGTTGT	1349		
Query 619	GGCTGTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGA	678		
Sbjct 1348	GGCTGTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGA	1289		
Query 679	TGCCCTCAGCTCGATGCGGTTACCAGGGTGTGCGCCCTCGAACTTCACTCGGCGCGGG	738		
Sbjct 1288	TGCCCTCAGCTCGATGCGGTTACCAGGGTGTGCGCCCTCGAACTTCACTCGGCGCGGG	1229		
Query 739	TCTTGTAGGTGCCGTCGTCCTTGAAGGAGATGGTGCCTCCTGGACGTAGCCTTCGGGCA	798		
Sbjct 1228	TCTTGTAGGTGCCGTCGTCCTTGAAGGAGATGGTGCCTCCTGGACGTAGCCTTCGGGCA	1169		
Query 799	TGGCGGACTTGAAGAAGTCGTGCTTCAATGTTGGTGGGGTAGCGGCTGAAGCACTGCA	858		
Sbjct 1168	TGGCGGACTTGAAGAAGTCGTGCTTCAATGTTGGTGGGGTAGCGGCTGAAGCACTGCA	1109		

Query 798 AGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAG 857
 |||
 Sbjct 2627 AGGATAAGCCACTCATAGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAG 2686

Query 858 AGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAANAAGATCACAGTCT 917
 |||
 Sbjct 2687 AGACAAAAGCTGAATTACATCAACTTTCCGGGGGTGAGAAGAANAAGATCACAGTCT 2746

Query 918 TCAAGGAGATCAGCTACACCACCTCCTTCTGTGCATCAACTCAGATGGNNTTCTAANGT 977
 |||
 Sbjct 2747 TCAAGGAGATCAGCTACACCACCTCCTTCTGTGCATCAACTCAGATGGGTTTCTAAGCGT 2806

Query 978 CATTCAAAAAGCTGCTGGGTAAATCCCGGGCTCTATAGCTCAGATCATTGTACAGTCTG 1034
 |||
 Sbjct 2807 CATTCAAAAAGCTGCTGGGTAAATCCCGGGCTCTATAGCTCAGATCATTGTACAGTCTG 2866

Query 1035 TACTG 1039
 |||
 Sbjct 2867 TACTG 2871

F2

Score	Expect	Identities	Gaps	Strand
1829 bits(990)	0.0	1004/1015(99%)	2/1015(0%)	Plus/Plus
Query 20	CTTC-AGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC			78
Sbjct 2403	CTTCAAGTTGTTTGATAGCCTCACCTTATTGGCCTCAGGAAGACTTATGTTCCACGGGCC			2462
Query 79	TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA			138
Sbjct 2463	TGCTCAGGAGGCCTTGGGATACTTTGAATCAGCTGGTTATCACTGTGAGGCCTATAATAA			2522
Query 139	CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG			198
Sbjct 2523	CCCTGCAGACTTCTTCTTGGACATCATTAATGGAGATTCCACTGCTGTGGCATTAAACAG			2582
Query 199	AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT			258
Sbjct 2583	AGAAGAAGACTTTAAAGCCACAGAGATCATAGAGCCTTCCAAGCAGGATAAGCCACTCAT			2642
Query 259	AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT			318
Sbjct 2643	AGAAAAATTAGCGGAGATTTATGTCAACTCCTCCTTCTACAAAGAGACAAAAGCTGAATT			2702
Query 319	ACATCAACTTTCCGGGGGTGAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA			378
Sbjct 2703	ACATCAACTTTCCGGGGGTGAGAAGAAGAAGATCACAGTCTTCAAGGAGATCAGCTA			2762
Query 379	CACCACCTCCTTCTGTGCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAAGCTGCT			438
Sbjct 2763	CACCACCTCCTTCTGTGCATCAACTCAGATGGGTTTCTAAGCGTTCATTCAAAAAGCTGCT			2822
Query 439	GGGTAATCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT			498
Sbjct 2823	GGGTAATCCCAGGCCTCTATAGCTCAGATCATTGTCACAGTCGTACTGGGACTGGTTAT			2882
Query 499	AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT			558
Sbjct 2883	AGGTGCCATTTACTTTGGGCTAAAAAATGATTCTACTGGAATCCAGAACAGAGCTGGGGT			2942
Query 559	TCTCTTCTCCTGACGACCAACCAGTGTTCAGCAGTGTTCAGCCGTGGAACCTTTTGT			618
Sbjct 2943	TCTCTTCTCCTGACGACCAACCAGTGTTCAGCAGTGTTCAGCCGTGGAACCTTTTGT			3002
Query 619	GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGCATCTTA			678
Sbjct 3003	GGTAGAGAAGAAGCTCTTCATACATGAATACATCAGCGGATACTACAGAGTGCATCTTA			3062
Query 679	TTTCTTGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT			738
Sbjct 3063	TTTCTTGAAAACTGTTATCTGATTTATTACCCATGAGGATGTTACCAAGTATTATATT			3122

```

Query 739 TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT 798
          |||
Sbjct 3123 TACCTGTATAGTGTACTTCATGTTAGGATTGAAGCCAAAGGCAGATGCCTTCTTCGTTAT 3182

Query 799 GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCATGGCACTGGCCATAGCAGC 858
          |||
Sbjct 3183 GATGTTTACCCTTATGATGGTGGCTTATTCAGCCAGTTCATGGCACTGGCCATAGCAGC 3242

Query 859 AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTGTGTTTATGAT 918
          |||
Sbjct 3243 AGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCATGACCATCTGTTTGTGTTTATGAT 3302

Query 919 GATTTTTTCNGGTCTGTTGGGTCAATCTCNNAACCATTGCATCTTGGNTGTCATGGNTTC 978
          |||
Sbjct 3303 GATTTTTTCAGGTCTGTT-GGTCAATCTCACAACCATTGCATCTTGGCTGTCATGGCTTC 3361

Query 979 AGTACTTCAGCATTCCACGATATGGATTTACGGCTTGCAGCNNNATGANTTTTT 1033
          |||
Sbjct 3362 AGTACTTCAGCATTCCACGATATGGATTTACGGCTTGCAGCATAATGAATTTTT 3416

```

F482 (Mutation highlighted)

Score	Expect	Identities	Gaps	Strand
1803 bits(976)	0.0	991/1002(99%)	1/1002(0%)	Plus/Plus
Query 17	GCGGATACTACAGAGTGCATCTTATTTCCCTTGAAAACCTGTTATCTGATTTATTACCCA			76
Sbjct 3038	GCGGATACTACAGAGTGCATCTTATTTCCCTTGAAAACCTGTTATCTGATTTATTACCCA			3097
Query 77	TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC			136
Sbjct 3098	TGAGGATGTTACCAAGTATTATATTTACCTGTATAGTGTACTTCATGTTAGGATTGAAGC			3157
Query 137	CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA			196
Sbjct 3158	CAAAGGCAGATGCCTTCTTCGTTATGATGTTTACCCTTATGATGGTGGCTTATTCAGCCA			3217
Query 197	GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA			256
Sbjct 3218	GTTCCATGGCACTGGCCATAGCAGCAGGTCAGAGTGTGGTTTCTGTAGCAACACTTCTCA			3277
Query 257	TGACCATCTGTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA			316
Sbjct 3278	TGACCATCTGTTTGTGTTTATGATGATTTTTTCAGGTCTGTTGGTCAATCTCACAACCA			3337
Query 317	TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT			376
Sbjct 3338	TTGCATCTTGGCTGTCATGGCTTCAGTACTTCAGCATTCCACGATATGGATTTACGGCTT			3397
Query 377	TGCAGCATAATGAATTTTTGGGACAAAACCTTCTGCCCAGGACTCAATGCAACAGGAAACA			436
Sbjct 3398	TGCAGCATAATGAATTTTTGGGACAAAACCTTCTGCCCAGGACTCAATGCAACAGGAAACA			3457
Query 437	ATCCTTGTAACATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC			496
Sbjct 3458	ATCCTTGTAACATGCAACATGTACTGGCGAAGAATATTTGGTAAAGCAGGGCATCGATC			3517
Query 497	TCTCACCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTGCC			556
Sbjct 3518	TCTCACCTGGGGCTTGTGGAAGAATCACGTGGCCTTGGCTTGTATGATTGTTATTTCC			3577
Query 557	TCACAATTGCCTACCTGAAATGTTATTTCTTAAAAAATATCTTAAATTGGATTCTAGA			616
Sbjct 3578	TCACAATTGCCTACCTGAAATGTTATTTCTTAAAAAATATCTTAAATTGGATTCTAGA			3637
Query 617	GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT			676
Sbjct 3638	GGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT			3697

```

Query 677 GTTTGCCCTCCCCGTCCTTCCTTGACCCTGGAAGGTGCCACTCCCACCTGTCCTTTCC 736
          |||
Sbjct 3698 GTTTGCCCTCCCCGTCCTTCCTTGACCCTGGAAGGTGCCACTCCCACCTGTCCTTTCC 3757

Query 737 TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTggggggt 796
          |||
Sbjct 3758 TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGT 3817

Query 797 ggggtggggCAGGACAGCAAGGGGGANGATTGGGAAGACAATAGCANGCATGCTGGGGAT 856
          |||
Sbjct 3818 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT 3877

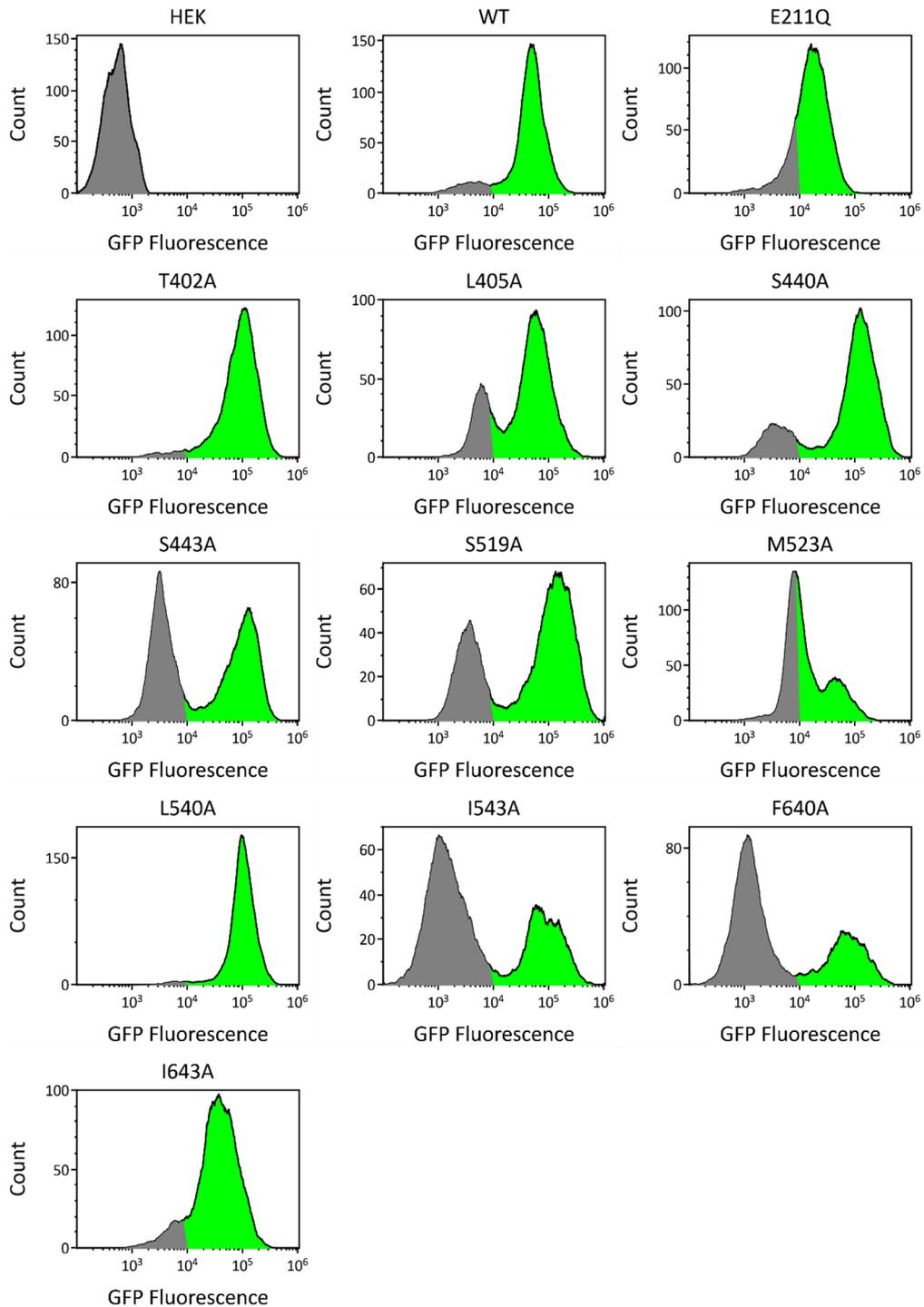
Query 857 GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC 916
          |||
Sbjct 3878 GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCC 3937

Query 917 CACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGGTGTGGTGGNTACGCGCAGCGTGACC 976
          |||
Sbjct 3938 CACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGGTGTGGTGGTTACGCGCAGCGTGACC 3997

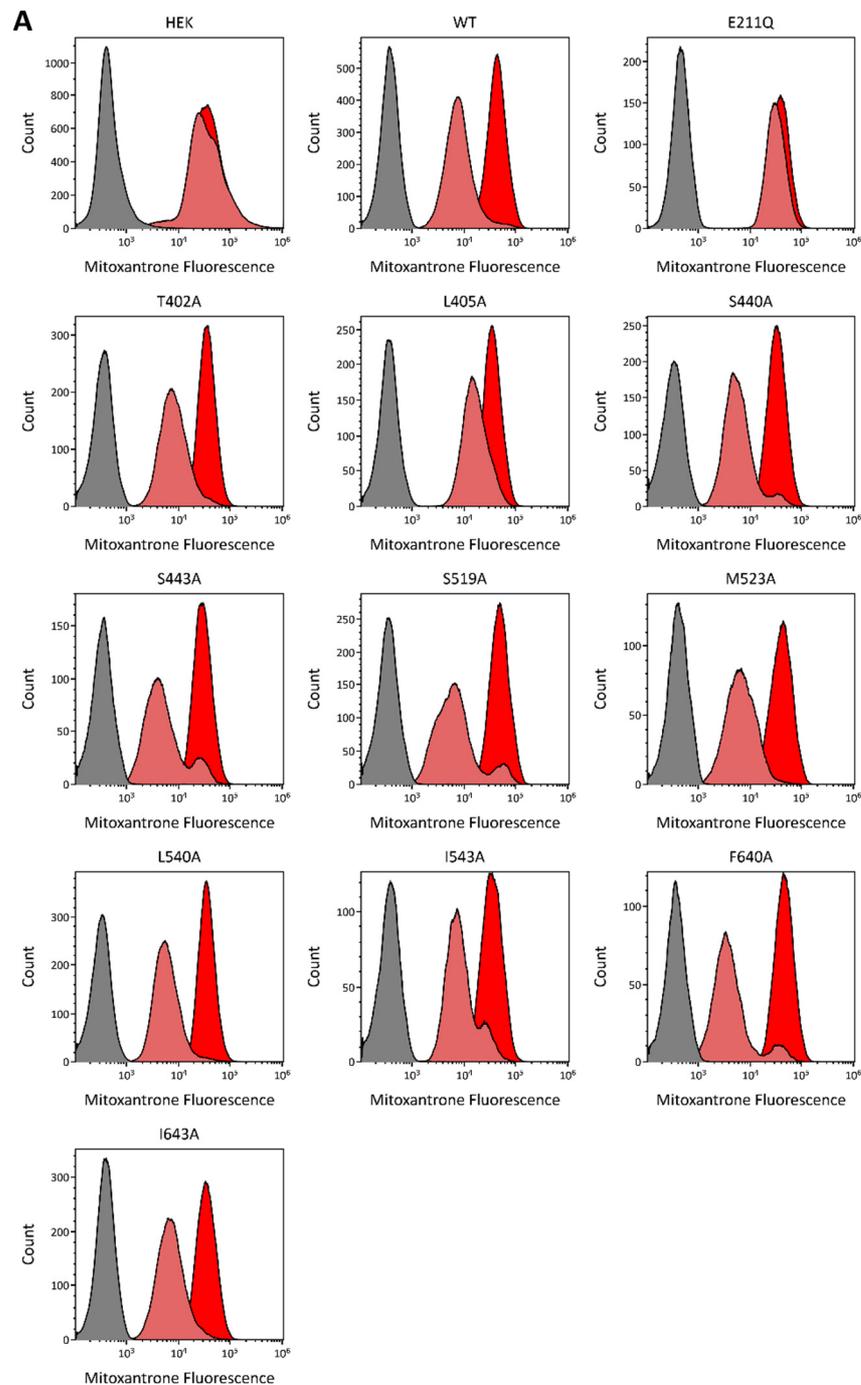
Query 977 GCTACACTTGCCAGCGCCN-AGCGCCNGCENNCTTTCGNTTTC 1017
          |||
Sbjct 3998 GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTC 4039

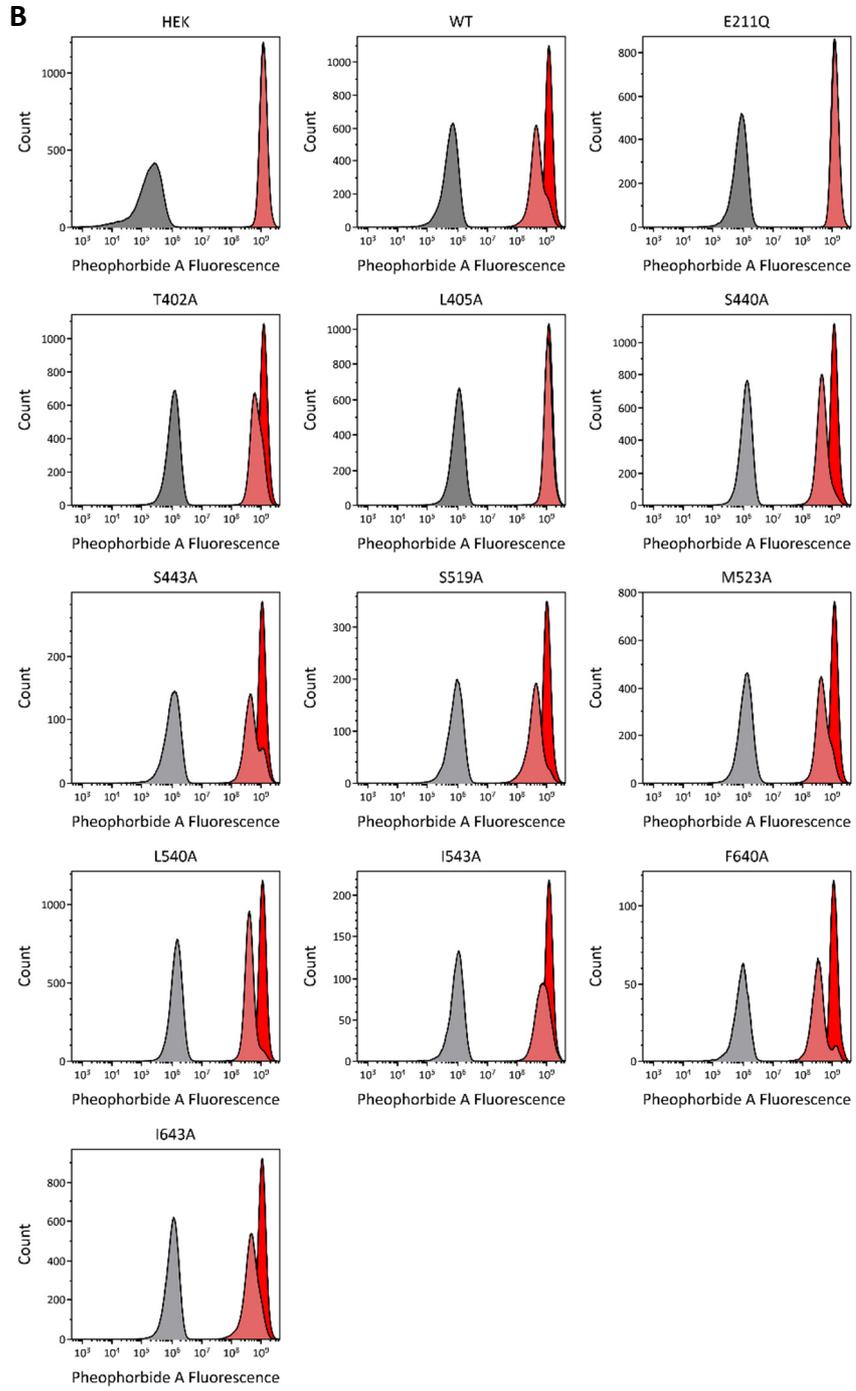
```

Appendix 2 ABCG2 mutants demonstrate a phenotype split in protein expression. Flow cytometric histograms for GFP fluorescence (sfGFP-ABCG2 expression) for ‘lateral slice’ mutants. Area shaded grey corresponds to expression gating for low expressing cells (excluded from analysis) and the area shaded green represents wild type like ABCG2 protein expression gating (denoted ‘high’). Data representative of N≥3 independent repeats, figures generated in Kaluza© (Beckman Coulter).

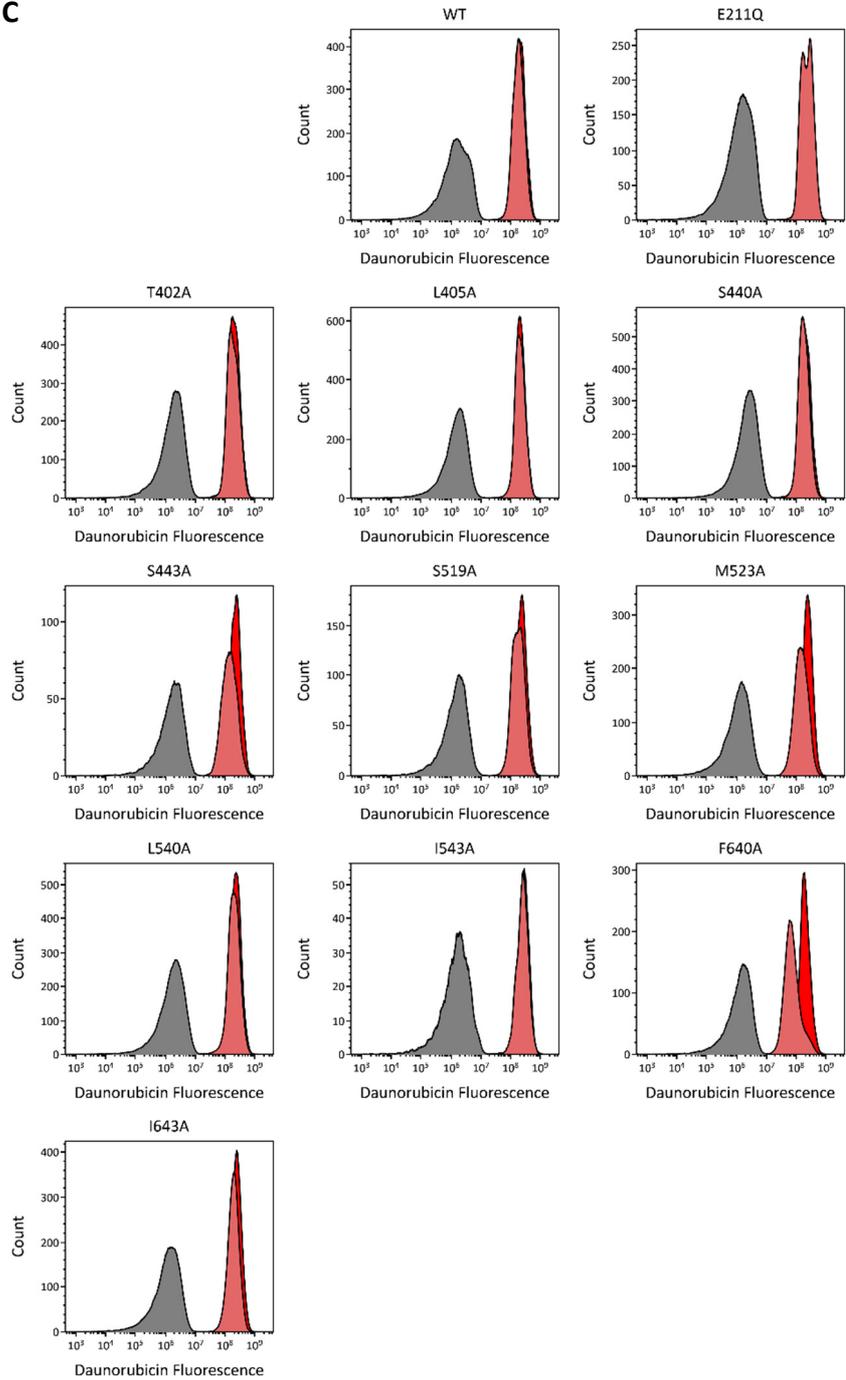


Appendix 3 ABCG2 mutants demonstrate functional perturbation. Cell lines expressing ABCG2_{WT} and mutant variants were incubated with either 10 μ M of mitoxantrone (A), pheophorbide A (B), or daunorubicin (C) in the presence (bright red) or absence (light red) of 1 μ M Ko143 (ABCG2 specific inhibitor). A left ward shift in MX fluorescence in the absence of Ko143 (light red) demonstrates ABCG2 drug efflux. Kaluza© single channel overlay histograms for drug fluorescence (gated for expression) are representative of merged data sets across N \geq 3 independent experiments.



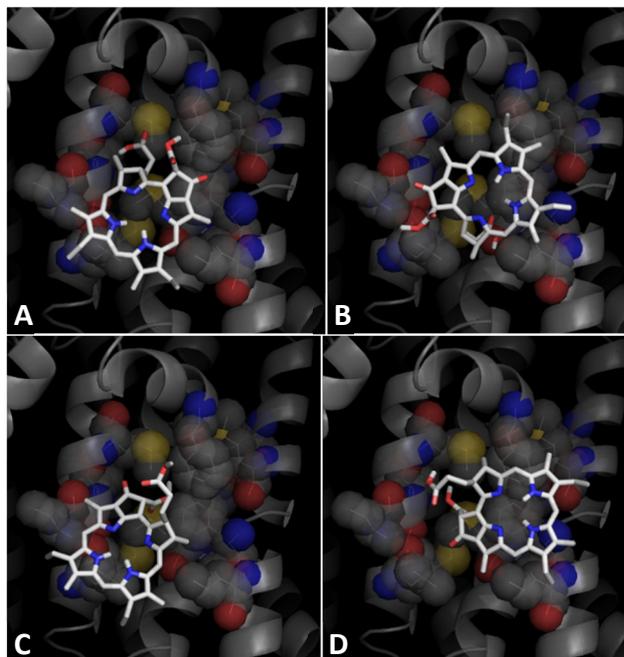


C



Appendix 4 Alternative binding poses for pheophorbide A and daunorubicin docking into ABCG2. A-D represent the top binding poses (approx. 5.5kcal/mol binding energy) for pheophorbide A and daunorubicin docked into ABCG2.

Pheophorbide A



Daunorubicin

