Methionine-Aromatic Interactions in ABCG2 Influence Mitoxantrone Efflux

Nur Amirah Hannan Binti Azmir,

BSc Hons.

Thesis submitted to the University of Nottingham
for the degree of Master of Research

Acknowledgements

I would like to extend my gratitude to Dr Ian Kerr for guiding me throughout this degree, and giving me the opportunity to develop my skills and make me a well-rounded scientist.

My biggest thank you to Deb Briggs for not only teaching me and helping me with my experiments when I needed it, but also being such a wonderful person who I hope we can continue to keep in touch.

Thanks to Dr James Mitchell-White and Parth Kapoor for helping me in the lab, and I will miss our weird discussions during lunch. From discussing apple varieties to figuring out the best obscure band name, and entertaining my love for animals (mostly cats) I enjoyed our little lunchtime chit-chat and hope we will always keep in touch.

Big thanks to Dr David Onion and Nicola Croxall (University of Nottingham Flow Cytometry Facility) as well as Seema Bagia and Tim Self (School of Life Science Imaging) for giving me the training, advice and use of equipment for this project.

Thank you to my family and friends both here in the UK and back home in Malaysia for giving me the support I needed when I made the decision to come here for my studies. They helped me when I felt homesick and when I needed that little bit of extra support to finish this degree.

Most importantly, my biggest thank you to Syaqil for being there for me through thick and thin. Even though we're miles apart from each other, I know we will always be there for each other. You helped shape me to become the person that I am today, and I really hope we continue to go through this life-long journey together.

Abstract

The ATP-Binding Cassette (ABC) transporters are one of many families of membrane transport proteins that utilise the hydrolysis of ATP to provide the energy required to actively transport substrates against a concentration gradient. Some ABC transporters have been implicated in cancer multidrug resistance (MDR), where some solid tumours and blood cancers develop reduced chemosensitivity to the antineoplastic agents used to treat them due to these MDR-type ABC transporters having a very diverse substrate specificity incorporating different chemistries. One of these ABC transporters implicated in human cancer MDR is ABCG2. The mechanism behind the ability for ABCG2 to recognise a wide variety of structurally dissimilar substrates is a key question that has yet to be fully answered, with many studies trying to identify both the location of drug binding sites and the translocation pathway for allocrite recognition and efflux.

In this study, an attempt was made to understand the role of a predicted drug-binding site on the surface of ABCG2. This site is rich in methionine-aromatic interactions and it was predicted that disruption of these interactions would influence drug transport. Mutagenesis of 12 residues (6 methionines and 6 phenylalanines) was performed and mutants successfully expressed and trafficked to the plasma membrane. In drug transport assays using the native substrate mitoxantrone, 3 of the 12 mutants showed a significant decrease in drug efflux compared to the wild type

From this work it was predicted that the pair of residues M541 and F545 are engaged in a methionine-aromatic interaction that is critical for drug recognition and efflux. Further future studies such as transport assays with

another native substrate and a non-native substrate can be performed to further validate this claim.

Declaration

I declare that the thesis "Methionine-Aromatic Interactions in ABCG2 Influence Mitoxantrone Efflux" 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.

Nur Amirah Hannan Binti Azmir

Student ID: 18023418

January 2020

(Total word count: 15,297)

Table of Contents

Chapter 1: Introduction 1

1.1 Membrane Transporter Proteins 1

1.2 ATP-Binding Cassette (ABC) Transporters 2

1.3 ABC Transporters and Multidrug Resistance 3

1.4 ABCG2 4

1.4.1 Overview of ABCG Subfamily 4

1.4.2 ABCG2 Structure and Function 5

1.4.3 Discovery of ABCG2 Structures and Further Discovery of Multidrug Binding Sites 7

1.5 Aim of Study 11

Chapter 2: Materials and Methods 12

2.1 Bioinformatics 12

2.1.1 Software, Protein Structure and Script Writing 12

2.2 Molecular Biology 12

2.2.1 Materials and Reagents 12

2.2.2 Designing of Mutant Constructs 13

2.2.3 Generation of Mutant Constructs 13

2.2.4 Restriction Digestion, Agarose Gel Electrophoresis and Plasmid DNA Visualisation 14

2.2.5 Transformation 15

2.2.6 Plasmid Purification 16

2.3 Cell Culture 17

2.3.1 Materials and Reagents 17

2.3.2 Cell Passage Maintenance 17

2.3.3 Transfection 18

2.3.4 Zeocin Selection of Stable Cell Lines 19

2.4 SDS-PAGE and Western Blotting 19

2.4.1 Cell Harvesting 19

2.4.2 Concentration Determination Assay 20

2.4.3 SDS-Polyacrylamide Gel Electrophoresis 20

2.4.4 Western Blotting 20

2.5 Confocal Microscopy 21

2.5.1 Materials and Reagents 21

2.5.2 Poly-L-lysine Treatment 21

2.5.3 Cell Seeding for Live Cell Imaging 22
2.5.4 Confocal Laser Scanning Image Acquisition 22
2.5.5 Cell Seeding for Confocal Plate Reader Image Acquisition 23
2.5.6 Confocal Plate Reader Image Acquisition 23
2.6 Flow Cytometry 23
2.6.1 Materials and Reagents 23
2.6.2 Fluorescence-based Drug Accumulation Assay 24
2.6.3 Flow Cytometry Data Acquisition 24
2.6.4 Flow Cytometry Data Analysis 25
Chapter 3: Results 26
3.1 Hypothesis Generation 26
3.2 Identification of Methionine-Aromatic Interactions at the Surface Site 28
3.3 Plasmid Construct Information 34
3.4 Construct Generation 36
3.5 Protein Expression 42
3.6 Protein Localisation 45
3.7 Protein Function Characterisation 48
3.8 Flow Cytometry Data Analysis 51
3.9 Summary 58
Chapter 4: Discussion 59
4.1 Methionine at Position 541 and Phenylalanine at Position 545 Share a Methionine-Aromatic Interaction that is Key to Allocrite Recognition 60
4.2 Tryptophan Substitution of Phenylalanine at Position 571 Introduces an Addition Methionine-Aromatic Interaction with a Nearby Methionine Residue 62
4.3 Methionine-Aromatic Interactions Play a Role in Allocrite Recognition through Binding Site Stabilisation 64
4.4 Maintenance of Stable Cell Lines in Higher Concentrations of Zeocin May Reduce Heterogeneity 65
References 70
Appendix 75

List of Figures

Figure 1.1: Structures of ABC Importers and Exporters.

Figure 1.2: The structure of human ABCG2.

Figure 1.3: Binding of a Ko143 derivative (MZ29) to a multidrug binding site in ABCG2.

Figure 1.4: *In silico* identification of several drug binding sites in ABCG2.

Figure 3.1: Identified methionine-aromatic residues in the apo-state structure of ABCG2.

Figure 3.2: Identified methionine-aromatic interactions differ in the ATP-bound state of ABCG2.

Figure 3.3: Visualisation of an attempt in identifying methionine-aromatic interactions in the internal groove of the apo-state structure of ABCG2.

Figure 3.4: Conservation of surface site residues and F571 across homologous ABCG2 structures in jawed vertebrates.

Figure 3.5: Methionine → Glutamine substitution mimics the structure of methionine sulphoxide.

Figure 3.6: Plasmid map of pcDNA-3.1/zeo(+)_sfGFP_ABCG2.

Figure 3.7: Schematic representation of mutant construct generation.

Figure 3.8 A, B, C and D: PvuII endonuclease digest.

Figure 3.9: Gradient PCR for Unsuccessful Mutants.

Figure 3.10: Confirmation of mutant codon replacement via Sanger sequencing.

Figure 3.11: Selection of stable cell lines carrying the methionine mutations.

Figure 3.12: Western blot analysis of stable cell line expression compared to the wild type.

Figure 3.13: Confocal image analysis of protein localisation.

Figure 3.14 A: Schematic representation of flow cytometry. B: Schematic of mitoxantrone efflux in the presence and absence of Ko143 in wild type expressing cells.

Figure 3.15: Single parameter histograms of mitoxantrone efflux in ABCG2 isoforms.

Figure 3.16: Measuring medians for all histograms to calculate the raw mitoxantrone efflux of the wild type protein.

Figure 3.17: Manual gating for GFP expression on a mixed population.

Figure 3.18: Relative mitoxantrone efflux profiles of all methionine mutants normalised to GFP expression.

Figure 3.19: Relative mitoxantrone efflux profiles of all phenylalanine mutants normalised to GFP expression.

Figure 4.1: Visualisation of the Methionine-Aromatic Interaction between F545 and M541 in the closed-conformation apo-state of ABCG2.

Figure 4.2: Epifluorescence microscopy images of stable cell lines grown in varying concentrations of zeocin.

Table 2.1: Forward Primers of Mutant Constructs.

Table 3.1: ABCG2 cryo-EM structures used for methionine-aromatic interaction parsing.

Table 3.2: Relative drug efflux data for methionine mutants.

Table 3.3: Relative drug efflux data for phenylalanine mutants.

Abbreviations

ABC	ATP-binding cassette transporter
ABCG2	ATP-binding cassette transporter subfamily G member 2
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
CMV	Cytomegalovirus
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ER	Endoplasmic reticulum
FBS	Foetal bovine serum
FSC	Forward scatter
GFP	Green fluorescent protein
HBSS	Hank's balanced salt solution
HEK 293 T	Human embryonic kidney 293 T cell line
kDa	Kilo Daltons
LB	Luria-Bertoni
LTα	Lymphotoxin alpha
MDR	Multidrug resistance
NBD	Nucleotide binding domain
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline-0.1% (v/v) Tween-20
PCR	Polymerase chain reaction
PEI	Polyethyleneimine
Pi	Orthophosphate
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
sfGFP	Superfolder green fluorescent protein

SSC	Side scatter
SV40	Simian vacuolating virus 40
TMD	Transmembrane domain
TMH	Transmembrane helix
TNFR1	Tumour Necrosis Factor Receptor 1
UV	Ultraviolet

Chapter 1: Introduction

1.1 Membrane Transporter Proteins

The term "membrane transporter proteins" covers a wide range of membrane proteins involved in the carrying of substrates - be it ions, small molecules or macromolecules such as large polypeptides - to and from the intracellular environment to the extracellular environment. Membrane transporters are involved in a wide variety of cellular reactions, be it importing substrates for further processing such as GLUT1 (glucose transporter 1) and its role in importing glucose for glycolysis, maintaining an electrochemical gradient to provide the electric potential needed to power and propagate an action potential as is the case with the sodium-potassium ATP pump in excitable cells, or removing xenobiotic substances from a cell to ensure cell survival. The transport of these substrates via these membrane transport proteins is either through facilitated diffusion or active transport. For facilitated diffusion, the substrates will move along a chemical gradient, while active transport requires energy to "pump" the substrate against a chemical gradient. The energy required by membrane protein pumps usually comes from the hydrolysis of ATP. One family of membrane transport proteins that makes use of this to pump out substrates is the ATP-binding cassette (ABC) transporter family.

1.2 ATP-Binding Cassette (ABC) Transporters

Since their initial discovery in the 1970s from attempts at understanding nutrient uptake in gram-negative bacteria (Berger and Heppel, 1974), many more different ABC transporters have since been discovered and characterised. As the name implies, these transporters require the hydrolysis of ATP to power the active transport of substrates against a concentration gradient. There are two major types of ABC transporters which depend on the direction of active transport: ABC importers and ABC exporters. ABC importers, which are further categories into Type I and Type II importers, are exclusively found in prokaryotic cells while ABC exporters are expressed in both prokaryotic and eukaryotic cells (Hollenstein et al., 2007). Both of these types of ABC transporters share conserved domains which allow for them to couple the hydrolysis of ATP to the transport of substrates.

The basic structure of an ABC transporter comprises of four domains: two nucleotide binding domains (NBDs) and two transmembrane domain (TMDs) (figure 1.1). The NBDs are cytoplasmic domains that are highly conserved amongst all the members of the ABC transporter family and comprise of the Walker A motif, responsible for initial nucleotide binding, the Walker B motif which provides a conserved glutamate residue that takes part in the nucleophilic attack on ATP via a water molecule, an LSGGQ or "ABC signature motif" that makes contact with the ATP molecule and the "Q-loop", which is believed to sense the γ-phosphate of the ATP molecule prior to hydrolysis (Hollenstein et al., 2007). In contrast, the TMDs are the least conserved domain and as such are responsible for providing the transporters a varied polyspecificity for substrates, which is fairly unique in membrane transporter proteins. The TMDs, in general, comprise of several membrane-spanning α-helices that differ in length and the number of helices between ABC importers

and exporters, and although they do not share many conserved residues, they do share a conserved topology depending on the type of ABC transporter (Wilkens, 2015, see figure 1.1). It is this domain that is responsible for the recognition and import/efflux of substrates, and the exact mechanism behind substrate recognition is still unknown, especially for ABC transporters that are implicated in multidrug resistance.

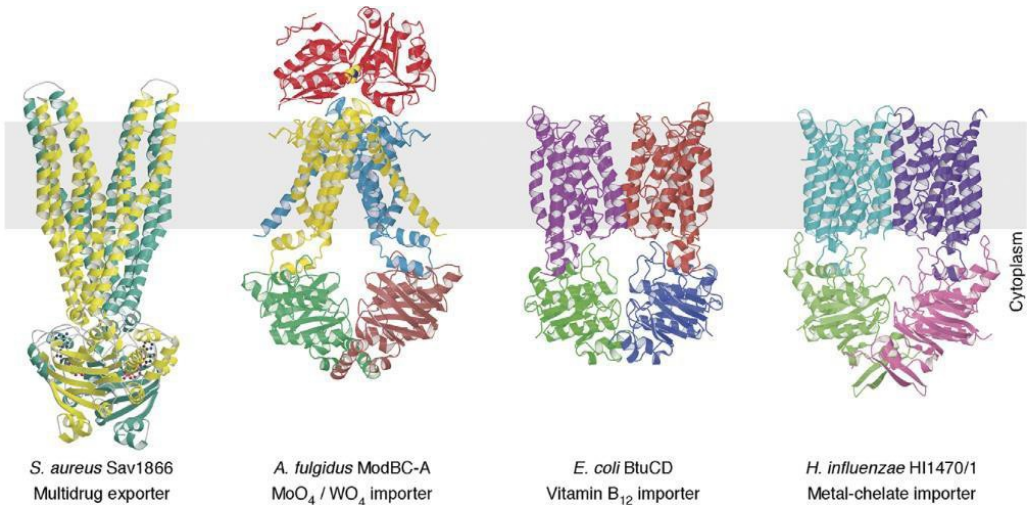


Figure 1.1: Structures of ABC Importers and Exporters. These structures show the two NBDs and two TMDs that make up a functional unit in ABC transporters as well as showing the difference in the transmembrane topology between Type I and Type II importers and exporters. Type I importers such as ModBC-A have an inward-facing topology, where the transmembrane pore is open to the cytoplasm. Type II importers such a BtuCD have an outward-facing topology where the transmembrane pore is open towards the extracellular environment. ABC exporters such as Sav1866 share a similar inward-facing topology as Type I importers. Diagram adapted from Hollenstein et al. (2007).

1.3 ABC Transporters and Multidrug Resistance

Multidrug resistance (MDR) is becoming an ever increasing problem in cancer therapeutics and drug development. It is known that a wide variety of factors lead to either inherent or acquired resistance to various antineoplastic agents which have little to no structural similarities to each other (some of these factors are summarised by Lage (2008) and Mellor and Callaghan (2008)),

but one of the most frequent factors that pop up is the presence of ABC exporter overexpression. It has been known that some human ABC transporters, of which there are 48 proteins, are involved in cancer MDR, such as ABCB1 (also known as P-glycoprotein; P-gp) which has been implicated in reduced chemosensitivity and increased survival of breast cancers (Lehne, 2005). Despite the understanding that these transporters are in- volved in MDR, the exact mechanism behind this polyspecificity for various antineoplastic agents is still unknown. In various studies over the years the human ABC transporters that have been identified and are implicated in cancer MDR are ABCB1, ABCC1 and ABCG2.

1.4 ABCG2

1.4.1 Overview of ABCG subfamily

The human ABC transporter subfamily G is composed of 5 proteins, ABCG1/G4, ABCG2 and ABCG5/G8 (Velamakanni et al., 2007). All members of this subfamily are "half-transporters", where the gene encoding these proteins only encodes for two of the four domains necessary for function. Four of the 5 proteins are believed to form heterodimers in order to create functional units while ABCG2 is known form a homodimer, although it has been observed to form higher order oligomeric structures (Wong et al., 2016a). All members of this subfamily are unique in that their structure is reminiscent of a bacterial importer due to the short size of their TMDs that do not extent into the cytoplasm (Taylor et al., 2017; Velamakanni et al., 2007), as well as the orientation of the NBDs and TMDs being N-terminal and C-terminal respectively (Velamakanni et al., 2007). ABCG1 is highly expressed in adipocytes and macrophages, and it is believed that it is

responsible for sterol transport. ABCG4 is also thought to redistribute cholesterol for HDL removal, but it is found to be exclusively expressed in nervous tissue with the exception of obligate ABCG1/G4 heterodimers in other tissues (Annilo et al., 2001). ABCG5/G8 is also a sterol transporter that is highly expressed in the liver, and was one of the first ABCG transporters to be successfully crystallised and its structure elucidated by Lee et al. (2016). ABCG2, although structures of it have been shown to have sterol-binding sites, is not known to be a sterol transporter (Taylor et al., 2017; Velamakanni et al., 2007). What is known, however, is that it is highly expressed in the apical membrane of enterocytes, at the blood-brain barrier and the placenta (Fetsch et al., 2006). It has also been found to be involved in multidrug efflux from its presence in some blood cancers and solid tumours that have reduced chemosensitivity to several antineoplastic agents such as mitoxantrone and methotrexate (Gillet and Gottesman, 2010; Nakanishi and Ross, 2012). Thus it is thought that due to its tissue localisation and role in reduced chemosensitivity in some cancers, it physiological role most likely is the removal of xenobiotics to ensure cell and tissue survival. With its interesting role in some solid tumours, overlapping native substrates such as methotrexate with ABCB1, it is believed that ABCG2 must be involved in cancer MDR (Bates et al., 2001). This then provides another target for drug design and development to target this protein to reduce MDR. However, the exact mechanism behind ABCG2's ability to efflux such a wide variety of antineoplastic agents which have little to no structural similarities to each other has yet to be determined.

1.4.2 ABCG2 Structure and Function

The structure of ABCG2 is tied to its function, and through several studies throughout the years to elucidate a full structure of the protein and to

determine the presence of a multidrug binding site, the understanding of its structure and function has been more comprehensive. The work of Clark et al. (2006) using radiolabelled ligands successfully showed the presence of a multidrug binding site in ABCG2, which helped provide a start in the identification of residues that could be involved in allocrite recognition (the word "allocrite" here would refer to the transported substrate and not ATP, the catalytic substrate). It is also found that the initial binding of the drug from this multidrug binding site prior to its translocation to the extracellular environment is tied to conformational changes that were transmitted from the binding of ATP to the NBDs, and that hydrolysis of ATP and the release of ADP and Pi lead to the translocation of the drug from a multidrug site to the extracellular environment (McDevitt et al., 2008). However, these find- ings did not reveal the location of a multidrug binding site in the protein.

A turning point in the field of ABCG2 multidrug binding site identification occurred when it was discovered that a mutation of the arginine residue at position 482 (R482) - a residue in transmembrane helix 3 (TMH3) - had led to a broadening of substrate recognition in ABCG2, where substitutions to either glycine or threonine (R482G/T) allowed for the efflux of non-native substrates such as daunorubicin and rhodamine-123 (Özvegy-Laczka et al., 2005; Robey et al., 2003). This coupled with the understanding of the presence of a multidrug binding site in ABCG2 has led to an explosion of exploratory studies focusing on residues at or near this site and their effects in allocrite binding and efflux.

1.4.3 Discovery of ABCG2 Structures and Further Discovery of Multidrug Binding Sites

Although these studies have provided a clue as to how ABCG2 is able to recognise and efflux such a wide range of substrates, the actual structure of ABCG2 had not been elucidated until fairly recently. Since 2017, various ABCG2 cryo-EM structures have now been deposited into the Protein Data Bank (PDB), with it now housing 8 structures in various conformations and with various bound ligands as of writing this thesis. This not only provided a clearer picture of what the protein actually looks like, it also provides a more detailed picture of what the residue substitutions at the multidrug binding site as well as substitutions of the NBDs would look like in affecting allocrite recognition and efflux. The first structure deposited by Taylor et al. (2017) was the first structure to show the protein in its simplest homodimeric structure (figure 1.2). This structure shows the short TMDs with an outward-facing topology reminiscent of Type I bacterial importers. It also shows the NBDs in constant contact with each other and the coupling helices with the intracellular loops that help transmit the conformational changes in the NBDs from the binding and hydrolysis of ATP to the TMDs, powering the translocation of substrates to the extracellular environment. Two distinct cavities are also observed, with one present in-between the monomers followed by a leucine-rich "plug" that blocks another smaller cavity leading to the extracellular environment. The location of R482 is close to the larger internal cavity, thus residues nearby R482 and the internal cavity were believed to be involved in multidrug binding and could possibly be the location of a multidrug binding site.

With more deposited structures in the PDB in 2018, it was proposed that both allocrites such as estrone-3-sulphate (E_1S) and inhibitory molecules such as derivatives of the ABCG2-specific inhibitor Ko143 and tariquidar

could bind to the large internal cavity, however only allocrites can move past the leucine plug and reach the second smaller cavity (Jackson et al., 2018; Manolaridis et al., 2018, see figure 1.3), providing further evidence of a site for both native substrate binding and efflux and inhibitory molecule binding, and thus providing further evidence that this large internal cavity is most likely a multidrug binding site. With this information, the next question to solve is how the allocrites gain access to this binding site.

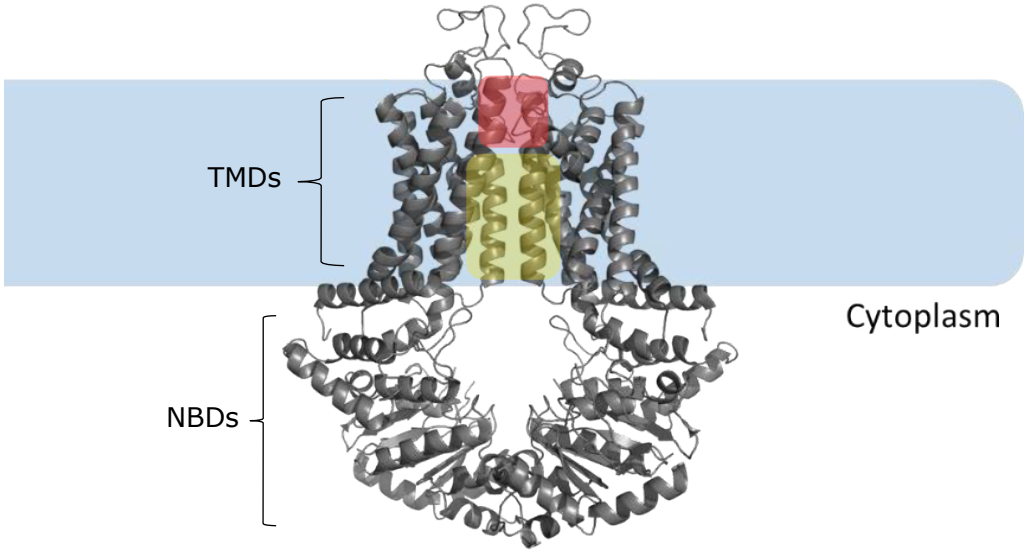


Figure 1.2: The structure of human ABCG2. The structure of ABCG2 is different compared to other human ABC transporters due to the shorter length of the TMDs, with the overall structure and TMD topology resembling a bacterial importer. The red square represents the location of the leucine plug, and the yellow rectangle represents the location of a putative multidrug binding site. The light blue rectangle is representative of the plasma lipid bilayer. The structure was visualised with PyMOL using the PDB structure 5NJ3, the first complete cryo-EM structure for ACBG2.

In an attempt to understand how allocrites gain access to the internal cavity, a molecular dynamics and docking study conducted by László et al. (2016) using a homology model of ABCG2 and native and non-native substrates as ligands was performed to understand how these substrates are able to access the internal cavity for efflux. Through their docking study it was dis- covered that 4 distinct sites with clustering of allocrites and non-native substrates were present in the TMDs and the locations of these sites are as follows: site 1 was closer to the NBDs and the coupling helices, and also in close proximity to the transmembrane pore facing the cytosol; site 2 was a site that is sandwiched between TMH1 and TMH4 and showed a distinct clustering of native substrates; site 3 was at the internal cavity below the leucine plug, and both native and non-native substrates were clustered in that location; site 4 was believed to be a transient site prior to allocrite re- lease to the extracellular environment as only a limited clustering of native substrates was present. The clustering of allocrites and non-native sub- strates and the location of these sites are visualised in figure 1.4. This then proposed a potential allocrite translocation pathway through the protein, although a massive question was still left unanswered: as all of the allo- crites as well as inhibitors are highly hydrophobic molecules it is unlikely that they would enter the multidrug binding site via site 1 (near the cytosol- facing transmembrane pore), *how do these allocrites gain access to the other sites if entrance via site 1 is unlikely?*

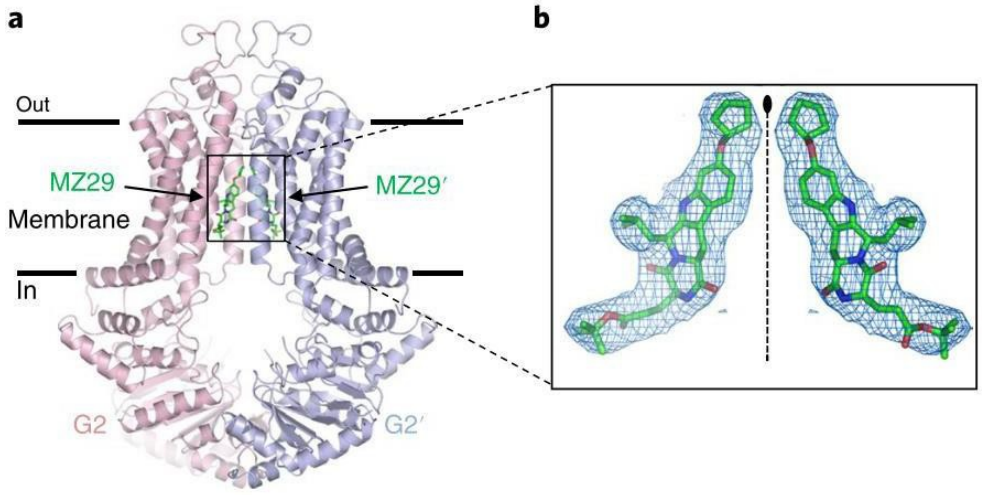


Figure 1.3: Binding of a Ko143 derivative (MZ29) to a multidrug binding site in ABCG2. The location for MZ29 binding is the same site where native substrates such as E_1S have been discovered to bind in cryo-EM structures of ABCG2 that have been deposited into the PDB. Diagram adapted from Jackson et al. (2018) using the structure 6ETI.

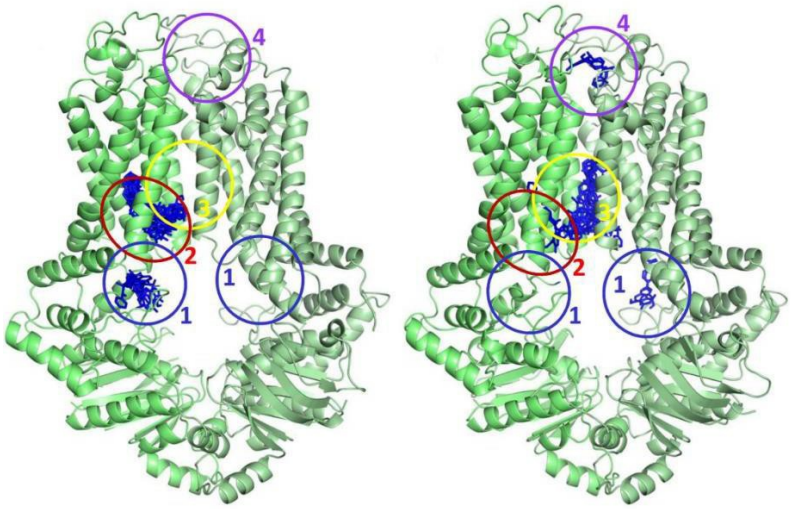


Figure 1.4: *In silico* identification of several drug binding sites in ABCG2. The main site of interest is Site 2 (labelled red) as this is the site that Dr Megan Cox has previously studied and found substitution of some residues at this site led to altered substrate recognition and efflux in human ABCG2. Diagram adapted from László et al. (2016).

An attempt to elucidate the mechanism of allocrite recognition prior to translocation to the internal multidrug binding site was conducted by members of the Kerr Lab by looking at residues at or near TMH3 and the large internal cavity. In several mutagenesis studies conducted by Dr Megan Cox, it was discovered that alanine substitutions to several residues present at or near the lipid-facing transmembrane helices of ABCG2, termed the "surface site", affected the recognition of the native substrates mitoxantrone and pheophorbide A and the non-native substrate daunorubicin (Cox et al., 2018). These alanine substitutions at key residues in TMH5a, 4 and 6b had led to varying drug efflux profiles, with some "gain of function" mutations leading to increased efflux capacity for native substrates and/or a gain in ability to efflux daunorubicin. This the led to the development of the "lateral slice hypothesis", whereby initial recognition of a substrates would occur at this surface site due to the substrates hydrophobicity allowing it to make its way through the lipid bilayer. From this initial recognition at the surface site, the substrate would then proceed to the internal cavity, and translocation of the substrate pass the leucine plug into the second external cavity for release to the extracellular environment will occur from conformational changes transmitted from the binding and hydrolysis of ATP from the NBDs to the TMDs.

From these findings, another question arises: *how do these key residues recognise substrates and allow translocation form the surface site to the internal cavity.* This question will be attempted in this study using a mutagenesis and drug efflux study using the native substrate mitoxantrone as the allocrite of interest.

1.5 Aim of Study

The aim is to build upon Dr Cox's previous work at understanding the mechanism behind allocrite recognition at this lipid-facing site of the TMDs, which she has given the term the "surface site". By conducted a mutagenesis study with several residues at this site and comparing the mutagenesis study results to the results obtained by Dr Cox, hopefully a better understanding of how the surface site is able to recognise native substrates such as mitoxantrone can be found.

Chapter 2: Materials and Methods

2.1 Bioinformatics

2.1.1 Software, Protein Structures and Script Writing

Visualisation and parsing was performed using PyMOL version 2.3 (Schrödinger, LLC). A custom script based on the basic workflow used by Weber and Warren (2019) was written using Python 3 (the entire script can be seen in Appendix 1). Protein structures used for the identification and visualisation of the methionine-aromatic interactions at the "surface site" were obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB). Sequence alignment was performed using HMMER version 3.3 (Eddy, 2011), and the BLOSUM62 alignment scoring matrix, based on the scoring matrices created and used by Henikoff and Henikoff (1992), was used to check conserved residues across ABCG2 protein residue sequences in 150 jawed vertebrates. Visualisation of these findings was done using the ggplot2 package in R. Sequence alignment and visualisation of the sequence alignment was performed with the help of Dr James Mitchell-White from the Kerr Lab.

2.2 Molecular Biology

2.2.1 Materials and Reagents

All molecular reagents were purchased from either New England Biolabs (NEB) or Promega unless otherwise stated.

2.2.2 Designing of Mutant Constructs

HEK293T was used as the main cell line and the plasmid used for generating the mutant constructs is p3.1zeo_sfGFP_ABCG2, previously made by Dr Kelvin Wong from the Kerr Lab. The backbone plasmid is a pcDNA™3.1/Zeo(+) mammalian expression vector with a "superfolder" Green Fluorescent Protein (sfGFP) fusion tag. The wild-type construct and all generated mutant constructs have three PvuII restriction enzyme recognition sites that cut the plasmid and give three distinct bands which are 4.5, 1.7 and 1.4 kbp in size, which were used to confirm the presence of the correct plasmid after plasmid purification. Further detail on this process is explained later in the results section.

2.2.3 Generation of Mutant Constructs

Site-directed mutagenesis based on the *in vitro* amplification of DNA by polymerase chain reaction (PCR) was done to create the mutants, and the forward mutant primers for the mutants are as listed in Table 2.1. These mutant primers were used to induce mutations in the plasmid and create the mutant constructs needed for transfection. The PCR reactions were set up in a final reaction volume of 50µl containing: template DNA (approximately 50ng), mutant primers (0.5mM), deoxynucleotides (dNTPs, 0.2mM), Pfu polymerase (1.5U) and the preferred Pfu reaction buffer. To incorporate the mutant codons into the template plasmid, template DNA denaturation, primer annealing and elongation was done with the following conditions per cycle: denaturation at 95°C for one minute, annealing at 55°C for 1 minute, extension at 72°C for 16 minutes; 18 cycles were performed and all products were held at 10°C after the final cycle had ended. The methylated template DNA was then digested using DpnI, which recognises and degrades methylated DNA. An incubation of 1 hour at 37°C with 20U of enzyme was

sufficient to digest the template DNA, and the digest was stopped by cooling the reaction to 4°C.

2.2.4 Restriction Digestion, Agarose Gel Electrophoresis and Plasmid DNA Visualisation

Plasmid DNA (200-500ng) was digested with 20U of PvuII in 15µl of the appropriate reaction buffer by incubating at 37°C for 1 hour, and the digest was halted by cooling the reaction to 4°C. 6x DNA loading dye (Thermo Fisher Scientific) was then added to the digested samples and they were loaded into 1% (w/v) agarose gels containing 10µg/ml ethidium bromide alongside a 1kb DNA ladder. Gel electrophoresis was done using 1x TAE buffer (90mM Tris, 90mM acetic acid and 2mM EDTA, pH 8.0) and gels were run at a constant voltage of 90-100V until the dye front has fully resolved. DNA bands were imaged under a UV transilluminator (Syngene).

Table 2.1: Forward Primers of Mutant Constructs

MUTATION	FORWARD PRIMER
M523Q	tcagccagttccCAggcactggccatag
M541Q	gcaacacttctcCAgaccatctgttttgtg
M636Q	gtggccttggcttgtCAgattgttattttcctc
M523L	tcagccagttccCTggcactggccatag
M541L	gcaacacttctcCTgaccatctgttttgtg
M636L	gtggccttggcttgtCTgattgttattttcctc
F545Y	cttctcatgaccatctgttAtgtgtttatgatg
F571Y	cttcagtactAcagcattcca
F640Y	gattgttatttATctcacaattgcctac
F545W	cttctcatgaccatctgttGGgtgtttatgatg
F571W	cttcagtactGGagcattcca
F640W	gattgttatttGGctcacaattgcctac

*highlighted codons denote a mutant codon, capital letters denote base changes from the wild-type sequence. The reverse primers are the reverse complement of the forward primers listed, and all sequences are given in a 5'→3'direction.

2.2.5 Transformation

Newly synthesised DNA was transformed into commercial DH5α competent cells (C2987I). 50µl aliquots of DH5α competent cells were thawed on ice. 5µl of the PCR amplicon was added to the competent cells. The competent cell mixture was then incubated on ice for 30 minutes, then cells were heat-shocked for 30 seconds at 42°C and incubated for a further 5 minutes on ice. Cells were supplemented with 950µl of Luria-Bertoni (LB) medium (1% (w/v) NaCl, 1% (w/v) tryptone and 0.05% (w/v) yeast extract) and incubated for 1 hour 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 100µg/ml) and finally incubated overnight at 37°C.

2.2.6 Plasmid Purification

Single colonies from LB-agar selection plates were inoculated into 5ml of LB medium supplemented with 100µg/ml of ampicillin. 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 500µl of the culture and adding 500µl of 30% (v/v) glycerol. The resultant glycerol stocks were subsequently 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 the NucleoSpin® Plasmid Miniprep Kit (Machery-Nagel) according to the manufacturer's protocol.

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

2.3 Cell Culture

2.3.1 Materials and Reagents

A mammalian protein expression system was used to express functional human ABCG2 for use in drug transport assays. HEK293T was used, as the cell line expresses a mutant version of the simian vacuolating virus (SV40) large T antigen, allowing for transfections with a high copy number plasmid containing an SV40 origin of replication.

All cell 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 otherwise stated.

2.3.2 Cell Passage Maintenance

Cells were grown in T25(cm^2) flasks with Dulbecco's Modified Eagle's Medium (DMEM, 4500mg/L-glucose, L-glutamine, sodium pyruvate and sodium bicarbonate) supplemented with 10% (v/v) foetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin (P/S, 100U/ml penicillin and 100U/ml streptomy- cin), and incubated at 37°C with 5% CO_2. Once cells reached 80-90% confluency, the medium was removed and cells were washed once with sterile phosphate-buffered saline (PBS). Cells were then detached from the plasticware by incubation with 0.5ml of 1x trypsin-EDTA for 1-5 minutes. Cells were then resuspended in DMEM and pelleted at 500xg for 5 minutes to remove trypsin. The pelleted cells were resuspended in 5ml of complete medium before being reseeded at a dilution of 1:10 of the original culture. Cells were passaged every 3 days to ensure a healthy, viable cell line was constantly maintained.

2.3.3 Transfection

Cells were seeded at 150,000 cells per well (determined using a haemocytometer) into a 6-well plate 24 hours prior to transfection. Two hours prior to transfection the media was replaced with reduced-serum medium (DMEM supplemented with 5% (v/v) FBS). Cells were transfected using linear polyethyleneimine (PEI) (10mM working stock solution dissolved in sterile PBS, based on the method first described by Boussif et al. (1995)) at a molar PEI nitrogen:DNA phosphorus ratio of 15:1. Complexes of PEI/DNA were made by the addition of PEI (9µl) to 2µg of DNA (1µg of DNA is approximately 3nM of phosphorus) and incubated for 15 minutes. 250µl of reduced-serum DMEM was then added to the preformed PEI/DNA complexes and subsequently added dropwise to the cells. Twenty-four hours post-transfection, the medium was replaced with complete medium. To confirm successful transfection and monitor stable cell line selection, all transfected cells were monitored daily through epifluorescence microscopy as successfully transfected cells will express sfGFP-tagged proteins and can be detected through this method. Epifluorescence microscopy images were obtained using the Zeiss Axiovert S100 inverted microscope with a Zeiss CP-ACHROMAT 10x/0.25 Ph1 air objective and a mercury arc-lamp for use in transmitted light imaging and fluorescence imaging. Images were taken with the Zeiss AxioCam MRm monochrome digital camera and processed with AxioVision version 4.8.2 SP3. Phase contrast as well as fluorescence images were taken for each cell line.

2.3.4 Zeocin Selection of Stable Cell Lines

Selection and maintenance of stable cell lines were done based on the methods described by Wong et al. (2016) and Cox et al. (2018). Transiently transfected cells were detached (1x trypsin/EDTA) from the 6-well plate and transferred to T25(cm^2) flasks containing 5ml complete medium. Following an overnight incubation (37°C, 5% CO_2), the cells were given fresh complete medium supplemented with 200µg/ml Zeocin™ (Invitrogen). Cells were transferred to T75(cm^2) flasks once fully confluent and Zeocin™ concentration maintained at 200µg/ml for 10-15 days, with media replacement every 2-3 days. Once healthy stable cell lines were observed, the concentration of Zeocin™ was lowered to 40µg/ml to maintain selection. For cryopreservation of stable cell lines, cells were stored in 90% (v/v) FBS, 10% (v/v) DMSO in liquid nitrogen at $1x10^6$ cells/ml.

2.4 SDS-PAGE Gel Electrophoresis and Western Blotting

2.4.1 Cell Harvesting

A confluent T25(cm^2) flask yielded an adequate number of cells to be able to asses protein expression by SDS-PAGE gel electrophoresis and western blotting. Cells were washed once with sterile PBS and detached with 0.5ml of 1x trypsin/EDTA incubated at 37°C, 5% CO_2 for 1-5minutes. Cells were pelleted via centrifugation (300xg) to remove trypsin and were subsequently washed with sterile PBS. Cells were then resuspended with ice-cold PBS, 10% (v/v) glycerol supplemented with EDTA-free protease inhibitor cocktail set III (Calbiochem) to prevent proteolysis. Cells were lysed by 2x 10 second burst sonication at 40% output (Microsonics Instruments).

2.4.2 Concentration Determination Assay

A modified Lowry protein assay (DC 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 0-10µg bovine serum albumin (BSA) constructed in parallel. All samples were analysed in duplicate.

2.4.3 SDS-Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Laemmli, 1970. Equal amounts of total protein lysate (20µg) 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 60-minutes; samples were then electrophoretically resolved on 8% (w/v) polyacrylamide gels alongside molecular weight protein markers (SeeBlue® Plus2 Prestained Standard, Novex) at 30mA/gel until the protein dye-front has fully resolved.

2.4.4 Western Blotting

Following SDS-PAGE, proteins were transferred (200mA for 2 hours, 4°C) onto a nitrocellulose membrane (Amersham) by electrophoresis as first described by Towbin et al., 1979. Membranes were then transiently stained with Ponceau S solution (0.1% (w/v) Ponceau S, 1% (v/v) acetic acid) to confirm efficient transfer, washed for 5 minutes with PBS-Tween (PBS-T, 0.1% (w/v) Tween-20), then incubated in blocking buffer (5% (w/v) skimmed milk powder in PBS-T) for 1 hour at room temperature to prevent non-specific binding ofantibody. Membranes were then incubated with the primary antibody BXP-21

(Calbiochem), at 1:1000 dilution (prepared in blocking buffer) for either 1 hour at room temperature, or (preferentially) overnight at 4°C. The membranes were then washed several times in PBS-T over a 15-20 minute period to remove any unbound primary antibody before incubation with a secondary antibody, rabbit anti-mouse horseradish peroxidase (DAKO) at 1:2000 dilution (prepared in PBS-T) for 1 hour at room temperature. Membranes were then washed several times in PBS-T, again over a 15-20 minute period, to remove any un- bound secondary antibody. Following the final wash, the membranes were in- cubated with enhanced chemiluminescence reagent (ECL, Supersignal West Pico, Thermo Fisher Scientific) and imaged using the LAS-3000 mini (Fujifilm).

2.5 Confocal Microscopy

2.5.1 Materials and Reagents

35mm^2 glass bottomed dishes with a No. 1.5 coverslip were purchased from MatTek® Corp. Black-framed 96-well tissue culture plates were obtained from Greiner Bio-One. All other reagents were obtained from Sigma-Aldrich unless otherwise stated.

2.5.2 Poly-L-lysine Treatment

To ensure efficient attachment of HEK293T cells, 35mm^2 glass bottomed dishes were coated with 0.1mg/mL poly-L-lysine and left to incubate at room temperature for at least 30 minutes prior to cell seeding. The dishes were then washed twice with sterile PBS to remove excess poly-L-lysine before cell seeding.

2.5.3 Cell Seeding for Live Cell Imaging

Cells were seeded at 100,000 cells per dish in poly-L-lysine treated 35mm^2 glass bottomed dishes and maintained in complete medium at 37°C, 5% CO_2 for 24 hours, with cell density determined through a haemocytometer. Thirty minutes prior to imaging, cells were washed twice with pre-warmed (37°C) imaging medium (phenol red-free Hank's Balanced Salt Solution, HBSS) and then maintained in imaging medium for the entire duration of confocal image acquisition.

2.5.4 Confocal Laser Scanning Image Acquisition

Confocal imaging was done using a Zeiss LSM 710 confocal laser scanning microscope with a Zeiss EC Plan-NEOFLUAR 40x/1.3 Ph3 oil immersion objective and argon laser. An oil immersion objective was used instead of a water immersion objective despite the imaging medium used having a similar refractive index as water due to its ability to take phase contrast images with transmitted light. The argon laser was set at 2% power to detect the sfGFP-tagged ABCG2 isoforms with a pinhole set to 1 AU (Airy unit), the excitation wavelength set to 488nm and the emission wavelength collected set between 500-530nm (EGFP channel). For any given dish seeded, 2 images were acquired at different areas with the following phase contrast parameters: 1024 x 1024 pixels and an image depth of 16-bits. This is to allow for high resolution image acquisition of protein localisation in a mixed cell population.

2.5.5 Cell Seeding for Confocal Plate Reader Image Acquisition

Some mutants that were shown to have a varied protein expression profile were grown in complete medium supplemented with different concentrations of Zeocin™ (40µg/ml, 50µg/ml, 80µg/ml, 100µg/ml) at 37°C with 5% CO_2 for 7 days, with a medium change every 1-2 days. Forty-eight hours prior to imaging, cells were seeded in a 96-well black-walled cell culture plate (Greiner Bio-One) at a density of 10,000 cells per well, with cell density determined using a haemocytometer. Thirty minutes prior to imaging, cells were washed with imaging medium and maintained in imaging medium as previously described.

2.5.6 Confocal Plate Reader Image Acquisition

Confocal imaging was done with the ImageXpress (IX) Ultra confocal plate reader with a Nikon Plan Fluor 20x/0.45 ELWD air objective, using the FITC (488nm excitation wavelength, 500-530nm emission wavelength) channel. Four images were obtained per well and each image was obtained with the following parameters: 1024 x 1024 pixels with an image depth of 16-bits. This was to ensure high resolution image acquisition of sfGFP-ABCG2. Images were then processed and analysed with MetaXpress software.

2.6 Flow Cytometry

2.6.1 Materials and Reagents

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

2.6.2 Fluorescence-based Drug Accumulation Assay

Cells were seeded at $1x10^6$cells/ml in phenol-red free DMEM and incubated with either DMSO (solvent control, 0.2% (w/v)) or mitoxantrone (MX, 10μm) in the presence or absence of the ABCG2 inhibitor Ko143 (1μM) at 37°C for 30 minutes with occasional agitation. Cells were then pelleted via centrifugation at 350xg for 5 minutes at 4°C and the supernatant discarded. Cells were then resuspended in either phenol-red free DMEM only (for samples that were incubated with DMSO, and MX only) or phenol-red free DMEM supplemented with inhibitor (for samples that were incubated with MX and Ko143), and were incubated at 37°C for 1 hour. Cells were further pelleted via centrifugation (350xg for 5 minutes at 4°C) and resuspended in phenol-red free DMEM prior to analysis by flow cytometry.

2.6.3 Flow Cytometry Data Acquisition

Data were collected using the Cytomics FC500 flow cytometer (Beckman Coulter) using the fluorescent parameters of MX and the appropriate fluorescence detection channel. For MX, which has an excitation wavelength of 635nm and an emission wavelength of 670nm, the FL4:633/75 channel was used. GFP fluorescence, representative of sfGFP-ABCG2 protein expression, was collected simultaneously to the drug fluorescent parameters and was determined using an excitation wavelength of 488nm and an emission wavelength of 526nm (on the FC500, the corresponding channel is FL1:523/40 (FITC/GFP)). Three independent experimental repeats were performed for each mutant.

2.6.4 Flow Cytometry Data Analysis

Data were initially analysed with Kaluza analysis version 1.5 (Beckman Coulter). Forward scatter versus side scatter (FSC vs SSC) was used to exclude debris from live cells. FSC versus the auxiliary parameter signal (AUX, a measure of the peak fluorescence signal in tandem with the integral fluorescence signal of the sample) gating was used to separate monodispersed cells from doublets. Single parameter histograms were used to identify fluorescent characteristics. Due to the nature of a heterogenous stable cell line, some GFP (protein expression) profiles demonstrated a clear split in the population distribution. Thus, a final parameter gating for GFP (protein) fluorescence was done for the wild type (WT) and all mutants that showed a clear population split.

The median fluorescent value for the solvent control was used as a baseline and was subtracted from the values obtained for both MX only or MX in the presence of inhibitor. The fractional difference in fluorescence between the "MX alone" or "MX + inhibitor" was calculated as a measure of raw efflux capacity. This value was then normalised for the expression level of the ABCG2 isoform compared to the wild type (based on GFP fluorescence), giving the relative efflux of MX with respect to GFP (and thus ABCG2) expression. The gating and normalisation ensured that data represented the function of the mutant isoforms and not their expression level.

The values for relative efflux were analysed using GraphPad Prism 8 (GraphPad Prism Software, Inc.) through a one-way ANOVA with a Dunnett's multiple comparisons test against the wild type protein to assess if drug efflux was perturbed in any mutant. Dunnett's multiple comparisons test was used as it allowed for the comparison of the mean relative drug efflux capacity of each mutant against the relative drug ef- flux capacity of the wild type (the baseline/control).

Chapter 3: Results

3.1 Hypothesis Generation

Previous work conducted by Dr Megan Cox from the Kerr Lab has identified the presence of a "surface site", a selection of amino acids from TM helices 4, 5a and 6b that are lipid-facing, that may be involved in substrate recognition prior to its transfer into the internal cavity of the ABCG2 homodimer in preparation for efflux. These residues are also part of a third proposed drug binding site found through molecular modelling and docking of ABCG2 with native and non-native substrates as conducted by László et al. (2016). Substitutions of some of the residues at the surface site (M523, M541, F545 and F640) to alanine have resulted in altered ABCG2 function, with some mutations allowing for the efflux of a non-native substrate (daunorubicin) or an enhanced/reduced efflux of native substrates (mitoxantrone and pheophorbide A). This provides further evidence that the surface site contributes to allocrite recognition and efflux. However, the exact mechanism behind this site's ability to recognise substrates is unclear.

It was proposed by Cox et al. (2018) that this alteration in allocrite recognition and efflux could be due to a potential disruption in unique methionine-aromatic interactions that are present at the surface site. The interaction between methionine, or more specifically the thioether group/sulphur atom, and aromatic residues has been identified in many proteins (Weber and Warren, 2019; Zauhar et al., 2000). However, the function of these interactions has yet to be fully elucidated. With the few studies that have been conducted with soluble proteins and through bioinformatics studies, it has been proposed that methionine-aromatic interactions are important in protein structure stabilisation, and in some cases may be important in the function of metalloenzymes

through the stabilisation of the metal ion cofactor (Lewis et al., 2016; Valley et al., 2012; Weber and Warren, 2018). The nature of this inter- action is truly unique: through a quantum calculation study conducted by Val- ley et al., 2012, they discovered that methionine and aromatic residues (phenylalanine, tyrosine and tryptophan) have an interaction energy that is similar to a salt bridge (1-3kcal/mol) but occurs at a much larger distance (5.0- 6.0Å), almost reaching the interaction distance of Van der Waals forces. This interaction between methionine and aromatic residues is further strengthened (by 0.5-1.5kcal/mol) through the oxidation of the thioether group, turning methionine into methionine sulphoxide (Met^{ox}) (Lewis et al., 2016). This, done through quantum calculations, molecular dynamics simulations and experimental methods showed that the oxidation of methionine residues in LTα ablated its ability to bind to TNFR1. This study showed that the methionine- aromatic interactions present between the methionine residues in LTα and the phenylalanine residues in TNFR1 were responsible for ligand-receptor binding between the two proteins, and oxidation of the methionine residues in LTα led to these residues forming intramolecular methionine-aromatic interactions, rotating the methionine residues at a ~30-60° angle, thus preventing these residues from forming any interaction with the phenylalanines in TNFR1 and leading to perturbed binding. This then implies that a disruption to the methionine-aromatic interactions does have an affect protein and ligand binding. Whether these unique interactions play a role in allocrite recognition and efflux in ABC transporters has yet to be determined.

Through this study, my aim was to determine experimentally if these proposed methionine-aromatic interactions affect the recognition and efflux of the ABCG2 native substrate mitoxantrone, and compare the results acquired to the previ- ous results obtained by Dr Cox in her previous studies.

3.2 Identification of Methionine-Aromatic Interactions at the Surface Site

Initial parsing of possible methionine-aromatic interactions in ABCG2 was done at the surface site by Cox et al. (2018), which consists of a patch of residues that are lipid-facing and are found in TM4, TM5a and TM6b as seen in figure 3.1. For the initial round of parsing, three structures of ABCG2 found in the Protein Data Bank (PDB) were used (PDB ID 5NJ3, 6ETI and 6HBU, see Table 3.1) and the distance between the methionine residues present at this surface site and any nearby aromatic residues was found using a custom Python script on PyMOL (see Appendix 1) as seen in figures 3.1 and 3.2. These three structures were used for the following reasons. 5NJ3 was used as it was the first cryo-EM structure of ABCG2 deposited into the PDB, thus this structure is used as a benchmark for initial methionine-aromatic parsing. 6ETI was used as the cryo-EM structure has a bound Ko143 derivative (MZ29) and, more importantly, is locked into a closed conformation apo-state by the presence of the 5D3 antibody, which is the conformation believed to represent the initial drug binding conformation. 6HBU was the last structure used as it has bound ATP in the NBDs, locking the homodimer in an open-conformation state, which would mean that there could be a change as to which residues had the methionine-aromatic interaction due to the transmitted conformational change from the NBDs to the TMDs when ATP is bound.

The parameters for identifying these interactions are derived from the criteria set by Valley et al. (2012) in their attempts at identifying methionine-aromatic interactions in soluble proteins and are as such: (1) the distance between the methionine and a nearby aromatic residue must be within a range of 5.0-6.0Å, (2) the sulphur atom in the thioether group present in methionine must be

pointing away from the plane of the benzene ring of the aromatic residues as seen in figures 3.1 and 3.2.

Table 3.1: ABCG2 cryo-EM structures used for methionine-aromatic interaction parsing.

PDB ID	Resolution (Å)	ABCG2 State	Bound molecules	Reference
5NJ3	3.78	Apo-state	NAG, 5D3	Taylor et al., 2017
6ETI	3.10	Apo-state	MZ29, 5D3	Jackson et al., 2018
6HBU	3.09	ATP-bound state	ATP, Mg^{2+}	Manolaridis et al., 2018

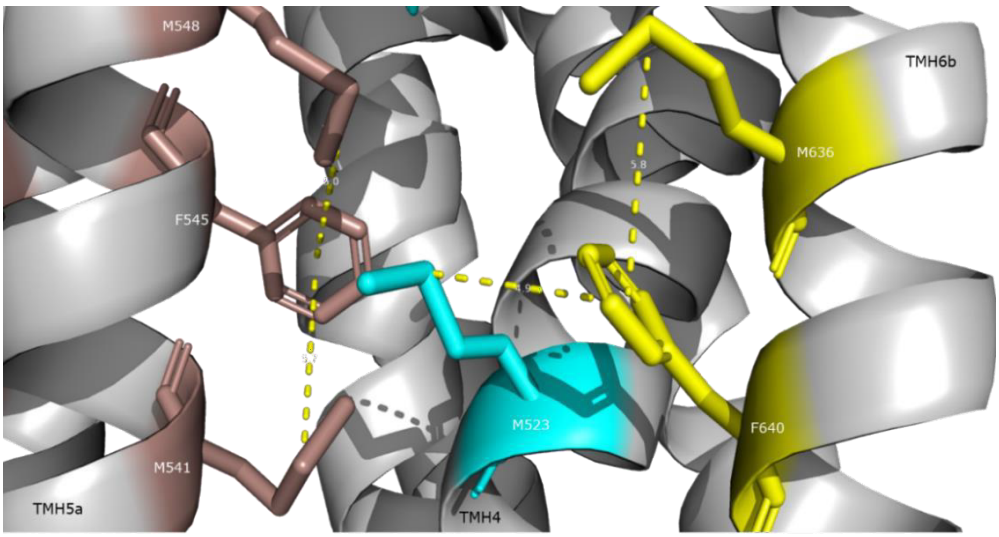


Figure 3.1: Identified methionine-aromatic residues in the apo-state structure of ABCG2. Residues in TMH5a are coloured dark salmon, residues in TMH4 in cyan, and residues in TMH6b in yellow. The proposed methionine-aromatic interactions are visualised using yellow dashes, with the methionine thioether orientation with respect to the plane of the benzene ring matching the parameters proposed by Valley et al. (2012). Parsing, identification and visualisation were done in PyMOL (Schrodinger, LLC) using a custom script with PDB structure 6ETI.

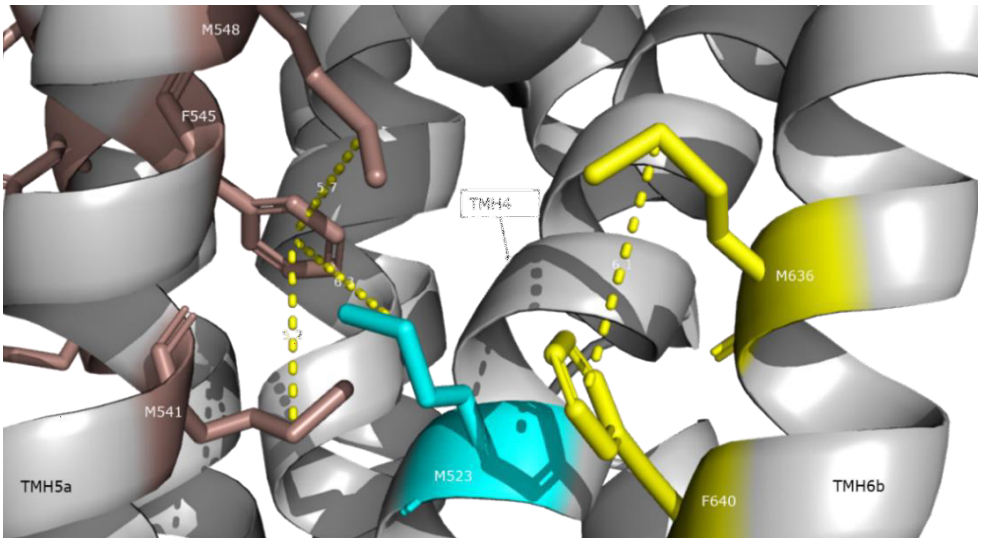


Figure 3.2: Identified methionine-aromatic interactions differ in the ATP-bound state of ABCG2. M523 appears to have formed an interaction with a different Phe residue when ABCG2 is in an open-conformation (M523-F545) compared to when the protein is in a closed-conformation (M523-F640) (see figure 3.1). Parsing, identification and visualisation were done in PyMOL using a custom script with the PDB structure 6HBU.

Attempts at parsing for methionine-aromatic interactions in the internal cavity did not produce any results. Although there were many phenylalanine residues present in the internal cavity, there were no available methionine residues in the internal cavity or in nearby helices that appear to have the ideal orientation to form these interactions. An example of the parsing and visualisation can be seen in figure 3.3.

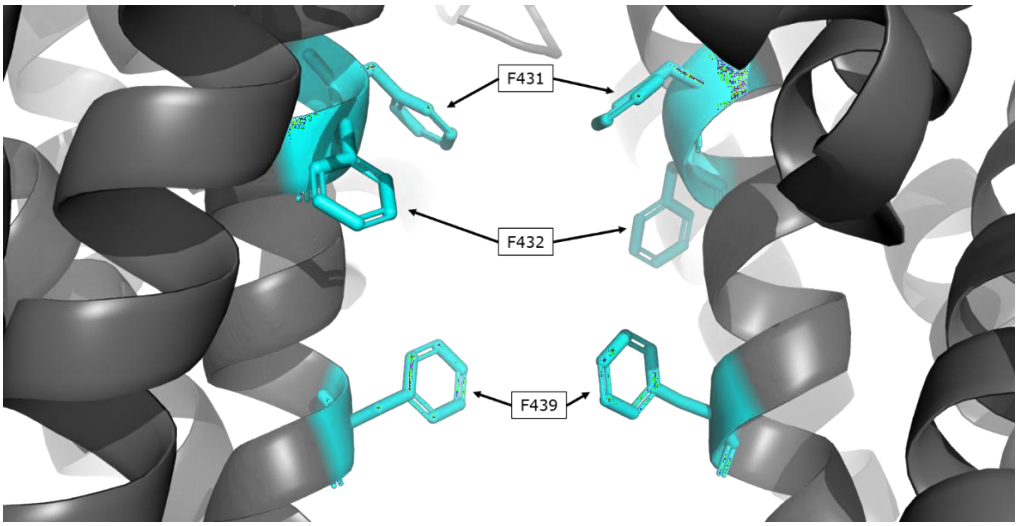


Figure 3.3: Visualisation of an attempt in identifying methionine-aromatic interactions in the internal groove of the apo-state structure of ABCG2. Although there are three prominent phenylalanine residues found in the internal cavity (residues coloured in cyan), there are no nearby methionine residues present in the internal cavity or in nearby helices that had the correct orientation or were within 5-6Å from the identified phenylalanine residues. Parsing, identification and visualisation were done in PyMOL using a custom script with the PDB structure 6ETI.

Thus, the phenylalanine and methionine residues identified at the surface site were carried forward for site-directed mutagenesis. Previously some of these surface site residues were substituted into alanine which, although is an aliphatic residue that would be well tolerated in the transmembrane helices and is unlikely to clash sterically with surrounding residues, does create additional empty space in the overall structure due to the drastic reduction in side-chain size compared to the wild-type residue. To further understand the function of ABCG2 through a mutagenesis study, ideally a series of both gross and more nuanced residue substitutions could help in understanding how both the size and additional/fewer inter- and intramolecular interactions could affect allocrite binding and efflux. This study has chosen the following residue changes which will be described in detail in the following paragraphs.

In the study carried out by Valley et al. (2012), it has been stated that phenylalanine-methionine has the highest prevalence amongst the ~15,000 PDB structures analysed, followed by tyrosine-Met and tryptophan-Met. Therefore, phenylalanine residues present at the surface site (F545 and F640) will be mutated to the aromatic residues tyrosine and tryptophan to see if a minor change in the side-chain structure (tyrosine; a benzene ring with an additional hydroxyl group) and a major change in side-chain structure (tryptophan; an indole ring) will affect drug binding and recognition. An additional phenylalanine residue, F571, a residue towards the extracellular side of TM4 just above the proposed surface site and in relatively close proximity to a methionine residue present at the surface site (M523), will also be mutated to tyrosine and tryptophan identify if a residue in close proximity to the surface site can disrupt the proposed methionine-aromatic interactions at the surface site.

The methionine residues present at the surface site (M523, M541, M636) will be mutated into leucine (L) and glutamine (Q). Leucine was chosen as it is an aliphatic residue which has a similar side-chain size with methionine, and glutamine was chosen to mimic the structure of Met^{ox}, but without the thioether group (figure 3.5). In a recent study by Walgenbach et al. (2018), a quantum calculations study predicted that a glutamine substitution for methionine residues led to an increase in the binding energy for some of the identified methionine-aromatic interactions in calmodulin (30-40% increase, although it is important to note that this increase is less than -1kcal/mol), hence why the glutamine substitution was made in this study.

In an attempt to identify if these residues were conserved amongst other ABCG2 homologues across different species, a sequence alignment was performed using the human ABCG2 FASTA sequence as the query sequence and the FASTA sequences of 150 jawed vertebrates using the free sequence

analysis software package HMMER. The scoring matrix of this sequence alignment was then performed using BLOSUM62 and the resulting findings were visualised in R with the help of Dr James Mitchell-White from the Kerr Lab. Residues which are highly conserved across multiple species would likely indicate that these residues most likely play an important role in either protein structure stabilisation and, more importantly for this study, may help play a role in allocrite recognition and possibly efflux. The result of these findings are shown in figure 3.4, with all 6 residues selected for mutagenesis confirmed to be highly conserved amongst the 150 ABCG2 homologous sequences, although out of all 6 residues F571 had the lowest conservation score (<75%).

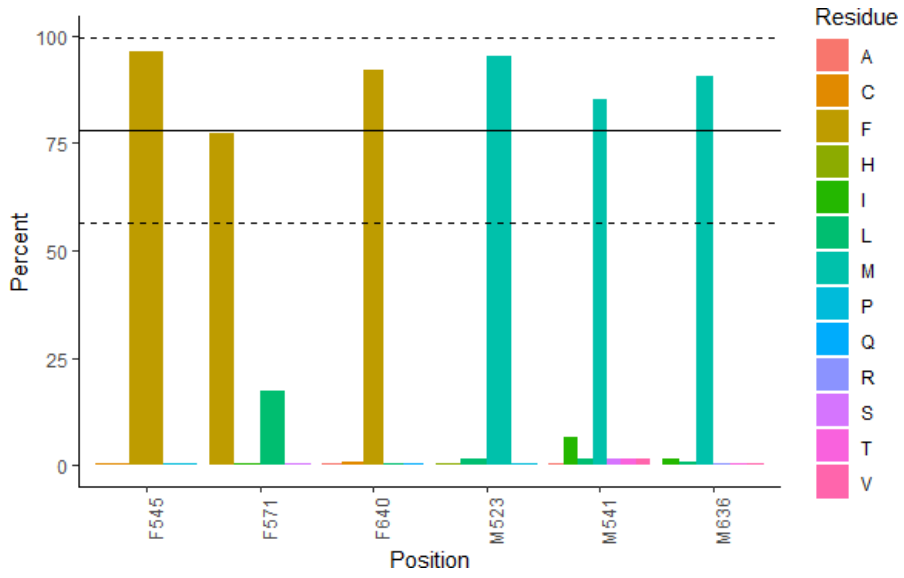


Figure 3.4: Conservation of surface site residues and F571 across homologous ABCG2 structures in jawed vertebrates. All residues in the surface site and F571 appear high conserved across 150 jawed vertebrate species, with a percentage value of ≥50% being deemed acceptable. The colours in the right-hand legend correspond to the amino acid (denoted in their single letter code) that is conserved and the higher the percentage, the higher the conservation score.

These residue substitutions listed above are made to disrupt the predicted methionine-aromatic interactions that may be present between these residues at the surface site. These residue substitutions may then affect allocrite recognition and therefore may lead to altered allocrite transport.

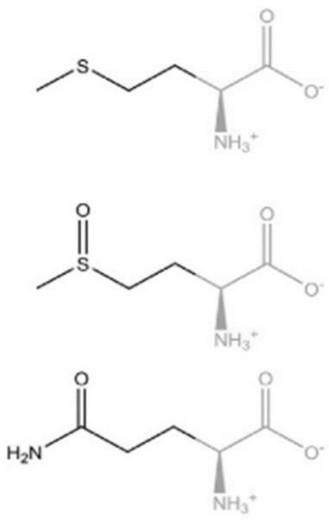


Figure 3.5: Methionine → Glutamine substitution mimics the structure of methionine sulphoxide. This amino acid substitution was performed due to the similarity in structure and electron dispersion between glutamine and methionine sulphoxide, which may affect the proposed methionine-aromatic interactions that have been identified. Diagram adapted from Walgenbach et al. (2018).

3.3 Plasmid Construct Information

The backbone plasmid used to create the mutant constructs study is a pcDNA™3.1/zeo(+) mammalian expression vector. This plasmid has the cytomegalovirus (CMV) and the simian vacuolating virus 40 (SV40) promoter region for high level expression of the gene of interest in a mammalian expression system and contains multiple cloning sites as well as an antibiotic resistance cassette (ampicillin) for use in selection of successfully transformed bacterial cells. The plasmid also contains the *Sh ble* gene, which provides the transfected eukaryotic cells with zeocin resistance, allowing for the selection of cells which stably express the plasmid (and in turn, the protein of interest).

The pcDNA-3.1zeo(+)_sfGFP_ABCG2 plasmid (figure 3.6) created by Dr Kelvin Wong from the Kerr Lab was engineered to include an enhanced GFP template containing additional "superfolder" mutations (M135T, V163A, S30R, Y30N) along with an A206K mutation to prevent self-dimerisation of the fluorescence tag (Pédelacq et al., 2006; Wong et al., 2016b; Zacharias et al., 2002). This A206K mutation was included as ABCG2 has been observed to form higher order oligomeric structures, with the simplest form being a dimer (Wong et al., 2016b). Thus this mutation in the superfolder GFP (sfGFP) is important to prevent the artificial dimerisation of an ABCG2 isoform, whose mutation had resulted in the inability to form a stable dimer when expressed. The sfGFP gene is placed in frame with the N-terminus of the ABCG2 cDNA allowing for easy visualisation and analysis through microscopy methods and flow cytometry respectively.

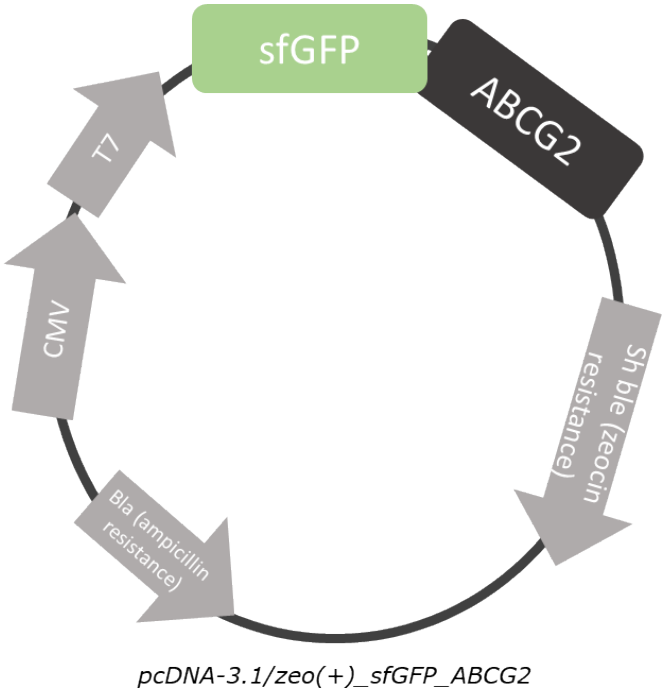


Figure 3.6: Plasmid map of pcDNA-3.1/zeo(+)_sfGFP_ABCG2. The plasmid contains a strong CMV promoter for enhanced gene expression, multiple cloning sites, and a bacterial resistance gene for genetic engineering. The plasmid also contains an sfGFP fusion tag for visualisation of the protein and *Sh ble* antibiotic resistance gene, providing the successfully transfected eukaryotic cell zeocin resistance for the selection of stable cells lines.

3.4 Construct Generation

Site-directed mutagenesis was performed to incorporate mutant codons into the template plasmid DNA based on the *in vitro* amplification of DNA via polymerase chain reaction (PCR) (Sambrook and Russell, 2001; summarised in figure 3.7). Briefly, the mutant codons were incorporated through the annealing of mutant forward and reverse primers to the template DNA and subsequently elongated to produce the mutant constructs. The template DNA is then removed using DpnI endonuclease, which recognises hemi-methylated DNA and degrades it. The resulting PCR amplicon would thus only contain the amplified DNA and carried forward to bacterial transformation.

These PCR amplicons were subsequently transformed into competent DH5α (*E. coli*) cells, and the plasmid DNA from resulting colonies was isolated and samples of these isolated plasmid DNA were digested using PvuII endonuclease restriction enzyme alongside the template DNA. The digested samples were then resolved in 1% agarose gel to confirm the presence of a pcDNA-3.1/zeo(+)_sfGFP_ABCG2 plasmid. The template DNA and isolated plasmid DNA from successful transformations will have 3 clear bands when digested with PvuII with a size of 4.5, 1.7 and 1.4 kbp (figure 3.8). In some lanes, an extra band with varying intensity between the putative mutant-containing plasmids can be seen, which is slightly less than 3kbp in size. This is seen as PvuII does exhibit occasional star activity leading to an altered recognition of restriction sites.

Of the intended mutant constructs designed, 8 were successfully generated. The 4 unsuccessful mutant primers were put through gradient PCR, using a range of annealing temperatures (52°C-61°C, with 3°C intervals) to determine if different annealing temperatures were needed to incorporate the mutant codons into the wild type plasmid. Samples of the PCR amplicons from these

different annealing temperatures were resolved on 1% agarose gel as seen in figure 3.9 after digestion with DpnI to see if any of the temperatures created any successfully amplified products.

For F545Y and F640W, the final annealing temperature used to create these mutant constructs was 52°C as this temperature created the highest band intensity on the 1% agarose gel, meaning that this temperature allowed for the most efficient amplification of plasmid DNA. F571Y showed the highest band intensity at 58°C while F571W showed the highest band intensity at 62°C (not shown). The PCR amplicon from these temperatures were then transformed into DH5α cells and the plasmid DNA was subsequently isolated.

Sanger sequencing was performed with all 12 isolated plasmid DNAs to confirm the presence of the mutant codons in the plasmid constructs as seen in figure 3.10. Successful insertion of the mutant codons at the correct locus in the ABCG2 gene allowed for the isolated plasmid DNA containing the mutant constructs to be carried forward for transfection into a mammalian expression system.

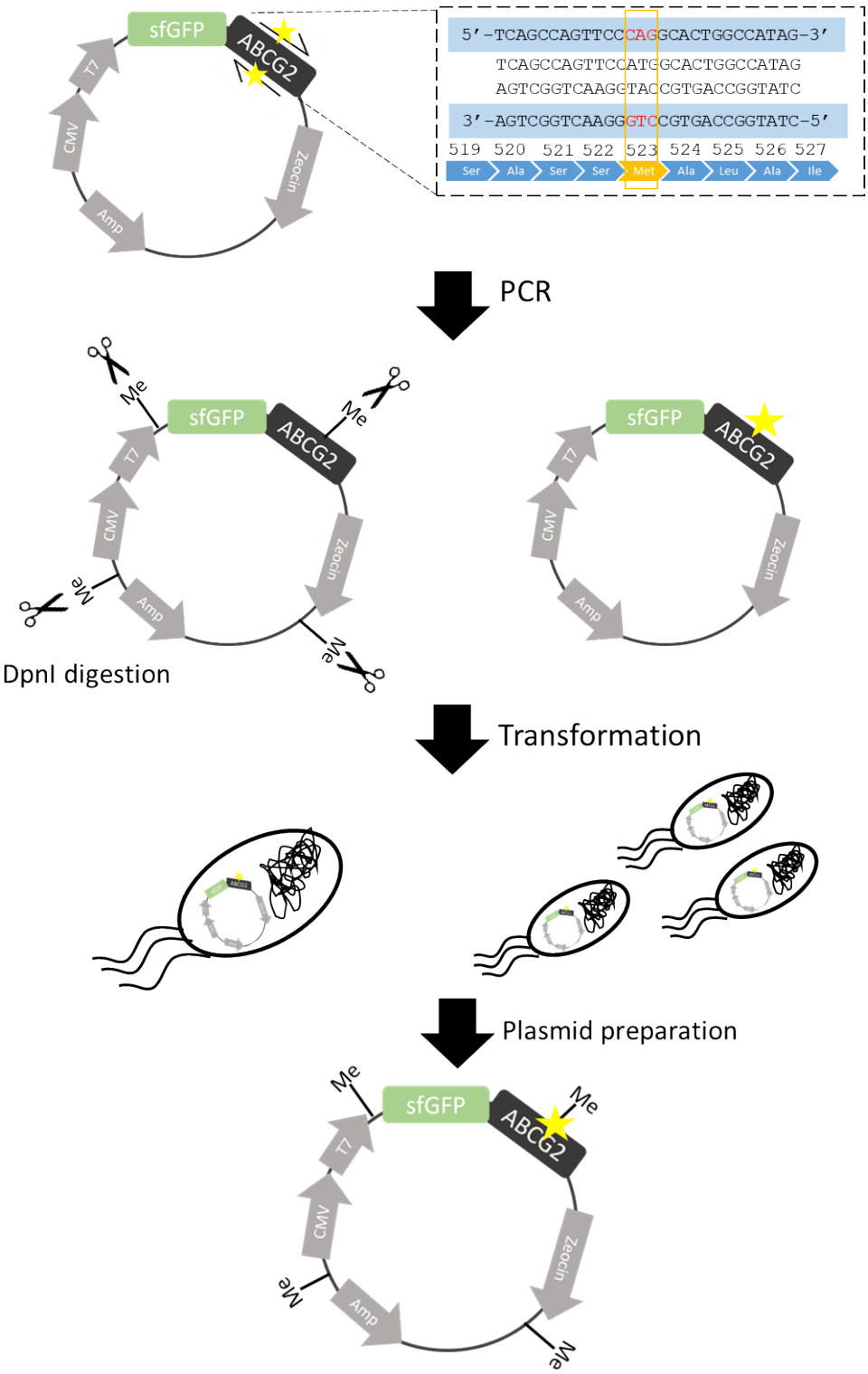


Figure 3.7: Schematic representation of mutant construct generation. Primers containing the mutant codons anneal to the template DNA via their complementary regions and PCR elongation incorporates the mutation. The methylated template DNA is removed through a DpnI digest and the PCR amplicons are subsequently transformed into competent DH5α cells. The plasmid DNA containing the incorporated mutation is then isolated.

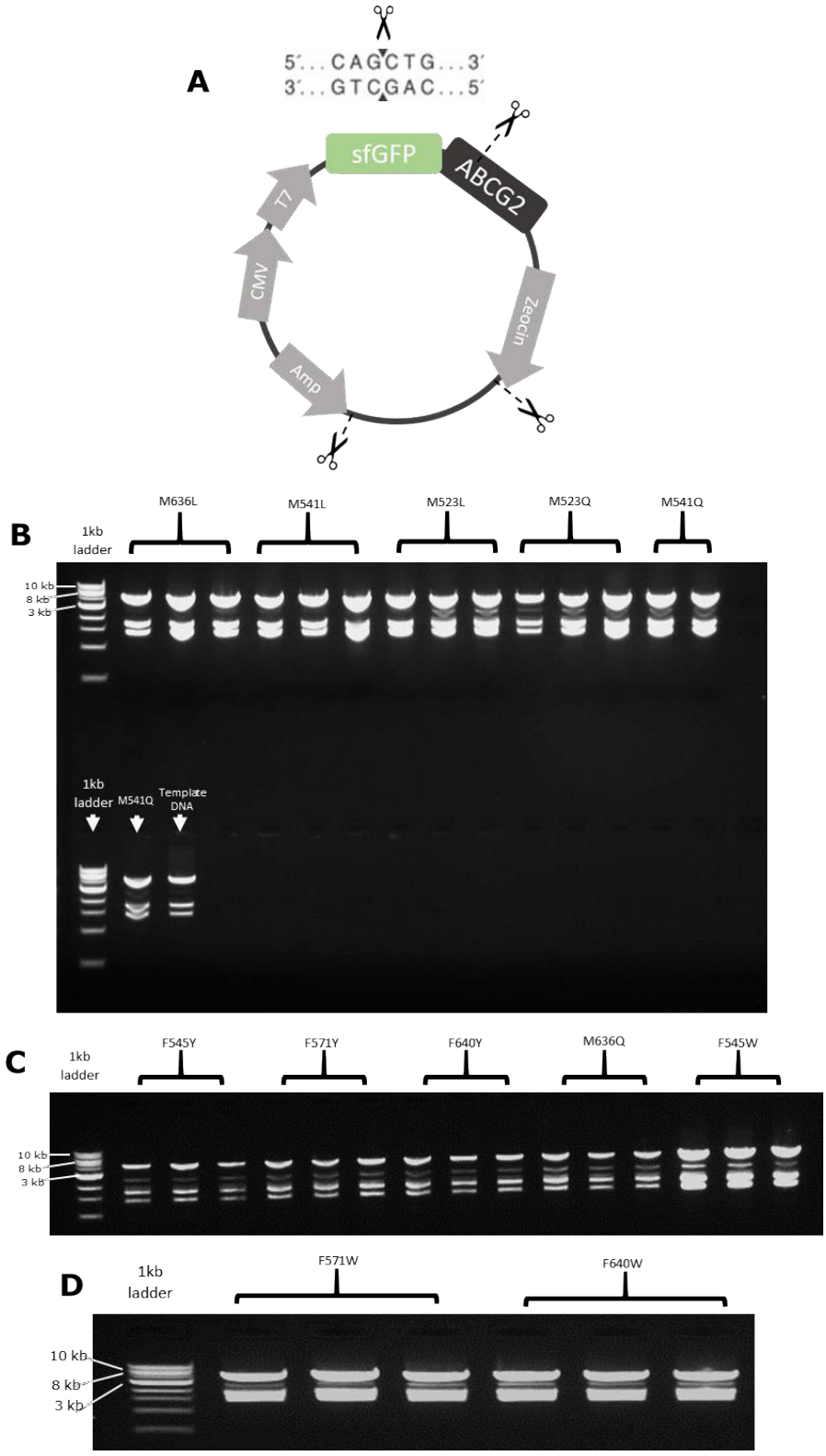


Figure 3.8 A: PvuII endonuclease digest. A schematic diagram showing the locations of the restriction recognition sites of PvuII on the plasmid **B, C and D:** Agarose gel analysis of mutant constructs digests. PvuII digests of the isolated plasmid DNA were performed to confirm successful transformation and confirm the correct backbone plasmid was used. No additional PvuII restriction recognition sites were found in any of the isolated plasmid DNA.

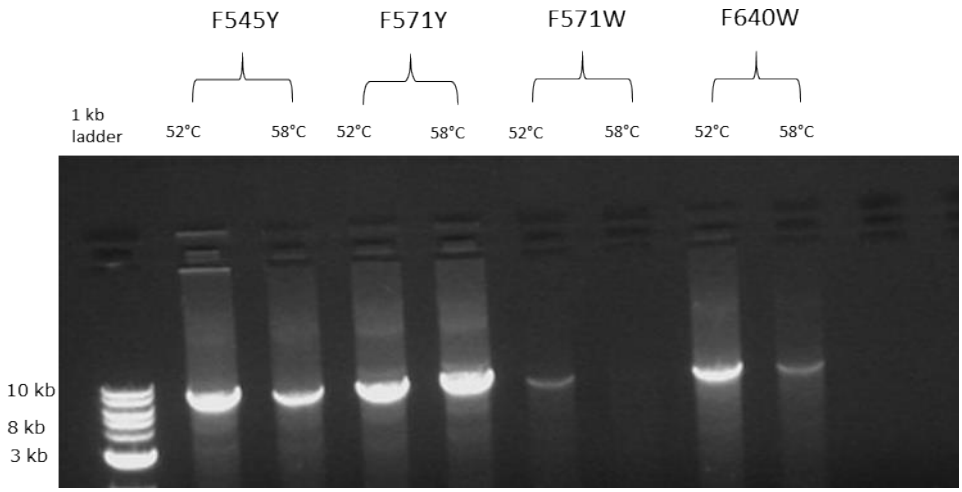


Figure 3.9: Gradient PCR for Unsuccessful Mutants. Four out of 12 mutants did not show any successful insertion of the mutant codons in Sanger sequencing, thus gradient PCR was performed. The annealing temperature was changed and a gradient starting from 52°C to 61°C with increments of 3 degrees was performed and the PCR amplicons were resolved on a 1% agarose gel (only a subset of conditions are shown for clarity). The temperature with the largest band was then carried forward for transformation and Sanger sequencing to confirm the presence of the mutant constructs. All bands are approximately 7kbp in size.

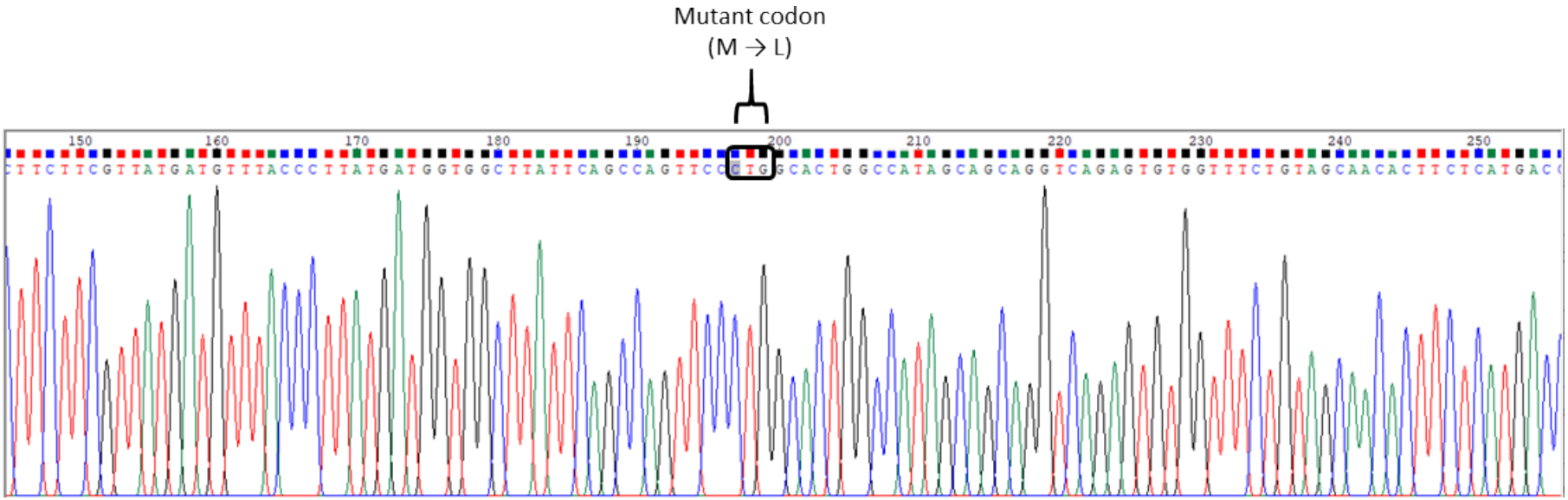


Figure 3.10: Confirmation of mutant codon replacement via Sanger sequencing. Sanger sequencing was performed using the primer *seq482* to confirm the successful introduction of the mutant codons, thus successful generation of the mutant constructs. This figure shows the successful mutation M523L. Full plasmid sequencing using the primer *seqF2, seqF0, seqR1* and *seqGFP-Rev* were performed with all the isolated plasmid DNA containing the successful mutant constructs to confirm the integrity of the plasmid is intact and no additional mutations further upstream or downstream of the intended loci were present.

3.5 Protein Expression

HEK 293 T cells, a human embryonic kidney cell line derivative, were used as the pcDNA3.1/zeo(+)_sfGFP_ABCG2 construct has previously been successfully expressed in this cell line by the members from the Kerr Lab (Cox et al., 2018; Haider et al., 2015; Wong et al., 2016b). This is due to the cell expressing a mutant SV40 large T antigen, which when combined with an expression vector that contains an SV40 promoter and origin of replication, allows for a high copy number expression of the gene of interest. As the wild type and mutant plasmid constructs have the SV40 promoter region intact, this would allow for successful overexpression of the sfGFP-ABCG2 isoforms and the wild type protein for visualisation of protein localisation and for functional assays.

All constructs were transiently transfected into HEK 293 T cells and subsequently underwent zeocin-mediated selection to create stable cells lines, using a concentration of 200µg/ml of zeocin supplemented into the complete growing medium. To validate transfection efficiency as well as monitor the selection of these stable cells line, cells were examined daily using epifluorescence microscopy by checking fluorescence as well as cell morphology changes. HEK 293 T cells appear as dense, round clusters when undergoing apoptosis due to external stressors present (in this case the high concentration of zeocin present), with the presence of dots indicating the presence of apoptotic bodies (i.e. dead cells). Once stable cell lines have been achieved and the zeocin concentration has been lowered, they will grow as a monolayer of squamous cells. The sfGFP-ABCG2 isoforms will fluoresce and can easily be detected using this method. Images of the cell lines taken from epifluorescence microscopy showing the presence of successfully transfected cells undergoing selection can be seen in figure 3.11.

Once stable cell lines were achieved and aliquots had been frozen in liquid nitrogen long-term storage, cells were harvested for use in SDS-PAGE and western blotting. Cells were lysed through sonication, and typically 20µg of total protein lysate was loaded into 8% polyacrylamide gels and resolved electrophoretically before being transferred onto a nitrocellulose membrane for western blotting.

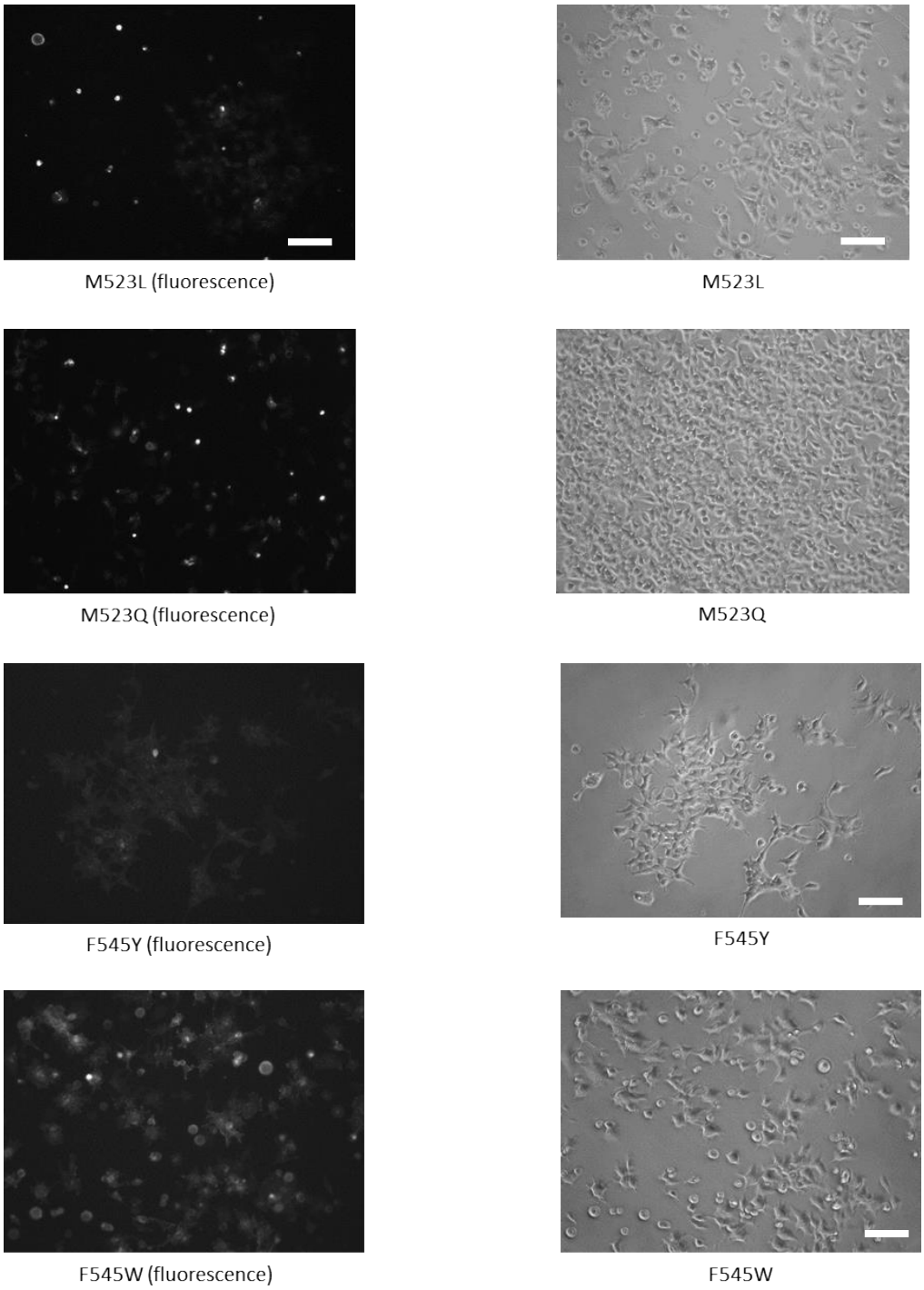


Figure 3.11: Selection of stable cell lines carrying the methionine mutations. Epifluorescence microscopy was used to confirm successful transfection of HEK 293 T cells and monitor the selection of stable cell lines. Transfected cells expressing sfGFP-ABCG2 will appear as fluorescent patches or dots as seen on the left hand images with the phase contrast images on the right, confirming the transfection efficiency was ≥50%. Scale bars are representative of 100μm for each image.

Western blotting of the sfGFP-ABCG2 isoform expression compared to the wild type protein was performed (figure 3.12).

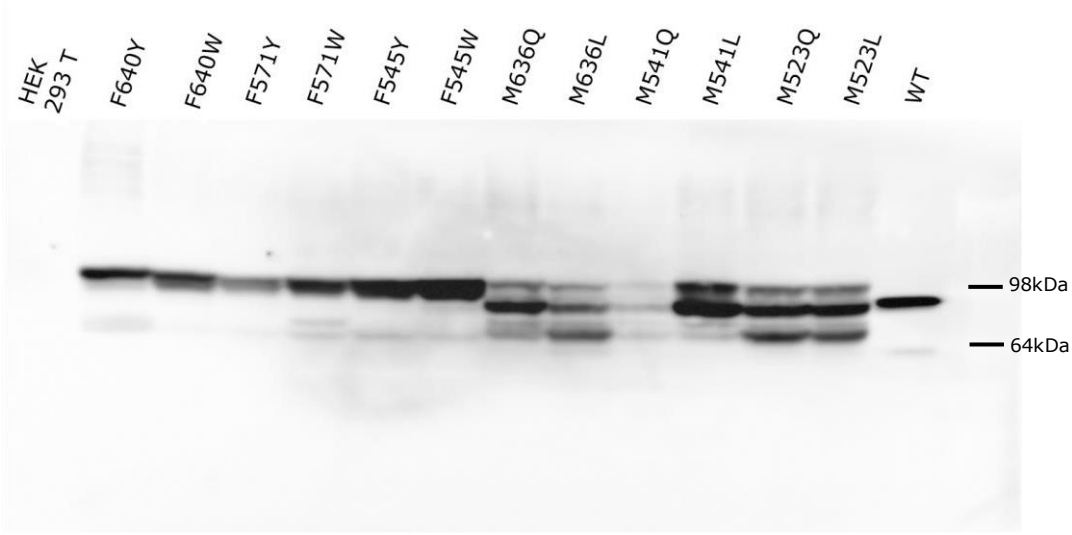


Figure 3.12: Western blot analysis of stable cell line expression compared to the wild type. Total protein lysates from the ABCG2 isoforms were resolved on an 8% polyacrylamide gel and a western blot was performed using BXP-21 to check ABCG2 expression against the wild type, with a negative control (untransfected HEK 293 T) loaded on the far left-hand side. sfGFP-tagged ABCG2 is approximately 100kDa in size. Three bands can be seen in some lanes likely as a result of slight protein degradation during the freezing process after cell harvesting.

All mutant isoforms are able to express full length ABCG2 with varying levels of expression compared to the wild type control. M541Q showed the lowest expression level compared to the all of the other isoforms followed by M636L and F571Y. F545W appears to have the highest expression level in comparison to the wild type, while the rest of the stable cell lines appear to have a similar expression level as the wild type protein. The extra bands present in some of the lanes (more so in the methionine mutant constructs) are likely due to protein degradation during the freezing process after cell harvesting, which had further degraded after sonication and denaturation of the protein during protein lysate preparation. This has previously been seen in mutagenesis studies conducted by Dr Megan Cox and noted by her in her thesis (unpublished).

3.6 Protein Localisation

Although western blotting can show whether these mutant constructs are able to express full-length ABCG2, it is not a particularly robust or quantitative measure of protein expression as it is unable to differentiate between total ABCG2 expression and surface expressing protein. Some mutants may be able to express a high amount of ABCG2 compared to the wild type, but may not be able to efflux any mitoxantrone due to internalisation of the protein in the cytoplasm. Therefore it is important to check if the presence of any mutations to the ABCG2 protein affects trafficking to the plasma membrane, thus potentially affecting its ability to function when a drug efflux assay is performed. This information would also provide information on whether any mutations affect protein folding as this process is intrinsically linked to post-translational release of the protein from the endoplasmic reticulum (ER) (Ellgaard et al., 2016). Analysis of ABCG2 isoform localisation was performed using laser scanning confocal microscopy. Live images of all stable cells line were obtained on 3 separate image acquisition sessions and at 2 randomly chosen areas of the seeded dishes to ensure that images representative of the cell population were obtained. Some images ac- quired from the confocal microscopy sessions showing the protein localisation of all ABCG2 isoforms and the wild type can be seen in figure 3.13.

All ABCG2 isoforms have shown successful localisation to the plasma membrane. This indicates that the mutations introduced in ABCG2 do not drastically affect protein folding and trafficking to the membrane. However, it is also seen that cells expressing either of the F640 mutants (F640W and F640Y) show pronounced intracellular vesicle formation containing the sfGFP-ABCG2 protein despite also showing clear localisation to the membrane.

The vesicles do not appear reticular in shape, arguing against protein misfolding at the ER, but their fairly small size (~1μm) and large intracellular quantity in some of the cells suggest some form of autophagy-mediated protein degradation, however this is merely speculation. Although it can be seen that the intracellular vesicles are also present in some of the other mutants such as F545Y and F571W, these were noted to be cells that were undergoing cell death. When these mutants were imaged at a much earlier time frame, these vesicles were not present. This is unlike the F640 mutants which, despite being imaged first in several sessions, still had pronounced vesicles present.

It is clear from both the western blot and confocal image analysisthat there are *varying levels* of protein expression in the ABCG2 isoforms compared to the wild type, although all mutants successfully traffic the protein to the plasma membrane. This could then lead to mutants showing *varying drug efflux* profiles due to the varying amount of mature ABCG2 that is being expressed and trafficked successfully to the plasma membrane. Therefore an assay that can quantify both functional transport and protein expression level is required to ensure accurate interpretation of the effects of mutations at the surface site.

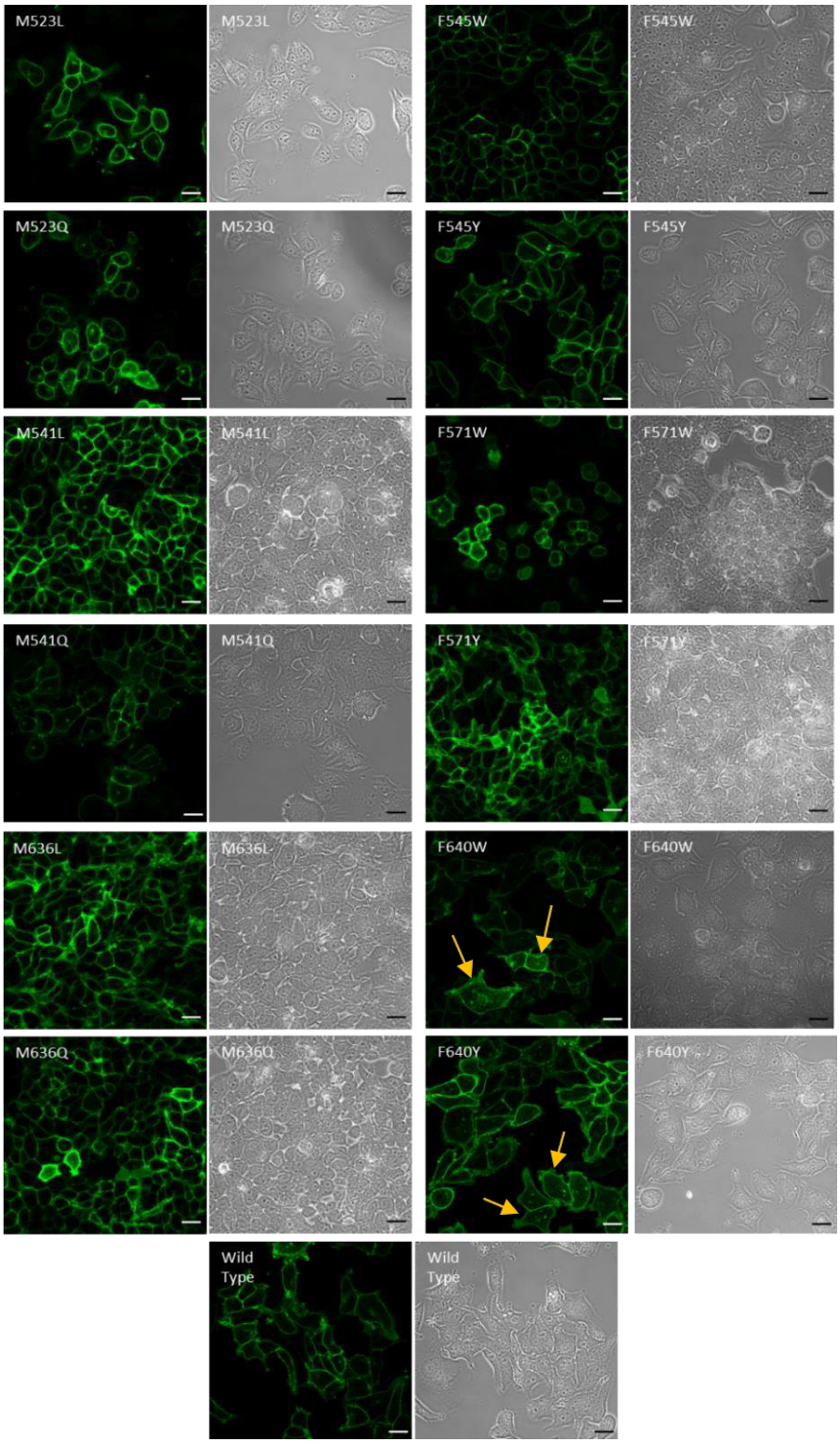


Figure 3.13: Confocal image analysis of protein localisation. All ABCG2 isoforms along with the wild type were imaged using laser scanning confocal microscopy on the Zeiss LSM 710, with fluorescence images on the left (EGFP channel, 488nm excitation and 500-530nm emission; 2% argon laser) and the phase contrast image (T-PMT channel) on the right. The scale bars on the bottom right-hand corner are representative of 20µm. All isoforms show clear membrane trafficking, although F640W and F640Y also show clear intracellular vesicles lined with GFP- tagged protein (yellow arrows).

3.7 Protein Function Characterisation

Determining the function of the ABCG2 isoforms was performed using a fluorescence-based drug accumulation assay via flow cytometry based on the methods described by Cox et al. (2018). By measuring the difference in cellular mitoxantrone fluorescence in the presence and absence of the ABCG2-specific inhibitor Ko143, the raw drug efflux can be calculated. By quantifying the GFP signal, the relative expression level of the mutant isoforms and the wild type can be factored into the overall analysis of function.

A schematic of how flow cytometry works and how it is used in this study can be seen in figure 3.14. Briefly, a cell suspension is streamlined until singular cells are streamed through a focused laser beam. The light that passes through the cell suspension is collected as the forward scatter signal (FSC), while the light that has reflected off the cell suspension is collected as the side scatter signal (SSC). The reflected light passes through several filters that will absorb and reflect certain wavelengths, giving the different channel signals that are used to measure mitoxantrone and GFP fluorescence. In this study, the FL1 and FL4 channels (measuring the GFP and mitoxantrone fluorescence respectively) were used and the signal collected from these channels were used to create single parameter histograms that show mitoxantrone efflux in the presence and absence of Ko143. These single parameter histograms for all ABCG2 isoforms along with the wild type and untransfected HEK 293 T can be seen in figure 3.15.

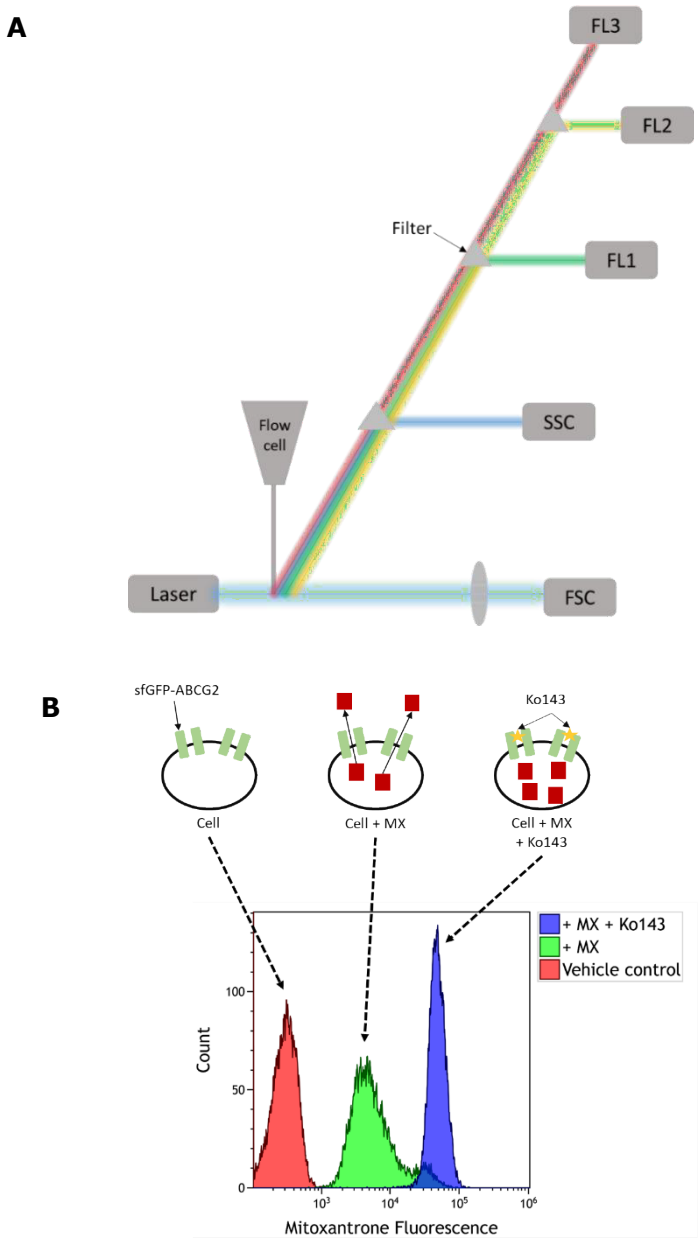


Figure 3.14 A: Schematic representation of flow cytometry. A cell suspension enters the flow cell where the cells are aligned into a single stream, and pass through a focused laser beam. The light from the laser that passes through the cell suspension is collected as forward scatter signal, and the light that is reflected off the individual cells will be collected as the side scatter signal and a series of fluorescence signals which depend on the filters used. In the flow cytometer machine used for this study, an additional AUX channel detector is present as the final channel (not shown). **B: Schematic of mitoxantrone efflux in the presence and absence of Ko143 in wild type expressing cells.** Cells were incubated with mitoxantrone (red squares) in the presence and absence of Ko143 (yellow stars). The difference in the fluorescence intensity as a result of the difference in the amount of mitoxantrone accumulated inside the cell is recorded and visualised as a single parameter histogram. This is then used for further calculations of raw efflux capacity for each mutant.

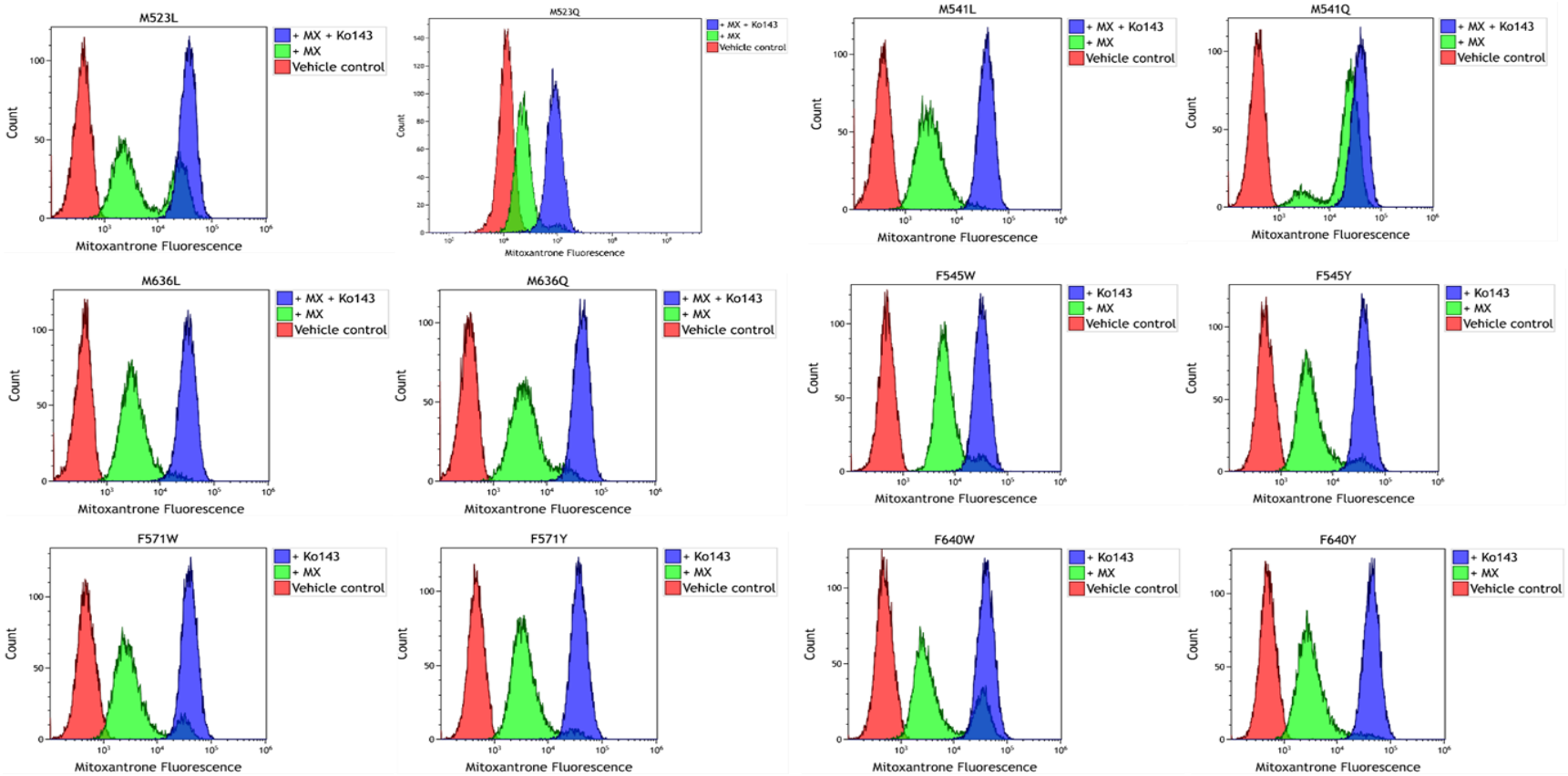


Figure 3.15: Single parameter histograms of mitoxantrone efflux in ABCG2 isoforms. Single parameter histograms using the FL4 channel show the drug efflux profile of all ABCG2 isoforms. The red histogram represents the autofluorescence of the cells (cells incubated with just the vehicle, DMSO), the green histogram represents the fluorescence of the cells with added mitoxantrone, and the blue histogram represents cell fluorescence with mitoxantrone supplemented with the ABCG2 inhibitor Ko143.

3.8 Flow Cytometry Data Analysis

It can be seen in figure 3.15 that there are varying mitoxantrone efflux profiles in the ABCG2 isoforms. Using these single parameter histograms for mitoxantrone fluorescence, the raw drug efflux of the entire cell population can be calculated. An example of the calculation will be shown in the following paragraphs.

Using the single parameter histogram for the wild type protein (sfGFP-ABCG2, figure 3.16), the medians for the DMSO vehicle control, the added drug and the drug in the presence of Ko143 were obtained. The median of the cells treated with mitoxantrone was subtracted from the mean of the DMSO control to remove the autofluorescence background signal (α). This was also done for the median of mitoxantrone in the presence of Ko143 (β).

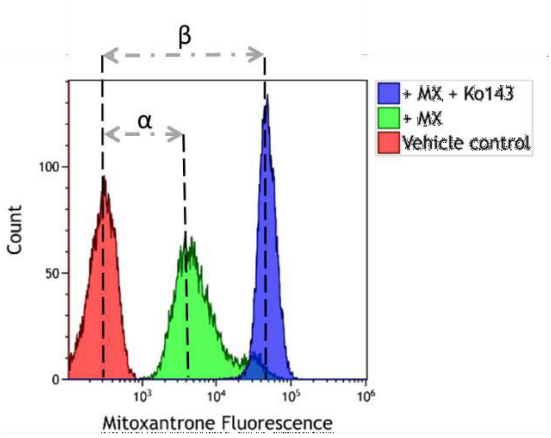


Figure 3.16: Measuring medians for all histograms to calculate the raw mitoxantrone efflux of the wild type protein. The median "autofluorescence" value of the cells represented by the DMSO vehicle control must be subtracted from the medians of cells treated with mitoxantrone in the presence and absence of Ko143. From there further calculations of the raw efflux of the cell line can be done.

From there, the difference between the mitoxantrone fluorescence with the presence and absence of Ko143 will be compared to the fluorescence of just mitoxantrone and the final value is given as a ratio of raw efflux capacity of the entire cell line.

$$\frac{\beta - \alpha}{\alpha} = raw\ drug\ efflux\ of\ entire\ cell\ population$$

It is important to note that as seen in the western blot and the protein localisation analysis, the ABCG2 isoforms do have varying protein expression levels, although all isoforms do have successful membrane trafficking. It can also be seen in the single parameter histograms for mitoxantrone efflux in figure 3.15 that some of the ABCG2 isoforms show two very distinct peaks (a clear example of these double peaks and population split can be seen in the single parameter histogram of mitoxantrone efflux for M523L in figure 3.15).

Indeed, for some of the ABCG2 isoforms there appears to be a population of cells comprised of high expressors of the protein that show clear efflux of mitoxantrone that is abolished/reduced significantly in the presence of Ko143. There also exists a population of cells that does not show any efflux of mitoxantrone, even in the absence of Ko143. In the case of M523L, two populations for cellular mitoxantrone fluorescence are seen in the absence of Ko143 with approximately equal frequency (figure 3.15, M523L panel, green histogram). Analysis of the GFP channel confirmed 2 populations for protein expression as well; a low expressing population related to the high mitoxantrone accumulation, and a high expressing population corresponding to the low mitoxantrone accumulation (data not shown). For M541Q the population split is ~10:90 in terms of high:low mitoxantrone accumulation. To exclude the non-transporting, low expressing cells from further analysis, another gate for GFP fluorescence was added to the analysis for all ABCG2

isoforms and the wild type. This final gating for high GFP expression was done by creating a dot plot of each mutant isoform along with the wild type protein, with the axes representing GFP and mitoxantrone fluorescence using the FL1 and FL4 channels respectively, and creating an elliptical gate around cell populations that showed clear mitoxantrone efflux in the absence of Ko143 and clear reduction/abrogation of efflux in the presence of Ko143. The median value for the gated populations was used (labelled as C for the wild type protein and D for the mutant isoform), an example of this final gating can be seen in figure 3.17.

The relative fold change in sfGFP expression of the mutant isoforms compared to the wild type protein is given as a ratio as seen below.

$$\frac{C}{D} = fold\ change\ in\ sfGFP\ (ABCG2)\ expression$$

From here, the raw mitoxantrone efflux values are normalised to the fold change in sfGFP expression for each mutant isoform.

$$\frac{C}{D} \times \frac{\beta - \alpha}{\alpha} = relative\ mitoxantrone\ efflux\ normalised\ to\ sfGFP\ expression$$

The relative mitoxantrone efflux values for all mutant isoforms and the wild type that have been normalised to sfGFP expression were then put through a one-way ANOVA to determine if there was any significant difference in mitoxantrone efflux for all ABCG2 isoforms compared to the wild type. The relative drug efflux for the methionine mutants as well as the phenylalanine mutants compared to the wild type and untransfected HEK 293 T can be seen in figures 3.18 and 3.19.

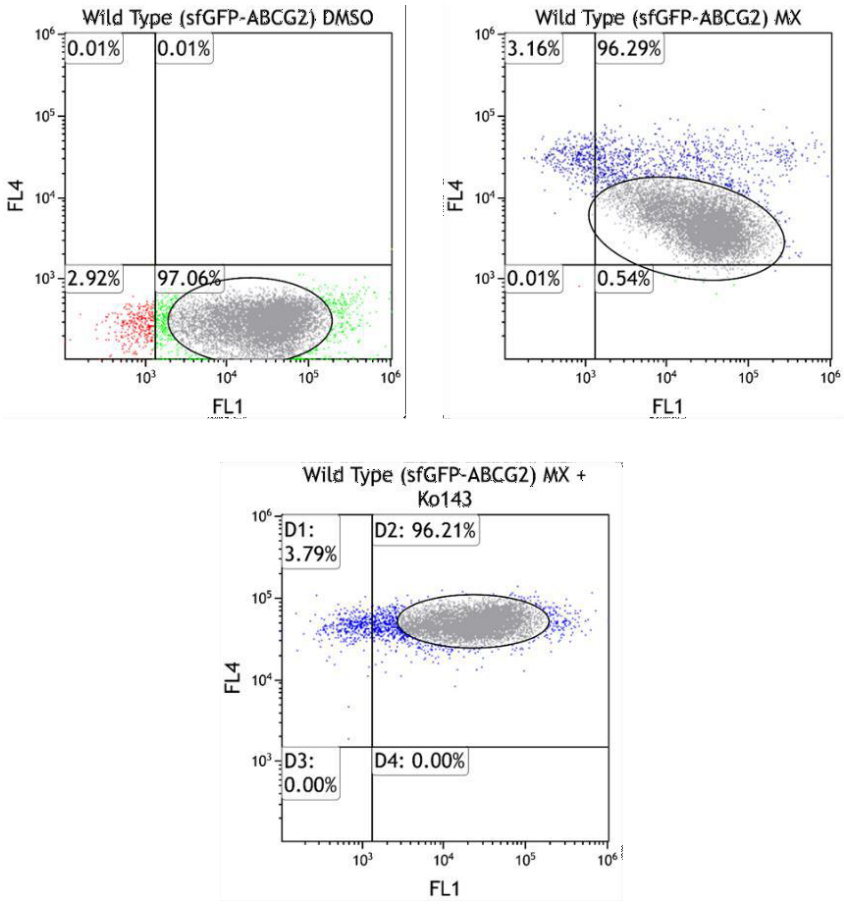


Figure 3.17: Manual gating for GFP expression on a mixed population. In many of the ABCG2 isoforms and the wild type isoform, a mixed population of cells with varying GFP (protein) expression as well as varying drug efflux capability. In this dot plot, each dot represents a single cell, with a total number of cells measured at 10,000. Each axis is representative of the fluorescence detected by the FL4 channel (representative of mitoxantrone fluorescence) and the FL1 channel (representative of GFP fluorescence), with each quadrant representative of the presence/absence of fluorescence for each channel. D1 is mitoxantrone positive and GFP negative, D2 is mitoxantrone positive and GFP positive, D3 is mitoxantrone negative and GFP negative and D4 is mitoxantrone negative and GFP positive. Gating for populations that show clear efflux of mitoxantrone and clear ablation of efflux in the presence of Ko143 was performed on isoforms that showed clear population splits from the initial single parameter histogram for mitoxantrone efflux (see figure 3.14). From here, the median GFP fluorescence values (from the FL1 channel) were obtained and used for further calculations.

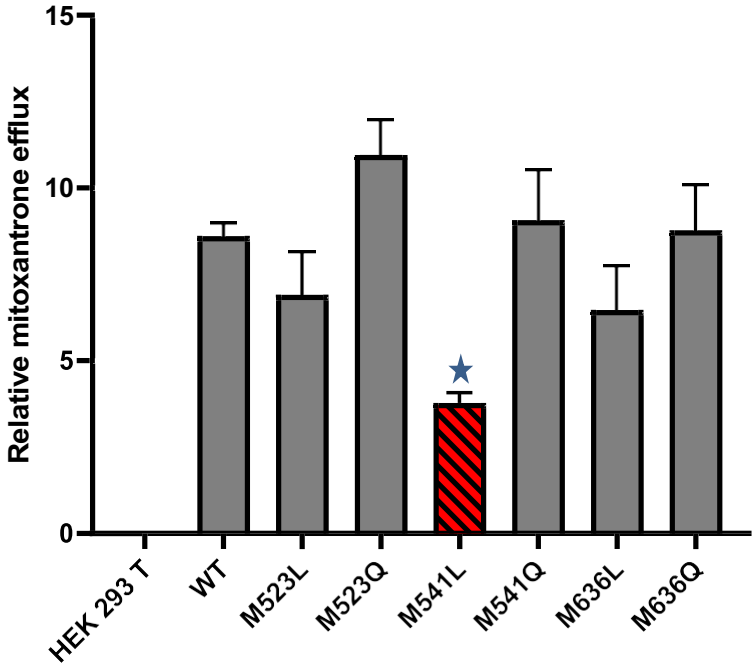


Figure 3.18: Relative mitoxantrone efflux profiles of all methionine mutants normalised to GFP expression. The raw drug efflux values were normalised to the GFP expression levels of each ABCG2 isoform and subsequently compared to the wild type (WT) and untransfected HEK 293 T using a one-way ANOVA. Red striped bars represent relative mitoxantrone efflux that is significantly different from the wild type. Data are representative of N=3 with a significant difference of $p \leq 0.05$ (denoted by the star above the histogram), error bars are representative of SEM.

As seen in figure 3.18, 1 out of the 6 methionine mutants had a significant difference in mitoxantrone efflux compared to the wild type. M451L shows a significantly reduced efflux capacity for mitoxantrone compared to the wild type (3.78, ± 0.30), which is a 2.3-fold decrease in relative efflux capacity for mitoxantrone, as seen in Table 3.2.

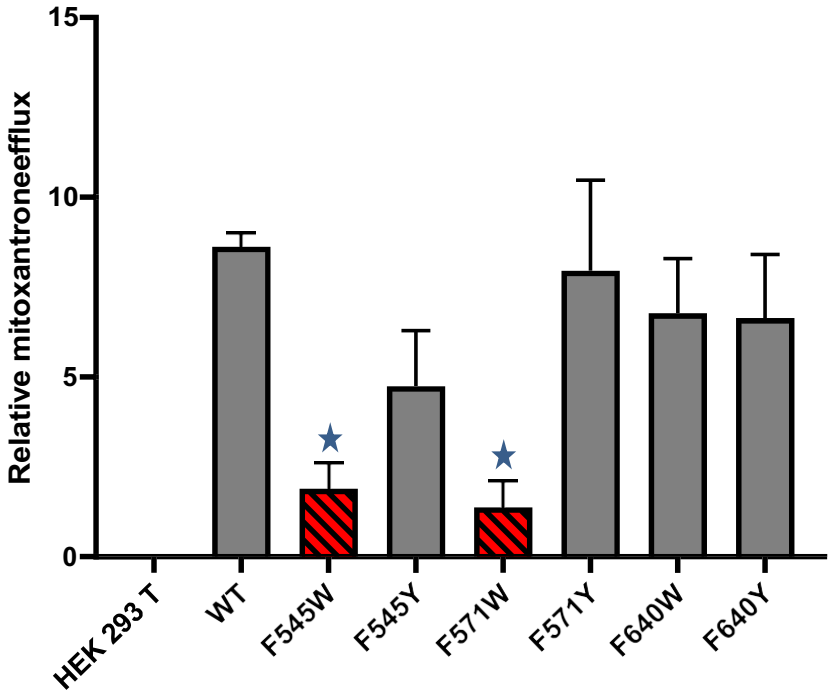


Figure 3.19: Relative mitoxantrone efflux profiles of all phenylalanine mutants normalised to GFP expression. The raw drug efflux values were normalised to the GFP expression levels of each ABCG2 isoform and subsequently compared to the wild type (WT) and untransfected HEK 293 T using a one-way ANOVA. Red striped bars represent relative mitoxantrone efflux that is significantly different from the wild type. Data are representative of N=3 with a significant difference of $p \leq 0.05$ (denoted by the star above the histogram), error bars are representative of SEM.

In figure 3.19, it can be seen that 2 out of the 6 phenylalanine mutants showed a significant difference in relative mitoxantrone efflux compared to the wild type. Both F545W and F571W show reduced efflux capacity for mitoxantrone (1.89, ± 0.71; 0.71, ± 0.25) as seen in Table 3.3. This is a 4.6-fold and 12-fold decrease in relative efflux capacity respectively. Thus it can be seen that 3 out of 12 mutants created showed reduced efflux capacity for mitoxantrone compared to the wild type.

Table 3.2: Relative drug efflux data for methionine mutants.

	MEAN	SEM
HEK 293 T	0.003	0.0005
WT	8.66	0.32
M523L	6.92	1.22
M523Q	10.96	1.00
M541L	3.78*	0.30
M541Q	9.08	1.45
M636L	6.47	1.26
M636Q	8.78	1.31

* Denotes p value ≤ 0.05.

Table 3.3: Relative drug efflux data for phenylalanine mutants.

	MEAN	SEM
HEK 293 T	0.004	0.0004
WT	8.62	0.37
F545W	1.89*	0.71
F545Y	4.75	1.54
F571W	0.71*	0.25
F571Y	7.95	2.52
F640W	6.77	1.52
F640Y	6.64	1.77

* Denotes p value ≤ 0.05.

3.9 Summary

This chapter aimed to identify potential methionine-aromatic interactions present at the surface site proposed by Cox et al. (2018) to be the site for allocrite recognition, and create single residue mutations that would help identify if these proposed methionine-aromatic interactions affect mitoxantrone efflux.

- Six methionine and six phenylalanine single point mutants were constructed based on the residues identified by Cox et al. (2018), and protein expression, protein localisation and protein function characterisation was performed for each mutant compared to the wild type.
- All mutants were confirmed to express full-length sfGFP-ABCG2 with the intended mutations at the correct residues.
- All mutants had varying expression levels compared to the wild type protein, and all mutants were successfully trafficked to the plasma membrane. Both F640 mutants were also observed to have pronounced protein-lined vesicles present intracellularly.
- One methionine mutant (M541Q) and two phenylalanine mutants (F545W and F571W) showed reduced relative mitoxantrone efflux capacity compared to the wild type protein.

Chapter 4: Discussion

Throughout the course of this study I had set out to look at whether the proposed methionine-aromatic interactions that are present at the surface site on ABCG2 played a role in mitoxantrone efflux. This is a part of a larger attempt at understanding the intramolecular interactions at the surface site that could be involved in allocrite recognition and efflux, and through the use of a mutagenesis study using more nuanced residue substitutions, it can be seen that some key residues at the surface site that are proposed to have a methionine-aromatic interaction have an effect in mitoxantrone efflux.

This discussion chapter will look at the results obtained from this study and compare these results with the results obtained by Dr Cox in her previous studies. It will also tie in these current studies conducted by other groups looking at ACBG2/ABC transporter multidrug binding sites and the mechanisms behind the recognition of substrates.

As seen in the previous chapter, all mutant isoforms were able to express full-length sfGFP-ABCG2 with varying expression levels and all mutant isoforms successfully trafficked protein to the plasma membrane, although the F640 mutants also showed pronounced vesicles lined with sfGFP-ABCG2 protein. Out of the 12 mutant isoforms constructed, one methionine mutants (M541L) and two phenylalanine mutants (F545W and F571W) showed a significant decrease in mitoxantrone efflux compared to the wild type protein after raw efflux was normalised to sfGFP-ABCG2 expression level. When comparing these results to the results obtained by Dr Cox with residue substitutions to alanine, it appears that there are differing mitoxantrone efflux profiles for some of the mutant constructs.

In Dr Cox's many mutagenesis studies, the M541A mutation did show a decrease in mitoxantrone efflux compared to the wild type, with the mutant having approximately half the relative efflux capacity for mitoxantrone compared to the wild type (M541A 1.58 ± 0.33; WT 3.51 ± 0.60). For the M541L isoform, there is also an approximately half-fold decrease in mitoxantrone efflux compared to the wild type protein, thus this result obtained in this study is fairly similar to the result obtained by Dr Cox. For the F545A isoform in her study, it is noted that it had no significant change in mitoxantrone efflux compared to the wild type (F545A 2.81 ± 0.67; WT 3.51 ± 0.60), while the F545W isoform has a significant reduction in mitoxantrone efflux with a 4.6-fold decrease in relative mitoxantrone efflux capacity. The F571A isoform had a slight decrease in mitoxantrone efflux (F571A 2.31 ± 0.3; WT 3.51 ± 0.60), while the F571W isoform showed the most significant reduction in mitoxantrone efflux capacity with a whopping 12-fold decrease in relative efflux compared to the wild type.

4.1 Methionine at Position 541 and Phenylalanine at position 545 Share a Methionine-Aromatic Interaction that is Key to Allocrite Recognition

It was observed that both M541L and F545W showed a pronounced reduction in relative mitoxantrone efflux capacity, and both M541 and F545 do show a methionine-aromatic interaction in the *in silico* parsing for these interactions in cryo-EM structures showing both the closed conformation apo-state and the open conformation ATP-bound state (the interaction in the closed- conformation apo-state can be seen in figure 4.1).

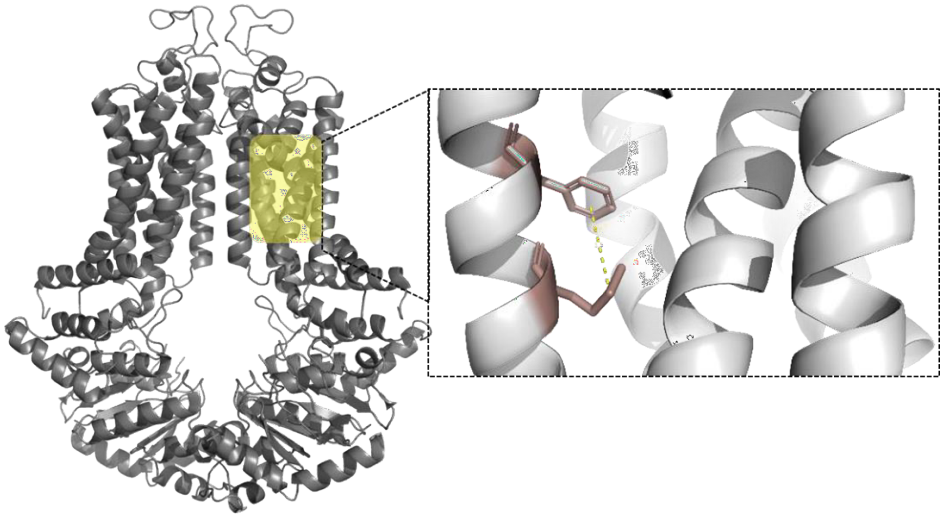


Figure 4.1: Visualisation of the Methionine-Aromatic Interaction between F545 and M541 in the closed-conformation apo-state of ABCG2. Both residues (labelled dark salmon) were observed to share an interaction between each other (yellow dashed lines) at the surface site (represented as the yellow box). This, combined with the mutagenesis and drug accumulation assay results suggest these residues as being key to the recognition of mitoxantrone. Visualisation was done with PyMOL using the structure 6ETI.

The difference in relative mitoxantrone efflux for F545A from Dr Cox's findings compared to the results obtained in this study appear to show that F545 may be a key residue in this interaction, as the M541A and M541L mutations both share similar results. The indole ring in the tryptophan side chain does have a stronger binding energy compared the benzene ring in the phenylalanine side chain as observed in a quantum calculation study performed by Valley et al. (2012). In this study it was shown that the binding energy of dimethyl sulphate (DMS; as a stand in for the methionine thioether group) with an indole ring was significantly stronger (-4.50 – -6.57 kcal/mol) compared to the binding energy of DMS with a benzene ring (-2.94 – -4.13 kcal/mol). From here it can be deduced that the substitution of the

phenylalanine residue at position 545 with a tryptophan would lead to a stronger interaction forming with M541. This could lead to a reduction in flexibility, reducing the capacity for mitoxantrone to fit into the surface site prior to it translocating to the internal cavity for efflux. The added steric bulk from the tryptophan substitution could also lead to steric clashing with the equally as bulky polycyclic mitoxantrone molecule, reducing its ability to bind with the preferred orientation to the surface site. This would then lead to a reduced efflux capacity as the mitoxantrone molecule cannot effectively bind to the surface site and pass beyond this site to the internal cavity.

One way to confirm whether these two residues and the proposed methionine-aromatic interaction between them could be part of allocrite recognition would be conducting the same drug accumulation assay with an M541L and F545W double mutant construct, and comparing the relative efflux of this construct with the single residue mutations along with the wild type.

4.2 Tryptophan Substitution of Phenylalanine at Position 571 Introduces an Additional Methionine-Aromatic Interaction with a Nearby Methionine Residue

The relative mitoxantrone efflux for both the F571W mutant from this study and the F571A mutant from Dr Cox's study is significantly reduced compared to the wild type. However, the F571W mutant produces the most significant reduction in relative mitoxantrone efflux with a 12-fold reduction in mitoxantrone efflux capacity. From the initial *in silico* parsing F571 did not appear to form any methionine-aromatic interaction with the nearest methionine residue (M636) neither F571 nor M636 meet the distance parameter set by Valley et al. to form a methionine-aromatic interaction,as the distance

between the two residues is beyond 6.0Å. However, the indole ring of the tryptophan residue substitution is believed to allow the thioether group in M636 to form a methionine-aromatic interaction as the added length of the side chain with the indole ring may allow for the benzene ring present in the side chain to reach the thioether group of M636. This would then lead to an additional methionine-aromatic interaction present at this site, and with the methionine-tryptophan interaction having a stronger binding energy as previously described, this is thought to create a more rigid surface site. This is would then lead to a reduced capacity for mitoxantrone to bind to this surface site, thus preventing it from translocation into the internal cavity for efflux. As with the F545 substitution with tryptophan, it is also believed that the added steric bulk from the indole ring would sterically clash with mitoxantrone and prevent it from properly binding at the correct orientation at the surface site.

In order to confirm whether this residue is key in allocrite recognition would be to perform a drug accumulation assay with the F571W mutant using a native substrate such as pheophorbide A and a non-native substrate such as daunorubicin, as it was noted in Dr Cox's findings that although the F571A mutant did not show any change in relative efflux of both of these substrates compared to the wild type it would be interesting to know if this mutation does confer the ability to efflux a non-native substrate and/or affect the ability to efflux another native substrate.

4.3 Methionine-Aromatic Interactions Play a Role in Allocrite Recognition through Binding Site Stabilisation

The methionine-aromatic interaction present between M541 and F545 is believed to be key in providing stability to the putative binding pocket at the surface site. This comes from the discovery that methionine-aromatic interactions do play a role in stabilising not only the overall protein structure during protein folding, but also providing stability in bindings sites for metal ion co-factors. In a study conducted by Weber and Warren (2018), it is believed that the presence of methionine-aromatic interactions in a protein structure provides protein stabilisation during protein folding as well as providing stabilisation of co-factors in metalloenzymes. This was discovered through high throughput parsing of PDB entries of metalloenzymes and observing the large presence of methionine-aromatic interactions at or near the binding site for an ion co-factor such as Zn^{2+}. From here it can be de- duced that a disruption of the methionine-aromatic interaction at this co- factor binding sites will lead to perturbed metalloenzyme activity which re- quires the co-factor to participate in the enzymatic reaction. Therefore, it can be deduced that a disruption to this methionine-aromatic interaction between M541-F545 would lead to destabilisation of the putative binding site at the surface site of ABCG2, and lead to a reduced capacity for drug binding prior to its translocation to the internal cavity.

Future work to confirm this conclusion would, as stated before, to generate a double M541L-F545W mutant and perform a drug transport assay with mitoxantrone along with another native substrate such as phophorbide A and a non-native substrate such as daunorubicin. The ability for this double mutant to transport either drug will give a clearer answer on this key interaction between the two residues and their effects on allocrite recognition and transport.

It was interesting to note that throughout this study, a few generated stable cell lines showed pronounced heterogeneity. Ideally a homogenous population would allow for a better representation of how a mutant construct would affect drug efflux as the entire population would have a similar protein expression level and thus a consistent drug efflux profile. However it was noted that some of the mutant isoforms as well as the wild type sfGFP- ABCG2 had cell populations with varying levels of protein expression, which resulted in multiple drug efflux profiles detected in the total population as seen in the previous chapter. This heterogeneity in some stable cell lines had previously been observed by members of the Kerr Lab, and an attempt to grow high expressing cells after fluorescence-assisted cell sorting (FACS) was performed by Parth Kapoor from the Kerr Lab for the wild type sfGFP- ABCG2 stable cell line. After initial cell sorting, the whole cell population was observed to be expressing protein at a consistent level via epifluorescence microscopy, however after several passages, the cell population began reverting into a mixed population (data no shown). It is thought that this heterogeneity could be due to the concentration of zeocin used to maintain the stable cell line.

Thus a preliminary study was conducted to determine whether increasing the concentration of zeocin throughout stable cell line maintenance reduced this heterogeneity in the stable cell lines.

4.4 Maintenance of Stable Cell Lines in Higher Concentrationsof Zeocin May Reduce Heterogeneity

In this short study two stable cell lines (wild type sfGFP-ABCG2 and F640Y) were grown in varying concentrations of zeocin (40µg/ml, 50µg/ml, 80µg/ml, 100µg/ml) for 7 days, with a medium change every 2-3 days.

These two cell lines were chosen as the wild type expressing cells showed a clear population split comprising of low expressors and high expressors as observed during flow cytometry data acquisition, while F640Y was chosen as it was one of a select few mutant constructs that showed consistent homogeneity during cell passage maintenance throughout the course of this study, and as such was used as a "control" for homogeneity in a cell population.

Throughout the 7 days of stable cell maintenance, it was observed that the cells growing at concentrations beyond 50µg/ml appeared to be going through a second round of selection as some cells began detaching from the flasks. When these flasks were observed under an epifluorescence microscope, it can be seen that the cells that have detached were cells that did not express any sfGFP-ABCG2 isoform, and thus did not carry and express the plasmid construct. It was also observed that the wild type isoform showed a slight increase in fluorescence when grown in concentrations beyond 50µg/ml when observed under widefield epifluorescence microscopy after 5 days of maintenance, while F640Y cell fluorescence stayed roughly the same throughout (figure 4.2). None of the cell lines had an altered morphology throughout this experiment.

After the 7-day period, cells were harvested and subjected to SDS-PAGE and western blotting to determine if protein expression levels were affected by maintenance at varying concentrations of zeocin (data not shown). It was observed that an increase in zeocin concentration did slightly affect the protein expression level, although this can only be seen in the F640Y isoform, which contradicts what was observed during epifluorescence microscopy image acquisition.

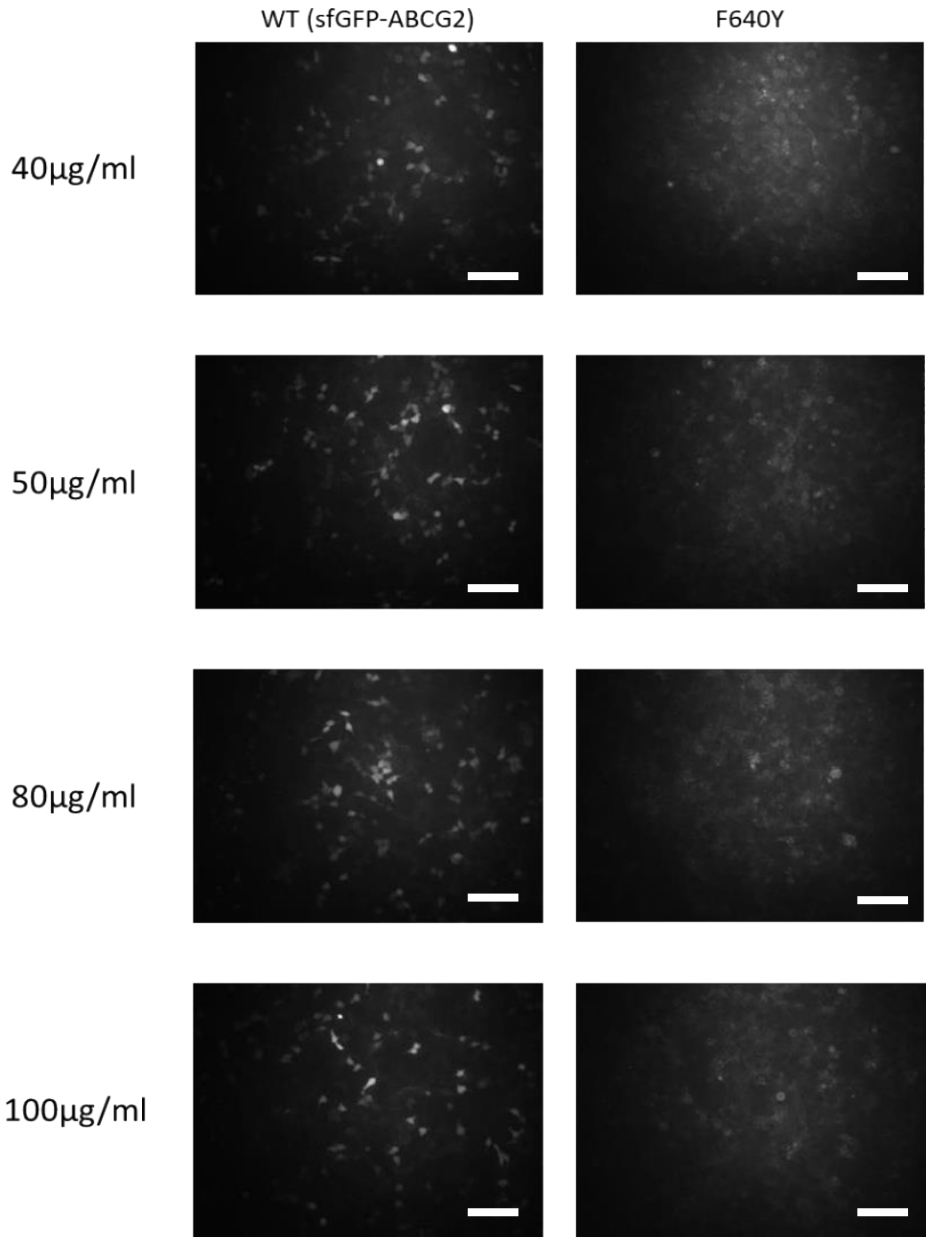


Figure 4.2: Epifluorescence microscopy images of stable cell lines grown in varying concentrations of zeocin. Images were obtained after 5 days of cell maintenance. F640Y showed no change in cell morphology nor fluorescence intensity when observed under a widefield epifluorescence microscope. The wild type sfGFP-ABCG2 cells did show a slight increase in cell fluorescence (an increase in the number of brightly fluorescing cells can be observed), and did not show any altered morphology throughout the short experiment. Cells were imaged using the Zeiss Axiovert S100 inverted microscope using a mercury arc-lamp for transmitted light imaging and widefield fluorescence imaging. Scale bars are representative of 200µm for each image.

However, this is merely a qualitative observation as the widefield epifluorescence microscope is not able to determine the fluorescence intensity on a cell-by-cell basis, and western blotting only shows the protein expression levels of the whole cell population. Therefore a more qualitative and sensitive measure of protein expression must be done in order to confirm whether an increase in zeocin concentration used for stable cell line maintenance could promote homogeneity in a stable cell line population. An attempt was done using a confocal plate reader, however since no nuclear staining was performed, the fluorescence intensity was not able to be determined on a cell-by-cell basis (data not shown).

For future work, both confocal image acquisition of fixed and live cells stained with a nucleic stain such as DAPI and a surface expression assay using flow cytometry can be performed on stable cells growing in these varying zeocin concentrations to determine whether increasing concentrations do affect the fluorescence intensity of the cell population on a cell-by-cell basis, and thus whether an increase in zeocin concentration during stable cell line maintenance truly does help reduce heterogeneity in a stable cell line.

References

Annilo, T., Tammur, J., Hutchinson, A., Rzhetsky, A., Dean, M., Allikmets, R., 2001. Human and mouse orthologs of a new ATP-binding cassette gene, ABCG4. Cytogenet. Genome Res. 94, 196–201. https://doi.org/10.1159/000048816

Bates, S.E., Robey, R., Miyake, K., Rao, K., Ross, D.D., Litman, T., 2001. The role of half-transporters in multidrug resistance. J. Bioenerg. Biomembr. 33, 503–511. https://doi.org/10.1023/A:1012879205914

Berger, E.A., Heppel, L.A., 1974. Different mechanisms of energy coupling for the shock-sensitive and shock-resistant amino acid permeases of Escherichia coli. J. Biol. Chem. 249, 7747–55.

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. 92, 7297–7301. https://doi.org/10.1073/pnas.92.16.7297

Clark, R., Kerr, I.D., Callaghan, R., 2006. Multiple drugbinding sites on the R482G isoform of the ABCG2 transporter. Br. J. Pharmacol. 149, 506–515. https://doi.org/10.1038/sj.bjp.0706904

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. 475, 1553–1567. https://doi.org/10.1042/BCJ20170923

Eddy, S.R., 2011. Accelerated Profile HMM Searches. PLoS Comput. Biol. 7, e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Ellgaard, L., McCaul, N., Chatsisvili, A., Braakman, I., 2016. Co- and Post-

Translational Protein Folding in the ER. Traffic 17, 615–638. https://doi.org/10.1111/tra.12392

Fetsch, P.A., Abati, A., Litman, T., Morisaki, K., Honjo, Y., Mittal, K., Bates, S.E., 2006. Localization of the ABCG2 mitoxantrone resistance- associated protein in normal tissues. Cancer Lett. 235, 84–92. https://doi.org/10.1016/j.canlet.2005.04.024

Gillet, J.-P., Gottesman, M.M., 2010. Mechanisms of Multidrug Resistance in Cancer, in: Methods in Molecular Biology (Clifton, N.J.). pp. 47–76. https://doi.org/10.1007/978-1-60761-416-6_4

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, e00241–e00241. https://doi.org/10.1042/BSR20150150

Henikoff, S., Henikoff, J.G., 1992. Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. 89, 10915–10919. https://doi.org/10.1073/pnas.89.22.10915

Hollenstein, K., Dawson, R.J., Locher, K.P., 2007. Structure and mechanism of ABC transporter proteins. Curr. Opin. Struct. Biol. 17, 412–418. https://doi.org/10.1016/j.sbi.2007.07.003

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. https://doi.org/10.1038/s41594-018-0049-1

Laemmli, U.K., 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature 227, 680–685. https://doi.org/10.1038/227680a0

Lage, H., 2008. An overview of cancer multidrug resistance: A still unsolved problem. Cell. Mol. Life Sci. 65, 3145–3167. https://doi.org/10.1007/s00018-008-8111-5

László, L., Sarkadi, B., Hegedűs, T., 2016. Jump into a New Fold—A Homology Based Model for the ABCG2/BCRP Multidrug Transporter. PLoS One 11, e0164426. https://doi.org/10.1371/journal.pone.0164426

Lee, J.Y., Kinch, L.N., Borek, D.M., Wang, Jin, Wang, Junmei, 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. https://doi.org/10.1038/nature17666

Lehne, G., 2005. P-glycoprotein as a Drug Target in the Treatment of Multidrug Resistant Cancer. Curr. Drug Targets 1, 85–99. https://doi.org/10.2174/1389450003349443

Lewis, A.K., Dunleavy, K.M., Senkow, T.L., Her, C., Horn, B.T., Jersett, M.A., Mahling, R., McCarthy, M.R., Perell, G.T., Valley, C.C., Karim, C.B., Gao, J., Pomerantz, W.C.K., Thomas, D.D., Cembran, A., Hinderliter, A., Sachs, J.N., 2016. Oxidation increases the strength of the methionine-aromatic interaction. Nat. Chem. Biol. 12, 860–866. https://doi.org/10.1038/nchembio.2159

Manolaridis, I., Jackson, S.M., Taylor, N.M.I., Kowal, J., Stahlberg, H., Locher, K.P., 2018. Cryo-EM structures of a human ABCG2 mutant trapped in ATP-bound and substrate-bound states. Nature 563, 426–

430. https://doi.org/10.1038/s41586-018-0680-3

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. https://doi.org/10.1111/j.1742-4658.2008.06578.x

Mellor, H.R., Callaghan, R., 2008. Resistance to chemotherapy in cancer: A complex and integrated cellular response. Pharmacology 81, 275–300. https://doi.org/10.1159/000115967

Nakanishi, T., Ross, D.D., 2012. Breast cancer resistance protein (BCRP/ABCG2): Its role in multidrug resistance and regulation of its gene expression. Chin. J. Cancer 31, 73–99. https://doi.org/10.5732/cjc.011.10320

Özvegy-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 - Biomembr. 1668, 53–63. https://doi.org/10.1016/j.bbamem.2004.11.005

Pédelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., Waldo, G.S., 2006. Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. 24, 79–88. https://doi.org/10.1038/nbt1172

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. https://doi.org/10.1038/sj.bjc.6601370

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory

Manual, 3th edn, Cold Spring Harbor Laboratory, New York,. Cold Spring, Harb. Lab. New York, 3th edn. https://doi.org/10.1128/AEM.71.8.4602

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, 504–509. https://doi.org/10.1038/nature22345

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. 76, 4350–4354. https://doi.org/10.1073/pnas.76.9.4350

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. https://doi.org/10.1074/jbc.M112.374504

Velamakanni, S., Wei, S.L., Janvilisri, T., Van Veen, H.W., 2007. ABCG transporters: Structure, substrate specificities and physiological roles - A brief overview. J. Bioenerg. Biomembr. 39, 465–471. https://doi.org/10.1007/s10863-007-9122-x

Walgenbach, D.G., Gregory, A.J., Klein, J.C., 2018. Unique methionine-aromatic interactions govern the calmodulin redox sensor. Biochem. Biophys. Res. Commun. 505, 236–241. https://doi.org/10.1016/j.bbrc.2018.09.052

Weber, D.S., Warren, J.J., 2019. The interaction between methionine and two aromatic amino acids is an abundant and multifunctional motif in proteins. Arch. Biochem. Biophys. 672, 108053. https://doi.org/10.1016/j.abb.2019.07.018

Weber, D.S., Warren, J.J., 2018. A survey of methionine-aromatic interaction geometries in the oxidoreductase class of enzymes: What could Met-aromatic interactions be doing near metal sites? J. Inorg. Biochem. 186, 34–41. https://doi.org/10.1016/j.jinorgbio.2018.05.008

Wilkens, S., 2015. Structure and mechanism of ABC transporters. F1000Prime Rep. 7, 1–9. https://doi.org/10.12703/P7-14

Wong, K., Briddon, S.J., Holliday, N.D., Kerr, I.D., 2016a. Plasma membrane dynamics and tetrameric organisation of ABCG2 transporters in mammalian cells revealed by single particle imaging techniques. Biochim. Biophys. Acta - Mol. Cell Res. 1863, 19–29. https://doi.org/10.1016/j.bbamcr.2015.10.002

Wong, K., Briddon, S.J., Holliday, N.D., Kerr, I.D., 2016b. Plasma membrane dynamics and tetrameric organisation of ABCG2 transporters in mammalian cells revealed by single particle imaging techniques. Biochim. Biophys. Acta - Mol. Cell Res. 1863, 19–29. https://doi.org/10.1016/j.bbamcr.2015.10.002

Zacharias, D.A., Violin, J.A., Newton, A.C., Tsien, R.Y., 2002. Partitioning of Lipid-Modified Monomeric GFPs into Membrane Microdomains of Live Cells. Science (80-.). 296, 913–916. https://doi.org/10.1126/science.1068539

Zauhar, R.J., Colbert, C.L., Morgan, R.S., Welsh, W.J., 2000. Evidence for a strong sulfur-aromatic interaction derived from crystallographic data. Biopolymers 53, 233–248. https://doi.org/10.1002/(SICI)1097-0282(200003)53:3<233::AID-BIP3>3.0.CO;2-4

Appendix

Appendix I

An example of the python script used is shown below. Do note that this is based on the basic workflow outlined by Weber and Warren (2019).

```
PyMOL>fetch 6eti
     please wait ...
ExecutiveLoad-Detail: Detected mmCIF
CmdLoad: loaded as "6eti".

PyMOL>set seq_view_format, 3
     Setting: seq_view_format set to 3.
PyMOL>set seq_view, on
     Setting: seq_view set to on.
PyMOL>color gray90
     Executive: Colored 9026 atoms.
PyMOL>select Phe545, resi 545
     Selector: selection "Phe545" defined with 22
atoms.
PyMOL>select Phe571, resi 571
     Selector: selection "Phe571" defined with 22
atoms.
PyMOL>select Phe640, resi 640
     Selector: selection "Phe640" defined with 22
atoms.
PyMOL>select Met541, resi 541
     Selector: selection "Met541" defined with 16
atoms.
PyMOL>select Met523, resi 523
     Selector: selection "Met523" defined with 16
atoms.
PyMOL>select Met636, resi 636
     Selector: selection "Met636" defined with 16
atoms.

PyMOL>show sticks, resi 545
PyMOL>show sticks, resi 571
PyMOL>show sticks, resi 640
PyMOL>show sticks, resi 541
PyMOL>show sticks, resi 523
PyMOL>show sticks, resi 636

PyMOL>wizard measurement
You clicked /6eti/A/A/MET`541/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`439/CA
You clicked /6eti/A/A/MET`541/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`545/CG
You clicked /6eti/A/A/MET`541/SD
     Selector: selection "sele" defined with 1 atoms.
```

```
You clicked /6eti/A/A/MET`523/CE
You clicked /6eti/A/A/MET`541/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`439/CA
You clicked /6eti/A/A/MET`541/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`545/CZ
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/MET`523/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`640/CE1
You clicked /6eti/A/A/MET`636/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`640/CG
You clicked /6eti/A/A/MET`636/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`571/CE1
You clicked /6eti/A/A/MET`636/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`571/CZ

PyMOL>wizard measurement
You clicked /6eti/A/A/MET`523/SD
     Selector: selection "sele" defined with 1 atoms.
You clicked /6eti/A/A/PHE`545/CE1
```