10-2023

MOLECULAR AND BIOCHEMICAL ANALYSIS OF CYTOCHROME P450 2C19 AND 2D6

Reema Saleous
United Arab Emirates University

UAEU
جامعة الإمارات العربية المتحدة
United Arab Emirates University

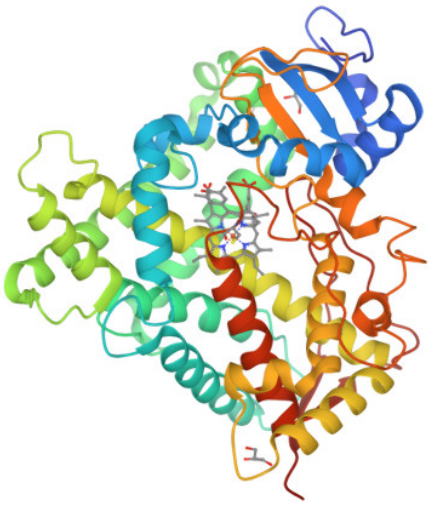
MASTER THESIS NO. 2023: 128

College of Medicine and Health Sciences

Department of Biochemistry and Molecular Biology

MOLECULAR AND BIOCHEMICAL ANALYSIS OF CYTOCHROME P450 2C19 AND 2D6

Reema Saleous



October 2023

United Arab Emirates University

College of Medicine and Health Sciences

Department of Biochemistry and Molecular Biology

MOLECULAR AND BIOCHEMICAL ANALYSIS OF CYTOCHROME P450 2C19 AND 2D6

Reema Saleous

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Medical Sciences (Biochemistry and Molecular Biology)

October 2023

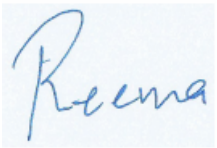
United Arab Emirates University Master Thesis
2023: 128

Cover: X-ray diffraction image of CYP2C19
(Photo: Reynald et al., 2012)

Print: University Print Service, UAEU 2023

Declaration of Original Work

I, Reema Saleous, the undersigned, a graduate student at the United Arab Emirates University (UAEU), and the author of this thesis entitled "*Molecular and Biochemical Analysis of Cytochrome P450 2C19 and 2D6*", hereby, solemnly declare that this is the original research work that has been done and prepared by me under the supervision of Prof. Bassam Ali, in the College of Medicine and Health Sciences at UAEU. This work has not previously formed the basis for the award of any academic degree, diploma or a similar title at this or any other university. Any materials borrowed from other sources (whether published or unpublished) and relied upon or included in my thesis have been properly cited and acknowledged in accordance with appropriate academic conventions. I further declare that there is no potential conflict of interest with respect to the research, data collection, authorship, presentation and/or publication of this thesis.

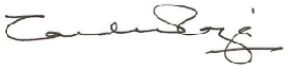
Student's Signature: 

Date: October 17, 2023

Approval of the Master Thesis

This Master Thesis is approved by the following Examining Committee Members:

1) Advisor (Committee Chair): Bassam R. Ali
   Title: Professor
   Department of Genetics & Genomics
   College of Medicine and Health Sciences
   Signature:  Date: October 17, 2023

2) Member: Haider Raza
   Title: Professor
   Department of Biochemistry and Molecular Biology
   College of Medicine and Health Sciences
   Signature:  Date: October 17, 2023

3) Member (External Examiner): Nathalie Khoueiry-Zgheib
   Title: Professor
   Department of Pharmacology and Toxicology
   Institution: American University of Beirut, Beirut, Lebanon
   Signature:  Date: October 17, 2023

This Master Thesis is accepted by:

Acting Dean of the College of Medicine and Health Sciences: Professor Ahmed Al-Suwaidi

Signature 

Date 22/12/2023

Dean of the College of Graduate Studies: Professor Ali Al-Marzouqi

Signature 

Date 22/12/2023

Abstract

Pharmacogenomics (PGx) is a relatively new field of study. It links genetics to pharmacology since it deals with the influence of the genetic makeup of the individual on their ability to respond to specific medications. Some of the most important genes in this field, dubbed very important pharmacogenes (VIPs), belong to the cytochrome P450 (CYP) superfamily of drug metabolizing enzymes. The two members of this family that are the main focus of this thesis are CYP2C19 and CYP2D6. They play major roles in the metabolism of numerous medications, and it is therefore imperative that variations within those genes in various populations are detected in an efficient and timely manner, as well as establishing their impact on the function of the encoded enzymes.

In this study, seven variants in *CYP2C19* that have been detected in the Emirati population were analysed to determine their potential effects on the enzyme's activity. Major aims of this thesis were to perform *in silico* analysis, generate the seven target variants in mammalian expression vectors, express them in mammalian cell lines, and measure the resulting enzyme's activities. The variants were successfully generated by site-directed mutagenesis using two expression vectors as templates and were expressed in COS and Hek293T cells. However, the enzymatic activity tests using lysates from those cells were inconclusive and, therefore, further analysis, perhaps using different kits, are needed to further establish the impact of those variants.

For CYP2D6, a long-range PCR-based technique was optimized and utilized to detect gene copy numbers using DNA extracted from blood samples isolated from psychiatric patients to determine their CYP2D6 metabolic status. In particular, this test was used to detect if the patients have a deletion star allele (*CYP2D6**5) or duplication of *CYP2D6*. The results indicate that this approach could be used to implement genetic testing to determine CYP2D6 copy numbers in patients requiring medication metabolized by this enzyme.

Keywords: Pharmacogenomics (PGx), cytochrome P450s (CYP), CYP2C19, CYP2D6, enzyme activity, personalized medicine.

Title and Abstract (in Arabic)

التحليل الجزيئي والكيميائي لإنزيمات السيتوكروم P450 و 2C19 و 2D6

الملخص

علم الأدوية الجينية (PGx) هو مجال دراسة جديد نسبيًا. يربط علم الأدوية بالوراثة لأنه يتعامل مع تأثير التركيب الجيني للفرد من حيث قدرته على الاستجابة لأدوية معينة. بعض من أهم الچينات في هذا المجال، والتي يطلق عليها اسم الچينات الدوائية المهمة للغاية (VIP) تنتمي إلى عائلة مستقبلات السيتوكروم P450 (CYP) من إنزيمات استقلاب الأدوية. العضوان الرئيسيان في هذه العائلة و اللذان تركز عليهم هذه الأطروحة هما CYP2C19 و CYP2D6. لديهما أدوارًا رئيسية في استقلاب العديد من الأدوية، ولذا فمن الضروري أن يتم اكتشاف التغيرات داخل هذه الچينات بطريقة فعالة وفي الوقت المناسب، وفهم تأثيرها على وظيفة الإنزيمات المشفرة.

في هذه الدراسة، تم تحليل سبعة متغيرات في CYP2C19 تم اكتشافها في السكان الإماراتيين لتحديد التأثير الذي لديها على نشاط الإنزيم. كان أحد الأهداف الرئيسية لهذا البحث هو إجراء تحليل حوسبي، والتعرف على سبعة متغيرات مستهدفة في نواقل الچينات، والكشف عنها في الخلايا الحيوانية، وقياس أنشطة الإنزيم الناتجة عنها. تم إحداث المتغيرات بنجاح عن طريق الطفرات الموجهة باستخدام نواقل چينية كنماذج وتم ظهورها في خلايا COS و Hek293T. ومع ذلك، كانت اختبارات النشاط الأنزيمي باستخدام lysates من تلك الخلايا غير قاطعة، وبالتالي يلزم إجراء مزيد من التحليلات، ربما باستخدام أدوات مختلفة.

بالنسبة إلى CYP2D6، تم تحسين تقنية PCR طويلة المدى واستخدامها لاكتشاف أعداد نسخ الجين باستخدام الحمض النووي المستخرج من عينات دم المريض لتحديد حالة الوضع الاستقلابي للإنزيم CYP2D6. تم استخدام هذا الاختبار بشكل خاص لتحديد ما إذا كان لدى المرضى أليل حذف نجمة (CYP2D6*5) أو تكرار CYP2D6. تشير النتائج إلى أن هذا النهج يمكن استخدامه لتطبيق الاختبارات الچينية لتحديد أعداد نسخ CYP2D6 في المرضى الذين يحتاجون إلى أدوية يتم استقلابها بواسطة هذا الإنزيم.

مفاهيم البحث الرئيسية: علم الأدوية الجينية (PGx)، إنزيمات السيتوكروم P450s (CYP)، CYP2C19، CYP2D6، نشاط الإنزيم، الطب الشخصي.

Acknowledgements

My thanks go to my advisor Dr. Bassam Ali for his guidance and assistance throughout my preparation of this thesis.

Special thanks go to my fellow lab members for their patience and support during this process.

Dedication

To my beloved parents. Thank you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADR	Adverse Drug Reactions
AKR	Aldo-Keto Reductases
BCA	Bicinchonic Acid
BSA	Bovine Serum Albumin
CNV	Copy Number Variations
CPIC	Clinical Pharmacogenetics Implementation Consortium
CYP	Cytochrome P450s
ER	Endoplasmic Reticulum
DMEM	Dulbecco's Modified Eagle Medium
FBS	Fetal Bovine Serum
MOE	Molecular Operating Environment
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PGx	Pharmacogenomics
Redox	Oxidation-Reduction
SDM	Site-directed Mutagenesis
UAE	United Arab Emirates
UV	Ultraviolet
VIP	Very Important Pharmacogenes
XL PCR	Extra-Long Polymerase Chain Reaction

Chapter 1: Introduction

1.1 Overview

Enzymes are very important proteins that catalyze and accelerate chemical reactions in cells of all living organisms. Without these proteins, essential chemical reactions would take very long periods of time and would therefore not be practical to sustain life. The importance of enzymology, or the study of enzymes, including the study of their structure and function, is illustrated in the effects of deficiencies or overexpression of some enzymes on living organisms, which often manifest in metabolic or other abnormalities and may lead to life-threatening diseases. There are seven main classes of enzymes in nature including oxidoreductases, transferases, hydrolases, lyases, isomerases, ligases, and translocases (Nelson et al., 2021). The cytochrome P450 superfamily of enzymes are classified as oxidoreductases. Therefore, the literature on this class will be reviewed and summarized in the following sections.

1.2 Oxidoreductases

Oxidoreductases are a large class of enzymes involved in oxidation-reduction (redox) reactions. They facilitate the transfer of electrons from a reductant, or electron donor, to an oxidant, or an electron acceptor. The types of oxidoreductases are separated based on the aspect of the redox reaction in which they play a role, as well as the reductants and/or oxidants involved. Oxidoreductases are classified into four groups: dehydrogenases, reductases, oxidases, and oxygenases (Frey & Hegeman, 2007). These groups will be reviewed briefly in the following subsections.

1.2.1 Dehydrogenases

As the name indicates, dehydrogenases are active in dehydrogenation reactions, where a hydrogen atom is removed from one molecule and added to another. There are two main types of dehydrogenases: aerobic and anaerobic dehydrogenases. Aerobic dehydrogenases are the ones where hydrogen is added to oxygen to create hydrogen peroxide as a byproduct. Anaerobic dehydrogenases are the ones where hydrogen is added to a coenzyme, such as NAD+, NADP+, or FAD (Vasudevan et al., 2013). While

these two types of dehydrogenases exist in nature, most tend to be of the anaerobic type (Nelson et al., 2021).

1.2.2 Reductases

Whereas dehydrogenases are responsible for the oxidation of the substrate, reductases are responsible for the reduction of the substrate. Aldo-keto reductases (AKRs) are an example of a superfamily of reductases involved in metabolizing many endogenous compounds, such as sugar and ketosteroids, and xenobiotics such as drugs and carcinogens (Penning & Drury, 2007).

1.2.3 Oxidases

Similar to aerobic dehydrogenases, oxidases catalyze the transfer of hydrogen to oxygen. The difference is that water is the byproduct of this reaction instead of hydrogen peroxide (Vasudevan et al., 2013). Urate oxidase is an example of an oxidase that catalyzes the transformation of uric acid to 5-hydroxyisourate and then to allantoin. It is absent in humans, as the levels of plasma urate are maintained by urine excretion. However, it is sometimes administered as a medication to treat elevated urate levels that, when elevated, can lead to cardiovascular diseases (Fetzner & Steiner, 2010).

1.2.4 Oxygenases

Unlike oxidases where oxygen is added to the byproduct, oxygenases add oxygen to the product (Frey & Hegeman, 2007; Nelson et al., 2021). There are two types of oxygenases: monooxygenases and dioxygenases. Monooxygenases, also called hydroxylases, add one oxygen to the product and the other forms water as the byproduct. On the other hand, dioxygenases add both oxygen atoms to the product (Vasudevan et al., 2013; Nelson et al., 2021). Some significant examples of monooxygenases are part of a superfamily of enzymes known as the cytochrome P450 group of metabolizing enzymes.

1.3 Cytochrome P450 enzymes

Cytochrome P450 enzymes (CYPs) are a superfamily of multi-function oxidases and monooxygenases. This group of proteins was discovered in 1958 by Garfinkel and

Klingenberg, and it was named a cytochrome because it contained a heme group that gave it redox capabilities. It was designated P450 because of its absorption band, while bound to carbon monoxide, was 450 nm (Garfinkel, 1958; Klingenberg, 1958). However, now it has been found that CYPs are heme-thiolate enzymes, as opposed to cytochromes, since the iron in the heme group has a bond with the sulfur of cysteine, and it was discovered that CYPs only transfer electrons to oxygen, not other atoms, like other cytochromes (Kahn & Durst, 2000; Lewis, 2001). There are 57 known human cytochrome P450s. Fifty of them are localized at the membrane of the endoplasmic reticulum (ER) and the remaining 7 are found in the mitochondria (Guengerich et al., 2016). Although these proteins are present in their highest concentrations in the liver, they are also expressed in other organs and tissues, such as the kidneys, pancreas, lungs, and brain (Lewis, 2001).

Each individual CYP is named using Arabic numbers to represent the family it belongs to, a letter to indicate the subfamily, and another Arabic number to distinguish the specific CYPs. To classify two proteins in the same family, they should have ≥40% amino acid similarity. To be in the same subfamily, the proteins should be within the same gene cluster and have ≥55% amino acid similarity. The last Arabic numeral is chosen in the discovery order (Nelson et al., 1996). As with other genes and proteins, when the name is italicized, it refers to the gene, while if it is not italicized, it refers to the protein (Manikandan & Nagini, 2018).

There are multiple terms used to classify CYPs. The most common way to differentiate between the different CYPs is by the role they play in detoxifying xenobiotics and those involved in the biosynthesis of endogenous compounds. Another way to refer to them is by defining them as moonlighting or non-moonlighting proteins. Moonlighting proteins are multifunctional, and these functions are typically not related or reliant on each other. Moonlighting CYPs are the ones involved in the biosynthesis of endogenous compounds. Non-moonlighting CYPs are the ones that detoxify xenobiotics (Manikandan & Nagini, 2018).

1.3.1 CYPs structure

Cytochrome P450s generally have similar structures made up of helices and folds. The majority of the folds are α-helical with a small number of β-sheets. They contain 400-500 amino acids and a heme prosthetic group located in the active site. The iron found in the heme group is in the ferric (Fe^{3+}) state rather than ferrous (Fe^{2+}). This form has two spin states: low spin and high spin. Low spin is when the 3d electrons are all paired, while high spin is when the 3d electrons are all unpaired. When the enzyme is substrate-free, and the iron is in low spin, a water molecule is used to stabilize it. Once a substrate needs to bind, the water molecule is removed (Manikandan & Nagini, 2018).

Figure 1.1 shows the structure of CYP2C19 based on the findings submitted to the Protein Data Bank by Reynald et al. (2012). While this report is specific to the structure of CYP2C19, other cytochrome P450s generally have similar structures.

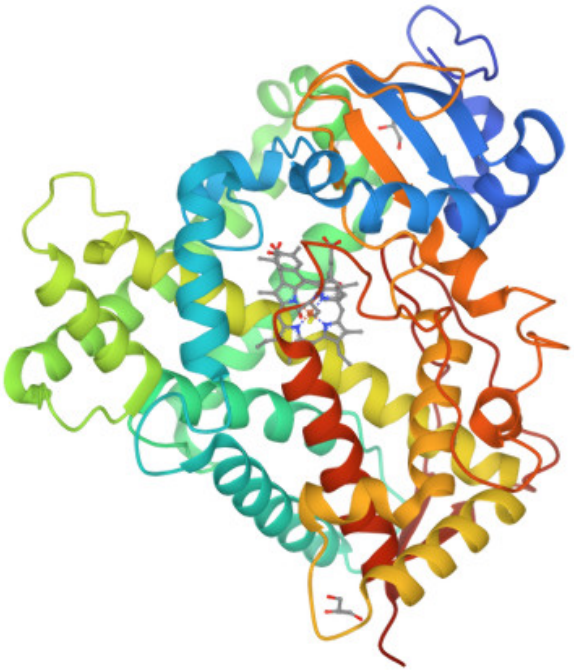


Figure 1.1: Structure of CYP2C19 from the Protein Data Bank.
The chains are each depicted in different colors starting from blue, which is the N-terminus, and ending at red, which is the C-terminus. The heme prosthetic group is represented by the stick representation in the center (Reynald et al., 2012).

1.3.2 CYPs function

Since cytochrome P450s are found throughout the body, they have many different functions. In the brain, CYPs help maintain cholesterol and vitamin A levels (Kuban & Daniel, 2020). In reproductive organs, on the other hand, CYPs synthesize and break down steroid hormones, while in the liver, CYPs play a role in bile acid biosynthesis and the metabolism of many xenobiotics including many drugs (Manikandan & Nagini, 2018).

The basic chemical reaction that occurs is that oxygen combines with the substrate to create a mono-oxygenated product and water. The coenzyme used in this reaction is typically NADH or NADPH. This reaction also requires redox partners, such as an iron-sulphur redoxin, a flavoprotein, or cytochrome b5 (Lewis, 2001).

Crucially, CYPs act in phase I reactions of drug metabolism. In contrast to phase II, phase I metabolism is non-synthetic; no new substance is created but a modified one that is easier for the body to excrete. This is usually done through oxidation, and to a lesser extent, through reduction or hydrolysis. Phase II metabolism, on the other hand, is synthetic. It involves combining the drug with an endogenous substance, such as sulfate or glycine, which leads to a polar product that can leave the body through the kidneys in urine or the liver in bile. While some drugs undergo phase I and II metabolisms, this is not the case in all, as some drugs only require one of the two phases to be excreted (Meyer, 1996). The two phases of drug metabolism are broadly illustrated in Figure 1.2.

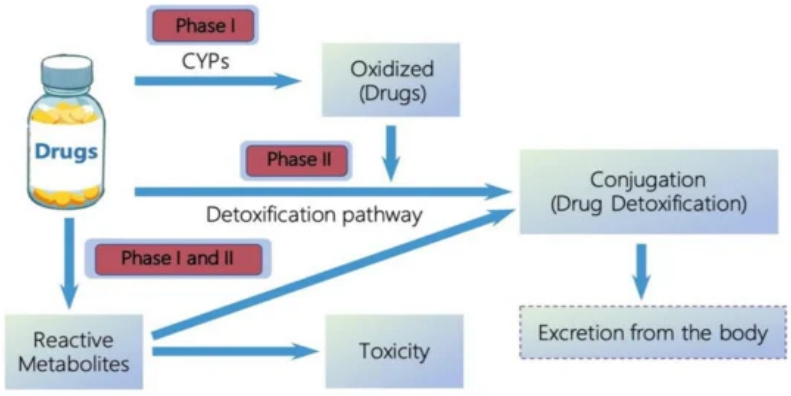


Figure 1.2: The general mechanism of phase I and phase II drug metabolism.
Phase I, commonly performed by CYPs, involves oxidizing the drug. Phase II detoxifies the drug, allowing it to exit the body (Zhao et al., 2021).

1.3.3 CYPs inhibition

One mechanism of CYPs inhibition is drug-drug interaction. If a patient takes more than one drug that can bind to the same enzyme, it will interfere with the enzyme's ability to metabolize one of the drugs. This can lead to above average levels of the drug in the patient's body, which increases the likelihood that they will develop adverse drug reactions (ADRs). Another form of inhibition, known as catalysis-dependent inhibition, is when a reactive species formed from drug metabolism interacts with moieties in the active site of the CYP. This would lead to the enzyme's inactivation (Manikandan & Nagini, 2018). Along with drugs, genetic variants can also affect the activity of these enzymes, leading to severe adverse reactions or impaired drug response. Due to the severity of some ADRs, it is crucial to study drug-drug interactions, and genetic variations that influence enzyme activity and efficacy. This field of study is known as pharmacogenetics, or pharmacogenomics when response is influenced by multiple genes.

1.4 Pharmacogenetics and pharmacogenomics

The phenomenon of different reactions to the same drugs based on genetic variability was recognized in the 1950s, which led to the beginning of a new field interchangeably called pharmacogenetics or, more recently, pharmacogenomics. Pharmacogenomics studies how one's genetic makeup affects drug response. In other words, it deals with the intersection between pharmacology and genetics. There are two possible pathways where genetic variation may exert its effects, pharmacokinetics and pharmacodynamics. Pharmacokinetics focuses on the absorption and transportation of the drug, its metabolic pathways, and how it is excreted from the body. Pharmacodynamics, on the other hand, focuses on the drug's effects on the body (Müller & Rizhanovsky, 2019).

Genetic variations in some genes can affect how efficiently a drug can be metabolized. This effect can either be an increase or a decrease in metabolizing ability. Individuals with an increased metabolizing ability are known as rapid or ultra-rapid metabolizers. In contrast, individuals can be classified to be poor metabolizers if they have reduced or abolished enzymatic activity. Determining the genetic variants in an enzyme-encoding gene can identify the individual's metabolizer status for a drug

metabolized mainly by this enzyme. It is essential to identify abnormal metabolizer statuses since increased or decreased levels of the drug can affect its response or increase the risk of developing ADRs (Pirmohamed, 2014). PharmGKB, the largest database curating pharmacogenomic variations and their effect on drug choice and dosing, designates the genes that code for enzymes with a crucial role in drug response as Very Important Pharmacogenes (VIP) (PharmGKB, n.d., b). Many members of the *CYP* family are categorized among VIPs.

1.4.1 The importance of CYPs

As mentioned in section 1.3, cytochrome P450s are a group of oxygenases and oxidases that play a role in phase I drug metabolism. The CYPs that typically affect human drug metabolism are in the subfamilies CYP1, CYP2, and CYP3. Members of these three subfamilies play roles in the metabolism of about 90% of medications (Isovoran et al., 2017). Due to the sheer number of life-saving drugs being used by individuals and the high percentage of them affected by *CYP* variations, it is imperative to study these genes and their encoded enzymes. Within the subfamily of CYP2, three main enzymes are the focus of pharmacogenomic research: CYP2C9, CYP2C19, and CYP2D6 (Cavallari et al., 2011). The focus of the current thesis is on two members of family 2, namely CYP2C19 and CYP2D6.

1.5 CYP2C19

CYP2C19 is involved in the metabolism of many commonly prescribed medications. The medications that have Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines based on *CYP2C19* genotypes are listed in Table 1.1 (Botton et al., 2021; Guidelines, 2021).

Table 1.1: List of medications with CPIC prescription guidelines for *CYP2C19* (Guidelines, 2021).

Name	**Medication Class**
Amitriptyline	Tricyclic antidepressant
Citalopram	Selective serotonin reuptake inhibitor
Clomipramine	Tricyclic antidepressant
Clopidogrel	Antiplatelet
Desipramine	Tricyclic antidepressant
Dexlansoprazole	Proton pump inhibitor
Doxepin	Tricyclic antidepressant
Escitalopram	Selective serotonin reuptake inhibitor
Esomeprazole	Proton pump inhibitor
Fluvoxamine	Selective serotonin reuptake inhibitor
Imipramine	Tricyclic antidepressant
Lansoprazole	Proton pump inhibitor
Nortriptyline	Tricyclic antidepressant
Omeprazole	Proton pump inhibitor
Pantoprazole	Proton pump inhibitor
Paroxetine	Selective serotonin reuptake inhibitor
Rabeprazole	Proton pump inhibitor
Sertraline	Selective serotonin reuptake inhibitor
Trimipramine	Tricyclic antidepressant
Voriconazole	Antifungal

Antidepressants, such as Citalopram and Escitalopram, and antiplatelet medications, such as Clopidogrel, are two of the most important medications influenced by CYP2C19. The metabolic pathways of Citalopram and Escitalopram are shown in Figure 1.3 and for Clopidogrel in Figure 1.4, illustrating the roles of CYP2C19 in their metabolism, among other CYPs.

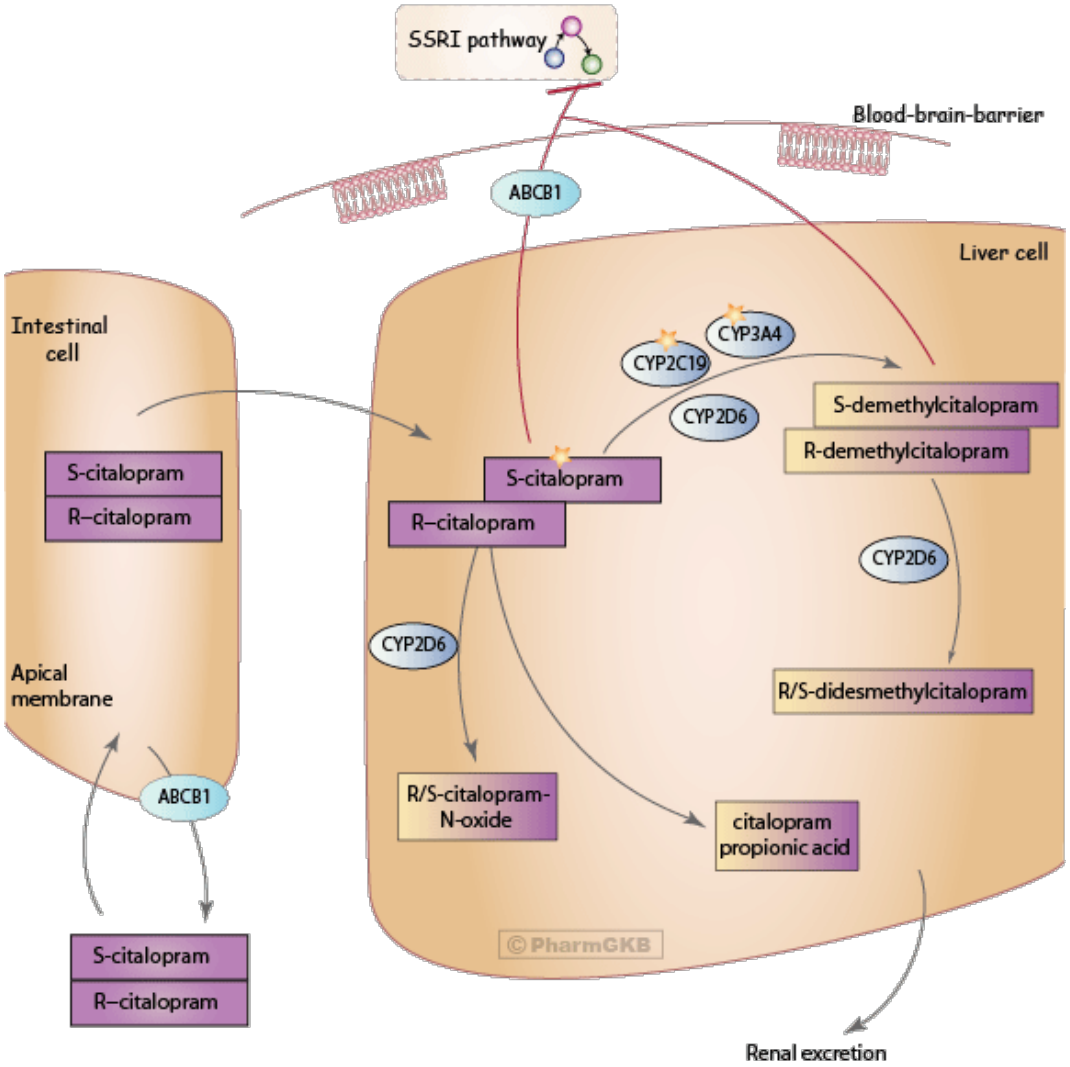


Figure 1.3: Depiction of Citalopram and Escitalopram metabolic pathway.
Citalopram is made up of its R/S enantiomers, while Escitalopram is made up of only the S enantiomer. CYP2C19, CYP2D6, and CYP3A4 catalyze both enantiomers into R/S-demethylcitalopram. CYP2D6 alone then converts it into R/S-didesmethylcitalopram, which can then be excreted from the body (Citalopram and Escitalopram Pathway, Pharmacokinetics, 2021).

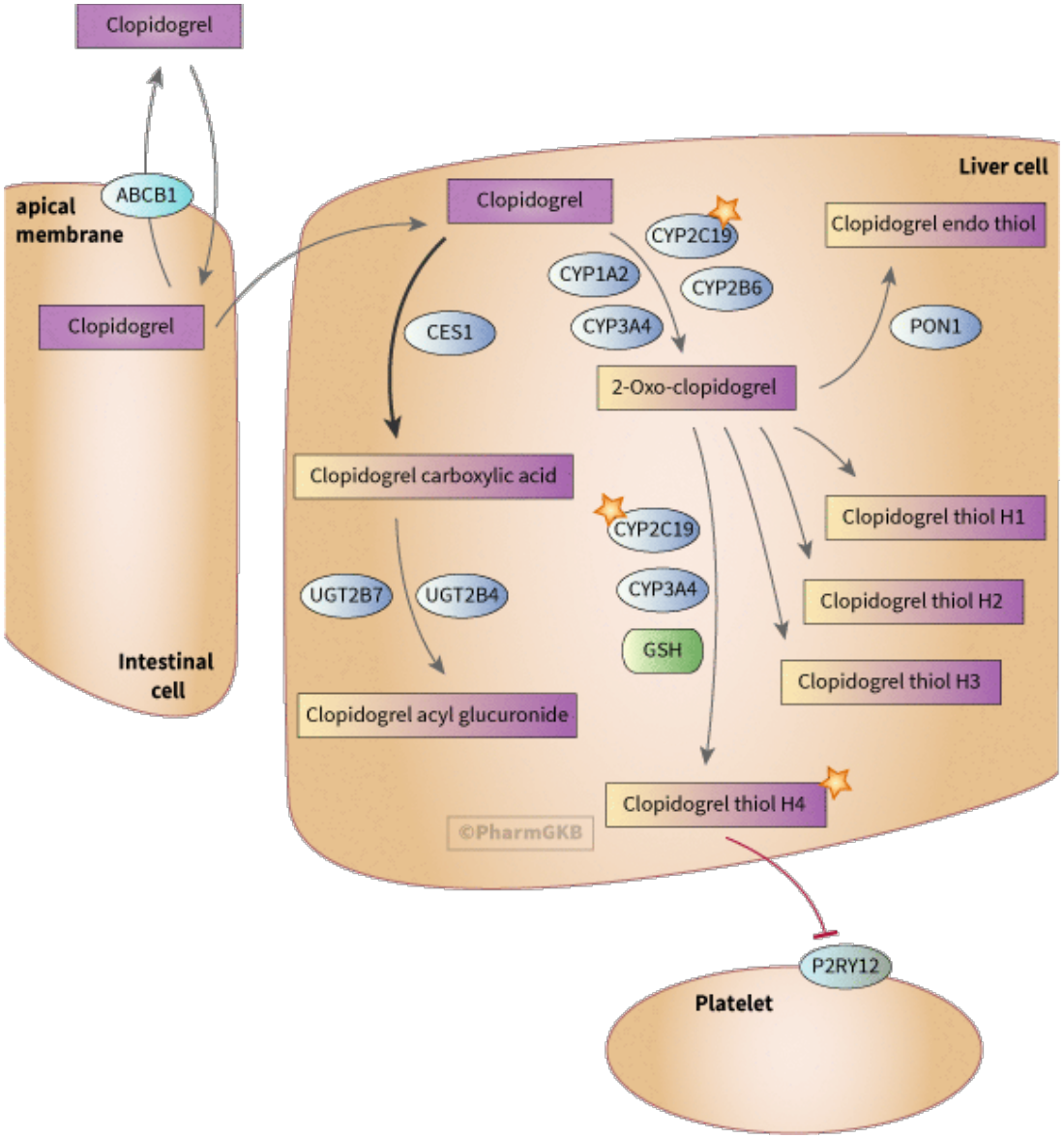


Figure 1.4: Clopidogrel metabolic pathway.

Clopidogrel is a prodrug that must be activated in the liver. This is done in two steps. The first step involves clopidogrel being converted into 2-oxo-clopidogrel by CYP1A2, CYP2B6, CYP2C19, and CYP3A4. The second step is when it is converted into the active metabolite. 2-oxo-clopidogrel is metabolized into cis- and trans-thiol diastereoisomers H1, H2, H3, and H4, as well as an endo-thiol isomer. Research of these products has shown that the active metabolite is H4, which is primarily formed by CYP2C19 (Clopidogrel Pathway, Pharmacokinetics, 2019).

CYP2C19 was first discovered in 1984 after it was noticed that healthy individuals were not able to properly metabolize the anti-convulsant drug mephenytoin (Küpfer & Preisig, 1984). *CYP2C19* encoding for this enzyme is one of four other cytochrome P450 genes located on chromosome 10q23.33. The other three genes are *CYP2C8*, *CYP2C9*, and *CYP2C18* (Goldstein & de Morais, 1994). Multiple alleles of *CYP2C19* have been identified and designated with the star allele nomenclature. Table 10

1.2 lists the major star alleles that have a known effect on the enzyme, as well as their resulting impact on enzyme function.

Table 1.2: List of *CYP2C19* star alleles with a known enzyme function assignment from CPIC (PharmGKB, n.d., a).

CYP2C19 Star Allele	**CPIC Function Assignment**
CYP2C19*1	Normal function
CYP2C19*2	No function
CYP2C19*3	No function
CYP2C19*4	No function
CYP2C19*5	No function
CYP2C19*6	No function
CYP2C19*7	No function
CYP2C19*8	No function
CYP2C19*9	Decreased function
CYP2C19*10	Decreased function
CYP2C19*11	Normal function
CYP2C19*13	Normal function
CYP2C19*15	Normal function
CYP2C19*16	Decreased function
CYP2C19*17	Increased function
CYP2C19*18	Normal function
CYP2C19*19	Decreased function
CYP2C19*22	No function
CYP2C19*24	No function
CYP2C19*25	Decreased function
CYP2C19*26	Decreased function
CYP2C19*28	Normal function
CYP2C19*35	No function
CYP2C19*36	No function
CYP2C19*37	No function
CYP2C19*38	Normal function

Due to the influence that CYP2C19 has on the effectiveness and response to multiple drugs, it is important to determine the effects that the variations in this gene may have on the enzyme's metabolizing ability. One way to do this is by using *in vitro* enzyme activity studies of expressed mutants. These studies are performed in laboratory settings using microorganisms or mammalian cells expressing the gene variants. Their goal is typically to define the effects of variants on the enzyme's activity and function. The cell lines most commonly used for CYP2C19 experiments are COS-1, COS-7, HepG2, Hek293T, and Hek293FT (Dai et al., 2015).

1.5.1 In vitro studies to evaluate the effects of variation on the enzymatic function of CYP2C19

In the work of Takahashi and colleagues, COS-7 cells were used to test the effects that the CYP2C19 star alleles *CYP2C19*1-CYP2C19*28* have on the enzyme's metabolizing ability in general, as well as specifically on clopidogrel (Takahashi et al., 2015). Site-directed mutagenesis was used to create the polymorphisms, followed by transfecting them into COS-7 cells for expression. Western blots were carried out to confirm the expression of CYP2C19 in each mutant sample. In order to determine how effectively clopidogrel is metabolized by each CYP2C19 star allele, the first product in the process, 2-oxo-clopidogrel was measured using liquid chromatography and mass spectrometry. Since clopidogrel is a drug that is highly sensitive to drug dosage changes, an *S*-mephenytoin 4'-hydroxylation assay was used to confirm if the results of the clopidogrel assay are specific to that drug. Out of the 21 variants tested, 8 of them were expressed in quantities too low to be tested. Seven of the remaining 13 variants were shown to have significantly decreased enzyme activity. These star alleles are *CYP2C19*5A*, *CYP2C19*5B*, *CYP2C19*6*, *CYP2C19*8*, *CYP2C19*16*, *CYP2C19*22*, and *CYP2C19*24*. Overall, this work confirmed previously reported findings, as well as determined the effects that less common polymorphisms have on clopidogrel metabolism (Takahashi et al., 2015). Due to how life-threatening cardiovascular conditions can be, the influence that CYP2C19 has on the effectiveness of clopidogrel is an important area of study. By repeating tests on previously reported variants and testing newly discovered ones, the risk of fatal and preventable adverse drug reactions can be reduced.

Similarly, Devarajan and coworkers focused on confirming the effects of a series of rare variants in *CYP2C19* and *CYP2C9* from the RIGHT study (Devarajan et al., 2019). The RIGHT study was conducted to determine the feasibility and impact that preemptive genetic testing can have in a clinical setting (Bielinski et al., 2014; Ji et al., 2016). Site-directed mutagenesis was used to recreate the target mutations in *CYP2C9* and *CYP2C19* plasmids. These mutations were then expressed in COS-1 cells. Western blots were done to confirm the presence of the proteins. Some cells were treated with proteasome inhibitors to determine if protein degradation could affect the results. Multiple enzyme assays were done to determine activity. The Vivid assay (ThermoFisher Scientific) was used for both cytochrome P450s. Tolbutamide assay was used for CYP2C9, and *S*-mephenytoin was used for CYP2C19. These two assays were performed using liquid chromatography and mass spectrometry. *In-silico* prediction analyses were also done using Polyphen version 2, PROVEAN, and SIFT software. The results showed that variants that had differing protein expression levels compared to the wild type measurements. This was also shown in the Vivid assays, as well as in the liquid chromatography and mass spectrometry experiments. In total, there were seven CYP2C9 mutations and seven CYP2C19 mutations tested. Five of the CYP2C9 variants showed decreased activity while the remaining two had similar levels to the wild type. For CYP2C19, six of the variants showed decreased activity, and one had similar activity to the wild type. The *in-silico* predictions had varying accuracy. For example, some variants that had significantly decreased levels were predicted to be benign (Devarajan et al., 2019). This work was able to determine the effects of previously reported variants. This experiment highlights the importance of reviewing studies focused on sequencing patients' samples.

1.6 CYP2D6

CYP2D6 is a very important enzyme in metabolizing more than 72 clinically-relevant drugs, such as antidepressants, antipsychotics, stimulants, and anti-malarial medications (Taylor et al., 2020). As other enzymes, ultra-rapid metabolizers taking medications metabolized mainly by CYP2D6 may not get benefit of treatment because drugs are metabolized too quickly. On the other hand, poor metabolizers are more likely to develop ADRs from the medication being in their system for too long. Due to these

risks, which are more sensitive when related to drugs in complex mental illnesses, it is imperative to determine these genetic variants before a prescription for the patient to receive the most benefits from taking their medication. However, sequencing CYP2D6 is very difficult and complicated by its highly polymorphic structure (Gaedigk, 2013). *CYP2D6* was discovered and named in 1988 by Gonzalez and colleagues while studying why people reacted differently to debrisoquine, an antihypertensive drug (Gonzalez et al., 1988). In 1989, Kimura and coworkers discovered that *CYP2D6* shares a gene locus with two other genes; *CYP2D7* and *CYP2D8* (Kimura et al., 1989). Not only do they share a gene locus, but they also share a long identical sequence. Variants only make the process of sequencing and cloning *CYP2D6* even more complicated. Some alleles are defined by large sections being deleted, duplicated, or even leading to a hybrid gene of *CYP2D6* and *CYP2D7* (Gaedigk, 2013). While it is difficult to sequence *CYP2D6*, it is not impossible.

In the work of Puaprasert and colleagues, patients from the Karen population on the Thailand-Myanmar border with *Plasmodium vivax* infection were tested for *CYP2D6* variations (Puaprasert et al., 2018). This population was selected due to the high rates of *P. vivax* infection in this area. These cases are treated using the anti-malarial drug primaquine, which is metabolized by CYP2D6. To detect *CYP2D6* deletions or duplications in DNA extracted from the blood samples collected, a tetra-primer extra-long PCR (XL PCR) technique was used (Puaprasert et al., 2018). This technique requires four primers; two are designed to flank the wild type *CYP2D6* gene and the other two are designed to flank variants (Medrano & de Oliveira, 2014). The products of the XL PCR were then reamplified before being sequenced. While the study size was small, with only 70 subjects, it was revealed that many carried a decrease-in-function allele. The small sample size limited the results of this study; however, the technique used proved feasibility (Puaprasert et al., 2018).

1.7 Pharmacogenetic and pharmacogenomics research and implementation in the United Arab Emirates (UAE)

The clinical implementation of pharmacogenomics potentially can increase the effectiveness of medications and decrease the risk of adverse drug reactions. While many studies analyzed the benefits of this practice in Western countries, very few studies have

been carried out in the Middle East, including in the UAE. There have been preliminary survey studies performed on medical professionals and students in the UAE to determine the level of understanding that they currently have. In the work of Rahma and co-workers, a survey was given to medical students to assess their level of knowledge and opinion on pharmacogenomics (Rahma et al., 2020). It was found that most medical students had a positive opinion of pharmacogenomics but needed a stronger understanding of it. The students did reflect a desire to learn more about pharmacogenomics and the benefits of genetic testing for patients. While there was a more optimistic view in this survey, it is essential for these opinions to be taken seriously and used for guidance in improving the education of this field so that it can be implemented fully in clinics (Rahma et al., 2020). A similar survey was also given to pharmacists practicing in the UAE. The results show that pharmacists do not have strong knowledge of pharmacogenomics and genetic testing, however, they are willing to learn and would like to be given reliable sources to use. While the pharmacists' raised some concerns in this study, such as their worry about the confidentiality of the genetic testing results, it is evident that most of them would like to have pharmacogenomics programs incorporated into their training (Rahma et al., 2020).

Overall, it has been demonstrated that there is a need for precision medicine and pharmacogenomics-based research studies in the UAE for effective implementation. It is even more important to genotype and identify unique variations in pharmacogenes, and determine their effects on the resulting enzymes, especially in small populations like Emiratis. In this context, Al-Mahayri and coworkers examined variants of 100 pharmacogenes in 100 healthy Emiratis (Al-Mahayri et al., 2020). This work identified common, rare, and novel variants in multiple very important pharmacogenes, including *CYP2C19* (Al-Mahayri et al., 2020). Seven of the rare or novel variants found in *CYP2C19* in that study were predicted, using *in silico* tests, to have a detrimental effect on the enzyme. These variants are shown in Table 1.3.

Table 1.3: Seven rare or novel variants found in *CYP2C19* in the Emirati population (Al-Mahayri et al., 2020).

rs ID	Location	Mutation (nucleotide)	Mutation (amino acid)	Rare/ novel	gnomad Frequency	References
rs572853437	96522626	c.C164G	p.T55S	Rare	0	Zhang et al., 2020
rs17878459	96534922	c.G276C	p.E92D	Rare	0.02769	Wang et al., 2011a, 2011b Zhang et al., 2020
rs145119820	96535152	c.G337A	p.V113I	Rare	0.0006918	Devarajan et al., 2019 Matimba et al., 2009 Zhang et al., 2020
rs577255883	96541719	c.G784A	p.D262N	Rare	0	Zhang et al., 2020
rs377674118	96541736	c.C801A	p.F267L	Rare	0	Zhang et al., 2020
rs748832049	96602641	c.C1009T	p.P337S	Novel	0	
rs372873637	96609683	c.A1159G	p.I387V	Novel	0	

Other studies determined some of the alleles and genotypes, as well as their frequencies, of other pharmacogenes in the UAE (Qumsieh et al., 2011; Al-Jaibeji et al., 2016; Al-Ahmad et al., 2017a, 2017b). Nevertheless, no follow-up studies analyzed the effects of these mutations on the CYP2C19 enzyme. This is an important piece of information to learn about these mutations with the aim to improve the practice of precision medicine in the country.

1.8 Aims and objectives

The main aim of the current thesis was to evaluate the impact of variation in two of the most important pharmacogenes, namely CYP2C19 and CYP2D6. The first section focuses on CYP2C19, and the effects of some novel or rare variations detected in Emiratis on the enzyme's activity. The second section aimed to optimize a protocol for detecting duplications and deletions in *CYP2D6* gene for psychiatric patients in the UAE requiring anti-depressants.

1.8.1 CYP2C19

As a follow-up of work carried out by Al Mahayri et al. (2020), specific CYP2C19 variants identified in Emirati genomes by next generation sequencing are to be examined. The criteria for selecting a variant included having *in silico* predictions of being "detrimental" to the enzyme product, being a rare or novel variant, and its potential effect on enzyme activity was either never studied before or has been studied with conflicting evidence. The primary objective of this part of the project was to determine the effects of selected CYP2C19 variants on enzyme activity, subcellular localization, and perform molecular modelling. The specific objectives were to generate the target variations in two CYP2C19 mammalian expression vectors, express them in COS-1, Hek293T, and HeLa cells, perform Western blots to confirm the presence of the protein, use immunofluorescence to detect the subcellar location of CYP2C19, use molecular modelling to view any differences in intramolecular bonds between the wild type protein and the mutated proteins, and attempt to use an enzyme assay to measure CYP2C19 enzyme activity with the variations.

1.8.2 CYP2D6

The aim of this part of the project was to optimize a protocol used for the detection of CYP2D6 structural variations (copy number variations), which is considered one of the main troublesome issues in pharmacogenomics laboratories. The objective was to optimize and test a simple long-range PCR-based technique used for CYP2D6 deletion and duplication detection, use agarose gel electrophoresis to analyze the results, and utilize the information for the implementation of pharmacogenetics-based treatment for individuals with depression.

Chapter 2: Methods

2.1 CYP2C19

The aim of this part of the study was to evaluate the potential effects of novel or rare variants found among Emiratis in CYP2C19 using *in silico*, as well as *in vitro*, approaches. The *in silico* approach included performing molecular modelling, and the *in vitro* approach involved the generation of variants by using site-directed mutagenesis with a mammalian expression vector as a template, followed by attempts at determining their effects on the enzymatic activity using *in vitro* enzyme assays. This was done by introducing the mutations using site-directed mutagenesis with the CYP2C19 cDNA cloned into mammalian expression vectors as templates. The WT and mutant constructs were then transiently transfected into COS-1 and Hek293T cells. To confirm the expression of the proteins, lysates from the transfected cells were tested using the Western blot technique. Finally, CYP2C19 enzyme activity testing was attempted to be measured using a commercially available CYP2C19 enzymatic assay kit. Immunofluorescence was also used to verify that the protein is localized at the endoplasmic reticulum of both COS-1 and Hela cells. Molecular modelling was used to establish if there are physical changes in the enzyme's structure due to the variations that could lead to a difference in the enzyme's activity.

The rare variants focused on in this project were chosen based on the results obtained by Al-Mahayri et al. (2020). They were chosen based on the lack of consistent data on their effects on CYP2C19 activity. The studied mutations are listed in Table 2.1.

Table 2.1: List of the studied mutations in *CYP2C19*.

rs ID	**Location**	**Mutation (nucleotide)**	**Mutation (amino acid)**
rs572853437	96522626	c.C164G	p.T55S
rs17878459	96534922	c.G276C	p.E92D
rs145119820	96535152	c.G337A	p.V113I
rs577255883	96541719	c.G784A	p.D262N
rs377674118	96541736	c.C801A	p.F267L
rs748832049	96602641	c.C1009T	p.P337S
rs372873637	96609683	c.A1159G	p.I387V

The following sections describe the materials and the main methodologies used.

2.1.1 Vectors

Two vectors harboring the CYP2C19 cDNA have been obtained and used in this thesis. The human untagged CYP2C19 clone in pCMV6-XL5 CYP2C19 plasmid has been purchased from Origene Inc. (NM_000769, Origene, Rockville, Maryland, U.S.A.). The pcDNA3.4 CYP2C19 untagged plasmid was a kind gift from Dr. Masahiro Hiratsuka of Tohoku University (Takahashi et al., 2015). The first vector was selected following the work of Devarajan and colleagues, in which they used enzyme activity assays to measure the activity of novel *CYP2C19* variations. To generate the mutations, the pCMV6-XL5 CYP2C19 human untagged clone was used (Devarajan et al., 2019). The structure and organization of this plasmid are shown in Figure 2.1.

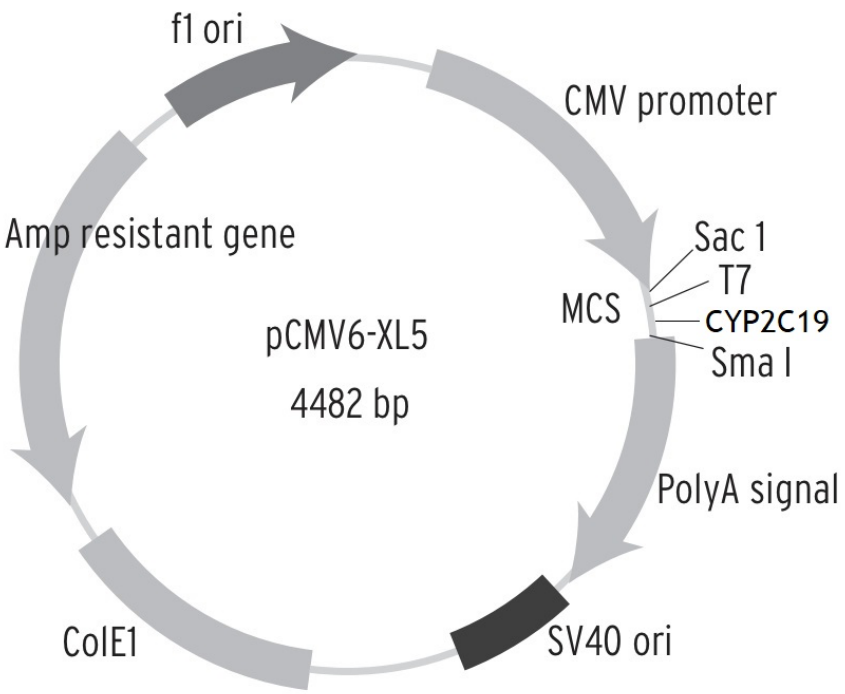


Figure 2.1: A map of the pCMV6-XL5 CYP2C19 human untagged clone sequence.
Abbreviations: CMV; Cytomegalovirus. MCS; Multiple cloning site. SV40 ori; Simian virus origin of replication. ColE1; Colicin E1. Amp; Ampicillin.

The pcDNA3.4 CYP2C19 untagged plasmid was also used because it has been used to successfully generate mutations, and in an enzyme activity assay (Takahashi et al., 2015). This plasmid's structure is shown in Figure 2.2.

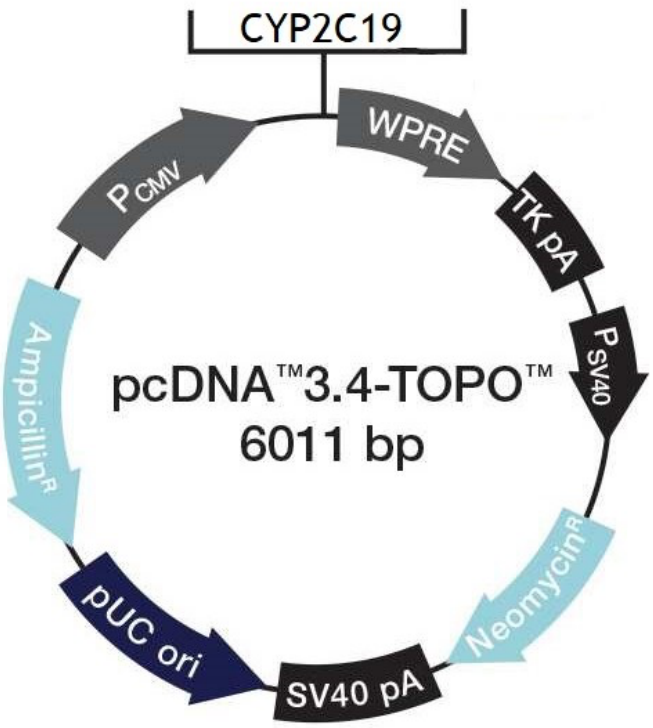


Figure 2.2: The vector map for the pcDNA 3.4 plasmid.
Abbreviations: WPRE; Woodchuck hepatitis virus posttranscriptional regulatory element. PSV40; Simian virus 40 promotor. SV40 pA; Simian virus 40 polyA signal. pUC; Plasmid cloning vector origin of replication. PCMV; Cytomegalovirus promotor (Takahashi et al., 2015).

2.1.2 Generation of variants by site-directed mutagenesis

The primers used for site-directed mutagenesis (SDM) were designed using the PrimerX website and the resulting primers are listed in Table 2.2. The designed primers were custom made by Metabion Inc (Metabion Inc., Steinkirchen, Germany).

Table 2.2: List of *CYP2C19* variants and their mutagenic primers.

Target	**Forward Primer**	**Reverse Primer**
c.C164G, p.T55S	5'-GCA AAT CCT TAA **G**CA ATC TCT CAA-3'	5'-TTG AGA GAT TG**C** TTA AGG ATT TGC-3'
c.G276C, p.E92D	5'-GAT CTT GGA GA**C** GAG TTT TCT GG-3'	5'-CCA GAA AAC TC**G** TCT CCA AGA TC-3'
c.G337A, p.V113I	5'-GGA TTT GGA ATC **A**TT TTC AGC AAT GG-3'	5'-CCA TTG CTG AAA AT**G** ATT CCA AAT CC-3'
c.G784A, p.D262N	5'-CAA CAA CCC TCG G**A**A CTT TAT TGA TTG-3'	5'-CAA TCA ATA AA**G** TTC CGA GGG TTG TTG-3'
c.C801A, p.F267L	5'-TTG ATT GCT T**A**C TGA TCA AAA TGG AGA-3'	5'-TCT CCA TTT TGA T**C**A GTA AGC AAT CAA-3'
c.C1009T, p.P337S	5'-AAA CCG GAG C**T**C CTG CAT GCA-3'	5'-TGC ATG CAG GAG **C**TC CGG TTT-3'
c.A1159G, p.I387V	5'-GGG CAC AAC C**G**T ATT AAC TTC C-3'	5'-GGA AGT TAA T**A**C GGT TGT GCC C-3'

The PCR step of the site-directed mutagenesis was performed by combining 19.5 µL of DNAse free water, 2.5 µL of Pfu buffer (Agilent, Santa Clara, California, U.S.A.), 1 µL of plasmid, 0.5 µL of dNTPs, 0.5 µL of the forward primer, 0.5 µL of the reverse primer, and 0.5 µL of the Pfu enzyme (Agilent, Santa Clara, California). The total volume in each tube was 25 µL. The SimpliAmp thermal cycler (ThermoFisher Scientific, Waltham, Massachusetts, U.S.A.) was set to the parameters shown in Figure 2.3.

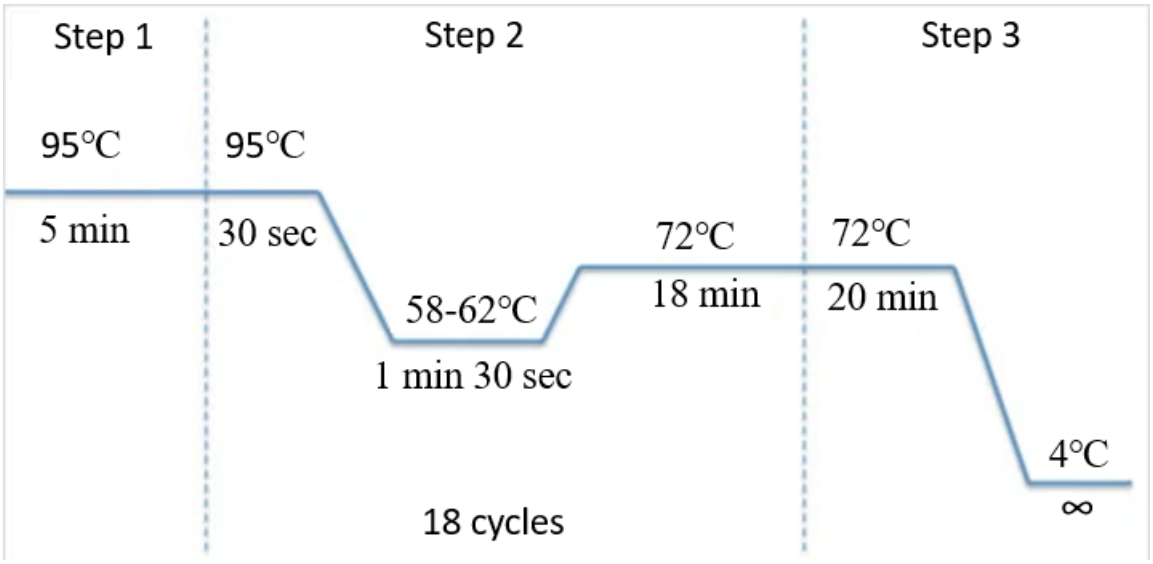


Figure 2.3: SimpliAmp thermal cycler parameters for SDM.

The primers used for Sanger sequencing to confirm the generation of the intended mutations are shown in Table 2.3.

Table 2.3: List of primers used for Sanger sequencing of *CYP2C19* mutations.

Targets	**Sequencing Primer**
c.C164G p.T55S c.G276C p.E92D	5'-GGA CTT TCC AAA ATG TCG-3' This is the VP1.5 primer (Origene, Rockville, Maryland, U.S.A.)
c.G337A p.V113I c.G784A p.D262N c.C801A p.F267L	5'-GTG AAG GAA GCC CTG ATT GA-3'
c.C1009T p.P337S c.A1159G p.I387V	5'-ACA CCA AGA ATC GAT GGA CA-3'

The sequence of the plasmids, location of the mutations, and the primers used are shown in Figure 2.4.

Figure 2.4: An illustration of the sequence of the CYP2C19 cDNA cloned in the expression vectors.

The nucleotides shown in red are the locations of the target mutations. The portions that are bold and underlined indicate the primers that are used in this experiment.

The next step is Dpn1 (Promega, Madison, Wisconsin, U.S.A.) digestion. This was done by adding 0.5 µL of Dpn1 enzyme to 10 µL of the PCR product. This mixture was then placed in a water bath set to 37°C for 4 hours.

The transformation was done by transferring 5 µL of the Dpn1 treated product to 100 µL of JM109 competent cells. This mixture was kept on ice for 30 minutes, with the tubes being gently tapped every 5 minutes. They were then placed in a water bath set to 42°C for 50 seconds and then immediately being put into ice again for 2 minutes. 100 µL of an LB broth and glucose mixture was added to each tube and then placed in a shaker set to 37°C and 150 rpm for 2 hours. They were then plated on ampicillin LB agar plates (Invitrogen, Waltham, Massachusetts, U.S.A.) and incubated for 18 hours at 37°C.

The colonies grown were then picked and placed into a tube containing 5 mL of LB broth and 5 µL of ampicillin. They were incubated in a shaker set to 37℃ and 150 rpm for 18 hours.

DNA extraction was done using the Qiagen Miniprep Kit (Qiagen, Hilden, Germany). The tubes were centrifuged at 4℃ and 4,000 rpm for 10 minutes. The supernatant was removed. The pellet was dissolved in 250 µL of P1 buffer and moved into an Eppendorf tube. 250 µL of P2 buffer was added and the tubes were inverted for 3 minutes. 350 µL of N3 buffer was added and the tubes were inverted again for 3 minutes. The tubes were then centrifuged at 4℃ and 13,000 rpm for 10 minutes. The supernatant was poured into a spin column and centrifuged for 1 minute. 750 µL of PE buffer was added and centrifuged for 1 minute. The flow-through was discarded, and the column was centrifuged again for 1 minute. 50 µL of water was added and allowed to sit at room temperature for 1 minute, followed by centrifuging for 1 minute. The DNA concentration of the product was then measured on the Implen Nanophotometer N60 (Implen, Westlake Village, California, U.S.A.).

2.1.3 Sanger sequencing

Sanger sequencing was performed on the SDM products to confirm the presence of the mutation. The BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Waltham, Massachusetts, U.S.A.) was used for this purpose. The machine used was the 3130xl Genetic Analyzer (Applied Biosystems, Waltham, Massachusetts, U.S.A.). The results were viewed using the Chromas software (Technelysium, Australia).

2.1.4 Cell culture

There were three main cell lines used: COS-1 cells, Hek293T cells, and HeLa cells. All of the cell lines were maintained in T25 flasks with Dulbecco's Modified Eagle Medium (DMEM; Invitrogen, Waltham, Massachusetts, U.S.A.) containing 10% Fetal Bovine Serum (FBS; Invitrogen, Waltham, Massachusetts, U.S.A.). Once the flasks reached 80-90% confluency, they were washed with 3 mL of phosphate-buffered saline (PBS; Invitrogen, Waltham, Massachusetts, U.S.A.). Trypsinization was done on the Hek293T cells for 1 minute and on the COS-1 and HeLa cells for 5 minutes using 0.5 mL of Trypsin for both. DMEM was then added once again to halt trypsinization. They

were then seeded into 6-well plates for the COS-1 and Hek293T cells and 24-well plates with coverslips for the COS-1 and HeLa cells.

For the COS-1 and Hek293T cells in 6-well plates, transfection began by adding 1 µg of the plasmid to 45 mL of Opti-MEM (Invitrogen, Waltham, Massachusetts, U.S.A.). This mixture was vortexed, spun down, and incubated at room temperature for 5 minutes. Next, 3 µL of FuGene HD (Promega, Madison, Wisconsin, U.S.A.) was added and mixed using a pipette. They were spun down and sat at room temperature for 12 minutes. They were then added into the wells and were incubated for 48 hours at 37℃. For COS-1 and HeLa cells in 24-well plates, the process is identical but with smaller volumes; 27.5 µL of Opti-MEM, 1.5 µL of FuGene HD, and 0.5 µg of plasmid.

Once both cell lines in 6-well plates were transfected, they were scraped, pelleted, and frozen at -80℃ until needed.

2.1.5 Western blotting

Western blotting was performed on COS-1 and Hek293T cells. The protein content in these cells is measured following the steps provided in the Bicinchonic Acid Protein Assay Kit (BCA; Pierce, Waltham, Massachusetts, U.S.A.). The lysis buffer used is Tris buffer with TritonX. The protein concentration needed for Western blot is 30 µg. The total amount that will be loaded into each well is 40 µL. In order to reach this amount, 8 µL of Laemmli buffer is added and the remainder is made up of Tris buffer with TritonX (Sigma-Aldrich, St. Louis, Missouri, U.S.A.). This mixture is then boiled at 98˚C for 5 minutes. These samples were loaded into premade gels (GenScript, Picataway, New Jersey, U.S.A.). The gel was run at 70V for 30 minutes before being increased to 120V for an hour. The transfer was done using PVDF membrane (Pierce, Waltham, Massachusetts, U.S.A.) at 100V for an hour on ice. Blocking was done using 5% milk. The primary antibody used to bind to the membrane was a CYP2C19 rabbit polyclonal antibody (Origene, Rockville, Maryland, U.S.A.). Anti-β-actin antibody was also used to provide a control. The membrane was incubated in a 5% milk and primary antibody mixture overnight on a roller at 4˚C. The membrane was then washed with TBST three times for 10 minutes each, followed by incubation with the secondary antibody. For the CYP2C19 antibody, an anti-rabbit antibody was used. For the β-actin

antibody, an anti-mouse antibody was used. The membrane was then washed thrice with TBST once again. The ECL Plus Kit (ThermoFisher Scientific, Waltham, Massachusetts, U.S.A.) was used to detect the bands on the membrane. The membrane was viewed using Typhoon FLA 9500 Imager. The images obtained were edited using the ImageJ software.

2.1.6 CYP2C19 enzyme activity assays

The CYP2C19 Enzyme Activity Assay Kit (Abcam, Cambridge, United Kingdom) was used to detect the activity of the mutated CYP2C19 enzymes. The mutated plasmids obtained with SDM were transfected into COS-1 and Hek293T cells. Once the protein concentration was detected, the concentrations used were between 20-250 µg. For this experiment, a 96-well plate was used, with each well containing 70 µL. Each well would contain either the sample, the inhibitor control (sample + inhibitor), background control, enzyme control, or enzyme-inhibitor control. To prepare the sample reaction mix, 2 µL of the NADPH Generating System was added to the sample, along with Assay Buffer to get a 50 µL sample reaction mix. An additional 20 µL of Assay Buffer is also added to get a 70 µL sample. For the inhibitor control, 50 µL of the sample is combined with 20 µL of the inhibitor. The background control is 70 µL of the Assay Buffer. The enzyme control consists of 25 µL of the Recombinant Human CYP2C19 Enzyme received in the kit, along with 45 µL of the Assay Buffer. The enzyme-inhibitor control is 25 µL of the Recombinant Human CYP2C19 Enzyme with 20 µL of the inhibitor and 25 µL of the Assay Buffer. Once all the wells are filled, the plate is incubated at 37ºC for 10-15 minutes. During the incubation period, the reaction mixture is made. For one well, the mixture is comprised of 0.12 µL of the CYP2C19 Substrate, 1 µL of the β-NADP+ stock, and 28.88 µL of the Assay Buffer, so a total of 30 µL of this reaction mix is added to one well. Once the reaction mix is added, the plate is taken to the microplate reader to get the Initial reading at Ex/Em = 406/468 nm. The plate is then incubated at 37ºC between readings, which were taken at 15-minute intervals for a total of 60 minutes.

2.1.7 Immunofluorescence confocal microscopy

Immunofluorescence was done on the transfected COS-1 and HeLa cells in 24-well plates. DMEM was removed and the coverslips were washed thrice with PBS. The cells were fixed using cold methanol. Following the fixation, the cells were washed thrice once again with PBS to wash off the methanol residue. Blocking was done using 3% Bovine Serum Albumin (BSA) for 30 minutes. The primary antibodies used were the CYP2C19 rabbit polyclonal antibody and goat polyclonal anti-calnexin. Calnexin is being used as a control since it is found in the endoplasmic reticulum. The cells were incubated with the primary antibodies at room temperature for an hour in a dark environment. They were then washed thrice with PBS before the secondary antibodies were added. The ones used were a red anti-rabbit antibody and a green anti-goat antibody. They were incubated once again at room temperature for an hour in the dark. The cells were washed thrice with PBS again before the coverslips were mounted onto a slide. The slide was then observed using confocal microscopy. Images were edited using ImageJ.

2.1.8 Molecular modelling

Molecular modelling was performed to view the protein, with and without the variations. The software used for this was the Molecular Operating Environment (MOE; Chemical Computing Group, Montreal, Canada).

2.2 Determination of CYP2D6 copy numbers using long-range PCR

2.2.1 Patient recruitment

Patients were recruited from Al Ain Hospital (Abu Dhabi, U.A.E.). The main requirement was that the patient had experienced an episode of unipolar major depression. If the patient had any other psychiatric diagnoses, this was not a reason for them to be excluded from this study. In total, fifteen patients were recruited.

2.2.2 Long-range polymerase chain reaction (PCR)

Long-range PCR was performed to amplify the amount of DNA from the psychiatric patient samples that were received. The primers used were provided by

Professor George Patrinos from the University of Patras in Greece and were designed to flank the regions of *CYP2D6* that indicate the presence of a duplication or deletion. These primers are shown in Table 2.4. Positive and negative controls were also provided by the collaborating laboratory.

Table 2.4: List of primers and mutations for *CYP2D6*.

Variation Type	**SNP**	**Forward Primer**	**Reverse Primer**	**Product Size**
Duplication	*CYP2D6*_XN_1 *CYP2D6*_XN_2	5'-TCC CCC ACT GAC CCA ACT CT-3'	5'-CAC GTG CAG GGC ACC TAG AT-3'	Wild type: 5.2 kb Duplication: 3.6 kb
Deletion	*CYP2D6**5_1 *CYP2D6**5_2	5'-TCC CCC ACT GAC CCA ACT CT-3'	5'-CAG GCA TGA GCT AAG GCA CCC AGA C-3'	Wild type: 3.3 kb Deletion: 4.85 kb

The PCR was performed by adding 3.5 µL of Long-Range PCR Master Mix (Biotech Rabbit, Berlin, Germany), 1.5 µL of Long-Range PCR Enhancer (Biotech Rabbit, Berlin, Germany), 1 µL of the forward primer, 1 µL of the reverse primer, and 3 µL of DNA. The total volume in each tube was 10 µL. The settings of the SimpliAmp thermal cycler are shown in Figure 2.5.

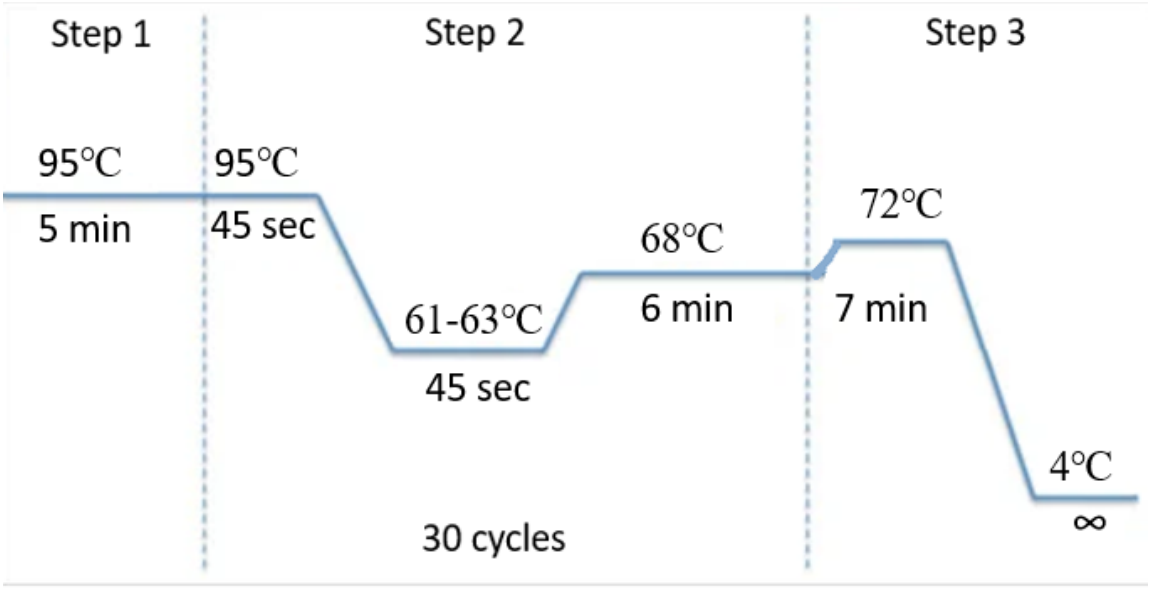


Figure 2.5: SimpliAmp thermal cycler parameters for long range PCR.

2.2.3 Agarose gel electrophoresis

Gel electrophoresis was used to confirm the success of amplification and to determine copy number variations through PCR product size discrimination. Due to the difference in the PCR product size of a duplication, deletion, and the wild type, it is possible to determine the CYP2D6 copy number variation through this technique. The gel used is 0.8% agarose with ethidium bromide, and 5 µL of the sample and dye were loaded in the gel, along with a 1kb DNA ladder (Qiagen, Hilden, Germany). The electrophoresis was run at 100 V for approximately an hour. Using UV light, the bands were then viewed using the Image Lab Software.

Chapter 3: Results

3.1 CYP2C19

3.1.1 The generation of the desired mutants confirmed by Sanger DNA sequencing

The results of the Sanger sequencing for the products of SDM confirmed the generation of all 7 target mutations in the pcDNA3.4 CYP2C19 untagged vector and 6 of the mutations in the pCMV6-XL5 CYP2C19 human untagged vector. The product sequence was compared to the wild type sequence using the NIH Blast Compare tool from NCBI to confirm that the desired mutations have been introduced. In addition, the Sanger sequencing chromatograms have been visually inspected for further confirmation. The chromatograms with the mutations are shown in Figure 3.1.

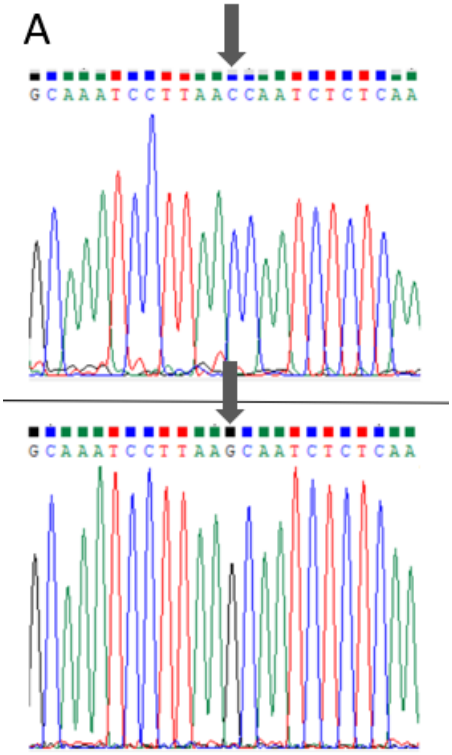


Figure 3.1: Chromatograms illustrating the inserted mutations.
The top chromatogram is for the wild type and the bottom chromatogram is for the mutation. A: c.C164G, p.T55S.

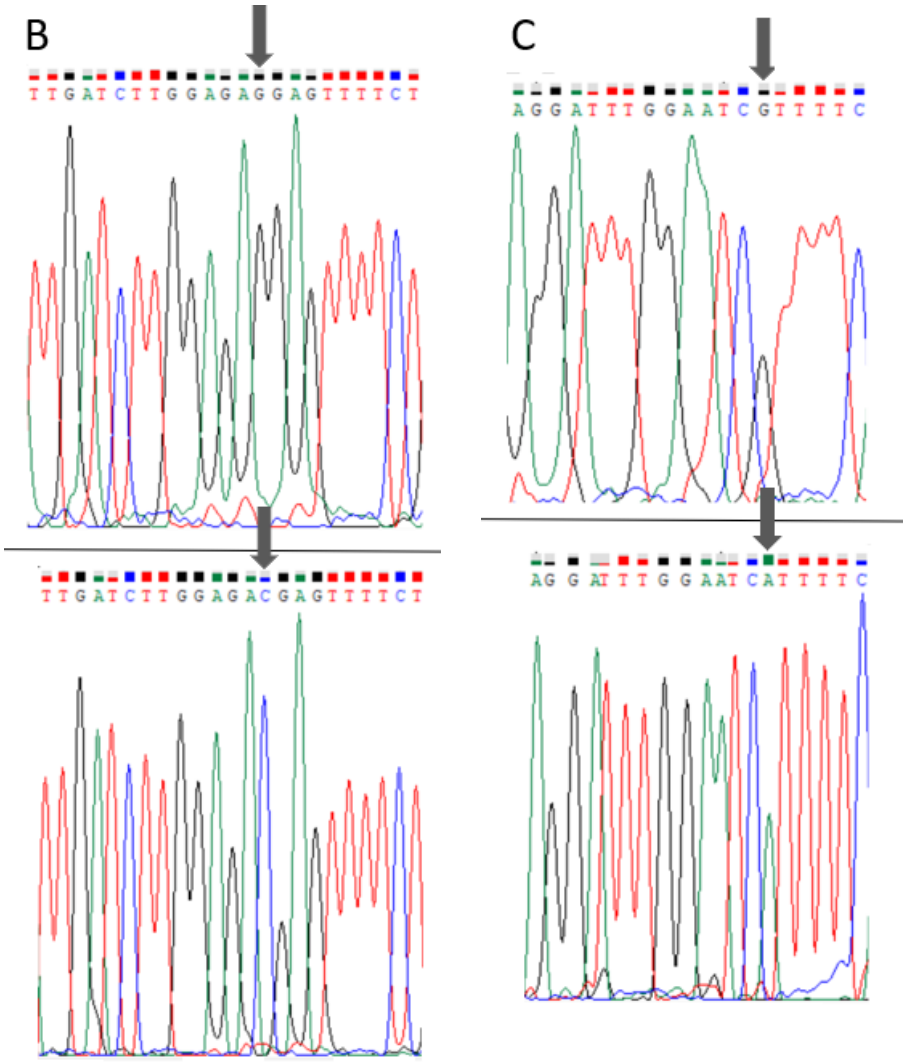


Figure 3.1: Chromatograms illustrating the inserted mutations. (Continued)
The top chromatogram is for the wild type and the bottom chromatogram is for the mutation. B: c.G276C, p.E92D, C: c.G337A, p.V113I.

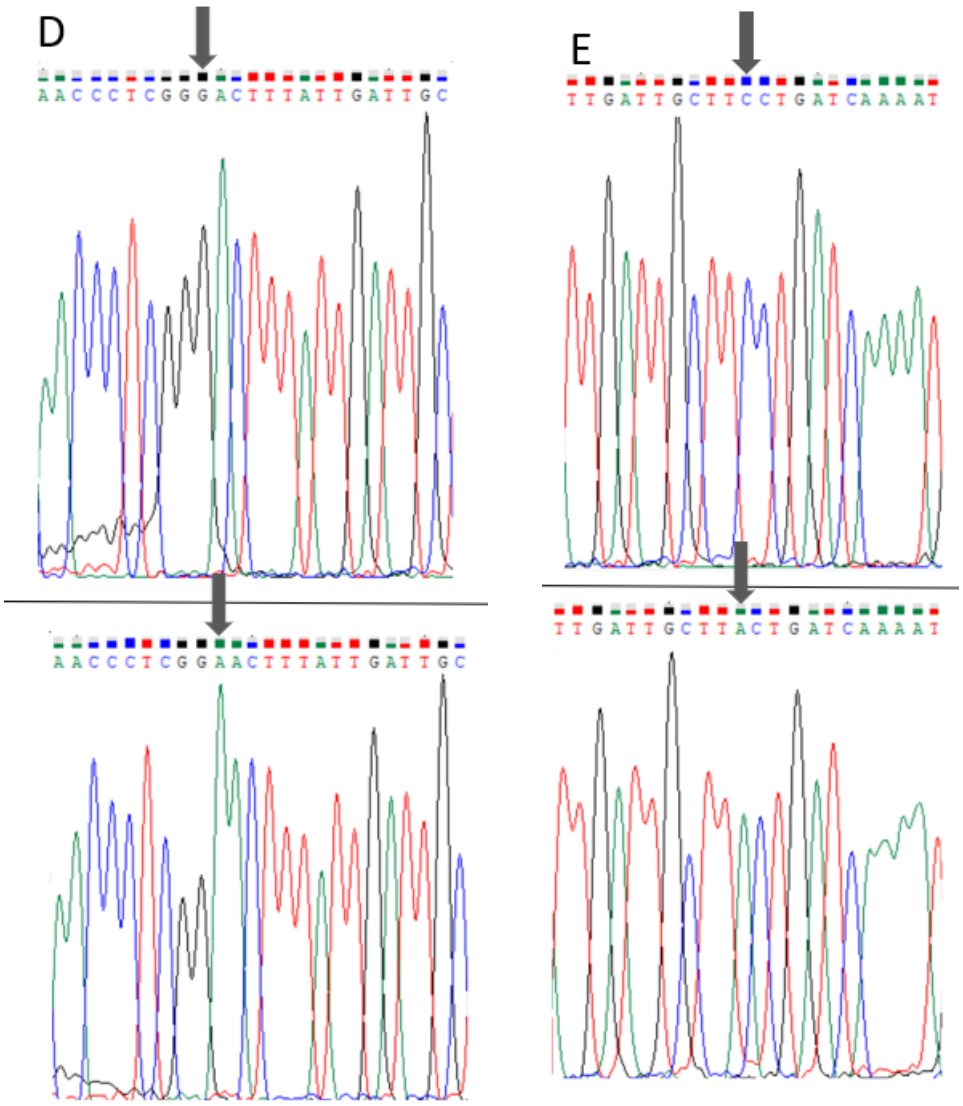


Figure 3.1: Chromatograms illustrating the inserted mutations. (Continued)

The top chromatogram is for the wild type and the bottom chromatogram is for the mutation. D: c.G784A, p.D262N. E: c.C801A, p.F267L.

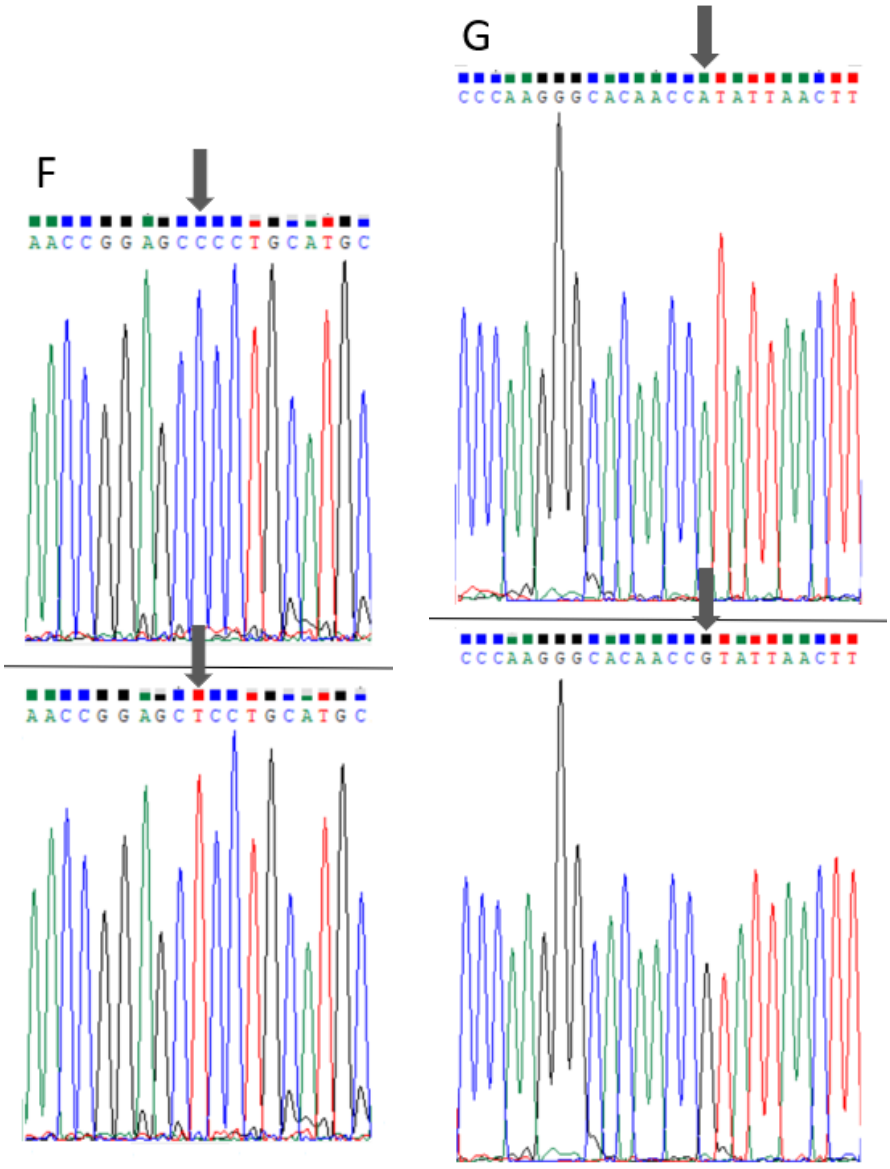


Figure 3.1: Chromatograms illustrating the inserted mutations. (Continued)
The top chromatogram is for the wild type and the bottom chromatogram is for the mutation. F: c.C1009T, p.P337S. G: c.A1159G, p.I387V.

3.1.2 Western blot confirmed that the mutant proteins are expressed in transfected mammalian cell lines

In order to confirm that the wild type and mutants are expressible in mammalian cells, they were transfected into two cell lines (COS-1 and Hek293T) with the various vectors, total lysates were prepared from the expressing cells, and then the expressed protein was analyzed by SDS-PAGE followed by Western blotting to detect the expressed proteins. The bands detected on transfer membranes of lysates from COS-1 cells transfected with pCMV6-XL5 CYP2C19 were probed with anti-CYP2C19

antibodies confirmed the presence of the CYP2C19 protein at the expected size of 55.93 kDa (Figure 3.2). This step was performed to confirm the presence of the CYP2C19 enzyme produced from the mutated gene. β-actin was also detected as a control to rule out human error in loading the samples. The target protein was detected at 42 kDa, which is the expected size. The images produced using the pCMV6-XL5 CYP2C19 human untagged clone transfected into Hek293T cells are shown in Figure 3.3 with similar results except that Hek293T cells seem to express endogenous CYP2C19 as evidenced by the presence of a band in the non-transfected cell samples. The results using the pcDNA3.4 CYP2C19 untagged plasmid are shown in Figures 3.4 and 3.5, showing similar results to pCMV6-XL5.

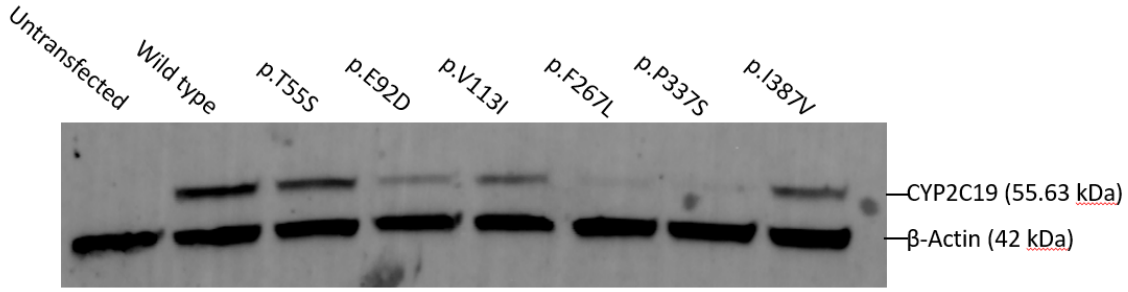


Figure 3.2: Western blot results using the COS-1 cell line.
Transfected using the pCMV6-XL5 CYP2C19 plasmid.

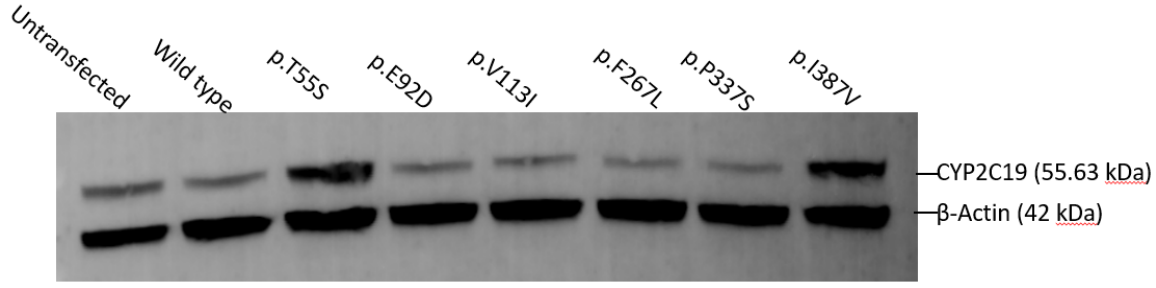


Figure 3.3: Western blot results using the Hek293T cell line.
Transfected using the pCMV6-XL5 CYP2C19 plasmid.

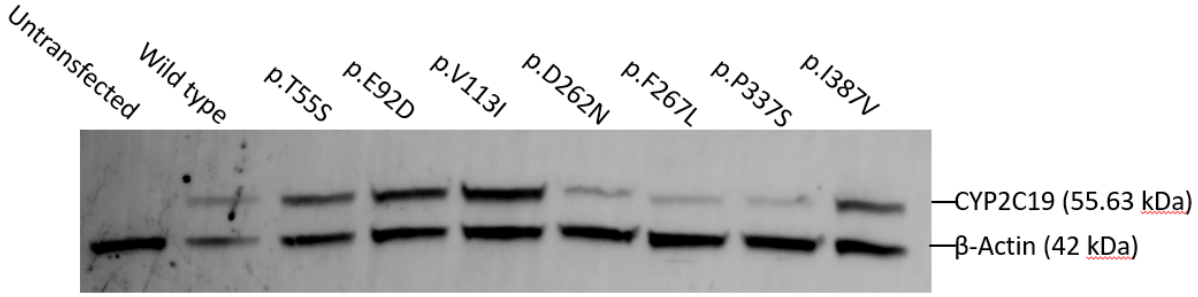


Figure 3.4: Western blot results using the COS-1 cell line.
Transfected using the pcDNA3.4 CYP2C19 plasmid.

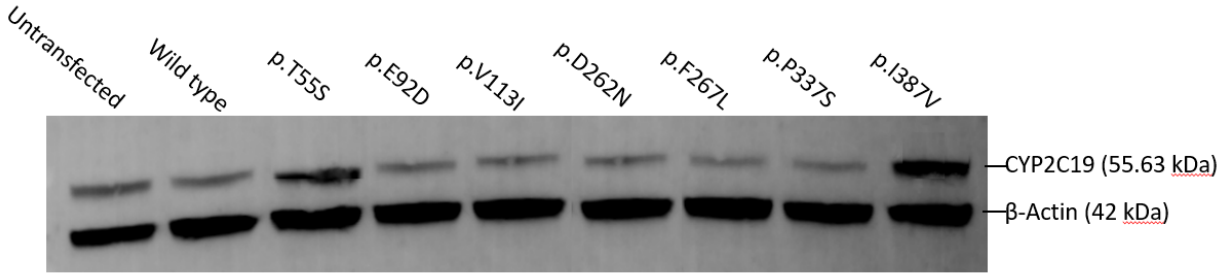


Figure 3.5: Western blot results using the Hek293T cell line.
Transfected using the pcDNA3.4 CYP2C19 plasmid.

The results confirm that no endogenous CYP2C19 protein is in the COS-1 cell line since the un-transfected cells do not have a band. On the other hand, there is endogenous CYP2C19 present in Hek293T cells. Regarding the mutations, apparently p.D262N, p.F267L, and p.P337S variants tend to express lower levels of CYP2C19 compared to the wild type, which may reflect either reduced stability or reduced expression in these samples. In comparison, p.T55S and p.I387V seem to express about the same or higher protein levels. In Figures 3.2, 3.3, and 3.5, the p.E92D and p.V113I variants have lower protein levels while in Figure 3.4, it has a greater protein concentration.

3.1.3 Immunofluorescence microscopy using the anti-CYP2C19 antibodies are inconclusive

To examine the sub-cellular localization of the exogenously expressed proteins, immunofluorescence microscopy on HeLa and COS-1 cells transfected with wild type in the two vectors was performed. The confocal microscopy results are inconclusive

because these antibodies show non-specific staining in all the cells, including the non-transfected cells. This suggests that the used antibodies are not suited for immunofluorescence microscopy. Although the imaged cells were transfected with the wild type and mutants, the data in Figure 3.6 are for the wild type only for illustration purposes. As shown in this Figure, the images appear to be quite grainy. There was a non-specific binding of the antibody since all of the cells are shown in both the wild type image and the Calnexin image. Therefore, conclusions cannot be drawn from this experiment due to the technical unsuitability of the antibodies for immunofluorescence.

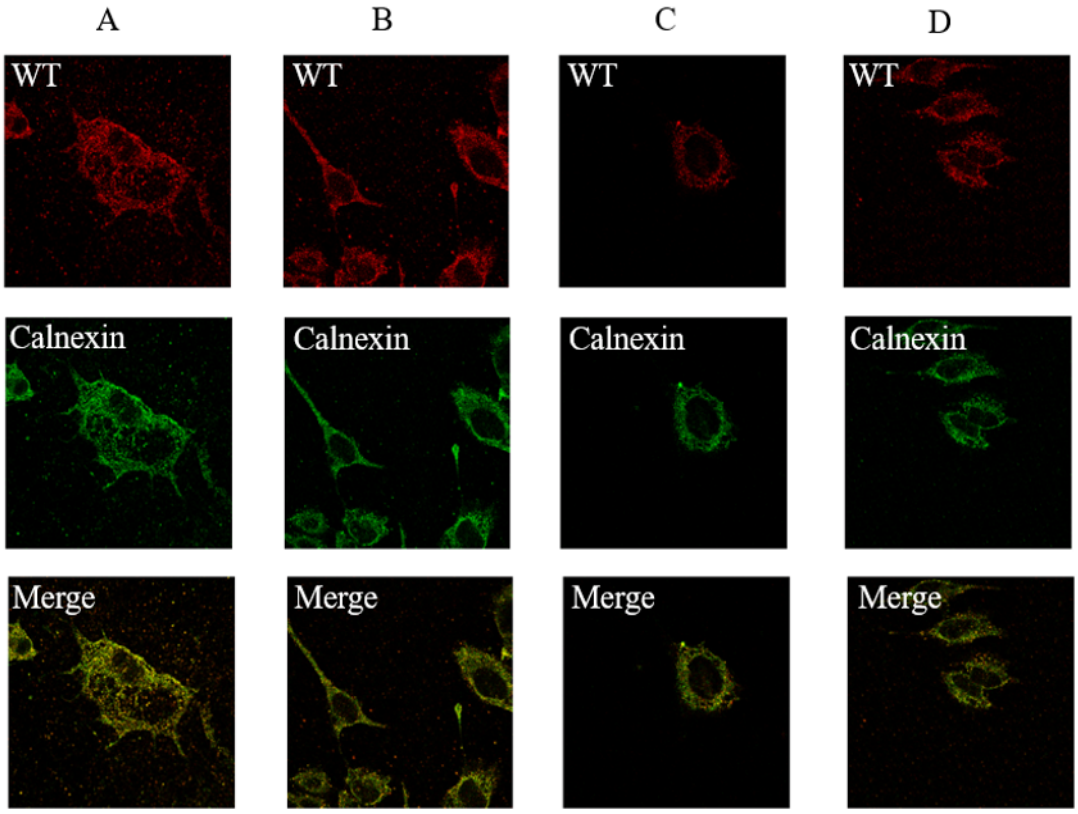
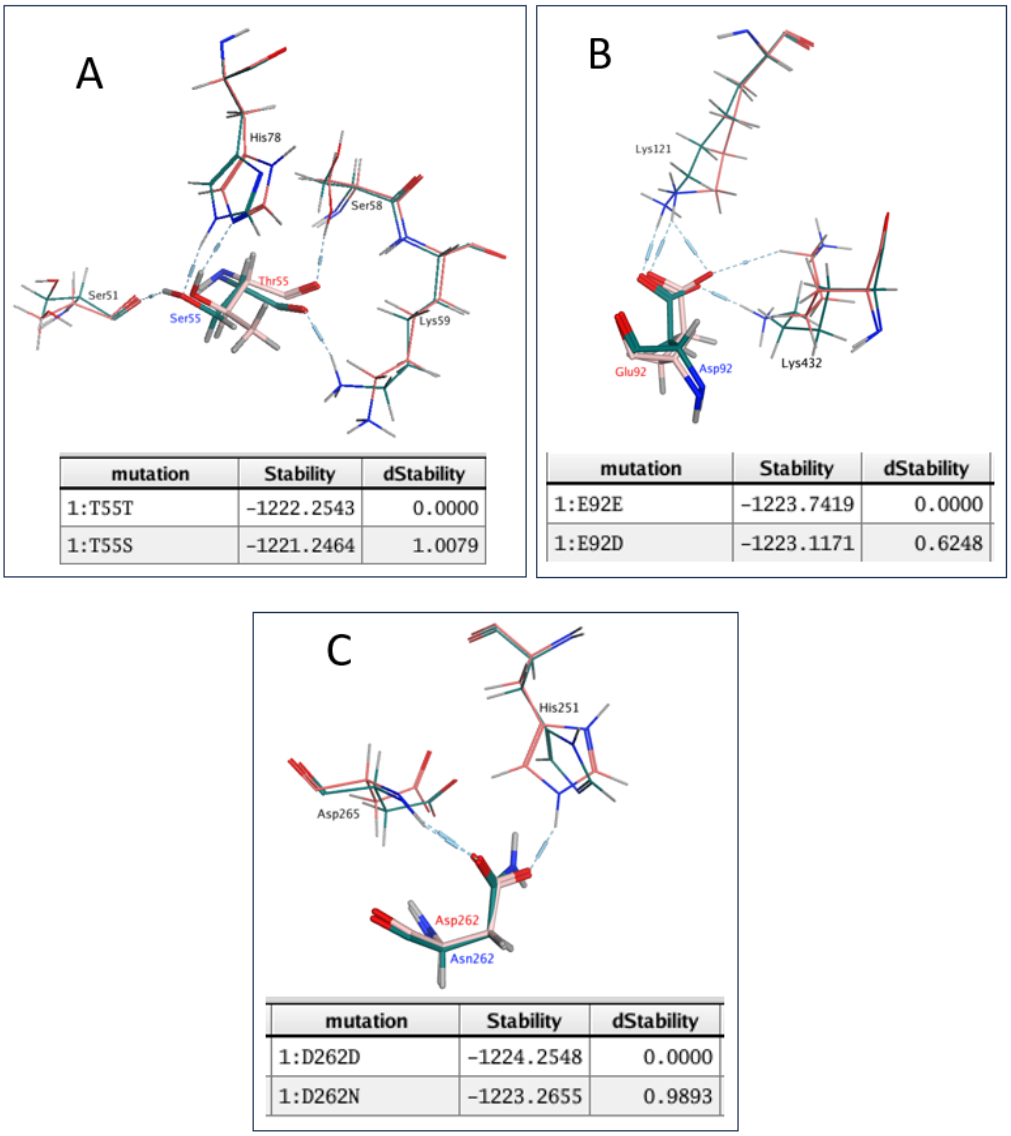


Figure 3.6: Representative immunofluorescence-stained COS-1 and Hela cells transfected with the pCMV6-XL5 CYP2C19 and pcDNA3.4 CYP2C19 plasmids to show CYP2C19 cellular localization.

Column A: COS-1 cells using the pcDNA3.4 CYP2C19 plasmid. Column B: Hela cells using the pcDNA3.4 CYP2C19 plasmid. Column C: COS-1 cells using the pCMV6-XL5 CYP2C19 plasmid. Column D: Hela cells using the pCMV6-XL5 CYP2C19 plasmid.

3.1.4 Molecular modelling

Once the mutations were inserted into the CYP2C19 PDB file obtained from the RCSB Protein Data Bank, the mutations were isolated and viewed in the MOE software. Based on molecular modelling of the studied variants, they are categorized into three groups: variants that lead to a change in bonding (Figure 3.7), variants that result in proteins with bonds similar to the wild type (Figure 3.8), and variants with no profound effect on the bonds (Figure 3.9). The stability scores are also shown in the Figures. If the Stability score is negative, then that means that the protein is sTable. The more positive the dStability score is, the less sTable it is compared to the wild type. All of the mutations were still sTable, according to the Stability score, however, the dStability score shows that they are all less sTable than the wild type.



mutation	Stability	dStability
1:T55T	-1222.2543	0.0000
1:T55S	-1221.2464	1.0079

mutation	Stability	dStability
1:E92E	-1223.7419	0.0000
1:E92D	-1223.1171	0.6248

mutation	Stability	dStability
1:D262D	-1224.2548	0.0000
1:D262N	-1223.2655	0.9893

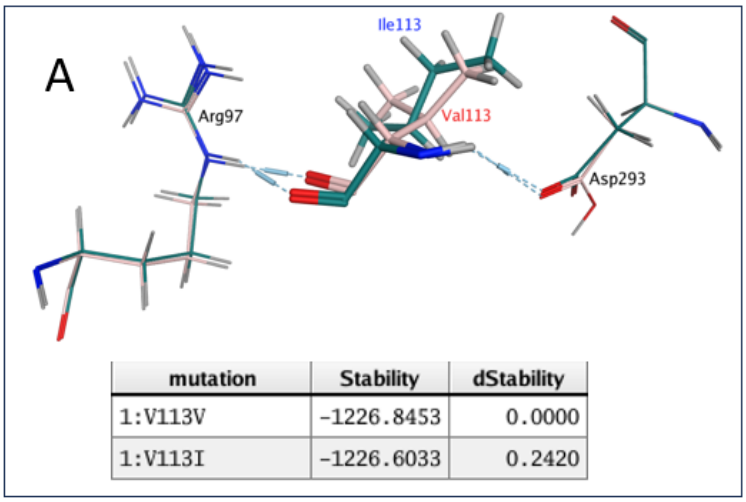
Figure 3.7: Molecular models of variants affecting the intramolecular bonding.

These images represent the superimposed models of wild type and mutation for comparison. The wild type is depicted as the pink molecule and the mutation is depicted as the green molecule.

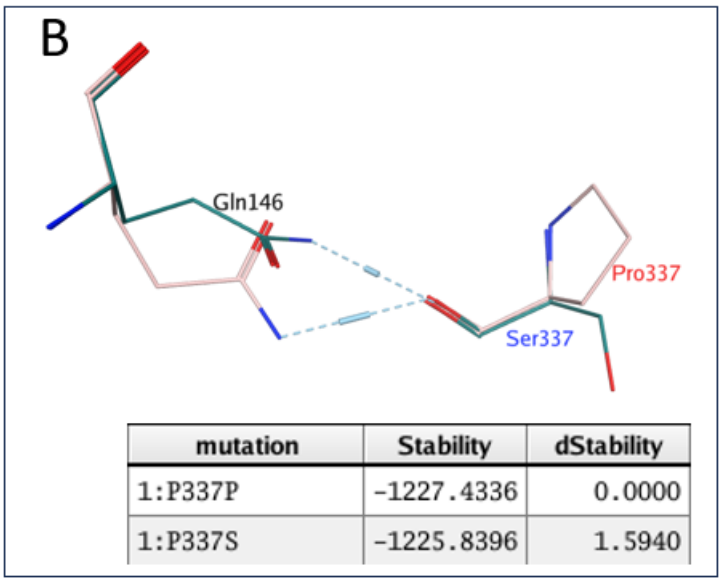
A: c.C164G, p.T55S. Threonine55 is connected by hydrogen bonds to Serine58 and Histadine78. Serine55 is connected by hydrogen bonds to Serine51, Lysine59, and Histadine78.

B: c.G 276C, p.E92D. Glutamic acid92 is connected by two hydrogen bonds to Lysine121 and 1 bond to Lysine432. Aspartic acid92 is connected by hydrogen bonds to Lysine121 and Lysine432.

C: c.G784A, p.D262N. Aspartic acid 262 is connected by hydrogen bonds to Histadine251 and Aspartic acid265. Asparagine262 is connected by a hydrogen bond to Aspartic acid265.



mutation	Stability	dStability
1:V113V	-1226.8453	0.0000
1:V113I	-1226.6033	0.2420



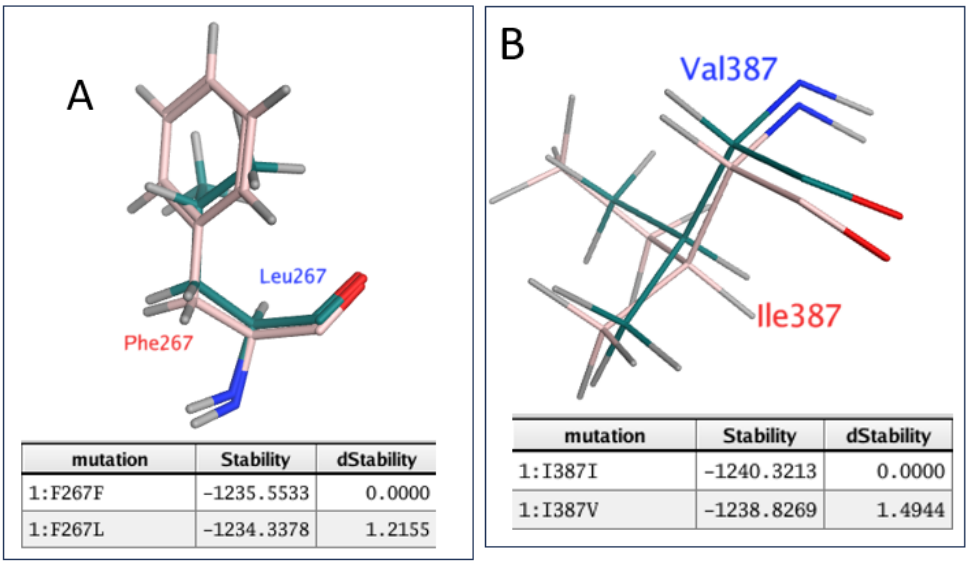
mutation	Stability	dStability
1:P337P	-1227.4336	0.0000
1:P337S	-1225.8396	1.5940

Figure 3.8: Molecular models of variants with the same bonds between the wild type and variations.

These images represent the superimposed models of wild type and mutation for comparison. The wild type is depicted as the pink molecule and the mutation is depicted as the green molecule.

A: c.G337A, p.V113I. Valine113 is connected by hydrogen bonds to Arginine97 and Aspartic acid293. Isoleucine113 is connected by hydrogen bonds to Arginine97 and Aspartic acid293.

B: c.C1009T, p.P337S. Proline337 is connected by a hydrogen bond to Glutamine146. Serine337 is connected by a hydrogen bond to Glutamine146.



mutation	Stability	dStability
1:F267F	-1235.5533	0.0000
1:F267L	-1234.3378	1.2155

mutation	Stability	dStability
1:I387I	-1240.3213	0.0000
1:I387V	-1238.8269	1.4944

Figure 3.9: Molecular models of variants without bonds affecting the folding of the protein.
A: c.C801A, p.F267L. B: c.A1159G, p.I387V.

3.1.5 Enzyme activity assay

The enzyme activity results are shown in time-lapse graphs, with the start of the reaction marked at 0 minutes. The end of the reaction is either 30 or 60 minutes. The results were calculated by subtracting the reading with the inhibitor from the one without it. Due to the results not presenting as the kit stated they would, multiple changes were made to attempt to optimize this kit. These changes included the use of different cell lines for expression, different protein concentrations, changing the lysis buffer, and using the recombinant CYP2C19 enzyme provided in the kit. All of the results from the listed changes were the same. These results are shown in graphs in Figures 3.10-3.17.

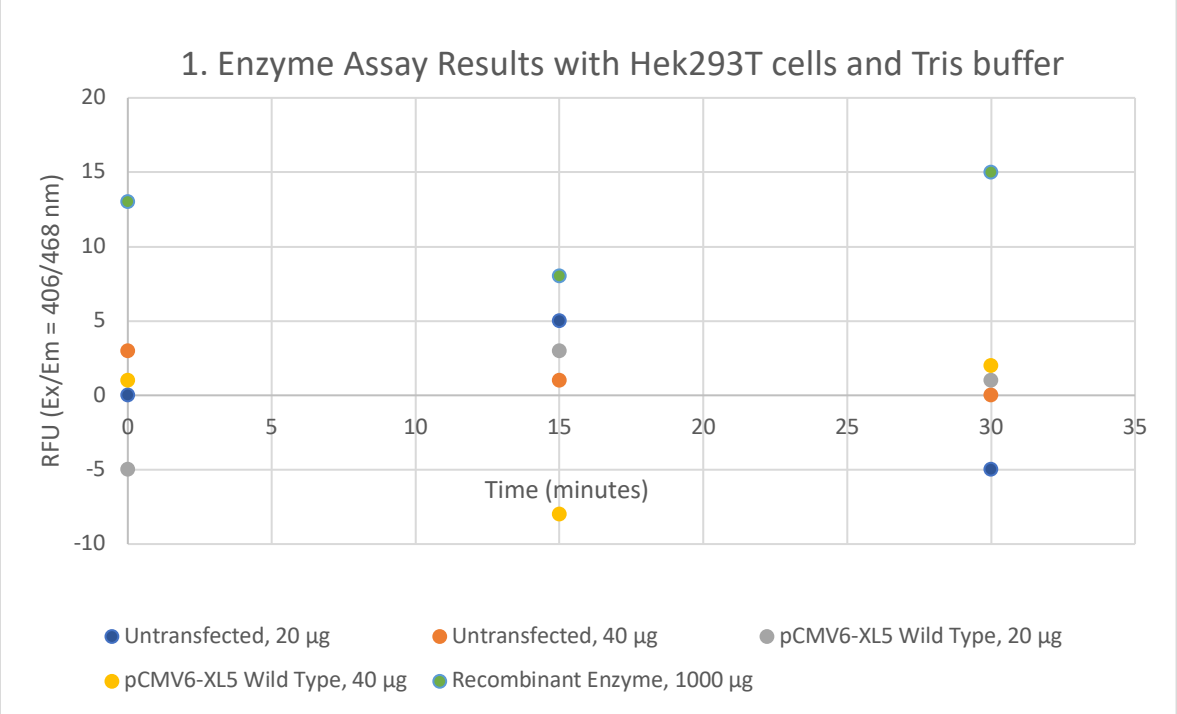


Figure 3.10: This graph shows the results received with Hek293T cells lysed using Tris buffer with TritonX.

The concentrations of protein used were 20 and 40 µg.

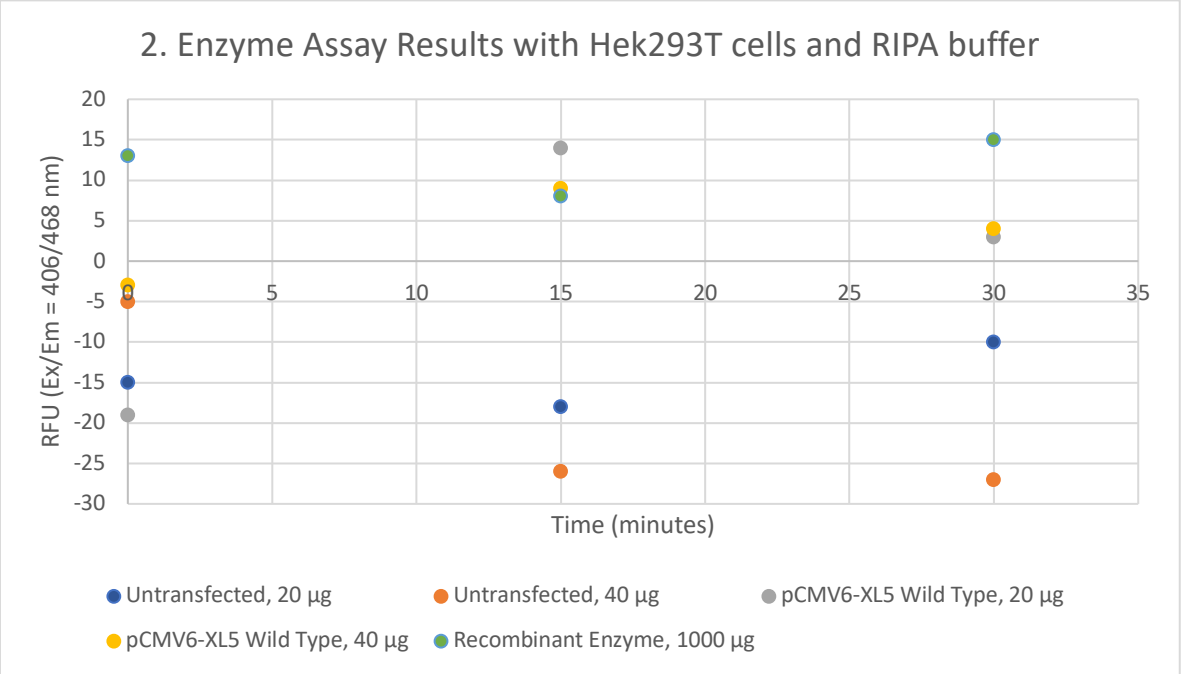


Figure 3.11: This graph shows the results received with Hek293T cells lysed using RIPA buffer.

The concentrations of protein used were 20 and 40 µg.

These were done to determine which lysis buffer should be used. The results were inconclusive. For the remainder of the results, Tris buffer with TritonX was used due to other literature citing that this buffer was used (Devarajan et al., 2019).

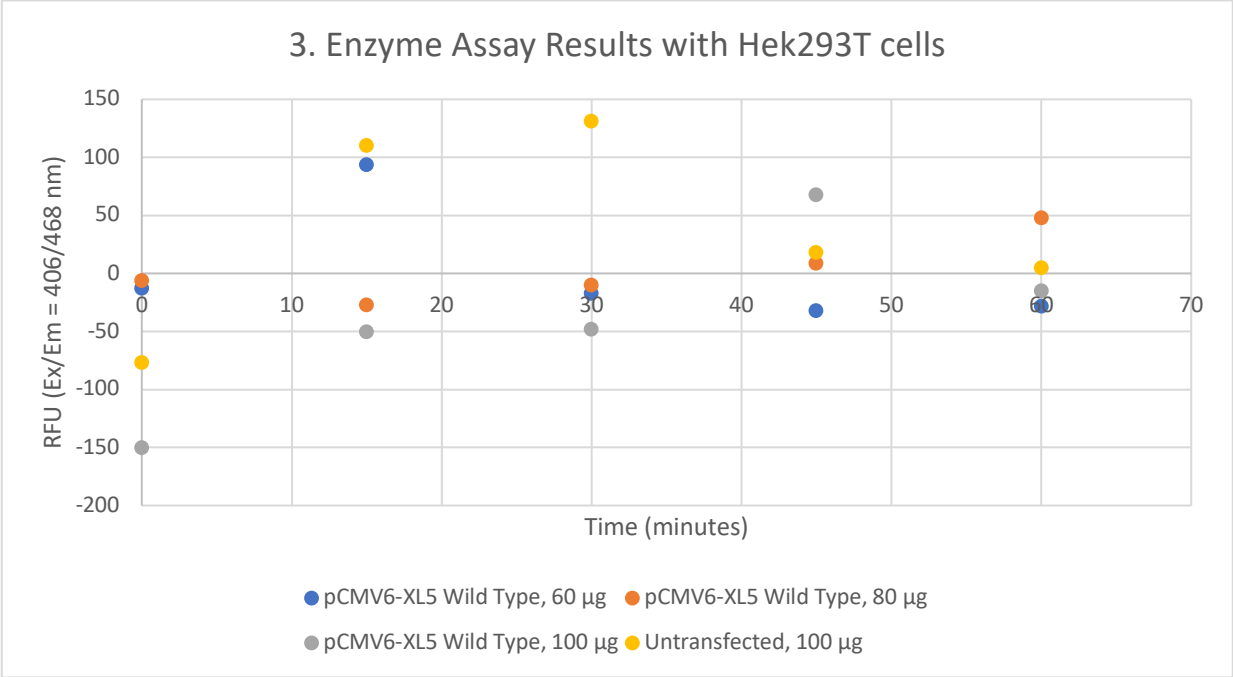


Figure 3.12: This graph shows the results using Hek293T cells.
The protein concentration was increased to 60, 80, and 100 µg.

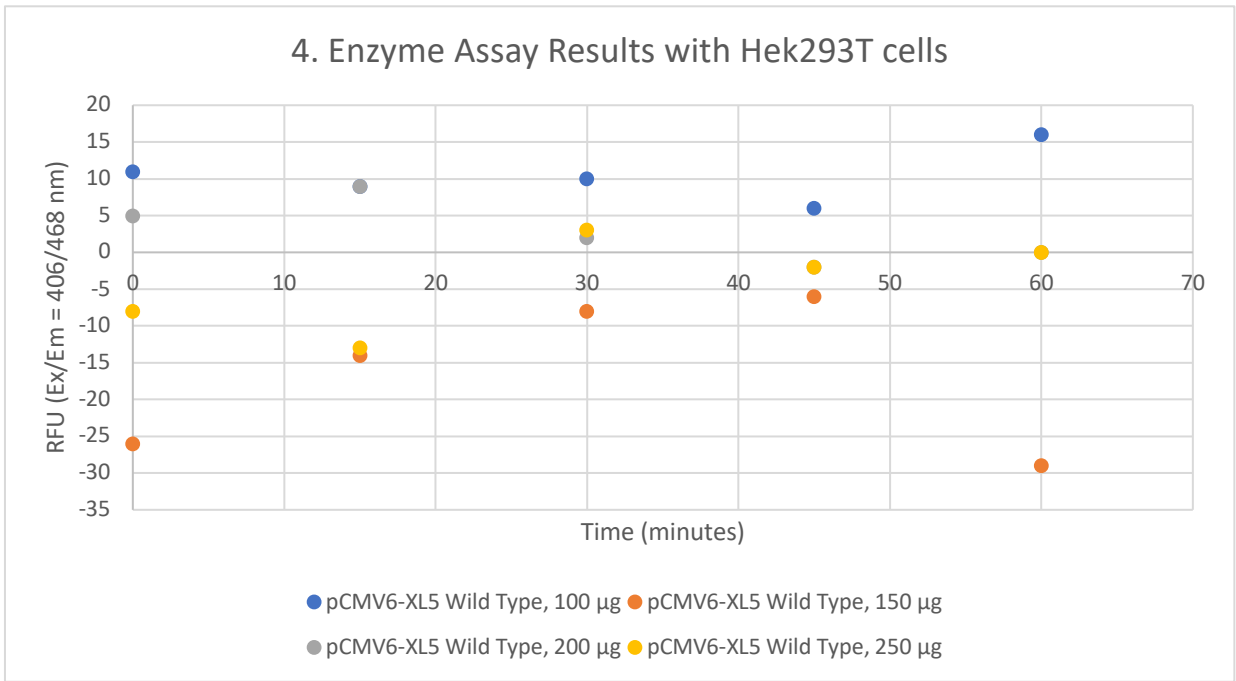


Figure 3.13: This graph shows the results using Hek293T cells.
The protein concentration was increased further to 100, 150, 200, and 250 µg.

The concentration of CYP2C19 in the samples were increased since it appeared that there was not enough protein to have the reagents in the Abcam kit react properly. The suggested concentration in the protocol was between 20 and 100 µg. However, it still does not seem to work properly.

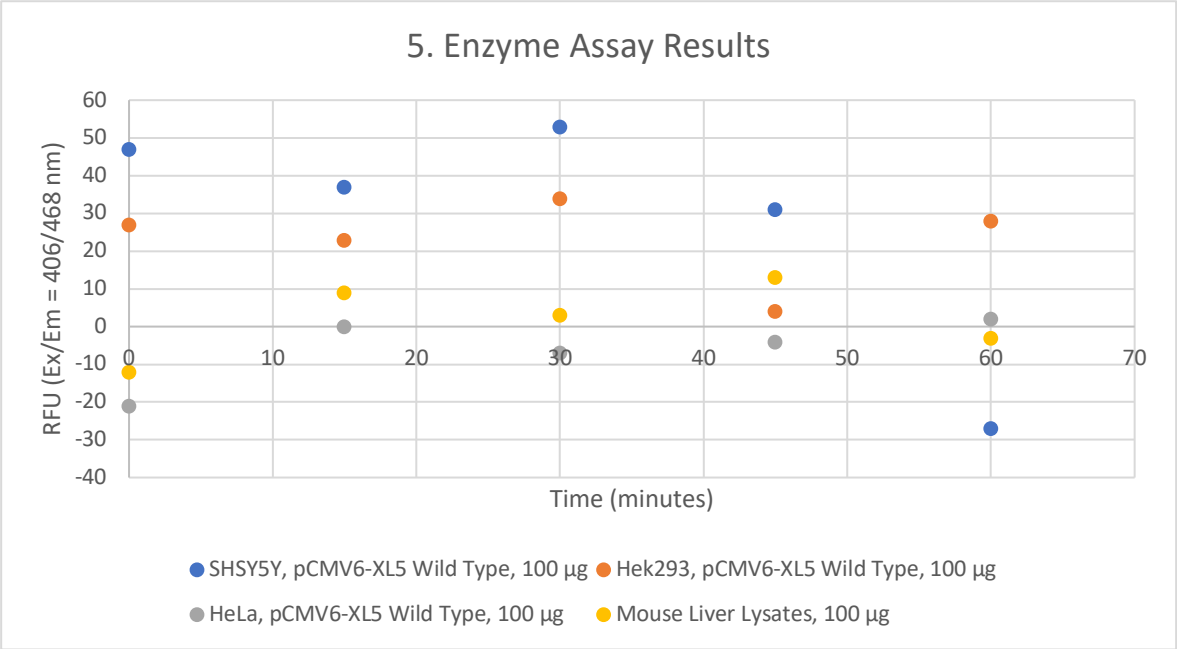


Figure 3.14: This graph shows the results using SHSY5Y, Hek293, and HeLa cells, as well as mouse liver lysates.

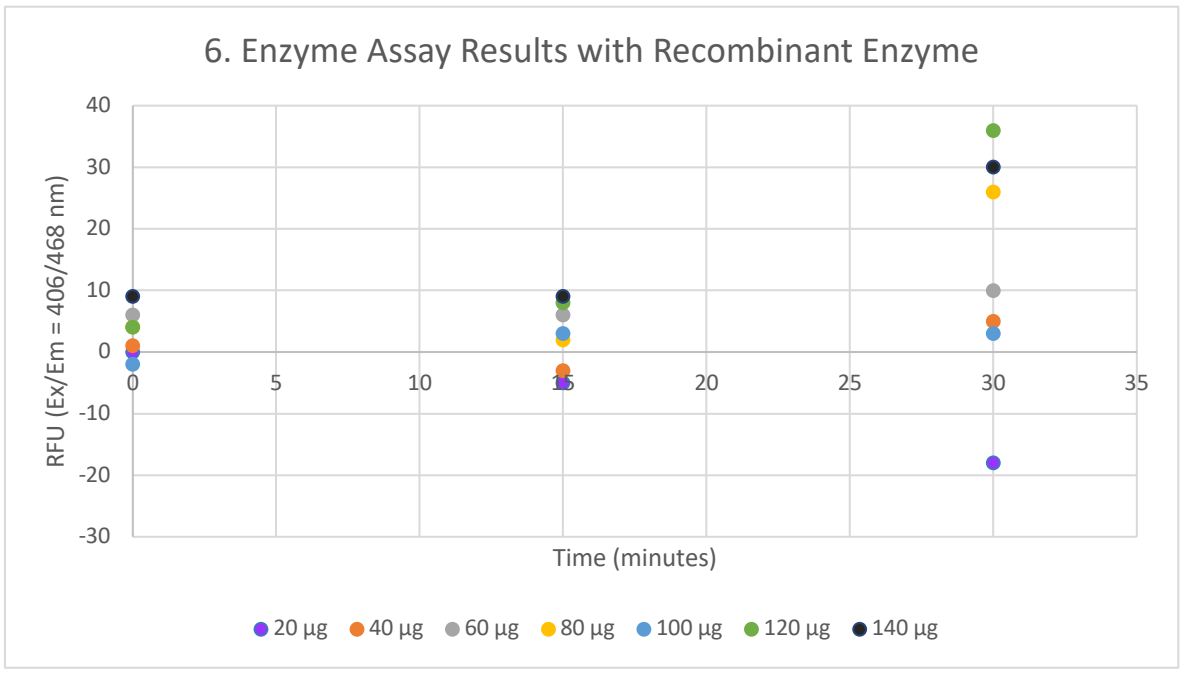


Figure 3.15: This graph shows the results using only the recombinant enzyme received in the kit, which was given to be used as a positive control.

Different cell lines were used to test if that was the issue causing the kit not to work, but the results remained the same as before. The positive control was also used with varying protein concentrations. The results were still not satisfactory and inconclusive.

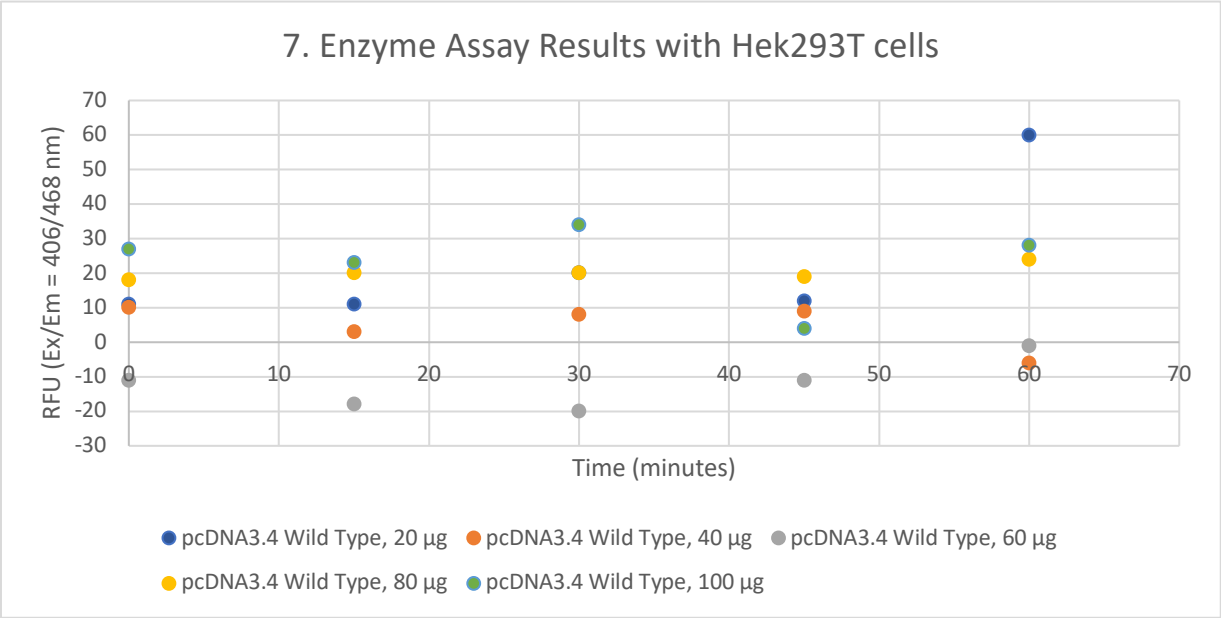


Figure 3.16: This graph shows the results using Hek293T cells with increasing concentrations from 20 to 100.

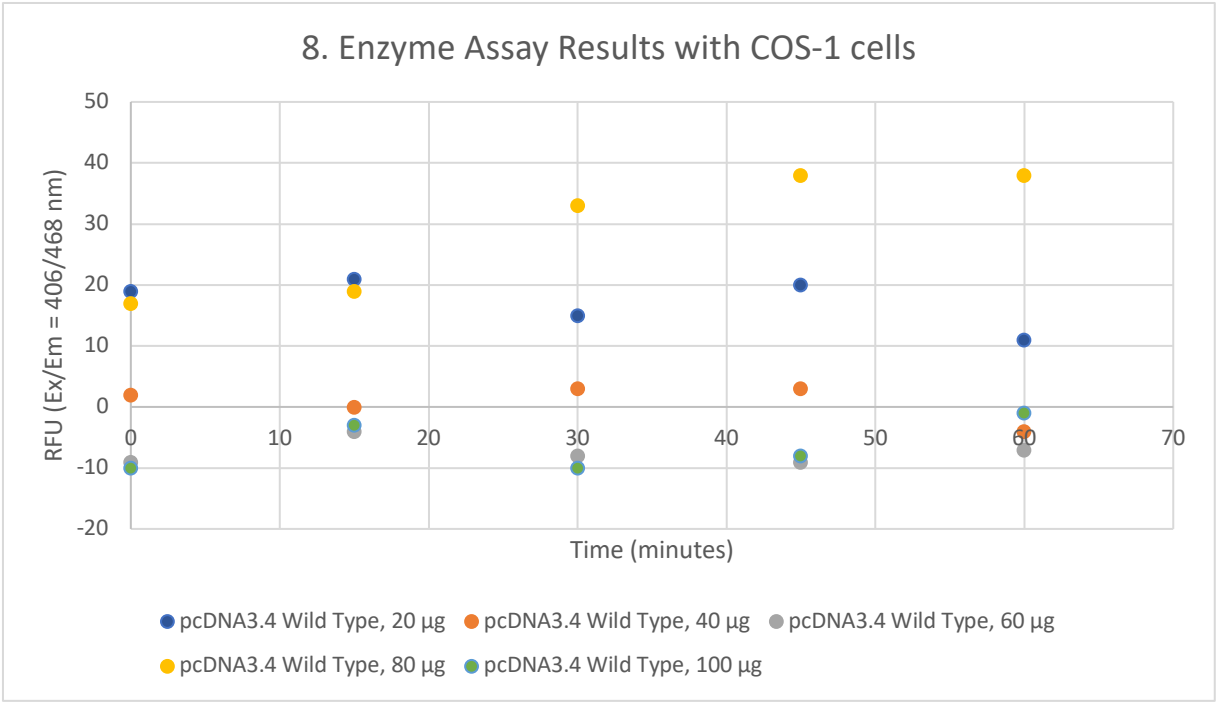


Figure 3.17: This graph shows the results using COS-1 cells with increasing concentrations from 20 to 100.

A different plasmid was used to determine if there could be an issue with the previously used one, as well as another cell line. The results still appear to remain the same.

3.2 The testing of an assay to determine CYP2D6 copy numbers in patients requiring antidepressants

3.2.1 Gel electrophoresis

The PCR results obtained from DNA extracted from patient's blood samples were run on agarose gels to determine if a *CYP2D6* structural variation is present, which could influence the patient's CYP2D6 metabolizer status. Positive and negative controls with duplications and deletions were also used. Each test set produced a PCR product that indicates carrying the deletion in one allele, two alleles, or none of the alleles, i.e., wild type for the deletion (referred to here as wild type deletion). The same applies to duplication, where the expected band representing the PCR product will be either compatible with the absence (i.e., wild type duplication) or presence of the duplication, however in a non-quantitative method (i.e., shows there are more than one copy of the gene without discriminating the number of extra copies present). The gels were viewed under UV light. Repeats were also done to confirm the results. A representative gel is shown in Figure 3.18.

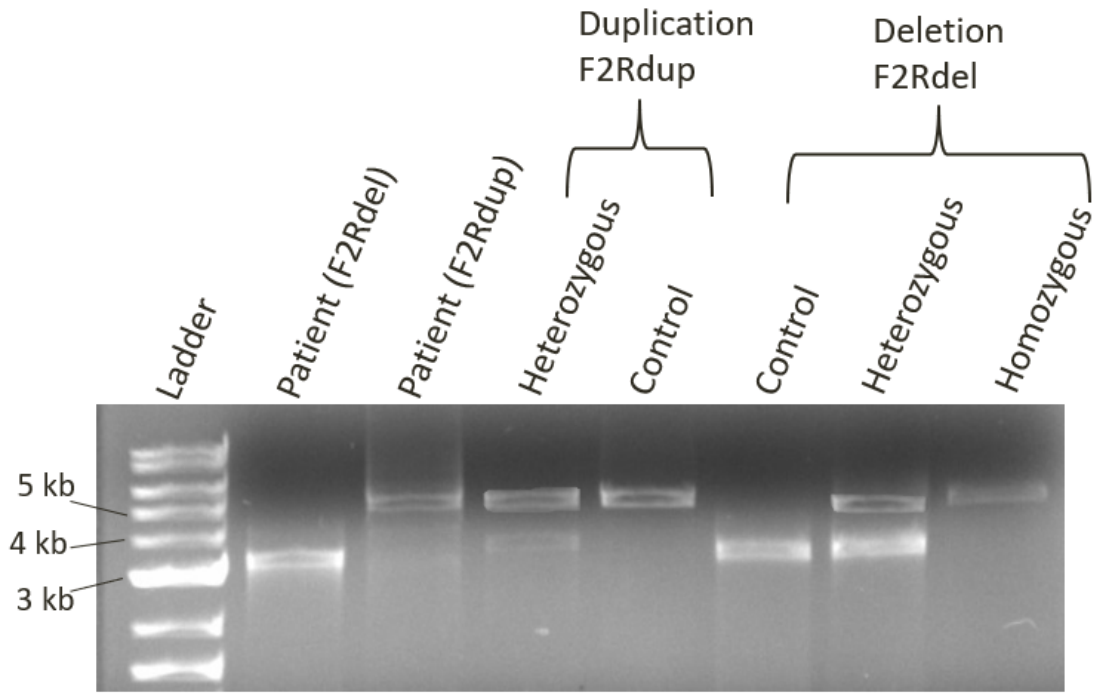


Figure 3.18: Representative agarose gel results.
This patient's (Ps015) results show that they do not contain a duplication or deletion.

The analysis of 15 samples revealed one carrier of a homozygous whole gene deletion (CYP2D6*5/*5), and one carrier of a heterozygous deletion. The rest of the samples (N=13) had bands representing carrying homozygous wild type. These gels prove that this PCR-based technique can be used to detect duplication and deletion in *CYP2D6*. The results of all 15 samples are shown in Table 3.1.

Table 3.1: Copy number variation results of all 15 patients.

Patient ID	Deletion (+/-)	Duplication (+/-)
Ps001	-	-
Ps002	-	-
Ps003	-	-
Ps004	-	-
Ps005	+	-
Ps006	-	-
Ps007	-	-
Ps008	-	-
Ps009	-	-
Ps010	-	-
Ps011	-	-
Ps012	-	-
Ps013	-	-
Ps014	+	-
Ps015	-	-

Chapter 4: Discussion

4.1 CYP2C19

Cytochrome P450s are a superfamily of enzymes that play both major and minor roles in numerous drugs' metabolism. As such, they tend to be extensively studied in pharmacogenomics. The two members of this family focused on within this thesis are *CYP2C19* and *CYP2D6*, which are responsible for metabolizing between 20% and 70% of all clinically used drugs (Ivanov et al., 2022). The current work aimed at two challenging topics in pharmacogenomics: (1) evaluating the effect of rare and population-specific variants in the very important pharmacogene *CYP2C19*, and (2) establishing an efficient and time-saving methodology for detecting troublesome structural variants (duplication and deletion) in *CYP2D6*.

Rare and novel variants contribute to significant differences in drug pharmacokinetics and are highly population-specific (Zhou & Lauschke, 2021; Siamoglou et al., 2022). The variants chosen to be studied in *CYP2C19* are c.C164G (p.T55S), c.G276C (p.E92D), c.G337A (p.V113I), c.G784A (p.D262N), c.C801A (p.F267L), c.C1009T (p.P337S), and c.A1159G (p.I387V). These variants were detected in the Emirati population and suggested by *in silico* analysis to have potential detrimental effects on the enzyme (Al-Mahayri et al., 2020).

In order to study the target variants, they were generated in two mammalian expression plasmids using site-directed mutagenesis (SDM). This method was chosen since it is a well-established technique used to introduce mutations into cDNA. It is also documented as being used for experiments with CYP2C19. In 2007, Hanioka and colleagues used SDM to generate variants that were detected in the Japanese population so that they could determine the effects they have on enzyme activity (Hanioka et al., 2007). For Devarajan and coworkers, this technique was used to replicate rare variants detected in *CYP2C9* and *CYP2C19* in the RIGHT study. The RIGHT study was performed by Mayo Clinic using 1,013 subjects to investigate the effectiveness of genetic testing so that prescriptions could be personalized for each patient. The study was dubbed the RIGHT study since the goal was to be able to prescribe patients with the right drug at the right dosage (Bielinski et al., 2014). In order to be able to study these

variants in depth, the SDM technique was implemented so that they could be tested *in vitro* (Devarajan et al., 2019). In another case, star alleles that were already discovered were studied by Takahashi and colleagues to confirm the effects that these amino acid changes would have on the enzyme's ability to metabolize clopidogrel. In order to study the 21 *CYP2C19* variants chosen, SDM was used to generate them (Takahashi et al., 2015). In the present study, SDM was successfully used to generate the seven target variants and Sanger sequencing proved this.

The cell lines used in this project are COS-1, Hek293T, and HeLa cells. COS-1 cells are African green monkey kidney cells. Hek293T cells are human embryonic kidney. HeLa cells are cervical cancer cells. The COS-1 and Hek293T cells were chosen for studying CYP2C19 enzyme activity since there is a precedent for these lines. There was an experiment conducted by Dai and colleagues in which they compared the protein expression levels, as well as *in vitro* functional analyses using CYP variants. The cell lines used with previously reported use in CYP experiments were Hek293T, HepG2, and COS-7, which is a variant of COS-1. A Hek293 cell variant, 293FT, was also introduced to determine its viability in CYP research. It was confirmed that COS-7 and Hek293T cells were the best lines to use from the established cell lines. It was also shown that 293FT cells showed high levels of protein expression and could be viable options for use in future CYP research (Dai et al., 2015). COS-1 cells are typically used when studying CYP2C19 due to it being relatively easy to culture and not endogenously expressing CYP2C19 (Devarajan et al., 2019; Goldstein et al., 1994; Gonzalez & Korzekwa, 1995). Similar to COS-1, Hek293T cells are also easy to maintain in cell culture and also have high transfection efficiency. This makes them ideal for producing variant enzymes and measuring their activity (Dai et al., 2015; Zhang et al., 2020).

According to the literature, the variants c.C164G (p.T55S) and c.G276C (p.E92D) were predicted as having no effect on the enzyme's activity based on *in silico* tests (Zhang et al., 2020). For variant c.G276C, there have also been enzyme activity assays performed that confirm this (Wang et al., 2011a, 2011b). The molecular modelling images for both variants showed that there was a difference in the bonds made compared to the wild type, but the Western blot bands were different for each. For p.T55S, it showed similar protein expression levels to the wild type. On the other hand, p.E92D

mostly had lower expression. For variant c.G337A (p.V113I), *in silico* tests performed predict that this variant would lead to a normal enzyme (Devarajan et al., 2019; Matimba et al., 2009; Zhang et al., 2020). While the Western blot band appears similar to the wild type, two enzyme activity assays were performed in previous literature that showed that this variant actually decreased enzyme activity. One assay was a modification of the Vivid CYP2C19 Blue Screening Kit (ThermoFisher Scientific, Waltham, Massachusetts, U.S.A.) that showed that the enzyme decreased in efficacy by about 20%. The second one was a liquid chromatography/mass spectrometry assay using *S*-mephenytoin as a substrate. The result of this showed a decrease in enzyme activity of about 90%. This discrepancy could be attributed to the substrate specificity of this enzyme, which implies that this variant may alter the binding site (Devarajan et al., 2019). Unlike the Western blot obtained from the published literature, the expression overall appeared to be lower than the wild type. For both c.G784A (p.D262N) and c.C801A (p.F267L), previously performed *in silico* testing predicted that they would damaging to the resulting enzyme. Western blotting showed that there was significantly less protein expressed for variant p.F267L compared to the wild type (Zhang et al., 2020). This result was also achieved in the Western blot performed in this project for p.F267L and showed that p,D262N also had lower protein expression. The remaining two variants, c.C1009T (p.P337S) and c.A1159G (p.I387V) were not mentioned in any previously published literature. The Western blot results showed that p.P337S had a lower expression level, similar to p.D262N and p.F267L, while p.I387V had the same expression as the wild type and p.T55S. A summary of the results is shown in Table 4.1.

Table 4.1: Summary of results obtained using Western blotting and molecular modelling compared to previously published literature.

Mutation	Western blot	Molecular modelling	Literature results	ANNOVAR
c.C164G p.T55S	Same or higher expression	Affected intramolecular bonding	No effect (Zhang et al., 2020)	Detrimental
c.G276C p.E92D	Varying expression	Affected intramolecular bonding	No effect (Wang et al., 2011a, 2011b; Zhang et al., 2020)	Detrimental
c.G337A p.V113I	Varying expression	Similar intramolecular bonding compared to the wild type	Decreased enzyme activity (Devarajan et al., 2019) No effect (Matimba et al., 2009; Zhang et al., 2020)	Detrimental
c.G784A p.D262N	Lower expression	Affected intramolecular bonding	Decreased enzyme activity (Zhang et al., 2020)	Detrimental
c.C801A p.F267L	Lower expression	No effect	Decreased enzyme activity (Zhang et al., 2020)	Detrimental
c.C1009T p.P337S	Lower expression	Similar intramolecular bonding compared to the wild type	No previously published literature	Detrimental
c.A1159G p.I387V	Same or higher expression	No effect	No previously published literature	Detrimental

4.2 CYP2D6

CYP2D6 is a very important pharmacogene that tends to influence psychiatry medications, such as antidepressants. This type of medication can have very serious repercussions to prescribing the wrong dosage, or even the wrong medication. In order to decrease the likelihood of this happening, genetic testing can be done to determine if there are any variations that could lead to a difference in enzyme activity. For this project, *CYP2D6* gene deletion (*CYP2D6**5) and duplication were focused on in

personalizing prescriptions for psychiatry patients with antidepressants. This part of the study aimed to establish a time-efficient method to detect these common structural variants.

To date, the *CYP2D6**5 allele has only been investigated in five countries within the Middle East. The frequency of this allele was between 0.98% and 3.3%. While this is a low number, there have been too few studies regarding this allele in the Middle East for it to be disregarded. This is an important allele to investigate since it is considered a null allele, which means it results in a nonfunctional CYP2D6 enzyme (Alali et al., 2022; Qumsieh et al., 2011). If a prescribed medication is primarily metabolized by CYP2D6, then a nonfunctional enzyme would be especially dangerous to the patient since it could lead to adverse drug reactions (ADRs), such as insomnia, headaches, and gastrointestinal dysfunction (Bousman et al., 2023). There are many commonly prescribed antidepressants that have CYP2D6 play a major role in its metabolic pathway or that act as CYP2D6 inhibitors. Its importance in drug metabolism is evident enough that there are CPIC guidelines for dosages, as well medications to avoid prescribing when a mutation in *CYP2D6* is detected (Gressier et al., 2015; Milosavljević et al., 2023; Nofziger et al., 2019). Due to the fact that CYP2D6 influences the efficacy of many commonly prescribed drugs, it is vital to create efficient methods for testing the presence of variations in *CYP2D6*.

Chapter 5: Conclusions

5.1 Summary

Pharmacogenomics is a relatively new field of study, especially in the Middle East. Two very important pharmacogenes belong to the cytochrome P450 superfamily, which are CYP2C19 and CYP2D6. Seven rare variants found in the Emirati population in CYP2C19 were focused on to determine the effects that they have on enzyme efficacy. Unfortunately, the kit used did not yield any accurate results. In CYP2D6, a PCR-based approach was used to detect a deletion and duplication in order to tailor medication concentrations to individual patients being treated with antidepressants. The results of this showed that this is a promising technique to be used in this type of genetic testing. Overall, pharmacogenomics research and personalized medicine should be implemented more in the Middle East in order to improve the quality of life for its citizens.

5.2 Future experiments

A different enzyme activity assay should be performed to determine the effects that these variants have on the resulting enzyme. The one that will be used here is the ThermoFisher Scientific Vivid CYP2C19 Blue Kit (ThermoFisher Scientific, Waltham, Massachusetts, U.S.A.). In addition to this, the Abcam Microsome Isolation Kit (Abcam, Cambridge, United Kingdom) will also be used to attempt to isolate the cytochrome P450s from other enzymes in the cell lysates that could be interfering with the enzyme activity readings.

5.3 Limitations

One limitation of this thesis was the use of a commercially available CYP2C19 enzyme activity assay. There appeared to be an issue with the sensitivity of the readings using this kit. Further optimization is required, or the use of other (possibly using home-made reagents) assays.

References

Al-Ahmad, M. M., Amir, N., Dhanasekaran, S., John, A., Abdulrazzaq, Y. M., Ali, B. R., & Bastaki, S. M. A. (2017). Genetic polymorphisms of cytochrome P450-1A2 (CYP1A2) among Emiratis. *PLOS ONE*, *12*(9), e0183424. https://doi.org/10.1371/journal.pone.0183424

Al-Ahmad, M. M., Amir, N., Dhanasekaran, S., John, A., Abdulrazzaq, Y. M., Ali, B. R., & Bastaki, S. (2017). Studies on N-Acetyltransferase (NAT2) Genotype Relationships in Emiratis: Confirmation of the Existence of Phenotype Variation among Slow Acetylators. *Annals of Human Genetics*, *81*(5), 190-196. https://doi.org/10.1111/ahg.12198

Alali, M., Ismail Al-Khalil, W., Rijjal, S., Al-Salhi, L., Saifo, M., & Youssef, L. A. (2022). Frequencies of CYP2D6 genetic polymorphisms in Arab populations. *Human Genomics*, *16*(1). https://doi.org/10.1186/s40246-022-00378-z

Al-Jaibeji, H. S., John, A., Al-Gazali, L., Soliman, K., Oulhaj, A., Katsila, T., Brand, A., Patrinos, G., & Ali, B. R. (2016). Allele and genotype frequencies of the two single nucleotide polymorphisms in the VKORC1 gene that are most important for warfarin treatment among Emiratis. *Hamdan Medical Journal*, *9*(1), 75-84. https://doi.org/10.7707/hmj.463

Al-Mahayri, Z. N., Patrinos, G. P., Wattanapokayakit, S., Iemwimangsa, N., Fukunaga, K., Mushiroda, T., Chantratita, W., & Ali, B. R. (2020). Variation in 100 relevant pharmacogenes among emiratis with insights from understudied populations. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-78231-3

Bielinski, S. J., Olson, J. E., Pathak, J., Weinshilboum, R. M., Wang, L., Lyke, K. J., Ryu, E., Targonski, P. V., Van Norstrand, M. D., Hathcock, M. A., Takahashi, P. Y., Mccormick, J. B., Johnson, K. J., Maschke, K. J., Rohrer Vitek, C. R., Ellingson, M. S., Wieben, E. D., Farrugia, G., Morrisette, J. A., … Kullo, I. J. (2014). Preemptive Genotyping for Personalized Medicine: Design of the Right Drug, Right Dose, Right Time - Using Genomic Data to Individualize Treatment Protocol. *Mayo Clinic Proceedings*, *89*(1), 25-33. https://doi.org/10.1016/j.mayocp.2013.10.021

Botton, M. R., Whirl-Carrillo, M., Del Tredici, A. L., Sangkuhl, K., Cavallari, L. H., Agúndez, J. A. G., Duconge, J., Lee, M. T. M., Woodahl, E. L., Claudio-Campos, K., Daly, A. K., Klein, T. E., Pratt, V. M., Scott, S. A., & Gaedigk, A. (2021). PharmVar GeneFocus: CYP2C19. *Clinical Pharmacology & Therapeutics*, *109*(2), 352-366. https://doi.org/10.1002/cpt.1973

Bousman, C. A., Stevenson, J. M., Ramsey, L. B., Sangkuhl, K., Hicks, J. K., Strawn, J. R., Singh, A. B., Ruaño, G., Mueller, D. J., Tsermpini, E. E., Brown, J. T., Bell, G. C., Leeder, J. S., Gaedigk, A., Scott, S. A., Klein, T. E., Caudle, K. E., & Bishop, J. R. (2023). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6, CYP2C19, CYP2B6, SLC6A4, and HTR2A Genotypes and Serotonin Reuptake Inhibitor Antidepressants. *Clinical Pharmacology & Therapeutics*, *114*(1), 51-68. https://doi.org/10.1002/cpt.2903

Cavallari, L., Jeong, & Bress. (2011). Role of cytochrome P450 genotype in the steps toward personalized drug therapy. *Pharmacogenomics and Personalized Medicine*, 123. https://doi.org/10.2147/pgpm.s15497

Citalopram and Escitalopram Pathway, Pharmacokinetics (2021, September 17). PharmGKB. https://www.pharmgkb.org/pathway/PA164713429. Access date: November 21, 2023.

Clopidogrel Pathway, Pharmacokinteics (2019, April 2). PharmGKB. https://www.pharmgkb.org/pathway/PA154424674. Access date: November 21, 2023.

Dai, D. P., Geng, P. W., Cai, J., Wang, S. H., Nic, J. J., Hu, J. H., Hu, G. X., & Cai, J. P. (2015). 293FT is a highly suiTable mammalian cell line for the in vitro enzymatic activity analysis of typical P450 proteins. *Pharmazie*, *70*(1), 33-7. https://doi.org/10.1691/ph.2015.4067

Devarajan, S., Moon, I., Ho, M.-F., Larson, N. B., Neavin, D. R., Moyer, A. M., Black, J. L., Bielinski, S. J., Scherer, S. E., Wang, L., Weinshilboum, R. M., & Reid, J. M. (2019). Pharmacogenomic Next-Generation DNA Sequencing: Lessons from the Identification and Functional Characterization of Variants of Unknown Significance in CYP2C9 and CYP2C19. *Drug Metabolism and Disposition*, *47*(4), 425-435. https://doi.org/10.1124/dmd.118.084269

Fetzner, S., & Steiner, R. A. (2010). Cofactor-independent oxidases and oxygenases. *Applied Microbiology and Biotechnology*, *86*(3), 791-804. https://doi.org/10.1007/s00253-010-2455-0

Frey, P. A., & Hegeman, A. D. (2007). *Enzymatic reaction mechanisms*. Oxford University Press.

Gaedigk, A. (2013). Complexities of CYP2D6 gene analysis and interpretation. *International Review of Psychiatry*, *25*(5), 534-553. https://doi.org/10.3109/09540261.2013.825581

Garfinkel, D. (1958). Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Archives of Biochemistry and Biophysics*, *77*(2), 493-509. https://doi.org/10.1016/0003-9861(58)90095-X

Goldstein, J. A., & de Morais, S. M. (1994). Biochemistry and molecular biology of the human CYP2C subfamily. *Pharmacogenetics*, *4*(6), 285-299. https://doi.org/10.1097/00008571-199412000-00001

Goldstein, J. A., Faletto, M. B., Romkes-Sparks, M., Sullivan, T., Kitareewan, S., Raucy, J. L., Lasker, J. M., & Ghanayem, B. I. (1994). Evidence That CYP2C19 is the Major (S)-Mephenytoin 4'-Hydroxylase in Humans. *Biochemistry*, *33*(7), 1743-1752. https://doi.org/10.1021/bi00173a017

Gonzalez, F. J., & Korzekwa, K. R. (1995). Cytochromes P450 expression systems. *Annual Review of Pharmacology and Toxicology*, *35*, 369-390. https://doi.org/10.1146/annurev.pa.35.040195.002101

Gonzalez, F. J., Skoda, R. C., Kimura, S., Umeno, M., Zanger, U. M., Nebert, D. W., Gelboin, H. V., Hardwick, J. P., & Meyer, U. A. (1988). Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. *Nature*, *331*(6155), 442-446. https://doi.org/10.1038/331442a0

Gressier, F., Verstuyft, C., Hardy, P., Becquemont, L., & Corruble, E. (2015). Response to CYP2D6 substrate antidepressants is predicted by a CYP2D6 composite phenotype based on genotype and comedications with CYP2D6 inhibitors. *Journal of Neural Transmission*, *122*(1), 35-42. https://doi.org/10.1007/s00702-014-1273-4

Guengerich, F. P., Waterman, M. R., & Egli, M. (2016). Recent Structural Insights into Cytochrome P450 Function. *Trends in Pharmacological Sciences*, *37*(8), 625-640. https://doi.org/10.1016/j.tips.2016.05.006

Guidelines (2021, March 26). CPIC. https://cpicpgx.org/guidelines/. Access date: November 21, 2023.

Hanioka, N., Hanioka, N., Tsuneto, Y., Saito, Y., Sumada, T., Maekawa, K., Saito, K., Sawada, J., & Narimatsu, S. (2007). Functional characterization of two novel CYP2C19 variants (CYP2C19*18 and CYP2C19*19) found in a Japanese population. *Xenobiotica*, *37*(4), 342-355. https://doi.org/10.1080/00498250601127038

Isvoran, A., Louet, M., Vladoiu, D. L., Craciun, D., Loriot, M. A., Villoutreix, B. O., & Miteva, M. A. (2017). Pharmacogenomics of the cytochrome P450 2C family: impacts of amino acid variations on drug metabolism. *Drug Discovery Today*, *22*(2), 366-376. https://doi.org/10.1016/j.drudis.2016.09.015

Ivanov, H. Y., Grigorova, D., Lauschke, V. M., Velinov, B., Stoychev, K., Kyosovska, G., & Shopov, P. (2022). CYP2C19 and CYP2D6 Genotypes and Metabolizer Status Distribution in a Bulgarian Psychiatric Cohort. *Journal of Personalized Medicine*, *12*(7), 1187. https://doi.org/10.3390/jpm12071187

Ji, Y., Skierka, J. M., Blommel, J. H., Moore, B. E., Vancuyk, D. L., Bruflat, J. K., Peterson, L. M., Veldhuizen, T. L., Fadra, N., Peterson, S. E., Lagerstedt, S. A., Train, L. J., Baudhuin, L. M., Klee, E. W., Ferber, M. J., Bielinski, S. J., Caraballo, P. J., Weinshilboum, R. M., & Black, J. L. (2016). Preemptive Pharmacogenomic Testing for Precision Medicine. *The Journal of Molecular Diagnostics*, *18*(3), 438-445. https://doi.org/10.1016/j.jmoldx.2016.01.003

Kahn, R.A., & Durst, F. (2000). Chapter Six - Function and evolution of plant cytochrome P450. *Recent Advances in Phytochemistry*, *34*, 151-189. https://doi.org/10.1016/S0079-9920(00)80007-6

Kimura, S., Umeno, M., Skoda, R. C., Meyer, U. A., & Gonzalez, F. J. (1989). The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene. *American Journal of Human Genetics*, *45*(6), 889-904.

Klingenberg, M. (1958). Pigments of rat liver microsomes. *Archives of Biochemistry and Biophysics*, *75*(2), 376-386. https://doi.org/10.1016/0003-9861(58)90436-3

Kuban, W., & Daniel, W. A. (2021). Cytochrome P450 expression and regulation in the brain. *Drug Metabolism Reviews*, *53*(1), 1-29. https://doi.org/10.1080/03602532.2020.1858856

Küpfer, A., & Preisig, R. (1984). Pharmacogenetics of mephenytoin: a new drug hydroxylation polymorphism in man. *European Journal of Clinical Pharmacology*, *26*(6), 753-759. https://doi.org/10.1007/BF00541938

Lewis, D.F.V. (2001). *Guide to Cytochromes P450: Structure and Function*. CRC Press.

Manikandan, P., & Nagini, S. (2018). Cytochrome P450 Structure, Function and Clinical Significance: A Review. *Current Drug Targets*, *19*(1), 38-54. https://doi.org/10.2174/1389450118666170125144557

Matimba, A., Del-Favero, J., Van Broeckhoven, C., & Masimirembwa, C. (2009). Novel variants of major drug-metabolising enzyme genes in diverse African populations and their predicted functional effects. *Human Genomics*, *3*(2), 169. https://doi.org/10.1186/1479-7364-3-2-169

Medrano, R. F., & de Oliveira, C. A. (2014). Guidelines for the tetra-primer ARMS-PCR technique development. *Molecular Biotechnology*, *56*(7), 599-608. https://doi.org/10.1007/s12033-014-9734-4

Meyer U. A. (1996). Overview of enzymes of drug metabolism. *Journal of Pharmacokinetics and Biopharmaceutics*, *24*(5), 449-459. https://doi.org/10.1007/BF02353473

Milosavljević, F., Bukvić, N., Pavlović, Z., Miljević, Ć., Pešić, V., Molden, E., Ingelman-Sundberg, M., Leucht, S., & Jukić, M. M. (2021). Association of CYP2C19 and CYP2D6 Poor and Intermediate Metabolizer Status with Antidepressant and Antipsychotic Exposure. *JAMA Psychiatry*, *78*(3), 270. https://doi.org/10.1001/jamapsychiatry.2020.3643

Müller, D. J., & Rizhanovsky, Z. (2020). From the Origins of Pharmacogenetics to First Applications in Psychiatry. *Pharmacopsychiatry*, *53*(04), 155-161. https://doi.org/10.1055/a-0979-2322

Nelson, D. L., Cox, M. M., & Hoskins, A. A. (2021). *Principles of Biochemistry*. Macmillan Learning.

Nelson, D. R., Koymans, L., Kamataki, T., Stegeman, J. J., Feyereisen, R., Waxman, D. J., Waterman, M. R., Gotoh, O., Coon, M. J., Estabrook, R. W., Gunsalus, I. C., & Nebert, D. W. (1996). P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*, *6*(1), 1-42. https://doi.org/10.1097/00008571-199602000-00002

Nofziger, C., Turner, A. J., Sangkuhl, K., Whirl-Carrillo, M., Agúndez, J. A. G., Black, J. L., Dunnenberger, H. M., Ruano, G., Kennedy, M. A., Phillips, M. S., Hachad, H., Klein, T. E., & Gaedigk, A. (2020). PharmVar GeneFocus: CYP2D6. *Clinical Pharmacology & Therapeutics*, *107*(1), 154-170. https://doi.org/10.1002/cpt.1643

Penning, T. M., & Drury, J. E. (2007). Human aldo-keto reductases: Function, gene regulation, and single nucleotide polymorphisms. *Archives of Biochemistry and Biophysics*, *464*(2), 241-250. https://doi.org/10.1016/j.abb.2007.04.024

PharmGKB. (n.d.). *CYP2C19 Named Alleles*. PharmGKB. https://www.pharmgkb.org/gene/PA124/haplotype. Access date: November 21, 2023

PharmGKB. (n.d.). VIPs: Very Important Pharmacogenes. PharmGKB. https://www.pharmgkb.org/vips. Access date: November 21, 2023.

Pirmohamed, M. (2014). Personalized Pharmacogenomics: Predicting Efficacy and Adverse Drug Reactions. *Annual Review of Genomics and Human Genetics*, *15*(1), 349-370. https://doi.org/10.1146/annurev-genom-090413-025419

Puaprasert, K., Chu, C., Saralamba, N., Day, N. P. J., Nosten, F., White, N. J., Dondorp, A. M., & Imwong, M. (2018). Real time PCR detection of common CYP2D6 genetic variants and its application in a Karen population study. *Malaria Journal*, *17*(1). https://doi.org/10.1186/s12936-018-2579-8

Qumsieh, R. Y., Ali, B. R., Abdulrazzaq, Y. M., Osman, O., Akawi, N. A., & Bastaki, S. M. A. (2011). Identification of New Alleles and the Determination of Alleles and Genotypes Frequencies at the CYP2D6 Gene in Emiratis. *PLOS ONE*, *6*(12), e28943. https://doi.org/10.1371/journal.pone.0028943

Rahma, A. T., Elbarazi, I., Ali, B. R., Patrinos, G. P., Ahmed, L. A., & Al Maskari, F. (2020). Genomics and Pharmacogenomics Knowledge, Attitude and Practice of Pharmacists Working in United Arab Emirates: Findings from Focus Group Discussions - A Qualitative Study. *Journal of Personalized Medicine*, *10*(3), 134. https://doi.org/10.3390/jpm10030134

Rahma, A. T., Elsheik, M., Elbarazi, I., Ali, B. R., Patrinos, G. P., Kazim, M. A., Alfalasi, S. S., Ahmed, L. A., & Al Maskari, F. (2020). Knowledge and Attitudes of Medical and Health Science Students in the United Arab Emirates toward Genomic Medicine and Pharmacogenomics: A Cross-Sectional Study. *Journal of Personalized Medicine*, *10*(4), 191. https://doi.org/10.3390/jpm10040191

Reynald, R. L., Sansen, S., Stout, C. D., & Johnson, E. F. (2012). Structural Characterization of Human Cytochrome P450 2C19. *Journal of Biological Chemistry*, *287*(53), 44581-44591. https://doi.org/10.1074/jbc.m112.424895

Siamoglou, S., Koromina, M., Hishinuma, E., Yamazaki, S., Tsermpini, E. E., Kordou, Z., Fukunaga, K., Chantratita, W., Zhou, Y., Lauschke, V. M., Mushiroda, T., Hiratsuka, M., & Patrinos, G. P. (2022). Identification and functional validation of novel pharmacogenomic variants using a next-generation sequencing-based approach for clinical pharmacogenomics. *Pharmacological Research*, *176*, 106087. https://doi.org/10.1016/j.phrs.2022.106087

Takahashi, M., Saito, T., Ito, M., Tsukada, C., Katono, Y., Hosono, H., Maekawa, M., Shimada, M., Mano, N., Oda, A., Hirasawa, N., & Hiratsuka, M. (2015). Functional characterization of 21 CYP2C19 allelic variants for clopidogrel 2-oxidation. *The Pharmacogenomics Journal*, *15*(1), 26-32. https://doi.org/10.1038/tpj.2014.30

Taylor, C., Crosby, I., Yip, V., Maguire, P., Pirmohamed, M., & Turner, R. M. (2020). A Review of the Important Role of CYP2D6 in Pharmacogenomics. *Genes*, *11*(11), 1295. https://doi.org/10.3390/genes11111295

Vasudevan, D. M., Sreekumari, S., & Vaidyanathan, K. (2013). *Textbook of Biochemistry for Medical Students*. JP Medical Ltd.

Wang, H., Kim, R. A., Sun, D., Gao, Y., Wang, H., Zhu, J., & Chen, C. (2011). Evaluation of the effects of 18 non-synonymous single-nucleotide polymorphisms of CYP450 2C19 onin vitrodrug inhibition potential by a fluorescence-based high-throughput assay. *Xenobiotica*, *41*(9), 826-835. https://doi.org/10.3109/00498254.2011.582893

Wang, H., An, N., Wang, H., Gao, Y., Liu, D., Bian, T., Zhu, J., & Chen, C. (2011). Evaluation of the effects of 20 nonsynonymous single nucleotide polymorphisms of CYP2C19 on S-mephenytoin 4'-hydroxylation and omeprazole 5'-hydroxylation. *Drug Metabolism and Disposition*, *39*(5), 830-837. https://doi.org/10.1124/dmd.110.037549

Zhang, L., Sarangi, V., Moon, I., Yu, J., Liu, D., Devarajan, S., Reid, J. M., Kalari, K. R., Wang, L., & Weinshilboum, R. (2020). CYP2C9 and CYP2C19: Deep Mutational Scanning and Functional Characterization of Genomic Missense Variants. *Clinical and Translational Science*, *13*(4), 727-742. https://doi.org/10.1111/cts.12758

Zhao, M., Ma, J., Li, M., Zhang, Y., Jiang, B., Zhao, X., Huai, C., Shen, L., Zhang, N., He, L., & Qin, S. (2021). Cytochrome P450 Enzymes and Drug Metabolism in Humans. *International Journal of Molecular Sciences*, *22*(23), 12808. https://doi.org/10.3390/ijms222312808

Zhou, Y., & Lauschke, V. M. (2021). Computational Tools to Assess the Functional Consequences of Rare and Noncoding Pharmacogenetic Variability. *Clinical Pharmacology & Therapeutics*, *110*(3), 626-636. https://doi.org/10.1002/cpt.2289

UAEU جامعة الإمارات العربية المتحدة United Arab Emirates University

UAE UNIVERSITY MASTER THESIS NO. 2023: 128

Pharmacogenomics is a relatively new field of study. Some of the most important genes in this field are *CYP2C19* and *CYP2D6*. In this thesis, seven variants in *CYP2C19* that have been detected in the Emirati population were analysed to determine their potential effects on the enzyme's activity. For *CYP2D6*, a long-range PCR-based technique was optimized and utilized to detect gene copy numbers using DNA extracted from blood samples isolated from psychiatric patients to determine their CYP2D6 metabolic status.

Reema Saleous received her Master of Medical Sciences from the Department of Biochemistry and Molecular Biology, College of Medicine & Health Sciences at UAE University, UAE. She received her Bachelor of Science in Biology from the College of Arts & Sciences, American University of Sharjah, UAE.

www.uaeu.ac.ae