CERTIFICATE

This is to certify that the dissertation entitled

"Method Development of Allele Specific PCR Assay for CYP2B6"

is the bonafide record of research work done by

MR. MOHD IQBAL BIN ZULKAFLI

during the period **21st of December 2008** to **30th of April 2009**

under my supervision.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Prof Rusli Ismail, without whose motivation and encouragement, I would not have considered a graduate career in pharmacogenomic research. I appreciate his vast knowledge and skill in many areas (e.g., vision, aging, ethics, interaction with participants), which have on occasion made me "GREEN" with envy. , I am grateful to him for enlightening me the first glance of research. Without him I may not even step my foot in Institute for Research in Molecular Medicine (INFORMM).

A very special thanks goes out to Miss NurFadhlina Musa, whose expertise, understanding, and patience, added considerably to my graduate experience. She provided me with direction and technical support throughout my project. It was though her, persistence, understanding and kindness that I completed my undergraduate degree and was encouraged to apply for graduate training. I doubt that I will ever be able to convey my appreciation fully, but I owe her my eternal gratitude. My sincere thanks also goes to Miss Fazni Razali for her encouragement, insightful comments, and hard questions.

Appreciation also goes to other members of Pharmacogentic group whose their advice and support was really helpful for my research: Kak Za, Faris, Mat, Firdaus, and Sanjaya

I also would like to thank my lab mate, Nazila Binti Talib for the stimulating discussion and for sleepless nights we were working together before deadlines. To my fellow friends, Hafiznur and Faizul who accompanied me through the whole night at the lab while I was doing my lab work. I also would like to acknowledge all my classmates (forensic science students 2005-2009) for the friendship, advice and support. I would never forget all the fun we have for the last four years.

I would also like to thank my family for the support they provided me through my entire life and in particular, I must acknowledge my future wife, 'Star', who is also my best friend, without whose love, encouragement and editing assistance, I would not have finished this thesis.

Table of content

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

ABSTRACT

CYP2B6 is one of the drug-metabolizing enzymes that catalyses many types of clinical drugs such as anti-cancer drugs, anti-malarial drugs and anti depressant drugs. Among other types of CYP isoforms, CYP2B6 is poorly characterized. It was initially thought that CYP2B6 constitute only small portion of total hepatic CYP (Shimada *et al.*, 1994) and expressed in low level (Mimura *et al.*, 1993 and Yamano *et al.*, 1989). Since most of the substrates of CYP2B6 are clinically important drugs, determination of CYP2B6 polymorphism in the population is important. We successfully developed a simple and specific method of Allele-Specific Multiplex Polymerase Chain Reaction (PCR) for the detection of CYP2B6 single nucleotide polymorphisms (SNP). DNA was extracted from blood and was subjected to a first PCR which amplify a portion of exon 1, exon 3 and 4 region of the gene. After that second PCR was used to detect polymorphism sites of CYP2B6 using primers that were allele specific. Sequencing was performed to validate the test results. As for conclusion we successfully developed simple and specific method for the detection of 62 A>T, 76 A>T and 157618 C>T polymorphism in CYP2B6 gene. Allele-specific PCR is proved to be specific with the ability to apply the nested technology where our method managed to separate CYP2B6 from the similar CYP2B7. We also managed to optimize multiplex reaction of two primers and the result was satisfying.

CHAPTER 1

INTRODUCTION

1.1 Pharmacogenomic

According to World Health Organization (WHO), drug is any substance that is used or intended to be used to give benefit to the recipient by modifying his/her physiological systems or pathological states. Nowadays, a lot of problem arises in giving the right medication to the patient. Studies from U.S. hospitals had found out that 6.7% of patients had developed adverse effect and 0.32% had fatal drug reactions (Lazarou, 1998). Fatal from drug reaction has caused approximately 100,000 deaths per year in the United States.

Effect of drugs is determined by how the interaction occurs between drugs and the body. The interactions are influenced by many factors. Several separate observation during studies in 1950s clearly indicated that drugs effects depends on the genetic constituent of the recipient. The observations include genetic variation of isoniazid acetylation (Hughes *et al.*, 1954), failing cholinesterase activity affecting succinylcholine action (Kalow, 1956), and primaquine-caused hemolysis due to deficiency of glucose-6-phosphate dehydrogenase (Beutler, 1959). These cases showed that each of the enzymes could vary in many ways because of different mutation.

With the advancement of molecular technology, researcher begins to study the effects of drug at the genetic level. Thus the concept of pharmacogenomic is born. One of the aims of pharmacogenomic is to create the basis of personalized medicine. This means that one person will be given dose and type of drug according to their genes. In order to reach the aim, the

genetic variant that affects drug action must be studied. One of the most frequently occurring genetic variants are single-nucleotide polymorphism (SNP) (Pennisi, 1998).

The human genome consists of approximately 3 billion basepairs, and SNPs occur on the average in approximately 1 per 1000 bases. From the total predicted 11-15 million of genetic polymorphism in human, the SNPs comprises for more than 90% of them. Using genomic methods, high-density maps of SNPs can be created. Thus by correlating SNPs and drug responses, the ability to predict drug efficacy or toxicity is possible within reasonable limits for any individual (Pfost *et. al.*, 2000).

Previously, the method that researcher has been used to investigate genetic variation and activity of drug metabolizing enzymes was by phenotyping. With the advent of new molecular technology, genotyping method has gained interest among researchers. Genotyping methodologies are easier to use than biochemical measurement in a clinical testing. It is more stable and not influenced by drug-drug interaction or food-drug interaction. Furthermore, only small amount of sample is required to complete the genotyping assay and the result can often be obtained more quickly than phenotyping method.

1.2 CYP 2B6

CYP2B6 is one of the drug-metabolizing enzymes that catalyses many type of clinical drugs such as anti-cancer drugs, anti-malarial drugs and anti depressant drugs. It is mainly expressed in liver, but also in the kidney, intestines, skin, brain and lung at lower extent (Gervot *et al.*, 1999; Miksys *et al.*, 2003). Consistently, it was responsible for the metabolism of about 3% of therapeutic drug (Rendic, 2002). The variability of CYP2B6 expression between individual

may due to the regulatory phenomena, but recent studies showed that it is also due to common genetic polymorphisms (Lang *et al.*, 2001).

Among other types of CYP isoforms, CYP2B6 is poorly characterized. It was initially thought that CYP2B6 constitute only small portion of total hepatic CYP (Shimada *et al.*, 1994) and expressed in low level (Yamano *et al.*, 1989). The expression level of CYP2B6 has been underestimated because of poor selectivity and sensitivity of previously used antibodides to detect CYP2B6 protein (Ekins and Wrighton, 1999). Consequently, further investigation on CYP2B6 expression revealed higher mRNA and protein levels than previously observed (Roy *et al.*, 1999).

Many of the CYP2B6 substrates are of clinically important drugs. Nevirapine and evafirenz for example are important drugs in HIV treatment. Cyclophosphamide is widely used as alkylating drug in cancer chemotherapy and immunosupression and CYP2B6 is the major enzyme in the bioactivation of the drug (Xie *et al.*, 2003). Ketamine and propofol are used as anesthetics agent and artemisinin is used as antimarial drug. The most specific substrate of CYP2B6, bupropion, is used as antidepresent and antismoking agent.

Since most of the substrates of CYP2B6 are clinically important drugs, determination of CYP2B6 polymorphism in the population is important. This is to prevent and possibility of developing toxicity among the patient and to give the best possible effect of the drug to the patient. It was also has been found out that ethnicity is an important variable contributing to inter-individual variability in drug metabolism, response and toxicity. (Chowbay *et al.*,2005; Evans *et al.*,2001; Xie *et al.*, 2001).

1.3 Genotyping CYP2B6

A lot of research has been conducted to study the genotype – phenotype relationship of CYP2B6 protein recently. It has been found that 1459C>T mutation (Lang *et al.*,2001), the 516G>T mutation (Ariyoshi *et al.*, 2001) and 785A>G mutation (Jinno *et al.*, 2003) can affect activity of the enzyme. There are also study of CYP2B6 allele frequency among population across the world such as Korean (Cho *et al.*, 2004), Japanese (Hiratsuka *et al.*, 2002), African and Asian (Klein *et al.*, 2005). Because more data on the expression and function of CYP2B6 are being collected, it is important to have a high resolution rapid genotyping procedure to correlate genomic and phenotypic data in large-scale studies.

This study will try to develop a simple genotyping method called Nested Allele-Specific Multiplex PCR to detect polymorphism of CYP2B6 gene. Allele-Specific PCR method is simple and very cost effective for laboratory routine especially in population studies (Zainuddin *et al.*,2003). This method is called nested because it is a two-step PCR method. The first PCR will amplify outer region contain the SNPs to be detected and the second PCR will amplify only the sequence of the desired SNP using the first PCR product as template.

Primers for the first and second PCR will be designed and the condition of each PCR reaction is optimized for each primer. After that this study will try to multiplex the primers so that more than one amplicon can be generated in one reaction. To ensure the specificity and accuracy of the method, the study will try to validate the method by sequencing. In this study, we aimed to develop multiplex PCR methods which can be used to detect 10 SNPs using only two steps. We hope that this will reduce labour, cost, and time for CYP2B6 genotyping.

CHAPTER 2

REVIEW OF LITERATURE

2.1 SNP Genotyping method

Since the idea of personalized medicine had being introduced, a lot of SNP genotyping method has been developed to study the genetic association of drug metabolizing enzyme. At the end of the year 2000, more than 1.5 million Single Nucleotide Polymorphisms (SNPs) has been found in human genome and is published to the public in a database. The availability of SNP map has provided a giant leap in the study of genotype-phenotype relation of drug metabolizing enzyme in human. Thus, genotyping of SNPs will most probably become a major part of every genetic association study, and the appropriate genotyping method is critical to the success of the study (Kwok, 2001).

According to Kwok (2001), an ideal genotyping method should be easy and quick to developed, the cost and time spent for optimization should be low, the reaction must be robust even with suboptimal DNA samples could produce reliable result, the assay should be automated and accurate, the reaction format must be flexible and scalable, and once optimized, the total cost per genotype should be low. To date, no such ideal method exists. Further improvements should be made in biochemistry, engineering and analytical software of the genotyping method.

In any genotyping method, amplification of the target DNA is crucial to increase the number molecule in the detectable level. Usually, for any typical DNA-based diagnostic test, the numbers of target molecule may range from ten to 10^5 (Whitcombe *et al.*, 1998). The

Invention of PCR was able to meet the need of having large number of DNA sample in pharamcogenomic research.

A PCR reaction is similar to DNA replication. In DNA replication, the double stranded DNA unwinds and an enzyme called DNA polymerase will add nucleotide on the unwound template producing entirely new DNA. The amplification was of the entire genome. As for PCR, it permits selective in vitro amplification of a particular DNA. This is where PCR is very useful because we can select the region of DNA where we wanted to amplify.

2.2 PCR assays for CYP 2B6

Several PCR assays had been applied in pharmacogentic studies for the detection of CYP2B6 polymorphisms. Each of the methods has their own pros and cons. Among the methods include Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP), Single-Strand Conformation Polymorphism (SSCP), PCR-Sequencing and Fluorescence Real-Time PCR.

2.2.1 RFLP

RFLP technique involves fragmenting DNA sample by means of restriction digesting using restriction enzymes. The enzymes will cut the DNA at specific site producing fragments with variable length. The fragments are then separated using gel electrophoresis and DNA ladder are used to estimate fragment size.

PCR-RFLP is the most commonly used method for SNP genotyping, but it is relatively time consuming (Hamajima, 2001). This is because the method usually involves first PCR, followed by restriction digestion which consist of incubation step of three to 24 hours, and lastly gel electrophoresis. Besides that it requires the use of restriction enzyme making it relatively expensive. Guan *et al.* (2006) used this method and detected 5 SNPs namely 64C >T, 516G >T, 777C >A, 785A >G and 1459C >T on 193 DNA samples of Han Chinese.

2.2.2 SSCP

As for SSCP, detection of mutation is done by analyzing conformation changes of the DNA molecule resulting from single base change of the sequence in the molecule itself. Under non-denaturing condition and reduced temperature, single-stranded DNA molecule assumes unique conformation that varies, depending on their nucleotide sequences. This conformational changes result in detectable differences in electrophoretic mobility of the molecule.

The disadvantage of this method is that electrophoretic behaviour of single-stranded molecule is unpredictable since it is very dependable on the temperature and condition of the electrophoresis (Tagu and Moussard, 2003). The rate of reannealing of DNA after denaturation during electrophoresis is very high especially if high concentration of DNA is loaded onto the gels (Selvakumar, 1997). Besides that, the formation of heteroduplex DNA from PCR products with similar sequence may occur frequently (Lee *et al.*, 1996). Another limitation of this method is that for fragment larger than 200bp, the mutation is not detectable (Tagu and Moussard, 2003).

2.2.3 Sequencing

Among all the methods for genotyping, sequencing method is the most accurate and produces reliable result. It involves a process called chain termination where dideoxynucleotide triphosohates (ddNTPs) is used in the reaction mixture. These chain-terminating dideoxynucleotides lack the 3' hydroxyl (OH) group needed to form phosphodiester bond between one nucleotide and another in DNA elongation. When ddNTP is incorporated, it seized further strand extension. The resulting variable length fragments is separated by size using electrophoresis (usually capillary electrophoresis) giving it the sensitivity to distinguish DNA fragment with single nucleotide difference. The only limitation of this method is that the instrumentation is very costly and not suitable for population studies (Zainuddin *et al.*, 2003). Its sensitivity and accuracy had commonly been applied as validation method by researchers to validate the genotyping method they developed. Lang *et. al* (2004) used this method and detected 5 novel SNPs namely 62A>T, 136A>G, 12820G>A, 13076G>A, and 21388T>A.

2.2.3 Real-Time PCR

Another popular method that had usually been used for genotyping SNP is Real-Time PCR (RT-PCR). RT-PCR is the modification of traditional PCR reaction, the advantage is that the PCR product can be monitored in real-time. So it skip the used of gel electrophoresis in order to analyze the result thus saves time. RT-PCR uses probes called fluorescent reporter that binds to the product formed and reports its present by fluorescence. The generated signals reflect the amount of product formed.

Although RT-PCR is relatively sensitive and does not susceptible to PCR inhibition, the cost for reagents is substantially high up to four-folds that of standard PCR procedure (Baric *et al.*, 2006). The number of amplicon to be detected will depends on the amount of fluorophores used. The problem is we need a method that can detect many amplicons at one time for population studies. Increasing the number fluorophores will increase the capital. Besides that, it needs a specialized real-time analyzer which is not available in many laboratories (Stark *et al.*, 2005). Xu *et. al* (2007) uses this method on 507 subject from southern China and the study intended only to detect two SNPs namely 64C>T and 516 G>T.

2.3 Allele-Specific PCR

A genotyping method that is quite reliable and simple to be carried out is the Allele-specific PCR method (Zainuddin *et al.*, 2003). The method still retains the basic principle of PCR, which utilizes the specificity of primer amplification. As we know, primer amplification in PCR reaction is specific, which means that it amplifies only template that match its sequence. The principle of this method is that each primer is previously designed in such a way that it will only amplify the sequence containing the SNP to be investigated. This method does not rely on the use of restriction enzymes and it is very specific with the ability to produce a consistent result (Jacob *et al.*, 2004).

Several researchers has successfully develop this method to genotype certain genes polymorphisms. Zainuddin *et al.* (2003) developed allele specific PCR method to detect five SNPs of CYP 2C9. The method specificity was increased by applying the principle of Nested PCR. In their method, they first design primer to amplify the exons which the alleles of

interest reside. After that they design a second PCR primer that has specific 3' end to differentiate single nucleotide change at the specific locus during PCR amplification where the first amplification product acts as template. They also design the second PCR primer so that to have similar annealing temperature and appropriate length for multiplexing reaction, which makes the method cause effective since multiple regions can be amplified in single reaction. Besides that, they used gel electrophoresis as method of amplicons detection which is simple and affordable compared to other detection methods such as matrix-assisted laser desorption/ionization time-offlight mass spectrometry (MALDI-TOF MS). Even though it may required high stringency and experience in gel reading, the problem is solved by employing two random blind reviewers for quality assurance.

In optimizing the reaction condition, Zainuddin *et al.* (2003) found out that the concentration of Magnesium Chloride, primers, DNA polymerase, and thermocycler's temperature and the duration of each PCR step has their own influence to the resulting product. The increment of concentration of $MgCl_2$ can increase the intensity of the desired band but at the same time it result in production of non specific product. They also tested the effect of DNA polymerase concentration on the product yield which becomes more unspecific when the concentration is increased. They also modified the annealing temperature in order to prevent unspecific product and false negative result. Their method is also reliable as they conduct validation test of the method on 40 DNA samples and five were sent for sequencing. Same method also has been applied by Muthiah *et. al* (2004) and detected three SNPs for CYP2C8.

Jacob *et al.* (2004) had also applied the same method on 135 DNA samples and detected seven SNPs for CYP2B6 gene plus five additional loci found within the SNP. The method was found to be robust, simple, rapid and specific where they also develop a two-step PCR.

However, compared to Zainuddin *et al.* (2003) and Muthiah *et.al* (2004), additional step was done by Jacob *et al.* (2004) to further confirm the specificity of their first PCR product. According to them, the amplification of first PCR product is subjected to cross-reaction with CYP2B7, the pseudogene of CYP2B6. This may lead to amplification of unspecific product.

In order to confirm the specificity, they designed a CYP2B7 positive primer and test it against the first PCR amplicon. If an amplicon was found not to be specific for CYP2B6, i.e., a product was generated in the CYP2B7-specific reaction, the initial amplification primers for that amplicon were redesigned. Once the amplicon appeared to be specific for CYP2B6, they then verify the amplicon using sequencing.

Contamination is common event in PCR reactions. Several researchers addressed this problem by applying controls experiments for their method development. For example Blievernicht *et. al* (2007) ensures no-template controls to be carried along in every plate of their multiplex PCR reaction to exclude contaminations. Jacob *et. al* (2004) on the hand uses internal control primers for both PCR reaction of their common and variant alleles. Muthiah *et.al* (2004) *et al.* utilizes both negative and positive control in their routine PCR batches as quality control to avoid assay error.

CHAPTER 3

OBJECTIVE OF THE STUDY

The objective of the study is to develop Allele-Specific PCR assay for genotyping CYP2B6 gene. The study will aim to optimize all of the components for the PCR reaction to produce simple and robust rapid genotyping procedure so it can used in large-scale genotype – phenotype correlation study of CYP2B6.

In this study, we aimed to develop multiplex PCR methods which can be used to detect 10 SNPs using only two steps. We hope that this will reduce labour, cost, and time for CYP2B6 genotyping.

CHAPTER 4
MATERIALS AND METHOD

4.1 Chemicals , Reagents and Instruments

The chemical and reagents used in this method are listed in **table 4.1** and the instruments used are listed in **table 4.2** (see appendix).

4.2 Reaction component of the PCR

Reaction component for all of the PCR reaction was divided into master-mix, primer-mix and the DNA template. The total reaction volume for the PCR is 25 µl per reaction. Formula to calculate the quantity of the stock reagent used was as follows:

(initial stock concentration) x (volume needed) = (final concentration) x (total reaction volume)

.

4.2.1 Master-mix

Master-mix constitutes 10 µl out of 25 µl of the total reaction volume. It consists of reaction buffer, Magnesium Chloride ($MgCl_2$), deoxynucleoside Triphosphate (dNTP) and DNA polymerase. The standard concentration of each of the master mix components are listed in **table 4.3**. This standard concentration was used as initial condition for all our PCR reaction.

Component	Concentration (working stock)	Concentration (in 25µl reaction volume)
Buffer	10 X	1 X
$MgCl_2$	50 mM	1.5 mM
dNTP	10 mM	0.2 mM
Taq	1 unit	1 unit

Table 4.3 – concentration of master-mix components

4.2.2 Primer-mix

Primer-mix consists of sets of primers (forward and reverse) for the PCR reaction. It constitutes 13 µl out of 25 µl of the reaction volume. One primer-mix can contain more than one set of primer for simultaneous amplification (multiplex). Working concentration of the primer mix is prepared at 5pmol/µl and the initial concentration in 25 µl reaction volume for all primers begins at 0.25 pmol/µl.