uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Russell Haywood Taylor

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biochemistry)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

A Guanine to Adenine Mutation -76BP from the Transcriptional Start Site Decreases Constitutive CYP1A2 Expression in a Novel Mouse Strain

TITRE DE LA THÈSE / TITLE OF THESIS

Mary Alice Hefford

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dennis Bulman

Sean Li

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

A GUANINE TO ADENINE MUTATION -76BP FROM THE TRANSCRIPTIONAL START SITE DECREASES CONSTITUTIVE CYP1A2 EXPRESSION IN A NOVEL MOUSE STRAIN.

By

Russell Haywood Taylor

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science, specialization in Biological Chemistry.

University of Ottawa

Faculty of Medicine

January 15, 2007

Supervisor _____
Dr. Mary Alice Hefford

Co supervisor _____
Dr. William Casley

Your file Votre référence
ISBN: 978-0-494-49283-3
Our file Notre référence
ISBN: 978-0-494-49283-3

A GUANINE TO ADENINE MUTATION -76BP FROM THE TRANSCRIPTIONAL START SITE DECREASES CONSTITUTIVE CYP1A2 EXPRESSION IN A NOVEL MOUSE STRAIN.

By

Russell Haywood Taylor

Chairpersons of the Supervisory Committee: Mary Alice Heford[1] and William L. Casley[2]
[1]Adjunct Professor, Department of Biochemistry, University of Ottawa
[2]Adjunct Professor, Department of Biology, University of Ottawa

ABSTRACT

Variations in the expression of drug metabolizing enzymes of the Cytochrome P450 superfamily are a principal cause of atypical reactions to therapeutics. The molecular mechanisms by which the metabolizing enzymes of the drug are regulated, and the effects of genetic variation on this regulation, are not completely understood. Cytochrome P450 1A2 (CYP1A2) is one such enzyme. The *APN* mouse strain has low expression of the CYP1A2 enzyme, relative to the *C3H/HeJ* strain. It was hypothesized that this difference in expression of the CYP1A2 was occurring as a result of a single nucleotide polymorphism at the *Cyp1a2* locus. This work has demonstrated that this variation in CYP1A2 expression occurs at the level of transcription and that a single nucleotide change 76 base pairs upstream of the transcriptional start site was critical for promoter function. The mechanism of constitutive CYP1A2 expression involves a previously unidentified *cis*-acting element in this region.

DEDICATION

This thesis is dedicated to my good friend Mike Channer, whose passion for life will always be remembered.

ACKNOWLEDGMENTS

I would like to thank my Thesis Advisory Committee; Dr. Odette Launeuville and Dr. Linda J. Bonen for their guidance and constructive feedback. From Health Canada, I would like to thank Dr. Aleksander Bliu, Mrs. Margeryta Nowakowska, Ms. Lori Lavigne, Mrs. Michèle Lemieux and especially Mrs. Carole A. Westwood, for her incredible patience and surgical skills. Also I would like to acknowledge my department supervisor Dr. Mary Alice Hefford and my research supervisor Dr. Bill Casley for their persistence and support. Lastly, I would like to thank my friends and family, particularly my parents Don and Charlotte Taylor for teaching their children that anything is possible when you believe in yourself.

"Science may set limits to knowledge, but should not set limits to imagination."

Bertrand Russell (1872 - 1970)

TABLE OF CONTENTS

LIST OF ABREVIATIONS

ADRs: adverse drug reactions

Ahr: aromatic hydrocarbon receptor

ANOVA: analysis of variance

AP-1: activator protein 1

APN: acetaminophen nonsusceptible

APS : acetaminophen susceptible

BLI: Bio luminescent imaging

C3-D: caffeine 3-demethylation

CYP1A2: human Cytochrome P450 1 A 2 gene

CYP1A2: human or mouse Cytochrome P450 1 A 2 protein

Cyp1a2: mouse Cytochrome P450 1 a 2 gene

CYP2B6: Cytochrome P450 2 B 6 protein

CYP2D6: Cytochrome P450 2 D 6 protein

EMSA: electrophoretic mobility shift assay

EVI1: ecotropic viral integration site 1

G6PD: glucose-6-phosphate dehydrogenase

hncDNA: heteronuclear complementary DNA

HNF-1: hepatic nuclear factor 1

hnRNA: heterogeneous nuclear ribonucleic acid

IVIS: *in vivo* imaging system

LB: Luria-Bertani

mRNA: messenger ribonucleic acid

MUT: *C3H/HeJ* -76 G>A mutant

MWM: molecular weight markers

NCBI: National Center for Biotechnology Information

NF1: nuclear factor 1

PCR: polymerase chain reaction

PDCD1: programmed cell death gene

PTH: parathyroid hormone

qPCR: quantitative polymerase chain reaction

QTL: quantitative trait locus

RUNX1: runt-related transcription factor 1

SLE: systemic lupus erythematosus

SNP: single nucleotide polymorphism

TBP: TATA box binding protein

TCDD: 2,3,7,8-tetrachlordibenzo-p-dioxin

TE: tris ethylenediaminetetraacetic acid

TSG: tumor suppressor genes

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Xenobiotic metabolism

A xenobiotic is a chemical which is not produced by the organism exposed to it. We live in a world full of xenobiotics. They range from environmental toxins such as 1,4-dimethylbenzene to chemotherapeutics such as acetaminophen. The metabolism and excretion of these chemicals is crucial in order to maintain homeostasis.

The mammalian xenobiotic metabolizing system was thought to have evolved as a result of heterotroph-auotroph interactions. Animals with detoxification systems had an advantage, in that they could graze on plants that produced toxic defensive secondary metabolites (Ehrilch and and Raven, 1964). Mammals, such as mice and humans, detoxify their chemical environment through complex array of enzymes found primarily in the liver.

Xenobiotic metabolism can be divided into two main reaction categories, phases I and II. Phase I is involved in the addition of a functional group via oxidation, reduction or hydrolysis. These chemical modifications can increase a xenobiotic's water solubility, which aids in the xenobiotic's excretion and/or renders it more amenable to phase II conjugation reactions. Phase II of xenobiotic metabolism involves addition reactions such as acetylation, sulphation, glucuronidation and addition of glutathione (Benet *et al.*, 1996).

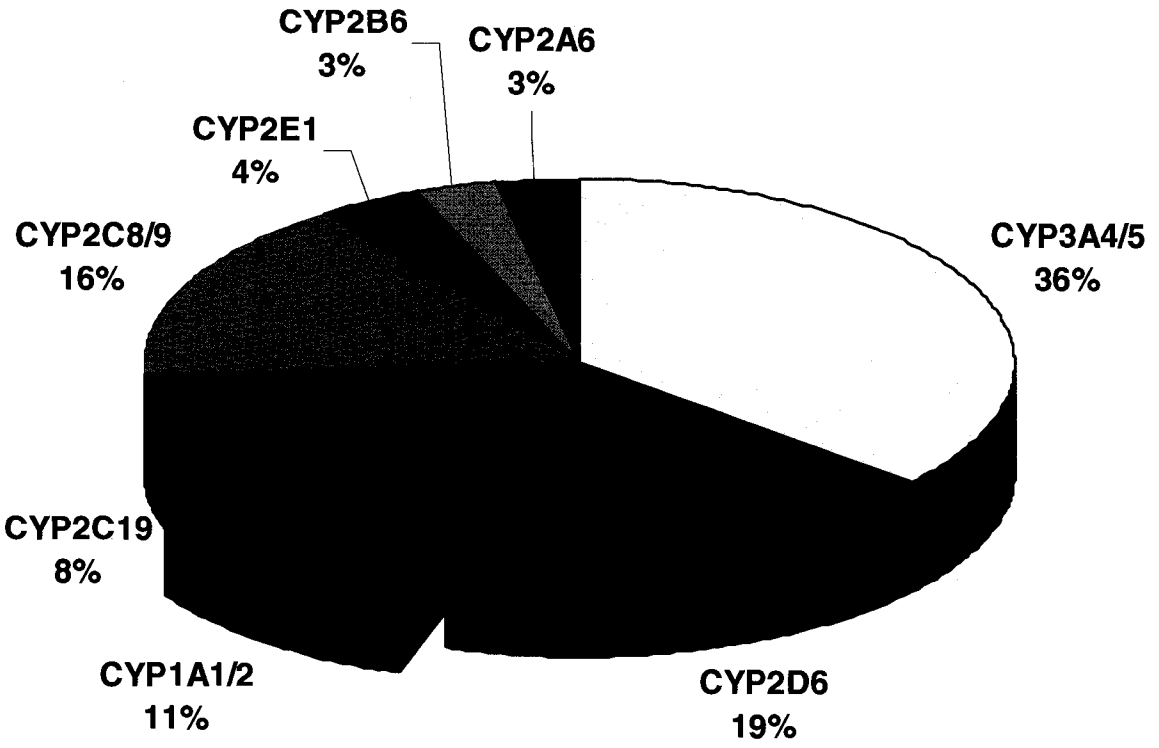
The rates at which different individuals can metabolize a given xenobiotic can vary dramatically within a population. The basis of this variation arises from the environmental and genetic influences on the multiple pathways through which xenobiotics are metabolized. Variation

in xenobiotic metabolism is of interest from both a toxicological and a pharmacological point of view. For a given procarcinogen, a fast metabolizing individual will have an increased susceptibility to cancer as a result of increased carcinogen activation. An individual with a slow xenobiotic metabolism for a specific therapeutic will be more prone to drug toxicity, due to increased bioavailability of the parent compound (slower clearance). Conversely, a fast metabolizing individual of the same drug may fail to show a therapeutic benefit due to lower bioavailability.

The contribution that variability in xenobiotic metabolism has towards adverse drug reactions is of interest. In 1998 the fourth highest cause of death in the United States was thought by some to be therapeutic drug complications (Lazarou *et al.*, 1998). Adverse drug reactions (ADRs) affect 10-20% of hospitalized patients and greater than 7% of the general population. More than 85% of these ADRs were dose-dependent (Gomes and Demoly, 2005). ADRs can be blamed, in part, on inappropriate individual xenobiotic metabolism profiling. Predictive genotyping may reduce these deaths by 10-20% per year (Ingelman-Sundberg, 2001).

In some cases, an individual's ability to metabolize a given xenobiotic is effected by a single gene variant. An example of this is seen in the Cytochrome P450 2D6 (CYP2D6) gene. This enzyme has many therapeutic substrates (Figure 1.1). These include: β-receptor antagonists, antiarrhythmic agents, morphine derivatives, antidepressants and antipsychotics. It has been estimated that 5-10% of Caucasions, 3.5-8% of African Americans and 1% of Asians carry a nonfunctional or unstable CYP2D6 enzyme. This slow drug metabolizing pheonotype can result in drug toxicities among these sub-populations (Gaedigk *et al.*, 2005). Screening for allelic variations of CYP2D6 can decreases drug toxicity for specific therapeutics.

Figure 1.1 The proportion of currently prescribed drug substrates metabolized by various enzymes in the Cytochrome P450 family. Adapted from Wrighton and Stevens (1992).


CYP2B6
3%
CYP2A6
3%
CYP2E1
4%
CYP2C8/9
16%
CYP3A4/5
36%
CYP2C19
8%
CYP1A1/2
11%
CYP2D6
19%

It is often the case, however, that an individual's ability to metabolize a given xenobiotic is affected by variations in multiple genes (Ozdemir *et al.*, 2005). An example of xenobiotic metabolism with multiple gene involvement has been observed in rats using Troglitazone, a potential diabetes type 2 mellitus therapeutic. Troglitazone was removed from human clinical trials because of its toxicity. This drug has been shown to interact with multiple xenobiotic metabolizing pathways. Variations in multiple genes involved in Troglitazone metabolism could, in part, be responsible for its toxicity (Vansant *et al.*, 2006). In this case, drug toxicity screening would be more complex than analyzing a single gene variant but would require a genetic profile of multiple genes.

A great deal of research has been done to produce new therapeutic compounds; however, there is considerably less information on the mechanisms regulating their metabolism. Greater understanding of these mechanisms may reduce the number of deaths due to ARDs as well improve therapeutic efficacy.

1.2 Cytochrome P450

Cytochrome P450 is a superfamily of heme-containing monoxygenases that encompass 57 functional genes in humans and 102 functional genes in mice (http://drnelson.utmem.edu/Genome.list.htm). The Cytochromes P450 are grouped into families and subfamilies based on amino acid homology. Cytochromes P450 with 40% or less homology to one another are labeled as different families, for example CYP1 and CYP2. Within these families, P450 enzymes that show 55% or greater homology with one another are delegated into the same

subfamily, for example: CYP2B and CYP2C. Individual isoforms of these enzymes are sequentially number, for example: CYP1A1 and CYP1A2 (Nelson *et al.*, 1993).

Many Cytochromes P450 enzymes play a large role in drug metabolism (Figure 1.1). They have diverse and often overlapping substrate ranges. These enzymes are primarily involved in phase I xenobiotic metabolism and catalyze the addition of polar groups (Eaton *et al.*, 1995). Briefly, xenobiotics form a complex by binding to the P450 enzyme which contains an iron atom in the ferric (Fe^{+3}) oxidation state. NADPH-Cytochrome P450 reductase converts NADPH to NADP+, which provides reducing equivalents that reduce the iron from a ferric to a ferrous (Fe^{+2}) state, allowing the complex to be oxidized by O_2. A second set of reducing equivalents are produced from NADPH-Cytochrome P450 reductase, which releases the reduced substrate from the complex (Champe *et al.*, 2005).

1.3 CYP1A2

This study focuses on CYP1A2, which is an enzyme involved in xenobiotic detoxification and drug metabolism of heterocyclic, aryalamines and nitroaromatic compounds specifically (Eaton *et al.*, 1995). CYP1A2 is a monooxygenase liver enzyme found within the Cytochrome P450 family. Though CYP1A2 is only involved in the metabolism of approximately 10% of prescribed drug (Figure 1.1), it is implicated in 75% of drugs causing ADRs (Phillips *et al.*, 2001).

The rationale for studying CYP1A2 regulation arises from the desire to understand the interindividual differences in drug response and susceptibility to environmental carcinogenesis. When the molecular mechanisms of regulation are known, one can begin to look for the basis of variation. The key to gene regulation lies, in part, in the DNA sequence.

The mechanisms that regulate the expression of the different Cytochromes P450 are not completely known. In the context of liver-specific expression, each P450 gene in the CYP enzyme family has been proposed to have a unique mechanism of expression (Gonzalez and Lee, 1996). Traditionally, CYP1A2 expression has been studied for its inducibility via the aromatic hydrocarbon receptor (Ahr) pathway. Ingestion of Ahr ligands, such as 2,3,7,8-tetrachlordibenzo-p-dioxin (TCDD), up-regulates the expression of many ligand response genes and has been shown to produce up to a 100 fold increase in CYP1A2 expression, depending on the ligand, species, dose and strain (Nebert *et al.*, 1993). However, CYP1A2 also acts on substrates that do not induce its expression, such as acetaminophen (Bray and Rosengren, 2001). Therefore, the mechanisms governing the basal expression of CYP1A2 are important for predicting xenobiotic metabolism.

In humans, extensive variation in CYP1A2 expression has been observed. It is difficult, however, to distinguish the genetic component of this variation from those due to environmental exposure. As of 2006, there are 34 alleles of *CYP1A2* identified including 9 predicted haplotyes (http://www.cypalleles.ki.se/). To date, human *CYP1A2* polymorphisms have been shown to affect this enzyme through the creation of splice variants and disruption of induction regulatory sties. (Allorge *et al.*, 2003; Nakajima *et al.*, 1999). No single nucleotide polymorphism (SNP) or haplotype in the CYP1A2 gene has yet been identified that can be used to provide unequivocal prediction of CYP1A2's basal expression (Jiang *et al.*, 2006).

Several regions of the *CYP1A2* gene in both humans and rodents have been identified as important to gene expression. In humans, E-box enhancer elements located in the -1987 to -3201 region are responsible for 60% of the basal CYP1A2 expression (Pickwell *et al.*, 2003). This region includes two Ahr binding sites one -2259 to -1987 and the other -2532 to -2423 (Quattrochi *et al.*,

1998), as well as -2352 to -2094 which contain binding sties for activator protein 1 (AP-1) and a for the liver specific transcription factor hepatic nuclear factor 1 (HNF-1) (Chung and Bresnick, 1997). In the proximal region of the human *CYP1A2* gene, the -72 to -31 region has been identified as important for basal expression. This region contains GC, CCAAT, and TATA Boxes (Chung and Bresnick, 1997).

In mice, a study found that the -1843 to + 53 region was insufficient for CYP1A2 expression in transfected hepatoma Hepa-1 cells (Owens and Nebert, 1990). In the mouse, basal CYP1A2 expression has been shown to be controlled, at least in part, by two enhancer elements in the -4401 to -4382 and from -4300 to -4292 regions. These elements are essential for CYP1A2 basal expression (Uchida *et al.*, 2002). The nuclear factor 1 (NF1) element found in humans, rats and mice has also been suggested to be required of for constitutive CYP1A2 expression (Zhang *et al.*, 2000).

CYP1A2 regulation is inherently challenging to study as the gene is not easily expressed in non-carcinogenic cell lines. Primary hepatocytes can, however, be used to study CYP1A2 expression (Roymans *et al.*, 2005). These cells are harvested and grown in culture for several days. However, the method of preparing and preserving primary hepatocytes is technically challenging and the expression levels of CYP1A2 begin to drop after harvesting and are undetectable within a few days. These preparations have not been recommended for long term CYP1A2 toxicological studies and must be interpreted with caution as they do not necessarily represent what occurs *in vivo* (Li *et al.*, 1999; Zhang *et al.*, 2006). For these reasons, an *in vivo* expression system is favored for CYP1A2 expression studies.

1.4 Pharmacogenetics

Pharmacogenetics can be viewed as the role of an individual's genetics in mounting a drug response. As the mechanisms with which genetics governs an individual's drug response become more apparent, one will be better able to exploit genomic tools for the use of personalized drug therapy.

Recent advancements in DNA sequencing technology are fueling the possibilities of pharmacogenetics. The cost (in 2001) of the initial draft of human genome was estimated at $500 million. However, the cost to sequence the rhesus macaque genome was a mere $22 million (Service, 2006). Advancement in cost effective sequencing technology has sparked a target cost of $1000 to sequence an individual's genome. Other DNA sequencing technologies include human SNP chips to aid in gene profiling (Ragoussis and Elvidge, 2006; Yap, 2002). These chips allow for metabolic profiling by analyzing multiple genes and their allele types. Technological advancements such as genome sequencing and SNP chips have made the use of genetics in pharmacology a clinical reality, albeit to a limited extent at present.

Pharmacogenetics is based on the premise that individual differences in susceptibility to ADRs or environmental procarcinogen activation have a genetic component. One of the first examples of a factor associated with variation in xenobiotic metabolism is glucose-6-phosphate dehydrogenase (G6PD) deficiency. This condition was first noted when WWII pilots who received their mandatory dose of primaquine suffered high incidents of hemolytic anemia when flying at moderate altitudes in unpressurized aircraft. Soldiers of African descent were more prone to this effect than Caucasians (Alving *et al.*, 1956). G6PD is an enzyme involved in dealing with oxidative stresses. Examples of oxidative stress include consumption of fava beans or

chemotherapeutics such as anti-malarials, sulfonamides, some analgesics and nitrofurans. Approximately 400 million people are G6PD deficient, a condition which usually presents with oxidative stress-induced hemolysis (Eichelbaum and Evert, 1996). There are estimated to be 50 genetic variants responsible for G6PD deficiency in humans, most of which result in a single amino acid mutation in the globin chain of the enzyme (Wiwanitkit, 2005). Though currently not routinely performed, DNA sequencing of the G6PD locus could be used to predict G6PD deficiency.

A recent example of applied pharmacogenetics has been proposed for smoking cessation therapies (Lerman and Calkins, 2006). The Cytochrome P450 2B6 (CYP2B6) enzyme is known to break down nicotine in the brain but is variably expressed in humans. Studies have shown that smokers with high levels of CYP2B6 are more likely to succeed at smoking cessation when using a nicotine replacement. Smokers with high activity CYP2B6 benefited from extended nicotine therapy. It has also been shown that variants in the dopamine receptor gene effect the outcome of cessation therapies (Lerman *et al.*, 2006; Lerman and Calkins, 2006). Also, CYP2A6 is the principle enzyme involved in nicotine inactivation as it converts 90% of nicotine to cotinine. Genetic variations of CYP2A6 have been shown by some to play a role in smoking addiction (Malaiyandi *et al.*, 2005). Since the genetic basis of the dopamine receptor, CYP2B6 and CYP2A6 are known, they can be used in the design and dosing of individual smoking cessation therapeutic regimes. There are many other examples of pharmacogenetics that may be used in clinical settings. The key to their success lies in understanding gene functions and the influences of genetic variation on these functions.

1.5 Gene regulation

Producing a protein from DNA involves many intermediate steps, all of which are potential sites of regulation. The first level at which protein production can be regulated is at the level transcription initiation. The initial stage of transcriptional control comes from the chromatin structure within the nucleus (Dillon, 2006). The expression of the coding portion of the gene is regulated by *cis*-acting elements or domains which are involved, to various degrees, in enhancing or repressing expression. Unlike an organelle, which sequesters elements using physical barriers, the nuclear environment is not divided. This means that *trans*-acting factors have access to all regions of the nucleus. Thus, the mechanism by which regions of DNA are sequestered from *trans*-acting factors is by way of chromosomal structure. One model proposes that multiple regulatory domains loop on each other to form a hub to which *trans* factors are attracted. This hub of regulatory elements acts to isolate the regulatory factors from both *cis* and *trans* elements (Dillon, 2006) .

There are essentially two types of DNA sequence regulatory elements found to control gene expression: core and non-core regulatory elements. Core sequence regulatory elements are the collection of elements that form the minimal biologically active promoter. They tend to be well conserved between genes and species. Examples of core regulatory sequences include the RNA pol II initiation site and TATA box. RNA pol II is the enzyme that produces all protein transcripts except for ribosomal and transfer RNA. RNA pol II requires multiple transcription factors in order to initiate transcription. They are known collectively as TFII's. TFII's form a holoenzyme complex with RNA pol II and bind the initiation motif in a single step. Of particular interest, TFIID, also know as the TATA box binding protein (TBP), binds to the core TATA box sequence, usually located 31 bp upstream from the transcriptional start site. The TATA box consensus sequence is TATAAAA with some variations.

Non-core sequence regulatory elements are much less conserved between species and vary drastically between genes. These short elements can either be proximal to the coding sequence or, as in the case for some enhancers, kilobases away (Uchida *et al.*, 2002). Examples proteins that bind these non-core regulatory elements are zinc finger domains, leucine zipper motifs, E-boxes, and helix loop helix motifs. Zinc fingers bind DNA and can interact with RNA pol II and TFII's. Non-core regulatory elements play a role in transcriptional repression. The combination and diversity of non-core elements allows for expression specificity among different tissues and specific responses to environmental stimuli. Transcription *in vivo* may require additional short regulatory elements in combination with the core elements. These elements are located at varying distances from the transcriptional start point (Pedersen *et al.*, 1999).

Tissue-specific transcriptions factors are a common mechanism that ensure genes are only expressed in appropriate tissues. An example of this is T protein, a protein which is responsible for mesoderm differentiation. Tissue specific expression of T protein activates genes involved in mesoderm differentiation (Kispert *et al.*, 1995). DNA methylation can also regulate gene transcription. CpG islands (which are often associated with a gene promoter) can be methylated, thereby reducing the transcriptional activity of the given gene. An example of this is seen during the inactivation of tumor suppressor genes (TSG) in neoplasm formation (Toyota and Issa, 1999). The tissue-specificity of CYP1A2 expression, which is primarily in the liver and, to a lesser extent, in nasal mucosal epithelium (Zhang *et al.*, 2000), is due to CpG methylation. This expression specificity is a result of gene hypomethlation in the liver and hypermethylation in other tissues (Jin *et al.*, 2004).

Gene expression can be regulated at the level of messenger ribonucleic acid (mRNA) degradation. The half life of a protein's transcript limits its protein production potential. An example of transcript regulation occurs in parathyroid hormone (PTH) production. A secondary effect to uremia is increased PTH synthesis. The uremic state increases the stability of the PTH mRNA transcript, allowing more PTH synthesis to occur and, in turn, greater PTH concentrations to be obtained (Yalcindag *et al.*, 1999). This transcript stabilization also occurs with CYP1A2, which is known to be post-transcriptionaly regulated in rodents. When Ahr ligands are present, there is a biological need to increase the level of some Cytochrome P450 enzymes to reestablish homeostasis. Ahr elements are believed to stabilize the mRNA in addition to increasing their rate of transcription resulting in increased CYP1A2 (Pasco *et al.*, 1988; Bigelow *et al.*, 1989).

Some mRNA's are not transcribed unless they are in specific region of the cell. This localization is hypothesized to regulate the cone growth of dendritic cells. In this case, protein synthesis occurs distally from the nucleus and must rely on mRNA transportation to regulate this protein synthesis (Crino and Eberwine, 1996). Therefore, mRNA transportation systems can be utilized as a means of controlling gene expression.

One of the last stages at which protein production can be regulated is at the level of the protein itself. Protein degradation is a mechanism used extensively by the cell to control gene expression. As an example, a protein can be targeted for destruction by ubiquitinylation. Ubiquitinylation increases a protein's affinity for the 20s protesome complex, which subsequently degrades it (Groll *et al.*, 2005). A gene whose protein has a greater affinity for ubiquitinylation will be expressed in lower concentrations due to protein degradation.

Thus, given the many levels involved in producing a protein from a gene, there are many potential sites for regulation of expression, a few examples of which are mentioned above.

1.6 Murine model

To investigate further the relationship that exists between drug response phenotypes and genomic sequences, a stable biological model is required. Since drug metabolism is influenced by genetic and environmental factors, one needs to control for environmental influences to observe the genetic effects. Controlling for environmental factors makes genetic studies in humans difficult.

Mice have frequently been used in toxicology studies because of their small size, high fecundity and for their similarities to humans in many metabolic processes. Mice have CYP1A2 enzymes similar to humans. First, both genes contain 7 exons, in which exon 1 contains non-coding mRNA sequence. Second, mice and humans are also similar in their response to aromatic hydrocarbons. Both organisms contain functional Ahr pathways, which are powerful inducers of gene expression. The promoters of mice and human *CYP1A2* display 68% homology between -150bp to the transcriptional start site and 80% homology between -841 to -758 in human and -1529 to -1439 in mice (Ikeya *et al.*, 1989).

Previously, an acetaminophen nonsusceptible (*APN*) inbred strain of mice was developed that displayed a significantly high tolerance to acetaminophen-induced hepatotoxicy relative to a coderived susceptible strain (Casley *et al.*, 1997b). These *APN* mice expressed low levels of CYP1A2 mRNA relative to an acetaminophen susceptible (*APS*) strains (Casley *et al.*, 1997a). Caffeine has been used as a probe drug to measure CYP1A2 expression in both mice and humans.

In mice, more than 80% of the caffeine 3-demethylation (C3-D) to paraxanthine has been attributed to CYP1A2 activity, using *Cyp1a2* knock-out mice (Buters *et al.*, 1996). The ratio of 1,7-dimethylxanthine/caffeine in serum has been used to measure CYP1A2 activity. Figure 1.2 shows the parent *APN* mouse strain, which was found to have a lowered C3-D index than other common laboratory strains including the *C3H/HeJ*(Casley *et al.*, 1997a). This phenotype had the characteristics of a complex genetic trait. The principal genetic determinants of the variation between *APN* and *C3H/HeJ* mouse strains were mapped as quantitative trait loci (QTL's) using an interval mapping approach (Casley *et al.*, 1999).

One QTL was mapped with strong a statistical significance as a 4.8cM interval on chromosome 9 (Casley *et al.*, 1999). This region on chromosome 9 contained the *Cyp1a2* locus. Preliminary sequence analysis of the *APN* and other common laboratory mice showed several polymorphic sites in the non-coding region of the gene, one of which was -76 bp from the transcription start site. It was hypothesized that this polymorphism accounted for the differences seen between *APN* and *C3H/HeJ*. It was further hypothesized that the constitutive expression of mouse CYP1A2 was regulated, in part, by an element in the proximal promoter region of the gene that includes this polymorphism.

Figure 1.2 The mouse strains used for genetic analysis. Above is the *APN* mouse and below is the *C3H/HeJ* mouse.

1.7 Objectives of this study

The overall objective of this study was to identify a mechanism governing the basal expression of CYP1A2 using genetic variation between mouse strains.

Specifically, the primary experimental objective was to ascertain the significance of a point mutation within the *Cyp1a2* promoter and determine if this polymorphism is the specific genetic variant underlying the previously identified QTL on chromosome 9, which, in part, results in lower constitutive CYP1A2 expression in *APN* mice.

The experimental approach was to first determine the level at which differential CYP1A2 expression was occurring, then to establish the functional significance of the -76 polymorphism to transcriptional activity.

CHAPTER 2: MATERIALS AND METHODS

2.1 Animal model

The *APN* mouse strain was originally derived from an outbred Swiss-Webster mice colony at Health Canada originating from Connaught Laboratories (Toronto, ON) in 1968. Breeding pairs were selected for non-susceptibility to acetaminophen toxicity and decreased C3-D (Casley *et al.*, 1997b). Congenic mouse strains containing *C3H/HeJ* chromosomal segments in an *APN* background were produced through phenotypic and marker-assisted selection by Ms. Yelena Markovich. The congenic intervals of strains *C1*, *C4* and *C9* encompassed the QTLs affecting C3-D, identified on chromosomes 1, 4 and 9 respectively. The *C3H/HeJ* and *BALB/c* strains were purchased from Charles River Laboratories, Inc. (Wilmington, MA, USA). The *C3H/HeJ*, *APN* and congenic strains were maintained at the animal care facility of the Health Protection Branch of Health Canada (Ottawa, ON). *BALB/c* male mice were purchased at 6 weeks of age and were housed individually for 2 weeks prior to injections at the animal care facility of the Health Protection Branch of Health Canada. All animals were fed irradiated Prolab 3000 rodent feed (Agway Inc. Portsmouth, RI, USA) and sterile water *ad libitum*. Mice were housed on sterile Sanichip maple chip bedding (P.W.I Industries, St. Hyacinthe, PQ). Animals were cared for in accordance with the guidelines of the Canadian Council on Animal Care.

2.2 Quantification of CYP1A2 mRNA

The cDNA preparation was done by Mrs. Margeryta Nowakowska. To sum, livers were harvested after cervical dislocation followed by decapitation from 4 of each; *APN*, *C3H/HeJ*, *C1*, *C4*, and *C9* male mice and stored in RNAlater (Ambion, Austin, TX, USA). RNA was extracted from ~20mg of liver using a spin column method (Qiagen RNeasy, Mississauga, ON) according to the manufacturer's protocol. A high capacity oligo dT primed cDNA synthesis kit (Applied Biosystems, Foster City, CA, USA) was used to produce cDNA from 2μg of total RNA in a 100μL final volume. The total RNA of each strain was quantified using a Heliose UV spectrophotometer (Waltham, MA, USA) and pooled at concentrations of 30ng/μL. Quadruplicate technical replicate reactions of each sample were performed in 25μL total volume that contained pooled cDNA template derived from 75ng RNA, 1xTaqman polymerase chain reaction (PCR) universal master mix and CYP1A2 (Mm00487224-m1) or β2-microglobulin mRNA (Mm00437762-m1) commercially available assays on demand (Applied Biosystems, Foster City, CA, USA). The assays were carried out on a ABI 7500 quantitative PCR (qPCR) Fast System (Applied Biosystems, Foster City, CA, USA) and the CYP1A2 mRNA levels were normalized to β2-microglobulin . The ABI 7500 Fast System software SDS v1.3 (Applied Biosystems, Foster City, CA, USA) was used to calculate the 95% confidence intervals of the means.

2.3 Quantification of CYP1A2 heterogeneous nuclear RNA (hnRNA)

The RNA from 5 male *APN* and 5 male *C3H/HeJ* animals was extracted from ~20mg of mouse liver as in 2.2. In order to measure hnRNA, a DNA-free RNA preparation was required. RNA samples were eluted with 50μL dH_20 and treated with DNase (Ambion, Austin, TX, USA) according to the manufacturer's specifications. Briefly, 9μL of isolated RNA solution was

incubated in 1xDNase I Buffer and 2U of rDNase I at 37°C for 25min. Samples were incubated for 2min at room temperature with 5μL of DNase inactivation reagent. RNA was recovered from the DNase inactivation resin by spinning the sample for 1.5min at 10Kxg and retaining the supernatant.

The DNA-free RNA was converted to cDNA using an oligo dT primed cDNA Synthesis System (Roche, Laval, QC) and a hnRT primer (Invitrogen, Carlsbad, CA, USA) specific for CYP1A2 intron 2 (Table 2.1). This provided templates for GAPDH mRNA and CYP1A2 hnRNA. The 20μL reaction consisted of 1μL (~60ng) of DNA-free RNA, 1μL dT12-18 , 1μM hnRT primer in 1x1st strand buffer, 10mM DTT 10mM dNTPs and 25U of AMV reverse transcriptase. Reverse transcription was performed at 70°C for 10min using a PTC-200 thermocycler (MJ Research, Watertown, MA, USA), then stored at -20°C.

The levels of CYP1A2 cDNA were measured using an ABI 7500 qPCR Fast System (Applied Biosystems, Foster City, CA, USA). These reactions were carried out in 20μL volumes with 1xTaqman PCR mix, 1μL of hncDNA template and either CYP1A2 (Mm00487227_m1) or GAPDH (loc14433Mm99999915_g1) assays on demand (Applied Biosystems, Foster City, CA, USA). The cDNA from 5 animals from each strain were assayed. The levels of CYP1A2 mRNA were normalized to GAPDH mRNA. Results are the means of 5 animals for each strain. The ABI 7500 qPCR Fast System software SDS v1.3 (Applied Biosystems, Foster City, CA, USA) was used to calculate the 95% confidence interval of the means.

Quantification of hnRNA was done using an Applied Biosystems SYBR green qPCR kit (Foster City, CA, USA). Reactions of 20μL were prepared using 1xSYBR green mix, 1μL of

cDNA template, 1μM of hnRNA specific-primers hnCYP1A2 F and hnCYP1A2 R (Table 2.1) (Invitrogen Carlsbad, CA, USA). GADPH expression was measured by similar qPCR conditions but with 1μM GAPDH F and GAPDH R primers (Table 2.1). The means and 95% confidence intervals were calculated using GAPDH normalized individuals. The correct product amplification was verified by agarose gel electrophoresis and pCR2.1 cloning with subsequent DNA sequencing.

Table 2.1 Oligonucleotides used in quantification of CYP1A2 hnRNA. Positions are based on Genebank (Benson *et al.*, 1994) data as follows:
[a]Gene bank accession # MMCY03, (Kimura *et al.*, 1984)
[b]Gene bank accession # M32599, (Sabath *et al.*, 1990)

primer name	sequence (5'-3')	position bound
hnRT	AGTGAGTCCTCAGCAGAAT	3091-3109[a]
hnCYP1A2 F	TCCCAGCTGCTGACTAGAAG	2891-2901[a]
hnCYP1A2 R	GGCCAGCCTGGGATTTAGA	2992-3010(c)[a]
GAPDH F	GGAGCCAAACGGGTCATCAT	nt 383-402[b]
GAPDH R	TCACGCCACAGCTTTCCAGA	nt 633-614(c)[b]

2.4 Cloning into pCR2.1 vector

The freshly amplified products were immediately ligated into the pCR2.1, TA cloning vector according to the manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). Ligation conditions were 2μL of the amplified product, 1xT4 Ligation buffer and 4U T4 ligase in 10μL final volume at 14°C for 16h. Transformation reaction consisted of 50μL chemically competent *E.coli* INVαF' cells (Invitrogen, Carlsbad, CA, USA) with 2μL ligation mix. Reactions were prepared on ice, heat shocked for exactly 30s at 42°C then immediately placed back on ice. 250μL of room temperature S.O.C. medium (Invitrogen, Carlsbad, CA, USA) was added to the transformed cells which were

subsequently plated on 50μg/mL kanomycine Luria-Bertani (LB) agar containing X-gal and grown at 37°C overnight.

Colonies which were positive by X-Gal selection were screened by PCR for the presence of the ligated products. Colonies were marked then a portion of the colony was picked with a 10μL pipette tip and placed in 1mL of dH_2O. This suspension was snap frozen in liquid N_2 and then thawed. Next, 2μL of the freeze thaw suspension was added to a 20μL PCR reaction containing 1xPCR buffer, 0.2μM dNTP (Roche, Laval, QC), 0.4μM M13 F/ M13 R oligonucleotides which flank the insertion site (Invitrogen, Carlsbad, CA, USA) (Table 2.3) and 1.5U *taq* polymerase (Roche). Amplification was carried out in a PTC-200 thermocycler using 94°C for 2min then 29 cycles of [94°C for 10s, 55°C for 30s and 72°C for 45s] (MJ Research, Watertown, MA, USA). The amplified products were resolved on an agarose gel impregnated with 1xSYBR safe (Invitrogen, Carlsbad, CA, USA).

Table 2.2 Oligonucleotides used in pCR2.1 screening assay.

primer name	sequence (5'-3')
M13 F	GTAAAACGACGGCCAGTG
M13 R	GGAAACAGCTATGACCATG

Colonies containing inserts of the correct size were grown in 4mL of 50μg/mL kanamycin Lb broth overnight at 37C°, 1mL of which was archived in 15% glycerol at -80 C°. Plasmid Mini-preparations were performed with 1mL of the remaining broth, using a 96 well Millipore Montage system (Billerica, MA, USA) according to the manufacturer's specification.

2.5 pCR2.1 sequencing

Sequencing reactions were performed on the purified pCR2.1-insert plasmids using the vector-specific M13 primers (Table 2.2). The 20μL sequencing reactions contained 2μL of Ready Reaction (ABI BigDye v3.0), 1xSequencing buffer 0.1μM of M13 F or R primer and ~300ng pCR2.1 plasmids. Reactions were performed in either thermocycler tubes or 96 well thermocycler plates using a PTC-200 thermocycler for 26 cycles of [1.0°C/s to 96°C , 96°C for 10s, 1°C/s to 50°C, 50°C for 5s, 1°C/s to 60°C and 60°C for 4min] (MJ Research, Watertown, MA, USA). Unincorporated dye label was removed using a Montage Seq96 kit (Millipore, Billerica, MA, USA) according to the manufacturer's specifications. Reaction products were eluted from the cleaning plate by shaking for 20min in elution buffer.

The sequencing fragments were resolved on a ABI 3100 sequencer equipped with ABI Prism data collection software v1.1 (Ann Arbor, MI, USA) using 80cm POP4 polymer filled capillaries (Applied Biosystems, Foster City, CA, USA). The sequence data were analyzed using Sequencher software v4.5 (Ann Arbor, MI, USA).

2.6 Genomic DNA isolation

Genomic DNA was isolated from mouse livers using a spin column (Qiagen DNeasey, Mississauga, ON) according to the manufacturer's protocol. To summarize, ~20mg of fresh liver pieces were added to 180μL ATL buffer, 20μL proteinase K and incubated at 55°C for 1 hour with occasional vortexing. Next, 200μL of AL buffer was added and incubated at 70°C for 10min. The reaction was cooled to room temperature and 200μL of anhydrous ethanol was added. This was loaded into a DNeasy mini spin column and spun at 8Kxg for 1min. The column was washed

with 2x250μL of AW1 buffer, and then 500μL AW2 buffer and air dried. DNA was eluted by incubating the resin in 200μL AE buffer for 1min, followed by centrifugation at 8Kxg for 1min. The extract was quantified spectrophotometricaly and stored at -20°C.

2.7 Reporter vector constructs

Reporter constructs were produced for the both the *APN* and *C3H/HeJ Cyp1a2* 5' regions via a two-step cloning process into pCR2.1 then pGL3-basic.

In the first step, a ~5Kb product spanning the 5' region and including exon I was amplified. The primers used in these amplifications had 5' modifications to add a 5' *MluI* or a 3' *XhoI* site (Invitrogen, Carlsbad, CA, USA) (Table 2.3). High fidelity amplifications were performed using HF 2 Advantage Polymerase (Clonetech, Mountain View, Ca, USA). Both reactions were performed in 20μL final volume and contained; 1xHF 2 buffer, 1xHF 2 dNTP's, 50ng of genomic DNA (from 2.6), 0.4μM of 18*MluI* F primer, 0.4μM of 5176*XhoI* R primer (Table 2.3) and 1xHF 2 *taq* polymerase. HF 2 cycling conditions were; 95°C for 15s, 30 cycles of [95°C for 15s and 64°C for 60s] then 64°C for 5min. Products were sized using 1xSYBR safe-impregnated (Invitrogen, Carlsbad, CA, USA) 0.5% agarose gel electrophoresis and immediately ligated into pCR2.1 vectors and cloned into *E.coli* (as in 2.4).

In the second step, the inserts were digested out of pCR2.1, gel purified and ligated into pGL3-Basic expression vectors from Promega (Madison, WI, USA). Double digests were performed on 1μg of both the pGL3 vector and the pCR2.1-insert vectors. The 50μL digests reactions contained 1xNEB buffer 2, 5μg BSA, 40U *XhoI* and 20U *MluI* restriction endonucleases

(New England Biolabs Ipswich, MA, USA). Digests were performed using a PTC-200 thermocycler (MJ Research, Watertown, MA, USA) by incubating at 37C° for 1h then 70C° for 5min.

The pCR2.1-insert digests were resolved in a 1% agarose gel impregnated with 1xSYBR safe (Invitrogen, Carlsbad, CA, USA). Two 100μg gel slices containing the *APN* and *C3H/HeJ Cyp1a2* 5' regions were purified using spin column from Qiagen DNeasy kit (Mississauga, ON) according to the manufacturer's protocols.

The linearized pGL3-Basic was dephosphorylated using Antarctic Alkaline phospatase (New England Biolabs Ipswich, MA, USA). The reaction was carried out by adding 5.5μL of 10xAntarctic alkaline phosphatase buffer and 10U of Antarctic alkaline phosphatase to the 50μL digestion reaction. The reaction was subject to a 37°C incubation for 1h followed by a 70°C incubation for 5min.

Ligation of pGL3-Basic to the purified 5' *Cyp1a2* fragment was done by combining 10μL of the lineraize pGL3-Basic, 10μL of the gel purified insert, 2.2μL of 10xLigation buffer and 400U of T4 ligase (New England Biolabs Ipswich, MA, USA). Ligation reactions took place at 20°C for 10min. Ligation mixtures were transformed into INVαF' chemically competent *E.coli* cells (as in 2.4).

Select colonies were tested for the presence of the ligated products, using the snap freeze method as described in 2.4, except that pGL3 vector-specific primers RV3 and GL2 (Promega,

Madison, WI, USA) were used (Table 2.3). The amplified products were resolved on a 1% agarose gel impregnated with 1xSYBR safe (Invitrogen, Carlsbad, CA, USA).

The sequence of the insertion sites of the pGL-*Cyp1a2* reporter constructs were verified by DNA sequencing (as in 2.5) using RV3 or GL2 (Table 2.3) as sequencing primers (Promega, Madison, WI, USA).

Table 2.3 Oligonucleotides used in reporter vector constructs. Restriction sites are shown in bold face type.

primer name	sequence (5'-3')
18*MluI* F	**ACGCGT**ATCTGGGTCTGGGAGGTATC
5176*XhoI* R	**CTCGAG**CTGCAGGGGAAATGGAAAGATG
RV3	CTAGCAAAATAGGCTGTCC
GL2	CTTTATGTTTTTGGCGTCTTCC

The control vectors; pGL3-Basic pGL3-Control and thymidine kinase promoter-driven renilla luciferase (pRL-TK) (Promega, Madison, WI, USA), were prepared by transfecting each control vector into INVαF' cells (as in 2.4). Once all the transformants were established, vectors were grown in batches of 100mL of 100μg/mL amplicillin LB broth. Plasmids were purified using columns from a Qiagen Maxi Endotoxin free prep kit, according to the manufacturer's instructions (Mississauga, ON). The final DNA pellet was dissolved in 200μL of Endotoxin-free tris ethylenediaminetetraacetic acid (TE) buffer.

A two step quantification approach was used for plasmid quantification. First, three separate 1 in 100 dilutions were made from the concentrated stocks. The accuracy of these dilutions was verified by gel electrophoresis. Samples were run through 1xSYBR safe (Invitrogen, Carlsbad, CA, USA) 0.5% agarose and imaged using a Kodak 2000R imaging station and Kodak 1D v3.6 (New Haven, CT, USA). The consistency of the dilutions was measured by band densitometry. The final DNA concentration was determined spectrophotometricaly. Each of the three dilutions was measured in triplicate and the means calculated. The molar mass of each vector construct was calculated and used to calculate their molar concentrations.

2.8 *In vivo* reporter construct delivery

Each animal received 5pmoles of one of the pGL3-derived vectors and 0.2pmol of a pRL-TK vector (Promega, Madison, WI, USA). Reporter vectors were stored and dissolved in 1xEndotoxin free TE (Qiagen, Mississauga, ON). Two minutes prior to injection, the vectors were diluted in room temperature Mirus TransIT® QR (Madison, WI, USA) delivery solution equivalent to 10% of the mouse body weight. Mice were anesthetized prior to injection with isofluorine. Tail vein injections were performed by Mrs. Carole A Westwood, an animal surgeon at Health Canada. Mice were warmed under a heat lamp for approximately 5 min prior to injection. A 27 gauge needle was used to enter the tail vein and the entire syringe volume was injected within 5-9s. Mice injections were randomized to ensure no fitness bias. Animal's whose injections were judged to be imperfect were removed from further analysis. Examples of imperfect injections include: the mouse died during/post procedure or the tail vein ruptured before the entire solution could be delivered. Vector expression assays were all performed 18h after injections.

2.9 *In vivo* expression assays

Mice were divided into groups of 3 and injected *IP* with 100μL of 30mg/mL D-luciferin (Molecular Imaging Products, Ann Arbor, MI, USA) in sterile saline and pen incubated for 5-15min. The animals were anesthetized using isoflurane and imaged for 5min using an *in vivo* Imaging System 50 (IVIS) from Xenogen (Alameda, CA, USA). Bio luminescent imaging (BLI) data were compiled and manipulated using IGOR pro 4.9A software (Alameda, CA, USA).

2.10 *Ex vivo* expression assay

Vector delivery for the *ex vivo* expression assay was the same as 2.8. After the 18h incubation, the mouse livers were harvested, gall bladder removed, and livers flash frozen in liquid nitrogen as described by Schuetz *et al* (2002). Livers were homogenized with an electric Polytron for 15s in 5mL of homogenization buffer; 0.1M Tris-HCL, pH 7.8, 2mM EDTA and 0.1% Triton X-100. Cell homogenates were diluted 1/5 in homogenization buffer and spun at 10Kxg for 30min at 4C°. Lysates were decanted into 1.5mL microcentrifuge tubes and stored at 4C°.

Lysates were assayed using a DYNEX Revelation 3.2 (Roodepoort, South Africa) or a Luminoskan Ascent (Ramsey, MI, USA) luminometers. Three 40μL technical replicates of each liver lysate were measured in a 96 well plate using a dual luciferase Promega Stop and Glow kit (Madison, WI, USA). Each assays consisted of; a 100μL injection of Luciferin reagent, detection of 10s, then a 100μL injection of Stop and Glow solution, and another detection of 10s.

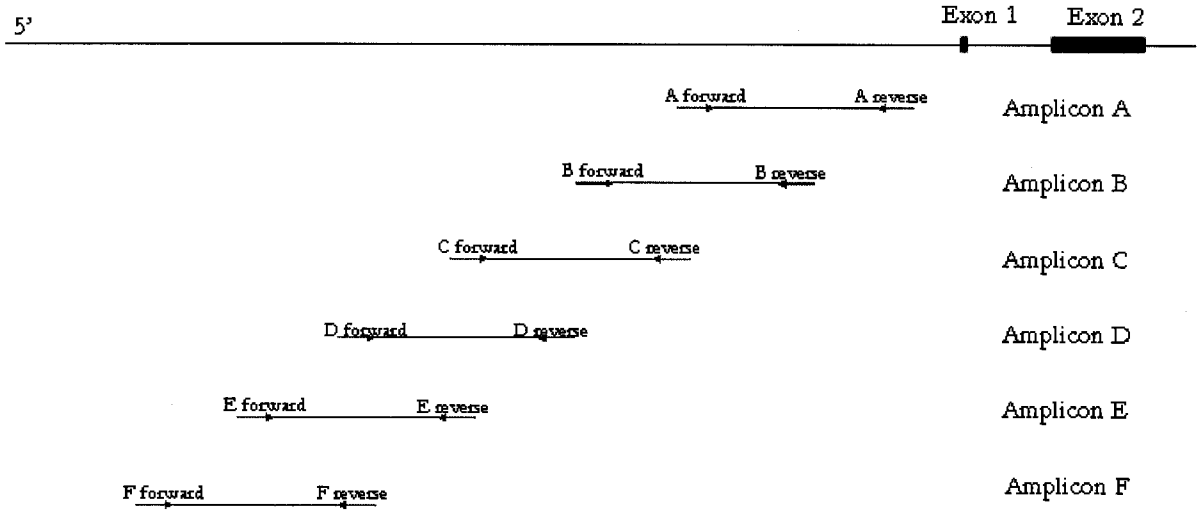
2.11 Data analysis

Dr. Aleksander Bliu, a Health Canada Statistician, was consulted for the appropriate statistical methods. The analysis set was defined as all subjects enrolled and successfully treated in the study. The initial calculations of the data were performed using Microsoft Excel. Means of the three luciferase and renilla readings were calculated for each sample. The luciferase/renilla ratio was calculated from these averages. Descriptive statistics of each vector group were calculated and included mean, minimum, maximum, range, and standard deviation. The hypothesis that reporter vector strengths differed was analyzed using Mann-Whitney two-sample test which was a non-parametric test for comparing two independent samples. First, the Kruskal-Wallis test, a non-parametric test analogue with one-way analysis of variance ANOVA, was used to test the overall difference among all groups. Given that this test showed an overall statistically significant difference amongst the vector groups (p value=0.0006), a Mann-Whitney two-sample test was used to test these differences. All tests were performed at 95% significant level. Calculations were performed using SAS v9.1 (Cary, NC, USA).

2.12 Promoter sequencing

Genomic DNA was extracted for sequencing from 3 *APN* and 3 *C3H/HeJ* mice (as in 2.6). Each animal's *Cyp1a2* gene was sequenced in an identical fashion. Amplification primers sets were synthesized based on known sequence data from the mouse genome database. Approximately 5.1kb of the *Cyp1a2* 5' region was divided into 6 overlapping fragments. These fragments were amplified, cloned into pCR2.1 as described in section 2.4, and sequenced (Figure 2.1).

Figure 2.1 Schematic of 5' *Cyp1a2* locus sequencing strategy. The black line represents 5.1kb of the *Cyp1a2* 5' locus with exons 1 and 2 being the black boxes. The green arrows show the binding positions of the amplifications primers (Table 2.4) relative to the *Cyp1a2* gene and the red lines represent the amplified products.


5'
Exon 1
Exon 2
A forward
A reverse
Amplicon A
B forward
B reverse
Amplicon B
C forward
C reverse
Amplicon C
D forward
D reverse
Amplicon D
E forward
E reverse
Amplicon E
F forward
F reverse
Amplicon F

2.13 Sequencing amplicons

High fidelity amplification reactions were performed using HF Advantage 2 polymerase (Clonetech, Mountain View, CA, USA) as described by the manufacturer's instruction. Amplicon reactions contained a forward and reverse primer pair for A, B, C, D, E or F (Table 2.4) designed using Clone Suite Manager 7 (Cary, NC, USA) and synthesized by Invitrogen (Carlsbad, CA, USA). Briefly, 20μL reactions containing HF 2 buffer, 1xHF 2 dNTP, 50ng of genomic DNA template (from 2.6), 0.4μM forward primer, 0.4μM reverse primer and 1xHF 2 advantage polymerase. HF amplification conditions were 30 cycles of [95C° for 15s and 68C° for 2min]. Amplicons were cloned into pCR2.1 vectors and sequenced as in 2.4 and 2.5.

Table 2.4 Oligonucleotides used to produce amplicons for *Cyp1a2* sequencing.

primer name	sequence (5'-3')	predicted amplicon (bp)
A forward	TGGCTGTCCTGGAACTCACTC	775
A reverse	TGGCTGGTCACATGGCATCTG	
B forward	TGTCTGCTTTCGGCTACGATG	799
B reverse	CAGAGGCAGGCAGATTTCTGA	
C forward	ACTCTTCCCACCACGTGAGTC	674
C reverse	GGACCTGGTCTGGACATGGAT	
D forward	GCCAGCCTGGTCTACATAGTG	680
D reverse	TGTTGAGACGGCTCAGTGAGT	
E forward	CATCCGCTGTTCCTCACTACC	683
E reverse	GGGAAGGAAGGCGCACCATTT	
F forward	CTTCATGTGCTCTGGCAGTTC	785
F reverse	GAGACCACTCGTGAGGCAATC	

2.14 Comparative sequencing of *Cyp1a2* 5' region

The sequenced *APN* and *C3H/HeJ* amplicons A to F were each assembled into contigs using Sequencher v4.5 (Ann Arbor, MI, USA). The chromatograms were manually examined for ambiguities and the sequence edited as appropriate. Gaps within the sequence were filled by resequencing the plasmids. The sequence was deemed to be complete when each of the 6 animals had double stranded coverage of the 5' region. The *APN* and *C3H/HeJ* sequences were analyzed for homology using Sequencher v4.5.

2.15 Database search

The National Center for Biotechnology Information (NCBI) mouse SNP database was used to identify previously discovered polymorphic regions of the *Cyp1a2* 5' region. This was done by manually searching the region proximal to the transcription start site. MatInspector (Higo *et al.*, 1999) was used to search the *Cyp1a2* 5' region for regulatory elements *in silico*.

2.16 Nuclear protein extraction

The *C3H/HeJ* nuclear protein extraction procedure was adapted from Gorski *et al* (1986). Mice were sacrificed by cervical dislocation and subsequent decapitation. Whole liver lobes (~2g per adult) were removed, immediately chilled and diced on ice. Subsequent, buffers, tubes and manipulations were at 4°C or on ice for the remainder for the extraction.

The tissue was added to 4.5mL extraction buffer (10mM HEPES, 25mM KCl, 0.15mM spermine, 0.5mM spermidine, 1.0mM EDTA, 10% glycerol, 2.5M sucrose, protease complete cocktail (Roche, Laval, QC) at pH 7.6) and placed in a 50mL homogenizer. The piston was stroked 10 times using a fixed electric drill and the extent of cell lyses was determined by

microscopic examination. When over 90% of the cells were lysed, extraction buffer was added until the total volume was 13mL. The homogenate was layered on a 4mL 2M-sucrose cushion in a Beckman ultracentrifuge tube. Tubes were then secured in a SWTi-41 rotor (Beckman, Fullerton, CA, USA) and spun at 4°C for 30min at 100Kxg. The supernatant was decanted and the pellet resuspended in 3.0mL of Lysis buffer (10mM HEPES, 100mM KCl, 0.1mM EDTA, 1mM DTT, 3mM $MgCl_2$, 10% glycerol, Roche Complete Protease cocktail, pH 7.6). Once dissolved 300μL of 4M ammonium sulfate was added, mixed on a rocker shaker for 30min, and then spun at 100Kxg for 1h to remove chromatin. The supernatant was saved and a 100μL aliquot was quantified using a Bradfrod total protein quantification assay (Biorad, Hercules, CA, USA).

Protein was precipitated by slowly dissolving ammonium sulfate in the supernatant until the final concentration was 0.33g/mL. This was spun at 100Kxg and the pellet was resuspended in Storage buffer (25mM HEPES, 100mM KCl, 0.1mM EDTA, 1mM DTT, 10% glycerol, 1mM PPT pH 7.6) to a concentration of 5μg/μL. The sample was dialyzed using a 10'000 MWCO dialysis cassette (Pierce, Rockford, IL, USA) three times for 1h against 1L of Storage buffer and stored in aliquots at –80°C.

2.17 DNase footprinting probe

The DNA probes were constructed using previously sequenced pCR2.1 vector constructs containing the proximal *APN* and *C3H/HeJ Cyp1a2* 5' locus. These constructs were prepared by Ms. Lori Lavigne. Amplification of the 310bp proximal region was performed using 1xPCR buffer, 0.2μM dNTP's (Roche, Laval, QC), 0.4μM of probe-R1, 0.4μM probe-F1 (Sigma-Genosys, Oakville, ON) (Table 2.5), 1.5U *taq* polymerase in 20μL final volume and the probe amplification

program was run using a PTC-200 thermocycler (95°C for 2min then 30 cycles of [94°C for 1min, 55°C for 1min and 72°C for 1min] finally 72°C for 10min) (MJ Research, Watertown, MA, USA). Footprinting vectors were produced by cloning these fragments into pCR2.1 vectors (as in 2.4). The sequence of these fragments was verified (as in 2.5).

Three fluorescently labeled probes were produced, all labeled 5'; (1) a FAM 310bp labeled *C3H/HeJ* probe, (2) a NED labeled *C3H/HeJ* probe and (3) a FAM labeled *APN* probe. DNA probes were produced in 100μL reactions containing 1xPCR buffer, 0.2μM dNTPs (Roche, Laval, QC), 0.4μM TA NED or TA FAM labeled primers (Applied Biosystems, Foster City, CA, USA), 0.4μM TA R primer (Invitrogen, Carlsbad, CA, USA) (Table 2.5), 10ng of vector template, and 1.5U *tac* polymerase. PCR amplifications were carried out using the probe amplification program (details above). Gel purification (Qiagen, Mississauga, ON) was used to purify the DNA probes according to the manufacturer's protocol.

Table 2.5 Oligonucleotides used to produce labeled *Cyp1a2* probes for DNase protection assay.

primer name	sequence (5'-3')
Probe R1	GGGAGTCAGTCCAGGACTTC
Probe F2	TGAACCCTGCCCTGTCTTTG
TA NED F	NED-CGCCAGTGTCGTGGAATTCG
TA FAM F	FAM- CGCCAGTGTGCTGGAATTCG
TA R	CGCCAGTGTGATGGATATCT

These initial probes were used for a second PCR reaction to increase the yield of probe and to further dilute any remaining unlabeled template. The conditions were the same as before, only

the products were purified using a Millipore Microcon-PCR clean up kit (Billerica, MA, USA) according to the manufacturer's protocol. Each probe was run in a 1xSYBR safe (Invitrogen, Carlsbad, CA, USA) 1% agarose gel to assess purity and size. Then, DNA probes were quantified using a spectrophotometer, aliquoted, and stored at –20°C.

2.18 DNase footprinting assay

The DNase protection assay was performed using a Core Footprinting system from Promega (Madison, WI, USA). The binding reactions were carried out in 50μL total volume which contained 25μL Binding Buffer, 40ng of labeled probes and 5μg of BSA or 5 to 25μg of hepatic nuclear protein extract (2.16). The reactions were incubated for 10min at 0°C and then for 1min at RT with of 50μL of Ca^{++}/Mg^{++} solution. DNase digestion was done by gently mixing in 3μL of RQ1 and incubating the mix for 1min at RT. The digestion was stopped by adding 90μL of the stop solution.

The restriction fragments were extracted using Phenol: chloroform: isoamyl alcohol. Briefly; 200μL of Phenol: chloroform: isoamyl was added, vortexed, centrifuged at 14Kxg for 10min and 170μL of upper aqueous phase was added to tube containing 500μL of 100% EtOH. Precipitation was done on ice for 20min, then centrifuged at 14Kxg for 10min. The pellet was washed twice with 70% EtOH by subsequent vortexing and centrifugation and then resuspended in 12μL of loading solution.

Prior to sequencing 1μL of TAMRA labeled 350bp molecular weight markers (MWM) was added to each digest reaction. The fragments were resolved on a ABI 3100 sequencer equipped

with ABI Prism data collection software v1.1 (Ann Arbor, MI, USA) using 80cm POP4 polymer filled capillaries (Applied Biosystems, Foster City, CA, USA).

2.19 Site directed mutagenesis of pGL-*C3H/HeJ* vector

Site directed mutagenesis of the -76 G>A was carried out using site-directed mutagenesis kit (Stratagene, Cedar Creek, Texas, USA). A 50μL reaction was set up containing 1xReaction buffer, 1x dNTP mix, 25ng of pGL-*C3H/HeJ* vector (2.3.2), 125ng of both MutA-F and MutA-R primers (Invitrogen, Carlsbad, CA, USA) (Table 2.6) and 2.5U of PfuTurbo DNA polymerase. The thermocycling conditions were; 95°C for 1min followed by 18 cycles of [95°C for 50s, 60°C for 50s and 68°C for 11min] then 68°C incubation for 7min.

The newly synthesized *C3H/HeJ* -76 mutant (*MUT*) vector was transfected into INVαF' cells (as in 2.4). These cells were grown on 100μg/mL amplicillin LB agar plates and screened for the mutated -76 A residue by DNA sequencing. The region containing the -76 site was amplified using a 20μL reaction which consisted of; 2μL freeze thaw colony DNA, 1xPCR buffer, 0.2μM dNTP(Roche, Laval, QC), 0.4μM Amp F and Amp R oligonucleotides (Invitrogen, Carlsbad, CA, USA) (Table 2.6) and 1.5U *taq* polymerase. Sequencing reactions were performed on each amplified product. Sequencing reactions were 20μL in volume and contained; 7μL amplified template, 1xSequencing buffer, 2μL Reaction Ready solution (Applied Biosystems, Foster City, CA, USA) and 0.2μM of Seq F or Seq R primer (Invitrogen) (Table 2.6). The rest of the sequencing method was identical to that in 2.5. A mutation positive colony was grown in a large batch 100μg/mL amplicillin LB broth and prepared for injection (as in 2.7)

Table 2.6 Oligonucleotides used in -76 G>A site directed mutagenesis and *MUT* plasmid sequencing. The -76 positions is indicated by the bold face type.

primer name	sequence (5'-3')
MutA-F	CCCAGGTGCCCGGGTCACTTGATAT**T**ATAGCTGGATGCTG
MutA-R	CAGCATCCAGCTAT**A**ATATCAAGTGACCCGGGCACCTGGG
Amp F	GGCTGTCCTGGAACTCACTC
Amp R	GTCCAGGGCTGAGAATTAAG
Seq F	TGCTGCACAAAGCTGAAAGG
Seq R	CTCGAGTTCTGGGTTCCGACG

2.20 Functionality of the -76 nucleotide

The *MUT* vector was tested for transcription function *in vivo* using *BALB/c* mice. The vector deliveries for the *MUT* vector were the same as in 2.8. The *MUT* vector was initially assayed *in vivo* using the IVIS system along with pGL-*C3H/HeJ* and pGL-*APN* vectors (as in 2.9). It was then assayed *ex vivo* as in 2.10 and the samples were tested for statistical significance as in 2.11.

CHAPTER 3: RESULTS

3.1 Quantification of CYP1A2 mRNA

Congenic mouse lines were used to confirm that CYP1A2 basal expression was regulated as a complex genetic trait. Mouse lines *C4* and *C9*, which contained fragments from *C3H/HeJ* chromosomes 4 and 9, respectively in an *APN* genetic background, showed an increase in relative CYP1A2 mRNA levels when compared to the *APN* mouse (Figure 3.1). The *C1* mouse line, however, did not show increased gene expression in spite of having elevated C3-D activity relative to *APN* (Figure 3.1). The sum of these relative *C4* and *C9* CYP1A2 mRNA levels (Figure 3.1) was equal to that of the *C3H/HeJ* mRNA within experimental error. The levels of relative *C4* and *C9* CYP1A2 mRNA suggest that the low CYP1A2 mRNA phenotype observed in the *APN* mouse strain relative to the *C3H/HeJ* mouse strain was the result of variation in at least 2 genetic loci located on chromosomes 4 and 9.

3.2 Quantification CYP1A2 hnRNA

CYP1A2 hnRNA levels were measured to investigate variations occurring as a result of differential transcription rates or premature mRNA degradation. The bottom panel of Figure 3.2 suggests the qPCR amplification reactions were specific for either GAPDH or CYP1A2 hnRNA products in both mice strains (lanes 4, 5 and 6, 7). Confirmation of the hnRNA product was obtained by DNA sequencing (data not shown). Lane 3 suggests the hnRNA preparation contained no contaminating *Cyp1a2* genomic template (Figure 3.2 bottom).

The ratios of CYP1A2 hnRNA and mRNA proved to be the same, within experimental error, for both mouse strains (Figure 3.2 Top).

Figure 3.1 The relative quantity of hepatic steady state CYP1A2 mRNA from 4 different mouse strains relative to *APN* evaluated by qPCR. Congenic strains are denoted by "C" and a number which corresponds to the *C3H/HeJ* chromosomes from which the congenic interval was derived in the *APN* background. Each bar represents the mean of 4 technical replicates of the pooled cDNA from 5 male mice. Error bars are standard errors of the means.

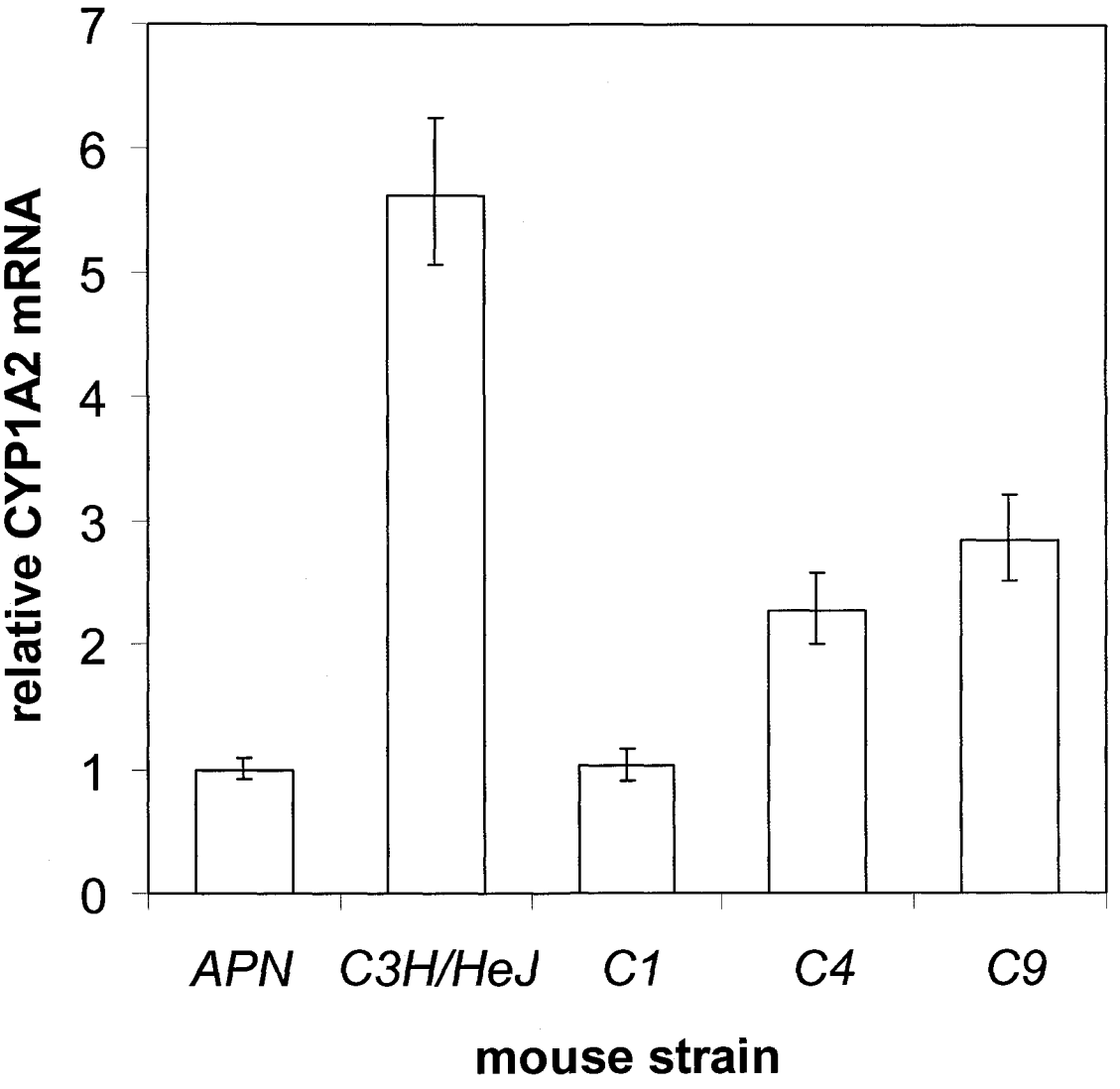
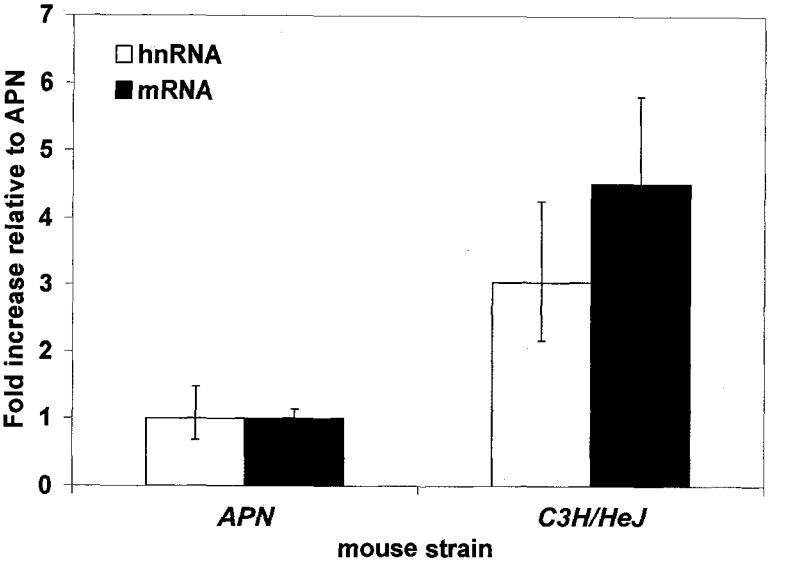
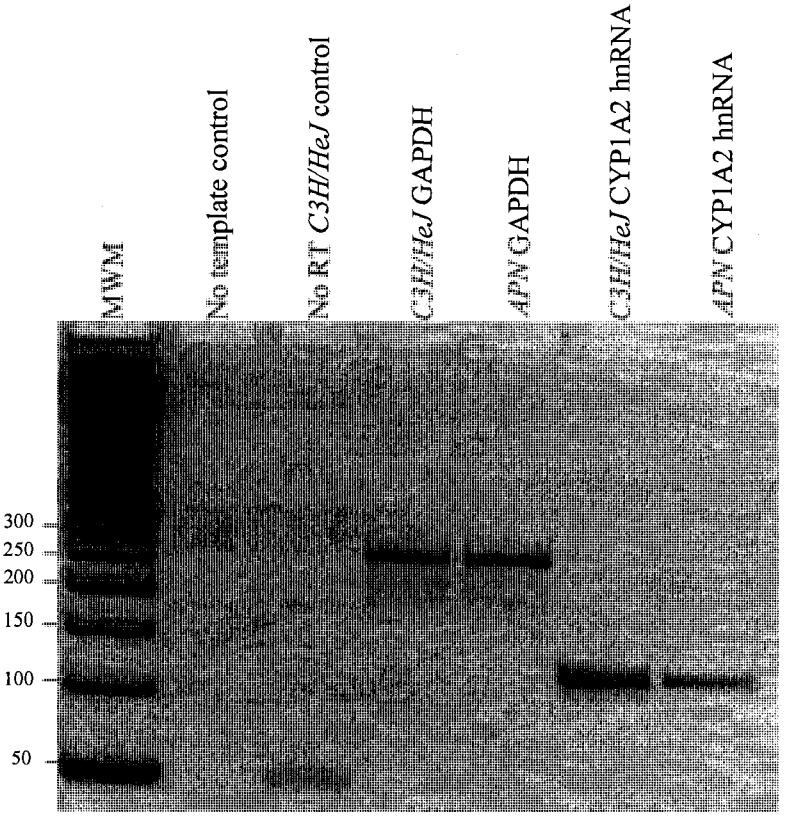

relative CYP1A2 mRNA
7
6
5
4
3
2
1
0
APN
C3H/HeJ
C1
C4
C9
mouse strain

Figure 3.2 Quantification of hepatic steady state CYP1A2 mRNA and hnRNA transcription products. Top; qPCR levels of *C3H/HeJ* CYP1A2 hnRNA and mRNA relative to *APN*. Shown are the averages of 5 animals with standard error of the means. Bottom; a digital photograph of a 1.0% agarose gel stained with SYBR Safe showing the products of the CYP1A2 hnRNA quantification reactions. Lanes 6 and 7 contain the hnRNA primer products, Lanes 4 and 5 contains the GAPDH primer amplification products, Lane 3 contains amplification products of the hnRNA primers in reverse transcriptase free cDNA prep, Lane 2 contains the amplification products of a template free hnRNA amplification, and Lane 1 is a molecular weight marker.


7
6
5
4
3
2
1
0
hnRNA
mRNA
Fold increase relative to APN
APN
C3H/HeJ
mouse strain


MWM
No template control
No RT C3H/HeJ control
C3H/HeJ GAPDH
APN GAPDH
C3H/HeJ CYP1A2 hnRNA
APN CYP1A2 hnRNA
300
250
200
150
100
50

The top panel of Figure 3.2 also suggests a decrease in the *APN* CYP1A2 hnRNA relative to the amount seen in the *C3H/HeJ* strain. Since hnRNA levels provide a 'snapshot' of a gene's transcriptional rate, these results suggest that the *APN* mouse strain had a lower rate of CYP1A2 transcription than the *C3H/HeJ*.

3.3 *In vivo Cyp1a2* promoter assay

An *in vivo* expression system was used to test the difference in the abilities of the *APN* and *C3H/HeJ* CYP1A2 promoters at driving expression. Firefly luciferase pGL3 reporter constructs containing the 5' region sequences for the *Cyp1a2* genes were prepared. Sequencing of the *APN* and *C3H/HeJ* 5' regions in the pGL3 vector constructs showed correct orientation of the inserts and the predicted restriction site (Figure 3.3).

When imaged, the *APN* reporter vector produced a brighter luminescence signal than that produced by the promoter-less pGL3-basic vector (negative control). The *APN* vector signal was also dramatically lower than that of the *C3H/HeJ* vector (Figure 3.4). This experiment was performed three times, each experimental replication showed the same trends.

Figure 3.3 Schematic of pGL3 vector construct containing 5Kb of *Cyp1a2* 5' promoter for transcription assay. Luc+ (green arrow) is the firefly luciferase reporter gene, Ampr (red arrow) represents the amplicilin resistance gene and the blue box indicates the *Cyp1a2* 5' region. Note the presence of exon 1 (orange box) and the -76 mutation within the 5' region as well as the *XhoI* and *MluI* restriction sites.

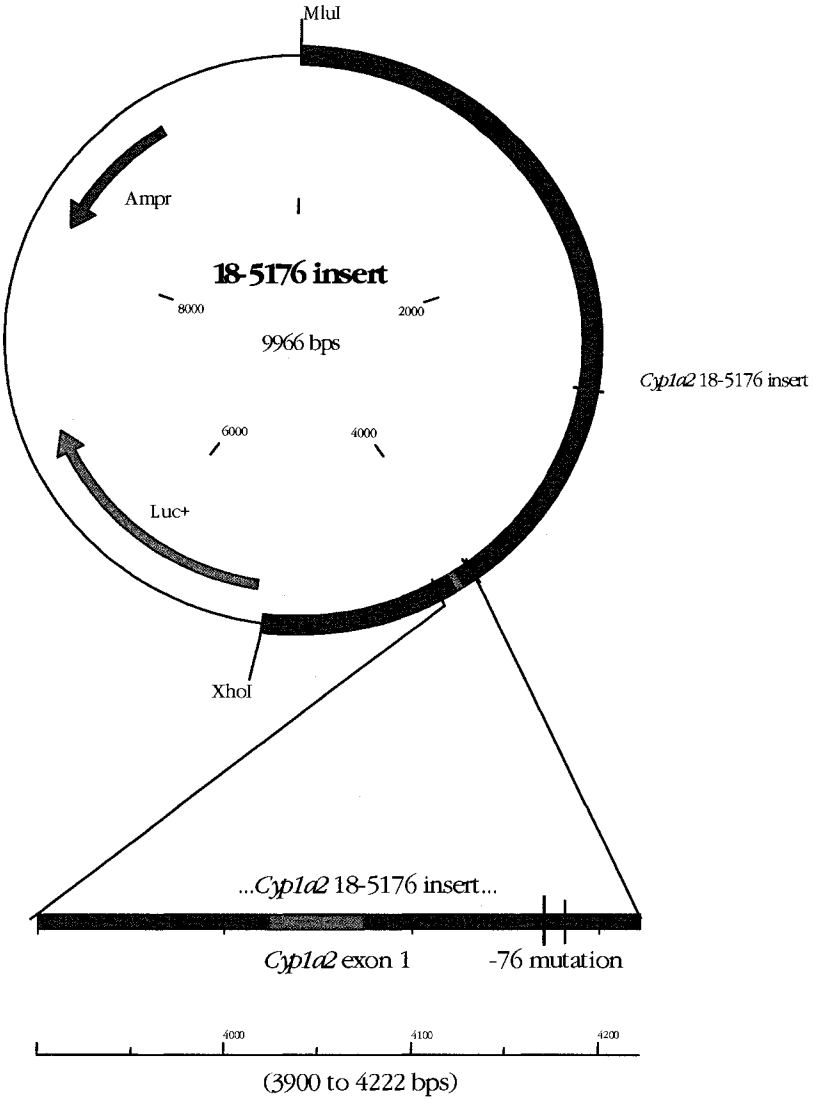
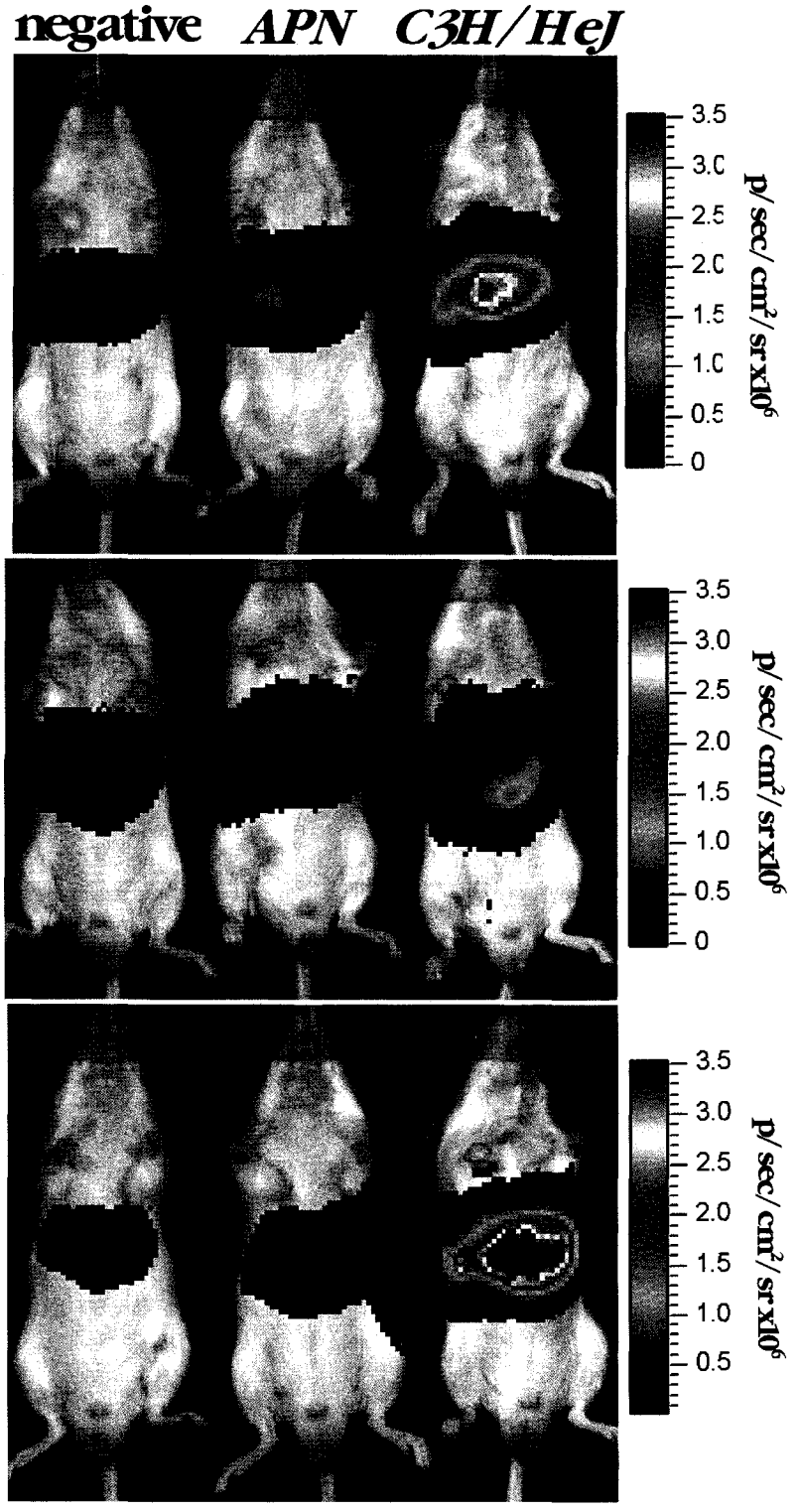

MluI
Ampr
18-5176 insert
8000
2000
9966 bps
6000
4000
Luc+
XhoI
Cyp1a2 18-5176 insert
...Cyp1a2 18-5176 insert...
Cyp1a2 exon 1
-76 mutation
4000
4100
4200
(3900 to 4222 bps)

Figure 3.4 *In vivo* image of *BALB/c* male mice containing *Cyp1a2* 5' promoter driven pGL3 vectors. Mice were injected with 100μL of 30mg/mL luciferin saline IP and imaged using IVIS 50 for 5min at high binning. *APN* and *C3H/HeJ* labeled mice were injected with pGL3 vectors driven by respective *Cyp1a2* 5' regions. The negative mouse contained pGL3-basic promoter-less vector.


negative APN C3H/HeJ
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
p/sec/cm²/sr x10⁶
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
p/sec/cm²/sr x10⁶
3.5
3.0
2.5
2.0
1.5
1.0
0.5
p/sec/cm²/sr x10⁶

3.4 *Ex vivo Cyp1a2* promoter assay

Figure 3.5 suggests the strength in driving transcription of the *APN Cyp1a2* 5' region to be significantly different, with a p value of 0.045, than the promoter-less pGL3-basic vector (negative control). The signal arising from the presence of the *APN* 5' region was almost twice as strong as the background signal. Figure 3.5 also demonstrates that the *APN Cyp1a2* 5' region was less active in driving transcription *in vivo* than the *C3H/HeJ Cyp1a2* 5' region with a p value of 0.00002. A full summary of the results including descriptive statistics for Figure 3.5 are presented in Table A.1 and Table A..2

3.5 Comparative sequencing of *Cyp1a2* 5' region

The initial comparative sequencing of the *Cyp1a2* gene including all exons and introns was performed by Ms. Lori Lavigne. Her data demonstrated several polymorphisms within the non-coding region of the gene. However, only a narrow region of the 5' promoter was sequenced. Figure 3.6 shows the expanded 5' region sequence in addition to the resequenced *Cyp1a2* gene with several sites of genetic variation. No variations were found within the exons or at the splicing sites which was consistent with the original data.

Figure 3.5 *Ex vivo* luciferase assay of *Cyp1a2* 5' promoter driven pGL3 vectors expressed in *BALB/c* male mice. The luciferase single was calibrated to a transfection control (TK-renilla vector). The *APN* signal was significantly greater than the negative ($p=0.045$) and less than the *C3H/HeJ* 5' region ($p=0.00002$). Error bars represent the standard error.

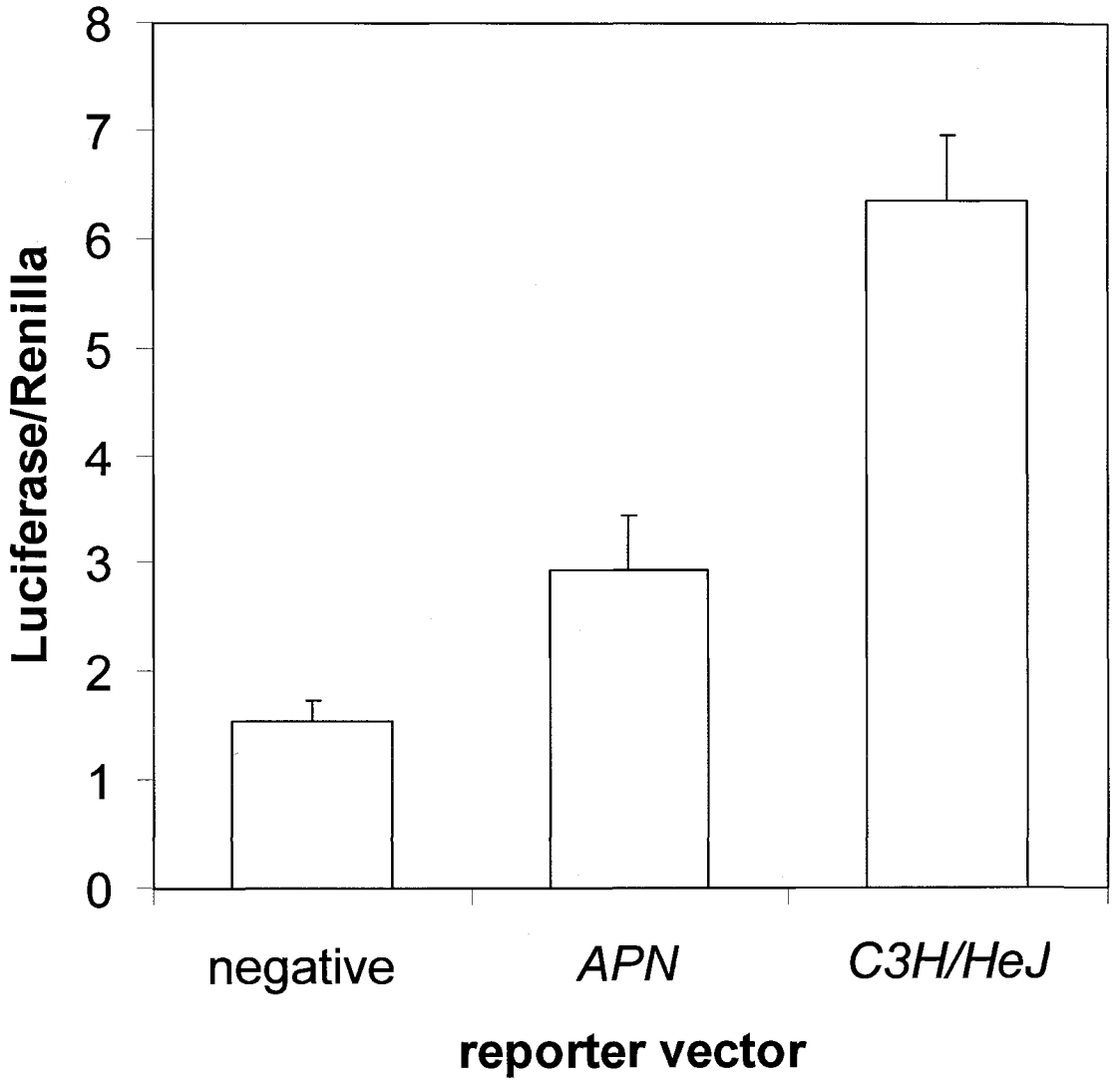
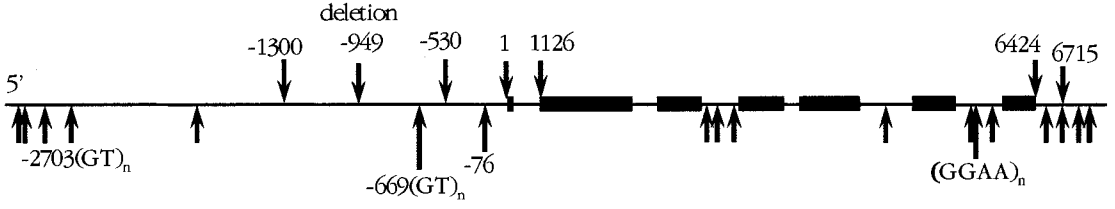

Luciferase/Renilla
8
7
6
5
4
3
2
1
0
negative
APN
C3H/HeJ
reporter vector

Figure 3.6 Schematic of the *Cyp1a2* gene highlighting the sequence variations between *APN* and *C3H/HeJ* mice. Exons are shown as thick black lines. The green arrows indicate polymorphisms unique to each strain. The purple arrows show regions known to have polymorphisisms, and the black arrows indicate two GT promoter repeats (-2703 and -669), a C3H sequence deletion (-949), the transcription start and stop sites (1 and 6715) and the translational start and stop sites (1126 and 6424).


deletion
-1300
-949
-530
1
1126
6424
6715
5'
-2703(GT)n
-669(GT)n
-76
(GGAA)n

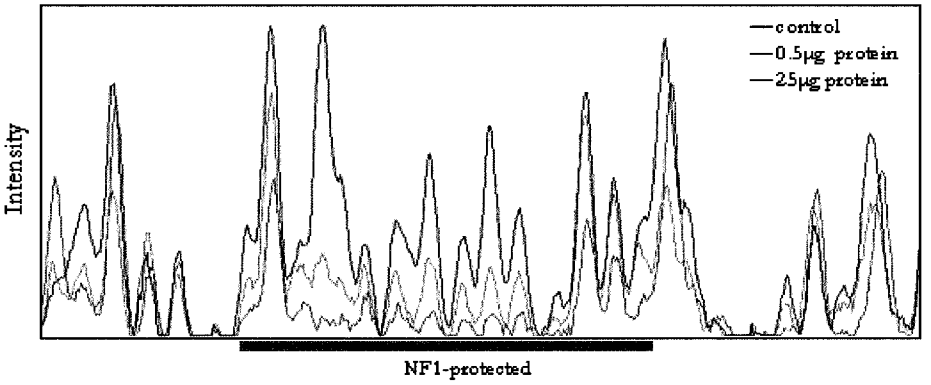
3.6 DNase footprinting assay

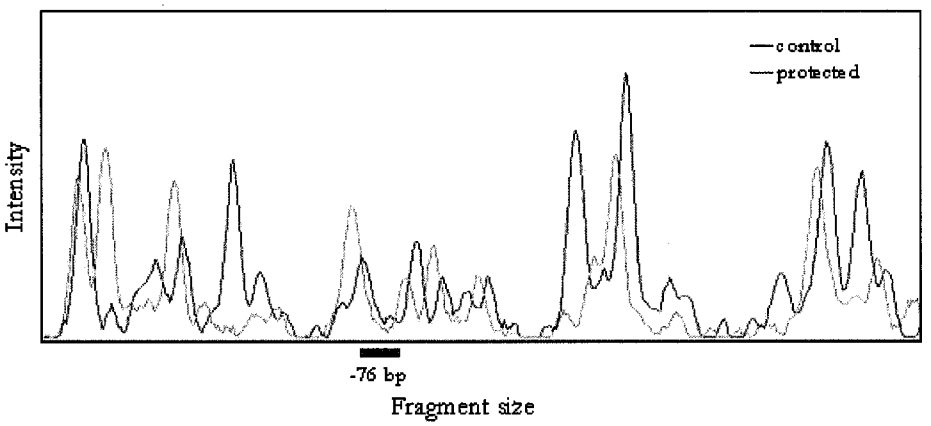
In vitro DNase footprinting was used in an attempt to test the binding of transcription factors to the *Cyp1a2* 5' region.

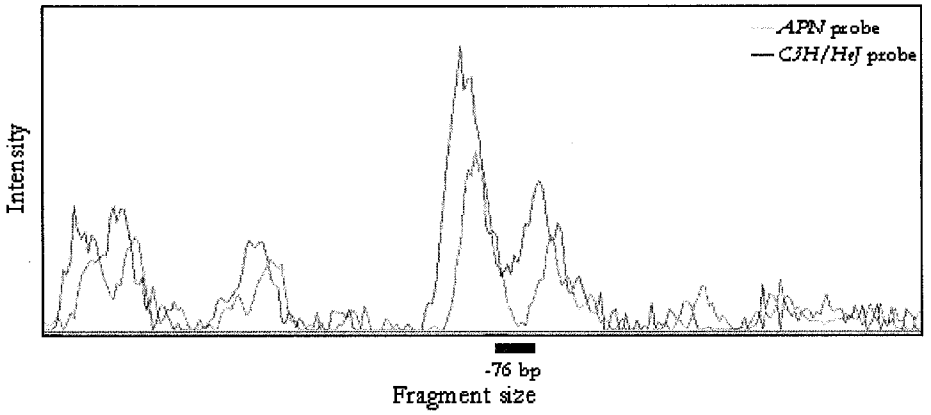
The capillary electrophoresis DNase footprinting technique was first validated by examining a region of the *Cyp1a2* 5' region known to bind by NF1. The top panel of Figure 3.7 highlights the NF1 binding region in the *Cyp1a2* 5' region. When nuclear hepatic protein was added in an increasing concentration, the NF1 region displayed increased protection against DNase digestion. The results of this control experiment demonstrate that the capillary electrophoresis system can be used to detect protein-DNA interactions using fluorescently labeled probes.

When the -76 region of the *C3H/HeJ Cyp1a2* 5' region was examined for DNase protection, though some prominent peaks had decreased, there was no continuous region of decreased peak amplitudes in the presence of hepatic nuclear protein (Figure 3.7 middle). Differential labeling of 5' regions allowed for simultaneous DNase sensitivity comparison between the *APN* and *C3H/HeJ Cyp1a2* 5' regions in a single electrophoretic run. The lower panel of Figure 3.7 suggests no prominent difference in DNase protection when the *APN* and *C3H/HeJ Cyp1a2* 5' regions were digested under identical conditions. The shift in electrophoretic mobility observed in middle and lower panel of Figure 3.7 arises from the differences in charge/mass ratios of the fluorescent labels.

Figure 3.7 Plot of fluorescently labeled *Cyp1a2* 5' promoter oligonucleotides subject to hepatic protein DNase protection assays. Top; shows protection occurring in the NF1 region of the *C3H/HeJ* oligonucleotide with two concentrations of hepatic nuclear protein. Middle; shows a typical protection assay at the -76 region of *C3H/HeJ Cyp1a2*. Bottom; shows a comparative protection assay of the *C3H/HeJ* and *APN* probes in the -76 region.


control
0.5μg protein
2.5μg protein
Intensity
NF1-protected
Fragment size


control
protected
Intensity
-76 bp
Fragment size


APN probe
CJH/HeJ probe
Intensity
-76 bp
Fragment size

3.6 Functionality of the region containing the -76 nucleotide

The biological significance of the -76bp SNP was assayed *in vivo* using hydrodynamically injected pGL3 reporter vectors. A vector was constructed that consisted primarily of *C3H/HeJ* genotype but with the *APN* genotype at -76bp (*MUT*). Sequence analysis confirmed the predicted G>A mutation (data not shown).

In vivo imagining was used to test the strengths of the *MUT* vector relative to the APN and *C3H/HeJ* vectors in *BALB/c* mice. When imaged, the *MUT* vector produced a luminescence signal lower than that produced by the *C3H/HeJ* vector but similar to that produced by the *APN* vector (Figure 3.8). This experiment was performed three times, each repetition showing similar trends.

A large scale *ex vivo* analysis was performed to further refine the estimate of the *MUT* promoter strength and to control for transfection variability. Figure 3.9 suggests the strength of the *MUT* vector to be nearly identical to the *APN* vector ($p=0.98$). The luminescent signals produced by these vectors were both significantly greater than the promoter less pGL3-basic (negative control), with p-values of 0.009 and 0.03, respectively. The *MUT* vector also showed a significant ($p=0.007$) decrease in promoter activity relative the native *C3H/HeJ* vector. This latter result demonstrates the importance of the -76bp in constitutive CYP1A2 expression. Moreover, it appears that the decreased level of CYP1A2 expression observed in the *APN* strain of mice relative to the *C3H/HeJ* strain stems, in part, from a G>A SNP -76bp from the transcriptional start site. A full summary of the results, including descriptive statistics for Figure 3.9, are presented in Table A.3 and Table A.4.

Figure 3.8 *In vivo* luciferase assay of a -76 mutant 4.5kb *MUT* (-76 G>A mutated *C3H/HeJ* CYP1A2 promoter) reporter vector expressed in *BALB/c* male mice. Mice were injected with reporter plasmids *APN*, *C3H/HeJ* and *MUT* and incubated for 18h. Imaging was done by injecting 100μL of 30mg/mL luciferin saline IP and scanned with IVIS 50 for 4min at high binning.

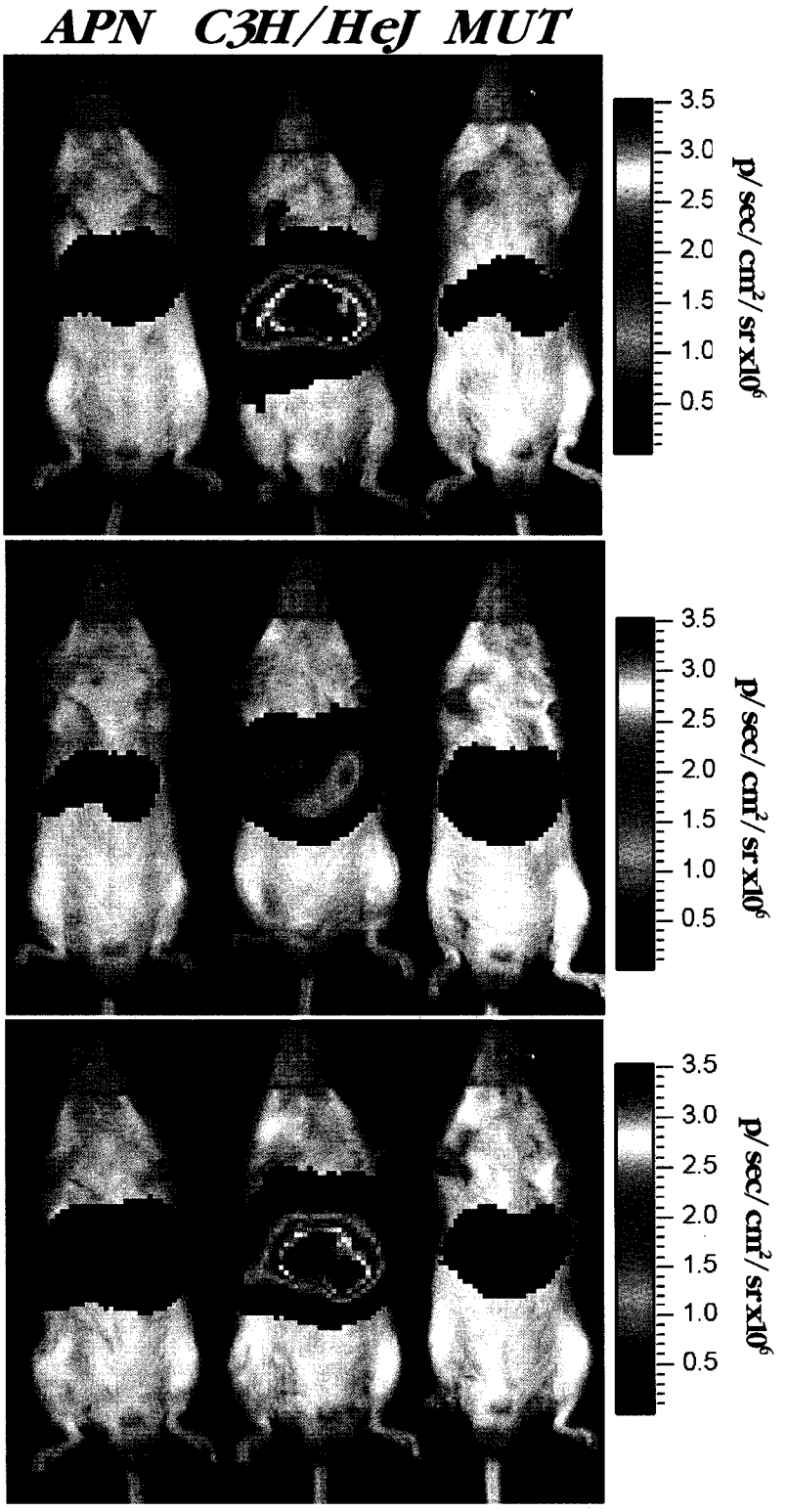
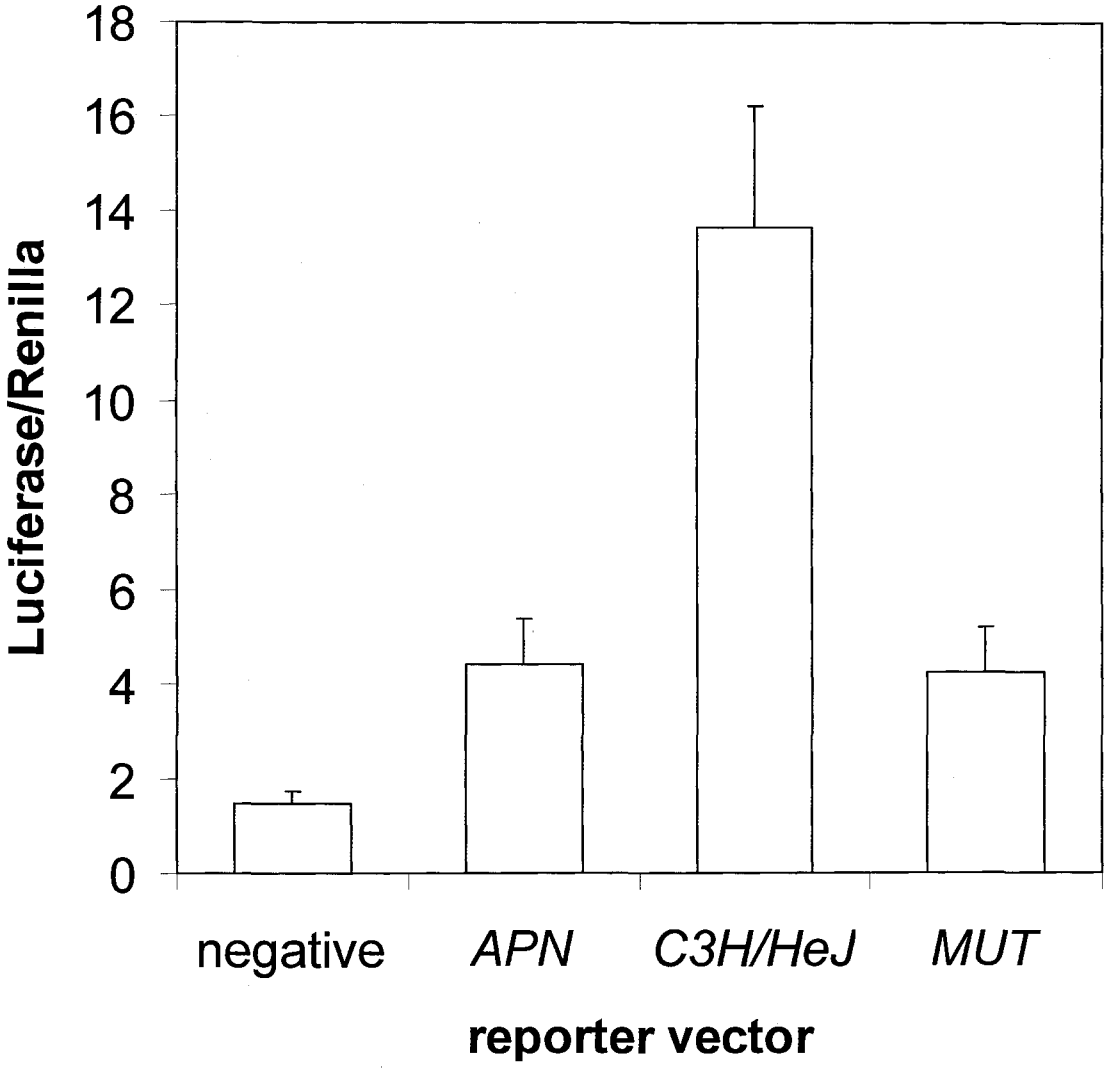

APN C3H/HeJ MUT
3.5
3.0
2.5
2.0
1.5
1.0
0.5
p/sec/cm²/sr x10⁶
3.5
3.0
2.5
2.0
1.5
1.0
0.5
p/sec/cm²/sr x10⁶
3.5
3.0
2.5
2.0
1.5
1.0
0.5
p/sec/cm²/sr x10⁶

Figure 3.9 *Ex vivo* dual luciferase assay of a -76 mutant *C3H/HeJ* (*MUT*) 4.5kb *Cyp1a2* 5' region driven pGL3 vector expressed in *BALB/c* male mice. The luciferase signal was calibrated to a transfection control (TK-renilla vector). The *APN* and *MUT* signals were statistically similar (p=0.98) but significantly greater than the negative (0.03 and 0.009 respectively). The *MUT* vector was significantly less than the native *C3H/HeJ* vector. (p=0.007). Error bars represent the standard error.


18
16
14
12
10
8
6
4
2
0
Luciferase/Renilla
negative
APN
C3H/HeJ
MUT
reporter vector

In silico analyses of the -76 *Cyp1a2 C3H/HeJ* and *APN* regions were performed to identify candidate *trans* factor binding sites using MatInspector (Higo *et al.*, 1999). The -76 base of the *C3H/HeJ Cyp1a2* gene was predicted to lie within the Ecotropic viral integration site 1 (EVI1) binding site gctat**G**ATA (Capitals are the core EVI1 sequence and bold face type indicates -76 base). The *APN Cyp1a2* gene promoter analysis did not predict a different binding factor in the -76 region.

CHAPTER 4: DISCUSSION

Xenobiotic metabolism profiling is of key importance to applied pharmacogenetics. Animal models that have similar metabolic pathways to humans are valuable tools for xenobiotic metabolism studies. Of particular value are animal strains with significant interstrain differences in xenobiotic metabolism rates for a given xenobiotic. These animal models are key to uncovering the genetic components of a given xenobiotic metabolism profile, which are often complex genetic traits.

The decreased C3-D index of the *APN* mouse strain relative to the *C3H/HeJ* mouse strain was demonstrated to be a complex genetic trait, involving loci on chromosomes 1, 4 and 9 (Casley *et al.*, 1999). Figure 3.1 suggests the QTL elements on *C9* and *C4* played an additive role in CYP1A2 expression. This additive effect of the *C9* and *C4* QTLs confirms that the decreased C3-D phenotype observed in *APN* mice by Casley *et al* (1997a) was a result of a complex genetic trait. Since the *Cyp1a2* gene itself was found on chromosome 9 and this QTL made the largest contribution to *Cyp1a2* expression (Figure 3.1), the *Cyp1a2* gene was a logical place to search for an underlying genetic variation responsible for its down regulation. At least one other locus (*C4*), however, was involved in the decrease of CYP1A2 expression. This decrease in expression is likely due to a *trans*-acting factor found of chromosome 4 which regulates *Cyp1a2* expression.

Despite the significant linkage of a QTL on chromosome 1 with decreased C3-D index (Casley *et al.*, 1999), the C1 mouse strain did not show an increase in CYP1A2 mRNA (Figure 3.1). This could be because the locus on C1 involves a different pathway of caffeine metabolism.

Recall that only 80% of C3-D occurs via the CYP1A2 enzyme (Buters *et al.*, 1996). The C1 locus may involve a non CYP1A2 C3-D caffeine metabolism pathway, which contributes to the overall complex phenotype.

The *APN* and *C3H/HeJ* mouse strains had previously been shown to have different levels of CYP1A2 protein expression (Casley *et al.*, 1997a). An experimental approach was devised to determine the level of gene expression at which this difference occurred. The results obtained demonstrate that this observed variation occurred as a result of differences in transcriptional rates. A polymorphism, -76 bp from the transcriptional starts site, was determined to be the underlying cause of different CYP1A2 enzyme levels.

There are many stages at which the production of a protein can be regulated. In order to narrow the search, experiments were performed to deduce where in protein expression the underlying variation in CYP1A2 expression occurred. When the mRNA levels of CYP1A2 were measured, the level in the *APN* CYP1A2 mRNA was almost 6 times lower than that of the *C3H/HeJ* mouse (Figure 3.1). Therefore, regulation at the protein level such as protein trafficking, missfolding and degradation were ruled out as the underlying cause for varied CYP1A2 basal expression. Since the rate of transcription was lower in *APN* relative to *C3H/HeJ*, based on the hnRNA analysis of Figure 3.2, mechanisms of transcriptional regulation were suspected to be the underlying cause of the varied CYP1A2 basal expression. Disruption of a key *cis* element in the 5' *Cyp1a2* locus was hypothesized.

The regulation of CYP1A2 is inherently difficult to study. The use of an *in vivo* system to assay expression of the *Cyp1a2* 5' region was ideal because the CYP1A2 enzyme is not expressed

easily in conventional non-carcinogenic cell lines (Zhang *et al.*, 2006; Li *et al.*, 1999). The hydrodynamic delivery system offered delivery of the reporter constructs primarily to the mouse liver, allowing *in vivo* assays to be performed. Since the liver is one of the primary tissues of *Cyp1a2* expression, transcriptional reporter constructs were assayed in conditions representative of cells *in vivo*.

The *in vivo* imaging system has been applied to the analysis of *Cyp1a2* expression before, and was successfully used to quantify induction of the gene when challenged by a various xenobiotics (Zhang *et al.*, 2004). In the present study, *in vivo* imaging of mice injected with reporter constructs allowed for constitutive transcriptional data to be acquired (Figure 3.4 and 3.8). These *in vivo* luciferase results were not normalized to a transfection control vector. Since *in vivo* transfection can be highly variable, to obtain statistical significance for differences between groups, one would need to use a prohibitively large sample size. To further validate the differences in *Cyp1a2* 5' promoter strengths and to control for transfection variability, large scale *ex vivo* analysis were performed (Figure 3.5 and 3.9). Firefly Luciferase vectors (pGL) from this experiment were normalized to a transfection control vectors (pRL-TK), thereby standardizing transfection variability.

Comparative sequence analysis demonstrated that the -76 base of the *Cyp1a2* 5' region was a guanine in the *C3H/HeJ* mouse as opposed to an adenine in the *APN* mouse (Figure 3.6). This site was of interest for several reasons: (a) the SNP database search showed there were no known polymorphisms in this region, (b) this polymorphism was not observed in 8 other mouse strains analyzed, including a Swiss Webster mice similar to the strain from which the *APN* mouse line was originally derived (Casley *et al.*, 1997a) (c) footprinting showed this region to be partially

protected when performed *in vivo* in the rat (Zhang *et al.*, 2000), (d) this region has not previously been investigated for the basal regulation of CYP1A2, (e) MatInspector (Higo *et al.*, 1999) predicted this region to bind an Ecotropic viral integration site 1 (EVI1) encoded factor, amino-terminal zinc finger domain. Therefore, the G>A SNP at -76bp was targeted as a candidate responsible for the variation in CYP1A2 expression.

Detection of a DNA-binding protein binding to the -76 region was unsuccessful when assayed *in vitro* using electrophoretic mobility shift assays (EMSA) (data not shown). The *In vitro* footprinting assays were used to detect protein-DNA interactions at the -76 region (Figure 3.7) as, in theory; mutation in this region could impair a transcriptional factor from binding and subsequently decrease the gene's function. Resolution of DNase footprint fragments by capillary electrophoresis provided increased sensitivity over traditional methods. The technical limitations of the EMSA and DNase assays are discussed in 4.4.

4.1 Proposed mechanism

Transcriptional down regulation of the *Cyp1a2* gene in this case appears to be a result of mutation in a critical *cis* element in the 5' promoter region. *In vivo* DNase protection of the highly homologous -76 region in the rat suggested a protein binds in this region (Zhang *et al.*, 2000). Here, I present two possible mechanisms by which this mutation may cause a decrease in CYP1A2 expression.

First, the -76 mutation could activate a cryptic transcriptional repressor site. The *in vivo* protection observed at -76 by Zhang *et al* (2000) could be a weakly binding transcriptional repressor which may have enhanced binding in the *APN* genotype. Polymorphisms are known to

be responsible for producing gain of repressor binding sites. An example of this occurs in the *BRCA1+* gene, in which a ACA polymorphism in the promoter region creates a binding site for a FAC1 transcriptional repressor (White *et al.*, 2006). Down-regulation of this gene is suspected to increase the susceptibility of developing hereditary breast cancer. In the case of *Cyp1a2*, *in silico* analysis of the 5' region for repressor binding sites did not show a known transcriptional repressor to bind to the -76 *APN* region. Therefore, it is not likely that the decreased *APN* levels in CYP1A2 expression arise from activation of a cryptic repressor site.

Second, the -76 mutation may disrupt the site of a transcriptional enhancer. In turn, this disruption would result in a decreased binding potential of the enhancing element and consequently decrease the rate of transcription. NF1 has been speculated to form a large multisubunit complex that could interact with -76 (Zhang *et al.*, 2000). The *APN* -76 mutation could impair the formation of such NF1 complex. The Nkx-2.5 factor also has a core binding site 4 base pairs away from the -76 mutation. However, this homeobox protein is normally only expressed during the foetal period (Kasahara *et al.*, 1998) so it is unlikely to play a role in adult basal CYP1A2 expression.

A GATA-like transcription factor may bind and regulate the -76 *Cyp1a2* region. In the *APN* mouse, the -76 A polymorphism interrupts a potential GATA-2-like binding site (Ko and Engel, 1993), which would decrease transcription. The GATA-2 transcription factors are normally present during foetal and not at the adult stage in the liver. However, GATA-4 which has a similar binding site to GATA-2 has been shown to be expressed in low levels of adult liver (Dame *et al.*, 2004). A member of the GATA family that is expressed in the adult stage might regulate the constitutive expression of CYP1A2.

A loss of a transcriptional activator binding site has been observed in other genes, some of which have been linked to serious diseases. An example of this occurs in the human autoimmune disease, systemic lupus erythematosus (SLE). In one case, SLE was linked to a 5' G>A promoter mutation within the programmed cell death gene (PDCD1). This SNP mutation disrupted a runt-related transcription factor 1 (RUNX1) binding site which lead to decrease PDCD1 expression (Prokunina *et al.*, 2002).

The -76 mutation may also disrupted the core of an *in silico* predicted EVI1 domain (gctat**G**ATA). EVI1 is an amino-terminal zinc finger domain, which is bound by an endogenously expressed EVI1 transcription factor. This transcription factor up-regulates transcription and is coded for by a locus on mouse chromosome 3 (MGI:95457). EVI1 has been shown to bind to several endogenous mouse gene promoters (Yatsula *et al.*, 2005).

The EVI1 name originates from the fact that the gene is a common site for retroviral integration. Retroviral integration can result in EVI1 over-expression. Over-expression of this transcription factor is often associated with oncogenesis, as it up-regulates many genes involved in cell function (Nucifora *et al.*, 2006). Thus, EVI1 is often classified as a protooncogene. EVI1 has also been demonstrated to be important for cell development and proliferation (Hoyt *et al.*, 1997; Perkins *et al.*, 1991). Expression of this transcription factor has been shown at low levels in the adult mouse liver (Perkins *et al.*, 1991). Here, it is speculated that EVI1 binds the -76 region of the *Cyp1a2* 5' region and enhances the gene's basal expression in the mouse liver. This speculation is based on four observations; (a) MatInspector has shown a high probability that the core EVI1 transcription enhancer could bind to the -76 site, (b) Zang *et al.* (2000) have shown this homologous region in the rat to be protected *in vivo*, (c) EVI1 is expressed at low levels in adult

mouse liver and (d) the *APN* phenotype reflects that of a disrupted EVI1 binding site in that EVI1 acts as a transcriptional enhancer.

4.2 Significance of findings

Several regions of the *Cyp1a2* gene in both humans and rodents have been identified for their importance in gene expression, most of which involve elements distal from the transcriptional start site (Chung and Bresnick, 1997; Pickwell *et al.*, 2003; Quattrochi *et al.*, 1998; Uchida *et al.*, 2002; Zhang *et al.*, 2000). These findings represent the first report showing the functional significance of the -76 region in constitutive CYP1A2 expression. They also suggest the presence of a transcription factor critical for basal CYP1A2 expression. This factor may be EVI1, which has not before been implicated in CYP1A2 expression activity. The only evidence linking EVI1 to the -76 *Cyp1a2* region is based on an *in silico* prediction. The role of EVI1 in CYP1A2 expression is highly speculative and remains to be demonstrated experimentally.

These data also represent proof of a specific allele underlying a QTL. The decrease in C3-D phenotype observed in the *APN* strain of mice was linked with extremely high likelihood (log-likelihood ratio >20) to a 4.8cM locus on chromosome 9 (Casley *et al.*, 1999). This locus contains, among other genes, the gene encoding the CYP1A2 enzyme, one of the primary enzymes in caffeine demethylation. The strong suggestive evidence for *Cyp1a2* as a candidate gene based on the correlation of gene function to phenotype is considered insufficient for a proof. Further testing of the candidate gene was required. Methods of replacing an allele with another are accepted by the complex trait community as proof of a QTL gene (Abiola *et al.*, 2003). Examples of accepted proofs include: transgenic animals in which the gene of interest is knocked in, polymorphisms in coding or regulatory proteins, gene function studies and *in vitro* functional

studies. This study represents a proof that the *Cyp1a2* gene is responsible for the QTL linked to decreased caffeine metabolism as it satisfies two of the above criteria. First, a variation in the candidate *Cyp1a2* gene function has been demonstrated *in vivo* using *Cyp1a2* 5' region constructs. Second, a polymorphism has been discovered which plays a direct role in transcriptional regulation. Therefore, these data fulfill the criteria to prove the link between low caffeine demethylation phenotype and the *Cyp1a2* gene.

Proof of the *Cyp1a2* gene involvement in a complex trait is noteworthy for two reasons. One of the most common and accepted approaches to prove a QTL is to produce the phenotype through a knock-in mutation of the candidate gene. As an example, a QTL for the complex genetic trait cholelithiasis has been linked to the Abcb11 gene in mice. This trait was linked to a locus on chromosome 2, *Lith 1*, which contained the hepatic bile salt transporter Abcb11. The link was proven to be partly responsible for cholelithiasis by hyperexpressing the Abcb11 gene in transgenic mice (Henkel *et al.*, 2005). Creating the transgenic mice required in these kinds of experiments is technically challenging and requires extensive resources.

The other challenge associated with proving gene linkages stems from the fact that multiple sequence differences found in a QTL analysis can individually have subtle phenotypical effects (Flaherty *et al.*, 2005). These subtle effects result in many more locus linkages than gene proofs. For example, in 2006, 426 genes with DNA sequence variation were reported to be linked to human obesity (Rankinen *et al.*, 2006); yet the individual effects of these genes on obesity remain largely unknown. Therefore, this proof of *Cyp1a2* involvement in mouse caffeine demethylation is of interest because it did not require the use of costly transgenic animals and because the identification of a single gene's contribution to a complex phenotype is relatively rare.

Congenics has frequently been used as a tool for dissecting the role individual QTLs have on a complex phenotype (Radcliffe *et al.*, 2006; Hoopes, Jr. *et al.*, 2006; Van Dijk *et al.*, 2006). In this study, the use of congenic mice containing the QTL regions linked to a decreased C3-D has allowed for isolation of the individual genetic factors. This isolation has assisted in confirming the link between the candidate locus and the phenotype. The congenic mice have also allowed the contributions and the mechanisms by which each locus affects the overall phenotype to be studied.

4.3 Broader implications

This work is significant because CYP1A2 plays an important role in toxicology, procarcinogen activation, and drug metabolism (Eaton *et al.*, 1995). These data suggest the presence of a mouse *Cyp1a2* regulatory site that has a dramatic effect on basal transcriptional rate. This regulatory site, however, has no apparent human homologue. Therefore, the mechanisms of basal CYP1A2 regulation may be different in humans than rodents. This difference in regulatory sites puts into question the appropriateness of the use of the rodent model in both therapeutic and toxicological studies involving CYP1A2 pathways.

4.4 Technical limitations

Initially, traditional EMSAs were performed to search for a novel DNA binding protein; however, these yielded inconclusive results (data not shown). This inconclusive EMSA could be due to technical limitations or a weak DNA-protein interaction.

Next, DNase Footprinting was performed on the *Cyp1a2* 5' region to detect DNA-protein interaction. The method of DNase capillary electrophoresis was superior to traditional gel

electrophoresis for several reasons. Firstly, the capillary electrophoresis DNase assay allowed for differential labeling of oligonucleoties and thereby permitted size standards to be run in tandem with the DNase digest fragments. Fragment sizes could be accurately calibrated to these standards. Differential labeling also allowed for digestion and resolution of multiple oligonucleotides in a single reaction. The *APN* and *C3H/HeJ Cyp1a2* 5' regions could be directly compared using this technique. Secondly, the capillary system had an increased resolving power over a broader size range and thus made it possible to assay larger probe fragments that differed by a single deoxyribo nucleotide. Finally, the data analysis of a capillary-based system allowed for accurate measurement of peak height, which permitted subtle differences in digest patterns to be quantified.

The capillary electrophoresis DNase method accurately demonstrated DNase protection in the known NF1 domain, but failed show clear DNA binding at the -76 polymorphic region of interest. The potential reasons for the inconclusive data obtained are several. First, the binding conditions may have been inappropriate to facilitate binding or the assembly of a complex at -76bp. This could have included unfavorable buffer pH or osmolarity, or unfavorable binding time and temperature conditions. *In vivo* binding conditions may, in fact, be the only conditions in which the transcription factor will bind -76. Second, the DNA binding factor may have been absent or inactive as a result of the extraction process. Mechanical manipulation of proteins is known to result in their denatuation and subsequent inactivation (Floury *et al.*, 2002). The putative binding protein may be more labile than the "control" NF-1. Also during nuclear lysis there was an extensive amount of DNA in solution. Any protein with DNA affinity present in low concentration at this point in the sample preparation could bind this DNA and be lost during chromatin extraction. Third, there may not be a transcription factor that binds to this site. This

third possibilty is unlikely because the rat, which has a nearly identical -86 to -63bp region to the mouse, displayed partial DNase protection *in vivo* at -76bp (Zhang *et al.*, 2000). Despite the suggestive evidence that a protein binds in -76 region, the -76 mutation could disrupt other characteristics of the DNA that do not involve transcription factor binding but do effect transcription. However, reports of other mechanisms of gene regulation by point mutations are not abundant in the literature. Improper binding conditions, transcription factor inactivation or lack of a transcription factor binding site are more likely reasons why the DNase protection assay was inconclusive at -76. Other data obtained, including *in silico* analysis and the *in vivo* and *in vitro* studies presented here, are certainly consistent with the idea that a transcription factor binding this -76 region does exist but that the experimental conditions used were not conducive to its detection.

4.5 Future questions

Future studies will address the function of the *Cyp1a2* -76 region. These studies could include the purification and identification of a transcription factor. For a protein purification approach to work, ideally, an *in vitro* system to assay DNA binding (such as an EMSA) must be set up. These DNA binding studies could involve arrays of binding buffer conditions and protein purifications. No doubt, this process could become very labor intensive.

If the problem lies with an inherent weak DNA protein interactions, an *in vivo* DNA-protein crosslinking approach might be employed (Russmann *et al.*, 1998). Such crosslinking processes usually involve using UV lasers to form covalent bonds between proteins where they interact with

DNA and can be applied to isolated hepatocytes. Complexes such as these could be isolated and used for footprinting or possibly amino acid analysis.

EVI1 should be investigated for its potential role in constitutive CYP1A2 expression. First the presence of the mature EVI1 protein in the adult *C3H/HeJ* and *APN* mouse livers should be confirmed. Next, the *in vivo* system could be used to test *Cyp1a2* 5' region strengths when EVI1 is up regulated. The nature of EVI1 is such that one may be able to knock in a promoter using a retrovirus. Then, siRNA could be used to knock down EVI1 expression. Gene knockout technologies could be used to produce an adult mouse lacking EVI1. Levels of CYP1A2 expression could be measured in this model.

Homology studies of this unknown transcription factor could provide detail into human CYP1A2 regulation. A homologous human protein could be used as a target in both pharmacogenetics and pharmacology therapies. One would want to know where this protein binds, if there are alleles in the population which disrupt its binding site and to what degree can they be used to predict a drug response. The difference in homology would also be interesting to examine from an evolutionary perspective.

4.6 Summary

The levels of expression of CYP1A2 are important in drug metabolism and are variable amongst humans. A novel point mutation has been identified at the –76 region of the *Cyp1a2* locus in the mouse that has been demonstrated to be in a critical region for constitutive CYP1A2 expression. This region may bind a transcription factor. Characterization of this factor will give insight into the regulatory mechanisms of CYP1A2, which may be extrapolated to the human

system and possibly provide insights into the use of genomic sequence variation in drug prescription.

REFERENCE

Abiola,O., Angel,J.M., Avner,P., Bachmanov,A.A., Belknap,J.K., Bennett,B., Blankenhorn,E.P., Blizard,D.A., Bolivar,V., Brockmann,G.A., Buck,K.J., Bureau,J.F., Casley,W.L., Chesler,E.J., Cheverud,J.M., Churchill,G.A., Cook,M., Crabbe,J.C., Crusio,W.E., Darvasi,A., de Haan,G., Dermant,P., Doerge,R.W., Elliot,R.W., Farber,C.R., Flaherty,L., Flint,J., Gershenfeld,H., Gibson,J.P., Gu,J., Gu,W., Himmelbauer,H., Hitzemann,R., Hsu,H.C., Hunter,K., Iraqi,F.F., Jansen,R.C., Johnson,T.E., Jones,B.C., Kempermann,G., Lammert,F., Lu,L., Manly,K.F., Matthews,D.B., Medrano,J.F., Mehrabian,M., Mittlemann,G., Mock,B.A., Mogil,J.S., Montagutelli,X., Morahan,G., Mountz,J.D., Nagase,H., Nowakowski,R.S., O'Hara,B.F., Osadchuk,A.V., Paigen,B., Palmer,A.A., Peirce,J.L., Pomp,D., Rosemann,M., Rosen,G.D., Schalkwyk,L.C., Seltzer,Z., Settle,S., Shimomura,K., Shou,S., Sikela,J.M., Siracusa,L.D., Spearow,J.L., Teuscher,C., Threadgill,D.W., Toth,L.A., Toye,A.A., Vadasz,C., Van Zant,G., Wakeland,E., Williams,R.W., Zhang,H.G., and Zou,F. (2003). The nature and identification of quantitative trait loci: a community's view. Nat. Rev. Genet. *4*, 911-916.

Allorge,D., Chevalier,D., Lo-Guidice,J.M., Cauffiez,C., Suard,F., Baumann,P., Eap,C.B., and Broly,F. (2003). Identification of a novel splice-site mutation in the CYP1A2 gene. British Journal of Clinical Pharmacology *56*, 341-344.

Alving,A.S., Carson,P.E., Flanagan,C.L., and Ickes,C.E. (1956). Enzymatic deficiency in primaquine-sensitive erythrocytes. Science *124*, 484-485.

Benet,L.Z., Koretz,D.L., and Sheiner,L.B. (1996). Pharmicokinetics: The dynamics of drug absorption, distribution and elimination. In Goodman and Gilman's The pharmacological basis of therapeutics, J.G.Hardman, L.E.Limbird, P.B.Molinoff, R.W.Ruddon, and A.Goodman Gilman, eds. (New York: McGraw-Hill), pp. 3-28.

Benson,D.A., Boguski,M., Lipman,D.J., and Ostell,J. (1994). GenBank. Nucleic Acids Res. *22*, 3441-3444.

Bigelow,S.W., Collins,A.C., and Nebert,D.W. (1989). Selective Mouse Breeding for Short Ethanol Sleep Time Has Led to High-Levels of Hepatic Aromatic Hydrocarbon (Ah) Receptor. Biochemical Pharmacology *38*, 3565-3572.

Bray,B.J. and Rosengren,R.J. (2001). Retinol potentiates acetaminophen-induced hepatotoxicity in the mouse: Mechanistic studies. Toxicology and Applied Pharmacology *173*, 129-136.

Buters,J.T.M., Tang,B.K., Pineau,T., Gelboin,H.V., Kimura,S., and Gonzalez,F.J. (1996). Role of CYP1A2 in caffeine pharmacokinetics and metabolism: Studies using mice deficient in CYP1A2. Pharmacogenetics *6*, 291-296.

Casley,W.L., Menzies,J.A., Girard,M., Larocque,L., Mousseau,N., Whitehouse,L.W., and Moon,T.W. (1997a). Differences in caffeine 3-demethylation activity among inbred mouse strains:

A comparison of hepatic Cyp1a2 gene expression between two inbred strains. Fundamental and Applied Toxicology *40*, 228-237.

Casley,W.L., Menzies,J.A., Mousseau,N., Girard,M., Moon,T.W., and Whitehouse,L.W. (1997b). Increased basal expression of hepatic Cyp1a1 and Cyp1a2 genes in inbred mice selected for susceptibility to acetaminophen-induced hepatotoxicity. Pharmacogenetics *7*, 283-293.

Casley,W.L., Menzies,J.A., Whitehouse,L.W., and Moon,T.W. (1999). Detection of quantitative trait loci affecting caffeine metabolism by interval mapping in a genome-wide scan of C3H/HEJ x APN F-2 mice. Drug Metabolism and Disposition *27*, 1375-1380.

Champe,P.C., Harvey,R.A., and Ferrier,D.R. (2005). Pentose Phosphate Pathway and NADPH. In Lippincott's Illistrated Reviews: Biochemistry, R.A.Harvey and P.C.Champe, eds. (Philadelphia: Lippincott Williams & Wilkins), pp. 143-154.

Chung,I.J. and Bresnick,E. (1997). Identification of positive and negative regulatory elements of the human cytochrome P4501A2 (CYP1A2) gene. Archives of Biochemistry and Biophysics *338*, 220-226.

Crino,P.B. and Eberwine,J. (1996). Molecular characterization of the dendritic growth cone: regulated mRNA transport and local protein synthesis. Neuron *17*, 1173-1187.

Dame,C., Sola,M.C., Lim,K.C., Leach,K.M., Fandrey,J., Ma,Y., Knopfle,G., Engel,J.D., and Bungert,J. (2004). Hepatic erythropoietin gene regulation by GATA-4. J. Biol. Chem. *279*, 2955-2961.

Dillon,N. (2006). Gene regulation and large-scale chromatin organization in the nucleus. Chromosome. Res. *14*, 117-126.

Eaton,D.L., Gallagher,E.P., Bammler,T.K., and Kunze,K.L. (1995). Role of Cytochrome P4501A2 in Chemical Carcinogenesis - Implications for Human Variability in Expression and Enzyme-Activity. Pharmacogenetics *5*, 259-274.

Ehrilch,P.R. and and Raven,P.H. (1964). Butterflies and plants: a study of coevolution. Evolution *18*, 586-608.

Eichelbaum,M. and Evert,B. (1996). Influence of pharmacogenetics on drug disposition and response. Clin. Exp. Pharmacol. Physiol *23*, 983-985.

Flaherty,L., Herron,B., and Symula,D. (2005). Genomics of the future: Identification of quantitative trait loci in the mouse. Genome Research *15*, 1741-1745.

Floury,J., Desrumaux,A., and Legrand,J. (2002). Effect of ultra-high-pressure homogenization on structure and on rheological properties of soy protein-stabilized emulsions. Journal of Food Science *67*, 3388-3395.

Gaedigk,A., Bhathena,A., Ndjountche,L., Pearce,R.E., Abdel-Rahman,S.M., Alander,S.W., Bradford,L.D., Rogan,P.K., and Leeder,J.S. (2005). Identification and characterization of novel sequence variations in the cytochrome P4502D6 (CYP2D6) gene in African Americans. Pharmacogenomics J. *5*, 173-182.

Gomes,E.R. and Demoly,P. (2005). Epidemiology of hypersensitivity drug reactions. Current Opinion in Allergy and Clinical Immunology *5*, 309-316.

Gonzalez,F.J. and Lee,Y.H. (1996). Cytochromes P450 .6. Constitutive expression of hepatic cytochrome P450 genes. Faseb Journal *10*, 1112-1117.

Gorski,K., Carneiro,M., and Schibler,U. (1986). Tissue-specific in vitro transcription from the mouse albumin promoter. Cell *47*, 767-776.

Groll,M., Bochtler,M., Brandstetter,H., Clausen,T., and Huber,R. (2005). Molecular machines for protein degradation. Chembiochem. *6*, 222-256.

Henkel,A., Wei,Z., Cohen,D.E., and Green,R.M. (2005). Mice overexpressing hepatic Abcb11 rapidly develop cholesterol gallstones. Mamm. Genome *16*, 903-908.

Higo, K, Ugawa, Y, Iwamoto, M, and Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database. 1999.
Ref Type: Computer Program

Hoopes,R.R., Jr., Middleton,F.A., Sen,S., Hueber,P.A., Reid,R., Bushinsky,D.A., and Scheinman,S.J. (2006). Isolation and confirmation of a calcium excretion quantitative trait locus on chromosome 1 in genetic hypercalciuric stone-forming congenic rats. J. Am. Soc. Nephrol. *17*, 1292-1304.

Hoyt,P.R., Bartholomew,C., Davis,A.J., Yutzey,K., Gamer,L.W., Potter,S.S., Ihle,J.N., and Mucenski,M.L. (1997). The Evil proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development. Mechanisms of Development *65*, 55-70.

Ikeya,K., Jaiswal,A.K., Owens,R.A., Jones,J.E., Nebert,D.W., and Kimura,S. (1989). Human CYP1A2: sequence, gene structure, comparison with the mouse and rat orthologous gene, and differences in liver 1A2 mRNA expression. Mol. Endocrinol. *3*, 1399-1408.

Ingelman-Sundberg,M. (2001). Genetic susceptibility to adverse effects of drugs and environmental toxicants. The role of the CYP family of enzymes. Mutat. Res. *482*, 11-19.

Jiang,Z., Dragin,N., Jorge-Nebert,L.F., Martin,M.V., Guengerich,F.P., Aklillu,E., Ingelman-Sundberg,M., Hammons,G.J., Lyn-Cook,B.D., Kadlubar,F.F., Saldana,S.N., Sorter,M., Vinks,A.A., Nassr,N., von Richter,O., Jin,L., and Nebert,D.W. (2006). Search for an association between the human CYP1A2 genotype and CYP1A2 metabolic phenotype. Pharmacogenet. Genomics *16*, 359-367.

Jin,B., Park,D.W., Nam,K.W., Oh,G.T., Lee,Y.S., and Ryu,D.Y. (2004). CpG methylation of the mouse CYP1A2 promoter. Toxicology Letters *152*, 11-18.

Kasahara,H., Bartunkova,S., Schinke,M., Tanaka,M., and Izumo,S. (1998). Cardiac and extracardiac expression of Csx/Nkx2.5 homeodomain protein. Circ. Res. *82*, 936-946.

Kimura,S., Gonzalez,F.J., and Nebert,D.W. (1984). Mouse Cytochrome-P3-450 - Complete Cdna and Amino-Acid-Sequence. Nucleic Acids Research *12*, 2917-2928.

Kispert,A., Koschorz,B., and Herrmann,B.G. (1995). The T protein encoded by Brachyury is a tissue-specific transcription factor. EMBO J. *14*, 4763-4772.

Ko,L.J. and Engel,J.D. (1993). Dna-Binding Specificities of the Gata Transcription Factor Family. Molecular and Cellular Biology *13*, 4011-4022.

Lazarou,J., Pomeranz,B.H., and Corey,P.N. (1998). Incidence of adverse drug reactions in hospitalized patients - A meta-analysis of prospective studies. Jama-Journal of the American Medical Association *279*, 1200-1205.

Lerman, C and Calkins, M W. Helping Smokers Quit Through Pharmacogenetics. LDI Issue Brief 11[6]. 6-1-2006.
Ref Type: Generic

Lerman,C., Tyndale,R., Patterson,F., Wileyto,E.P., Shields,P.G., Pinto,A., and Benowitz,N. (2006). Nicotine metabolite ratio predicts efficacy of transdermal nicotine for smoking cessation. Clinical Pharmacology & Therapeutics *79*, 600-608.

Li,A.P., Gorycki,P.D., Hengstler,J.G., Kedderis,G.L., Koebe,H.G., Rahmani,R., de Sousas,G., Silva,J.M., and Skett,P. (1999). Present status of the application of cryopreserved hepatocytes in the evaluation of xenobiotics: consensus of an international expert panel. Chemico-Biological Interactions *121*, 117-123.

Malaiyandi,V., Sellers,E.M., and Tyndale,R.F. (2005). Implications of CYP2A6 genetic variation for smoking behaviors and nicotine dependence. Clinical Pharmacology & Therapeutics *77*, 145-158.

Nakajima,M., Yokoi,T., Mizutani,M., Kinoshita,M., Funayama,M., and Kamataki,T. (1999). Genetic polymorphism in the 5 '-flanking region of human CYP1A2 gene: Effect on the CYP1A2 inducibility in humans. Journal of Biochemistry *125*, 803-808.

Nebert,D.W., Puga,A., and Vasiliou,V. (1993). Role of the Ah Receptor and the Dioxin-Inducible [Ah] Gene Battery in Toxicity, Cancer, and Signal-Transduction. Annals of the New York Academy of Sciences *685*, 624-640.

Nelson,D.R., Kamataki,T., Waxman,D.J., Guengerich,F.P., Estabrook,R.W., Feyereisen,R., Gonzalez,F.J., Coon,M.J., Gunsalus,I.C., Gotoh,O., Okuda,K., and Nebert,D.W. (1993). The P450

Superfamily - Update on New Sequences, Gene-Mapping, Accession Numbers, Early Trivial Names of Enzymes, and Nomenclature. Dna and Cell Biology *12*, 1-51.

Nucifora,G., Laricchia-Robbio,L., and Senyuk,V. (2006). EVI1 and hematopoietic disorders: history and perspectives. Gene *368*, 1-11.

Owens,R.A. and Nebert,D.W. (1990). Expression of the Chloramphenicol Acetyltransferase (Cat) Reporter Gene by the Murine Cypla-2 (Cytochrome-P3450) Promoter in Hepatoma-Cell Cultures. Biochemical and Biophysical Research Communications *172*, 1109-1115.

Ozdemir,V., Kalow,W., Tothfalusi,L., Bertilsson,L., Endrenyi,L., and Graham,J.E. (2005). Multigenic Control of Drug Response and Regulatory Decision-Making in Pharmacogenomics: The Need for an Upper-Bound Estimate of Genetic Contributions. Current Pharmacogenomics 53-71.

Pasco,D.S., Boyum,K.W., Merchant,S.N., Chalberg,S.C., and Fagan,J.B. (1988). Transcriptional and Post-Transcriptional Regulation of the Genes Encoding Cytochrome-P-450C and Cytochrome-P-450D Invivo and in Primary Hepatocyte Cultures. Journal of Biological Chemistry *263*, 8671-8676.

Pedersen,A.G., Baldi,P., Chauvin,Y., and Brunak,S. (1999). The biology of eukaryotic promoter prediction - a review. Computers & Chemistry *23*, 191-207.

Perkins,A.S., Mercer,J.A., Jenkins,N.A., and Copeland,N.G. (1991). Patterns of Evi-1 Expression in Embryonic and Adult Tissues Suggest That Evi-1 Plays An Important Regulatory Role in Mouse Development. Development *111*, 479-487.

Phillips,K.A., Veenstra,D.L., Oren,E., Lee,J.K., and Sadee,W. (2001). Potential role of pharmacogenomics in reducing adverse drug reactions - A systematic review. Jama-Journal of the American Medical Association *286*, 2270-2279.

Pickwell,G.V., Shih,H., and Quattrochi,L.C. (2003). Interaction of upstream stimulatory factor proteins with an E-box located within the human CYP1A2 5'-flanking gene contributes to basal transcriptional gene activation. Biochem. Pharmacol. *65*, 1087-1096.

Prokunina,L., Castillejo-Lopez,C., Oberg,F., Gunnarsson,I., Berg,L., Magnusson,V., Brookes,A.J., Tentler,D., Kristjansdottir,H., Grondal,G., Bolstad,A.I., Svenungsson,E., Lundberg,I., Sturfelt,G., Jonssen,A., Truedsson,L., Lima,G., Alcocer-Varela,J., Jonsson,R., Gyllensten,U.B., Harley,J.B., Alarcon-Segovia,D., Steinsson,K., and Alarcon-Riquelme,M.E. (2002). A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. Nat. Genet. *32*, 666-669.

Quattrochi,L.C., Shih,H., and Pickwell,G.V. (1998). Induction of the human CYP1A2 enhancer by phorbol ester. Archives of Biochemistry and Biophysics *350*, 41-48.

Radcliffe,R.A., Bludeau,P., Asperi,W., Fay,T., Deng,X.S., Erwin,V.G., and Deitrich,R.A. (2006). Confirmation of quantitative trait loci for ethanol sensitivity and neurotensin receptor density in

crosses derived from the inbred High and Low Alcohol Sensitive selectively bred rat lines. Psychopharmacology (Berl).

Ragoussis,J. and Elvidge,G. (2006). Affymetrix GeneChip system: moving from research to the clinic. Expert. Rev. Mol. Diagn. *6*, 145-152.

Rankinen,T., Zuberi,A., Chagnon,Y.C., Weisnagel,S.J., Argyropoulos,G., Walts,B., Perusse,L., and Bouchard,C. (2006). The human obesity gene map: The 2005 update. Obesity *14*, 529-644.

Roymans,D., Annaert,P., Van Houdt,J., Weygers,A., Noukens,J., Sensenhauser,C., Silva,J., Van Looveren,C., Hendrickx,J., Mannens,G., and Meuldermans,W. (2005). Expression and induction potential of cytochromes P450 in human cryopreserved hepatocytes. Drug Metab Dispos. *33*, 1004-1016.

Russmann,C., Stollhof,J., Weiss,C., Beigang,R., and Beato,M. (1998). Two wavelength femtosecond laser induced DNA-protein crosslinking. Nucleic Acids Res. *26*, 3967-3970.

Sabath,D.E., Broome,H.E., and Prystowsky,M.B. (1990). Glyceraldehyde-3-phosphate dehydrogenase mRNA is a major interleukin 2-induced transcript in a cloned T-helper lymphocyte. Gene *91*, 185-191.

Schuetz,E., Lan,L.B., Yasuda,K., Kim,R., Kocarek,T.A., Schuetz,J., and Strom,S. (2002). Development of a real-time in vivo transcription assay: Application reveals pregnane X receptor-mediated induction of CYP3A4 by cancer chemotherapeutic agents. Molecular Pharmacology *62*, 439-445.

Service,R.F. (2006). Gene sequencing - The race for the $1000 Genome. Science *311*, 1544-1546.

Toyota,M. and Issa,J.P.J. (1999). CpG island methylator phenotypes in aging and cancer. Seminars in Cancer Biology *9*, 349-357.

Uchida,Y., Yano,A., Kumakura,S., Sakuma,T., and Nemoto,N. (2002). Enhancer elements in the mouse Cypla2 gene for constitutive expression. Biochemical and Biophysical Research Communications *297*, 1297-1301.

Van Dijk,S.J., Specht,P.A., Lazar,J., Jacob,H.J., and Provoost,A.P. (2006). Synergistic QTL interactions between Rf-1 and Rf-3 increase renal damage susceptibility in double congenic rats. Kidney Int. *69*, 1369-1376.

Vansant,G., Pezzoli,P., Saiz,R., Birch,A., Duffy,C., Ferre,F., and Monforte,J. (2006). Gene expression analysis of troglitazone reveals its impact on multiple pathways in cell culture: a case for in vitro platforms combined with gene expression analysis for early (idiosyncratic) toxicity screening. Int. J. Toxicol. *25*, 85-94.

White,K.K., Wilson,C., Belloni,D.R., Booker,J., Silverman,L.M., Gardner,K.L., Tsongalis,G.J., Bowser,R., and Coleman,W.B. (2006). A novel +ACA BRCA1 promoter polymorphism creates a binding site for the FAC1 transcriptional repressor. Faseb Journal *20*, A1324-A1325.

Wiwanitkit,V. (2005). Comparison for functional aberration of G-6-PD deficiency variants with exon 10 mutations. Hematology *10*, 261-263.

Wrighton,S.A. and Stevens,J.C. (1992). The human hepatic cytochromes P450 involved in drug metabolism. Crit Rev. Toxicol. *22*, 1-21.

Yalcindag,C., Silver,J., and Naveh-Many,T. (1999). Mechanism of increased parathyroid hormone mRNA in experimental uremia: roles of protein RNA binding and RNA degradation. J. Am. Soc. Nephrol. *10*, 2562-2568.

Yap,G. (2002). Affymetrix, Inc. Pharmacogenomics *3*, 709-711.

Yatsula,B., Lin,S., Read,A.J., Poholek,A., Yates,K., Yue,D.X., Hui,P., and Perkins,A.S. (2005). Identification of binding sites of EVI1 in mammalian cells. Journal of Biological Chemistry *280*, 30712-30722.

Zhang,J., Zhang,Q.Y., Guo,J., Zhou,Y., and Ding,X. (2000). Identification and functional characterization of a conserved, nuclear factor 1-like element in the proximal promoter region of CYP1A2 gene specifically expressed in the liver and olfactory mucosa. J. Biol. Chem. *275*, 8895-8902.

Zhang,W., Moorthy,B., Chen,M., Muthiah,K., Coffee,R., Purchio,A.F., and West,D.B. (2004). A Cyp1a2-luciferase transgenic CD-1 mouse model: responses to aryl hydrocarbons similar to the humanized AhR mice. Toxicol. Sci. *82*, 297-307.

Zhang,Z.Y., Pelletier,R.D., Wong,Y.N., Sugawara,M., Zhao,N., and Littlefield,B.A. (2006). Preferential inducibility of CYP1A1 and CYP1A2 by TCDD: differential regulation in primary human hepatocytes versus transformed human cells. Biochem. Biophys. Res. Commun. *341*, 399-407.

CONTRIBUTIONS OF COLLABORATORS

Mrs. Margeryta Nowakowska prepared the cDNA from several mouse strains. This cDNA was used for messenger CYP1A2 RNA quantification. Ms. Lori Lavigne performed initial comparative sequencing of the *Cyp1a2 APN* and *C3H/HeJ* genes which included the -76 polymorphisim. Her proximal *Cyp1a2* 5' region vector constructs were used for the DNase footprinting probe templates. All reported tail vein animal injections were performed by Mrs. Carole Westwood. She also prepared a set of reporter vectors for injection from glycerol stocks. Dr. Bill Casley and Mrs. Carole Westwood performed IP luciferase injections for the IVIS. All animals were cared for in the Animal resources division of Health Canada by the animal care personnel. This included food, water and bedding changes. The *APN*, *C3H/HeJ* and congenic strain breeding colonies were produced by Ms. Yelena Markovich and maintained by Mrs. Carole Westwood. Dr. Aleksander Bliu was consulted after the *ex vivo* data were collected for advice on the appropriate statistical test to use. He also performed the calculations for these tests.

APPENDIX

A.1 Appended data

Table A.1 Descriptive statistics for *ex vivo* transcriptional assay (Figure 3.5).

vector	n	mean	min	max	range	SD
negative	12	1.34	0.09	2.50	2.40	0.72
APN	12	2.73	0.48	5.40	4.92	1.67
C3H/HeJ	11	6.45	3.28	9.50	6.21	1.85

Table A.2 Mann-Whitney two-sample test for *ex vivo* transcriptional assay (Figure 3.5).

	p-value		
vector comparison	normal approximation	t-approximation	exact test
APN-C3H/HeJ	0.0006	0.0025	0.00002
negative-*C3H/HeJ*	0.0001	0.0006	0.00001
negative-*APN*	0.0464	0.0452	0.0449

Table A.3 Descriptive statistics for *ex vivo* transcriptional assay (Figure 3.9).

vector	n	mean	min	max	range	SD
negative	13	1.47	0.08	3.69	3.60	0.88
APN	12	4.42	0.09	9.44	9.35	3.31
C3H/HeJ	11	13.66	2.36	29.2	26.8	8.62
MUT	12	4.25	1.47	10.6	9.10	3.27

Table A.4 Mann-Whitney two-sample test for *ex vivo* transcriptional assay (Figure 3.9).

	p-value	
vector comparison	normal approximation	t-approximation
negative-*APN*	0.018	0.026
negative-*C3H/HeJ*	0.0001	0.0006
negative-*MUT*	0.0043	0.0087
APN-C3H/HeJ	0.0079	0.0138
APN-MUT	0.9770	0.9772
C3H/HeJ-MUT	0.0028	0.0068

CURRICULUM VITAE

Russell H. Taylor

Education

2005-present McGill Montreal, Quebec
Dentariae Medicinae Doctor, D.M.D Program (in progress)

2004-present University of Ottawa Ottawa, Ontario
Master of Science, M.Sc Program (in progress)
Biological Chemistry

2001-2004 University of Ottawa Ottawa, Ontario
Baccalaureate of Science, B.Sc
Honors Biology

1998-2000 Heritage Collage Hull, Quebec
Diplôme d'études collégiales, DEC
Pure and Applied Science

Professional experience

2003-2005 University of Ottawa Ottawa, Ontario
Teaching assistant – "Entomology 3233"
- developed statistical analysis tool

2003-2004 Health Canada Ottawa, Ontario
Summer student

2002 University of Ottawa Ottawa, Ontario
Summer student

1999 Heritage Collage Hull, Quebec
Teaching assistant – "Electromagnetism"

Publications

Taylor,R.H., Fournier,S.M., Simons,B.L., Kaplan,H., Hefford,M.A. (2005). Covalent protein immobilization on glass surfaces: Application to alkaline phosphatase. Journal of Biotechnology *118*, 265-269

Awards	Deans Honor list, University of Ottawa
Professional memberships	Canadian Dental Association, CDA
Languages	English and French