Table of Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

A	Adenine
ACO	Acyl coA oxidase
AP	Adaptor Primer
ATP	Adenosine triphosphate
BSA	Bovine Serum Albumin
bp	Base pair
C	Cytosine
CTP	Cytosine triphosphate
CYP (CYP4A)	Cytochrome P450 (P450 4a subfamily)
DEHP	Di (2-ethylhexyl) phthalate
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
cDNA	Complementary DNA
DR1	Direct Repeat (of AGGTCA) spaced b y 1 nucleotide
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra acetic acid (disodium salt)

EET	Epoxyeicosatetraenoic acid
G	Guanine
GSP	Gene Specific Primer
GTP	Guanosine triphosphate
HETE	Hydroxyeicosatetraenoic acid
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	Kilobase
kDa	Kilodalton
KAc	Potassium acetate
LB	Luria-Bertani (medium)
MCP	Methylclofenapate
$MgCl_2$	Magnesium Chloride
NH_4Ac	Ammonium acetate
PCR	Polymerase Chain Reaction
PPAR	Peroxisome Proliferator Activated Receptor
PPRE	Peroxisome Proliferator Response Element
RNA	ribonucleic acid

mRNA	messenger RNA
SDS	Sodium Dodecyl Sulphate
SRS	Substrate Recognition Sites
T	Thymine
TBE	Tris-borate-EDTA buffer
TEMED	N, N, N', N',-tetramethylethylenediamine
TESS	Transcription Element Search Software
TMACl	TetraMethyl Ammonium Chloride
TTP	Thymidine triphosphate
UHP	Ultra High Purity
UTR	Untranslated Region
WWW	World Wide Web
X-gal	5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside

Amino acids are indicated by their standard single letter or three letter abbreviations where appropriate.

ACKNOWLEDGEMENTS

I wish to acknowledge the guidance and encouragement of my supervisor, Dr David Bell, who has been a constant source of support and direction throughout my PhD. I would also like to acknowledge the help of Alex Bell, Neill Horley and Declan Brady for dosing the animals when it was required for my work. I must also express my thanks to Sharon Kuo for kindly giving me access to her sequence data for analyses.

Special thanks must also be extended to my colleagues and friends, Alex "Big Hands" Bell, Neill "You wouldn't mind!" Horley and Munim "The Pimp" Choudhury, for stimulating scientific discussions in the coffee bar. Also, thanks to my housemate, Alex "What's going on then?" Gage for endless supplies of Blue Riband from Sainsbury's to keep me going with the thesis writing.

NOTE

This work was performed in the course of my PhD studies at the University of Nottingham, and is my original work.

I wish to make clear that the sequence of exons 10-12 from *Cyp4a12*, and exons 6-12 of *Cyp4a14*, was determined by Sharon Kuo. The experimental procedures on mice were performed by Declan Brady, Neill Horley and Alex Bell, under the aegis of their personal licences.

Chapter 1 INTRODUCTION

Section 1.1 The Cytochrome P450 System

Section 1.1.1 Biochemistry

The cytochrome P450 was initially identified by Klingenberg (1958) and Garfinkel (1958) as a specialized cytochrome that absorbed light maximally at 450nm when complexed with carbon monoxide. After many years of research into these proteins, the cytochrome P450 (CYP) system is now known to constitute a superfamily of enzymes, possessing an iron protoporphyrin IX as the prosthetic group and having a molecular weight of 45-55 kDa. Membrane-bound within the endoplasmic reticulum, these enzymes are key enzymatic components of the microsomal mixed function mono-oxygenase system (MFO). Microsomal cytochromes P450 enzymes are the terminal electron acceptors from NADPH and the closely associated NADPH cytochrome P450 reductase in this electron-transport chain. Cytochromes P450 may be found in other organelles as well as the ER. In the adrenal glands, P450 proteins are also found in the mitochondria, where they accept electrons via adrenodoxin and adrenodoxin reductase instead.

Most P450-mediated metabolism catalyzes the oxidative transformation of a large number of endogenous and exogenous substrates by the insertion of a single atom of oxygen, derived from O_2, into a substrate. Depending on the substrate, different reactions can occur, including oxidative and reductive dehalogenation, N-hydroxylation and N-oxidation, oxidative deamination, S-, N- and O- dealkylation and aliphatic and aromatic hydroxylation. As a result of the multiplicity of P450s and their diverse and overlapping reactions which they catalyze, the cytochrome P450 system is thus able to metabolize a wide variety of chemicals ranging from endogenous lipids to xenobiotics. For example, P450s can convert fatty acids and to their ω and/or ω-1hydroxylated form (by the CYP4 family) (Gibson *et al.*, 1982; Bains *et al.*, 1985), or be involved in drug metabolism (by the CYP2D subfamily) (eg Distlerath *et al.*, 1985). The catalytic cycle of the cytochrome P450 is shown in Figure 1.1. Initially, the substrate binds to a site on the enzyme close

to the haem to interact with oxygen (Step 1). The iron atom is then reduced to the ferrous state by an electron transferred from NADPH via the reductase (Step 2). Molecular oxygen then binds the P450-substrate complex (Step 3), after which, a second electron is transferred from NADPH via P450 reductase to the complex (Step 4). Finally, the complex then rearranges with insertion of one atom of oxygen into the substrate to yield a product (Steps 5 and 6) and water or hydrogen peroxide.

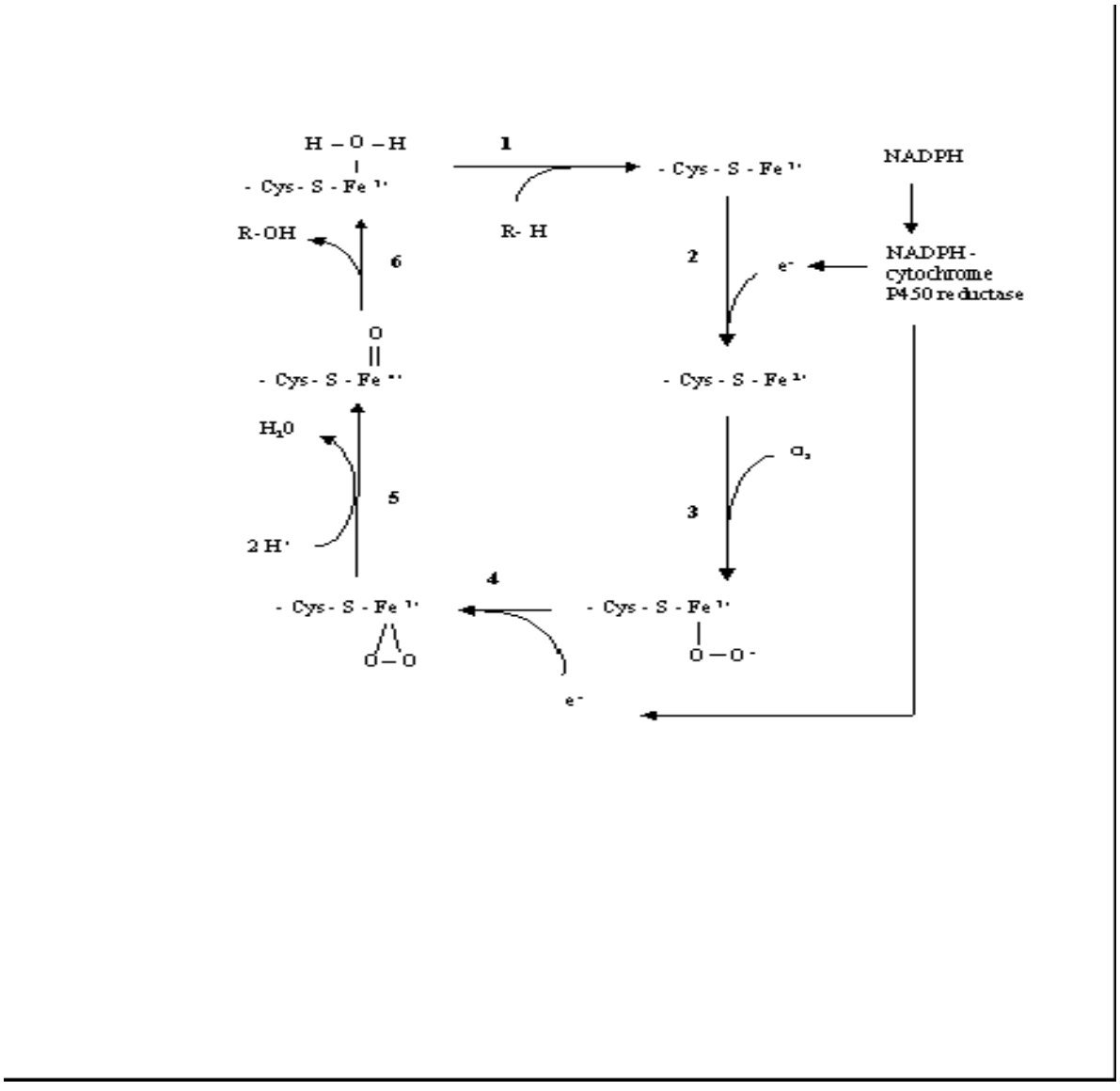


Figure 1.1 The catalytic cycle of the cytochrome P450. Step 1. Substrate binds to a site on the enzyme close to the haem to interact with oxygen. Step 2. Iron atom is reduced to the ferrous state by an electron transferred from NADPH via the reductase. Step 3. Molecular oxygen binds the P450-substrate complex. Step 4. A second electron is transferred from NADPH via P450 reductase to the complex. Steps 5/6. The complex rearranges with insertion of one atom of oxygen into the substrate to yield a product.

Section 1.1.2 Nomenclature

Nelson and coworkers (1996) have collated a nomenclature system that is constantly reviewed and updated, based on the protein sequences of P450 proteins and the divergent evolution of the superfamily. P450 genes have been identified and sequenced from plants, animals, bacteria and yeast, indicating a profound importance of this gene superfamily throughout many different phyla. To date, 481 P450 genes have been reported. These genes have been described in 85 eukaryote and 20 prokaryote species. Of 74 gene families so far described, 14 families comprise 26 mammalian subfamilies,. Each subfamily usually represents a cluster of tightly linked genes widely scattered throughout the genome.

In the nomenclature, recommendations for naming a P450 gene include the italicized root symbol 'CYP' ('Cyp' for the mouse and *Drosophila*), denoting cytochrome P450, an Arabic number designating the P450 family, a letter indicating the subfamily when two or more subfamilies are known to exist within that family, and an Arabic numeral representing the individual gene. For example, the primary hepatic drug-metabolizing enzymes currently consist of 4 gene families designated 1 through 4. Thus, '*Cyp4a12*' refers to the cytochrome P450 gene number 12 of subfamily 4A (genes of all other species are indicated by italicized and upper-cased names eg, 'CYP4A2' of the rat). Also, nonitalicized and all capital letters 'CYP4A12' represents the cDNA, mRNA and protein for all species. A P450 protein sequence from one gene family is usually defined as having more than 40% amino acid identity to another in the same family with 3 exceptions : (i) the CYP2D subfamily includes the most distant members of the CYP2 family (ii) the cockroach CYP4C1 sequence which is only 32-36% identical to mammalian CYP4A and CYP4B enzymes (Bradfield *et al.*, 1991), is assigned to family 4 on the basis of the presence of a 13-amino acid peptide found exclusively in CYP4As (Bradfield *et al.*, 1991) and (iii) two mitochondrial genes CYP11A1 and CYP11B1 are included in the same family even though the enzyme sequences are only 34%-39% identical. In genes examined thus far, genes within a defined

subfamily have been found to be nonsegregating within the same 'gene cluster', an example of which is the CYP2D subfamily (Matsunaga *et al.*, 1990). Different subfamilies within a gene family, however, are not necessarily genetically segregating distinct entities. For example, the CYP1A and CYP1B are located on different chromosomes.

Section 1.1.3 Evolution of Cytochrome P450

The first ancestral gene underwent duplication around 1360 million years ago and gave rise to cytochrome P450s found in 2 different cellular organelles, the mitochondria and the endoplasmic reticulum (Nelson and Strobel, 1987). The early history of P450 proteins appears to be related to the metabolism of cholesterol. The earliest eukaryotic P450 sequences yet known are of the mitochondrial CYP11A1 protein which catalyze cholesterol side chain cleavage to produce pregnenolone, the precursor of all steroid hormones (Nelson and Strobel, 1987). Related to these P450s are those that metabolize fatty acids such as the CYP4 family of lauric acid hydroxylases. The cholesterol- and fatty acid-metabolizing P450s may have been involved in the maintenance of membrane integrity of early eukaryotes (Nebert and Gonzalez, 1985). The formation of the endogenous steroid-synthesizing P450s, CYP17, CYP19 and CYP21 subfamilies and the catabolic enzymes then followed about 900 million years ago.

In eukaryotic genomes, the existence of gene families which contains more than one identical member can be explained by the repeated duplication of specific DNA sequences over time. Speciation can then occur to form two lines, each containing the same number of members. In divergent duplication, each of these members may then continue to diverge to become less and less similar to each other. Due to the possible occurrence of deleterious mutations, duplication of a functional gene may thus provide a buffer against such disadvantageous events and can be effective in increasing its frequency in a population.

Many other supergene families exist in eukaryotic genomes. A well studied example of this is the

mouse β-globin gene family, which is physically clustered on the chromosome (eg. Jahn *et al.*, 1980). Cytochrome P450 gene members, like the CYP2D subfamily of genes, have also found to be located on the same chromosome, forming the CYP2D cluster (Matsunaga *et al.*, 1990). Another recent example of a gene family arrangement is the genomic organization of the P450 cluster on human chromosome 19. This contains CYP2A, 2B and 2F subfamilies on a 350kb stretch of DNA (Hoffman *et al.*, 1995). It is therefore possible that other P450 subfamilies like CYP4A could be similarly colocalized in the genome.

Gene conversion, referring to a nonreciprocal recombination event in which a segment of one gene replaces the corresponding segment of a related gene (Baltimore, 1981), also serves to enhance the diversity of P450 genes, like the CYP2D subfamily. Gene conversion can also reestablish 100% similarity in a portion (or all) of a gene that had been considerably diverged from its neighbour. Therefore, gene conversion is an important molecular mechanism contributing to the evolution of species-specific genes. Four of five rat CYP2D genes characterized appear to have undergone gene conversion, leading to regions of exons and introns having high similarities relative to segments of the gene that have apparently not undergone gene conversion (reviewed in Gonzalez and Nebert, 1990). Gene conversion events have maintained homogeneity in and around the cysteine residue, which serves as the fifth ligand to haem iron and is important in the oxygenative functions of all P450 enzymes (Nebert *et al.*, 1987). This is also suggested by the sequence similarities of exons 8 and 12 in the rat CYP4A1 and CYP4A2 genes. Both these genes are about 85% identical to each other in these regions while only 66.5% identical at exon 6. Interestingly, exon 8 encodes for a highly conserved 13 amino acid sequence that is characteristic of the CYP4 family of genes (Bradfield *et al.*, 1991) while exon 12 encodes the amino acids making up the haem-iron binding domain.

Section 1.1.4 Structure of P450 Proteins

Thus far, the 3-dimensional structure is not known for any eukaryotic P450 because these membrane proteins have resisted crystallization for x-ray diffraction studies. However, the tertiary structure is known only for several bacterial P450 proteins : that of the soluble CYP101 ($P450_{cam}$) from *Pseudomonas putida* (Poulos *et al.*, 1987) and CYP102 (BM-3) from *Bacillus megaterium* (Ravichandran *et al.*, 1993; fig 1.2). There is however an obvious limitation to using bacterial P450 proteins as a structural framework for homologous modelling of eukaryotic P450s:

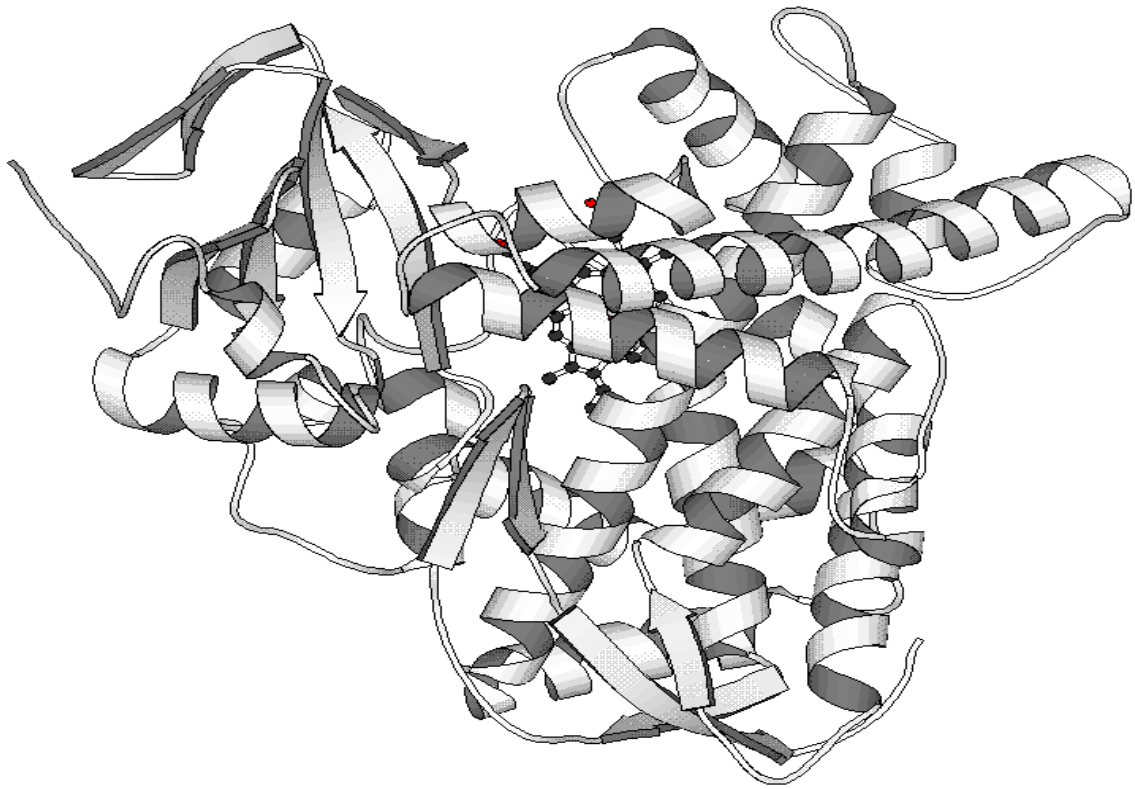


Figure 1.2 The 3-Dimensional structure of CYP102. The protein structure was deduced by MOLSCRIPT program (Kraulis, 1991) and using the protein sequence of Ravichandran and coworkers (1993).

eukaryotic P450s are membrane bound while bacterial P450 proteins are soluble, which might suggest very different structures. Nevertheless, Nelson and Strobel (1989) have aligned and predicted the secondary structure of numerous P450 proteins and used the P450cam crystal structure as a guide. Their work concluded similarities exist between bacterial and eukaryotic P450 proteins, that the helical hairpin at the NH_2 terminal of $P450_{cam}$ is maintained in all P450 proteins and the bulk of the enzyme would be exposed to the cytosolic side of the endoplasmic reticulum.

Many P450 proteins function to convert hydrophobic substrates to more hydrophilic derivatives that can be easily eliminated from the body. Because of the hydrophobicity of many P450 substrates, the membrane environment is quite suitable for P450 enzymes which are inserted into the lipid bilayer by their amino-terminal hydrophobic segments called "signal sequences", which are then cleaved off (Sabatini *et al.*, 1982). An acidic residue that is near the initiation methionine is followed by about 14 to 20 residues of hydrophobic amino acids and then by several basic residues (High and Dobberstein, 1992). Mitochondrial P450s have an amphipathic N-terminal signal sequence rich in positively charged amino acid residues that allows the protein to bind to the mitochondrial surface at contact sites between the inner and outer membranes (Omura and Ito, 1991), and have an "extrapeptide" 37 to 56 amino acids long, a conserved region of periodic arginine and lysine residues serves to transport the P450s across the outer membrane and into the mitochondrial compartment (Morohashi *et al.*, 1984).

It appears that only sequences immediately surrounding 2 key residues in P450s are absolutely conserved : the C-terminal cysteine residue, and a highly conserved threonine in the I helix involved in oxygen binding (Lesk and Chothia, 1985; Chothia and Lesk, 1980). The C-terminal Cys residue is part of an antiparallel β turn beginning at Phe (F) and ending at Cys (C) and is part of the highly conserved consensus sequence F(G/S)XGX(H/R)XCXGX(I/L/F)A, known as the haem-binding domain at the active site, which can serve as a signature for a P450 protein

(Kalb and Loper, 1988; Gonzalez, 1989). The threonine in the I helix points into the active site of the enzyme and together with a Glu or Asp just N-terminal to it, is conserved. Other conserved regions have been identified whose functional significance is less clear. These include an N-terminal proline-rich region, Proline-Proline-Glycine-Proline, that may serve to join the membrane-binding N-terminus to the globular region of the P450 (Black and Coon, 1986). Other conserved residues of note are the Glycine at position 95, a Glycine/Proline at position 86 and a Trp at position 150 (Gotoh, 1989). In the CYP4 family, there also exists a 13-residue peptide, EVDTFMFEGHDTT, that is wholly conserved around position 315 in the I helix (Bradfield *et al.*, 1991).

The molecular mechanisms underlying the broad but specific metabolic capacities of mammalian P450 systems are very poorly understood due to the lack of a crystal structure of such a mammalian protein. The substrate binding sites of the bacterial CYP101 have previously been identified by Poulos and coworkers (1987) and by inferences with alignments with this protein, Gotoh (1992) has deduced six putative substrate binding sites, SRS 1 to 6, in the CYP2 proteins. Since then, much work has gone to prove that the regions which were crucial for substrate recognition were in or around these six previously assigned sites. For example, using amino acid substitions within these SRS', He and coworkers (1994) were able to provide experimental evidence for the accuracy for the assignments of these sites in CYP2B1.

Section 1.1.5 Xenobiotic Induction of P450 genes

Some P450 genes are expressed constitutively, while others are inducible. Xenobiotic-metabolizing enzyme induction usually enhances detoxification. Thus, induction is a protective mechanism (Porter and Coon, 1991) and most often occurs at the level of transcription (Gonzalez, 1989). The inducible P450 genes represent interesting systems which can be studied whereby molecules can regulate and enhance transcription of specific genes. Classical xenobiotic inducers

of cytochrome P450s are well studied and include the polycyclic aromatic hydrocarbons, phenobarbital, steroids and peroxisome proliferators.

Section 1.1.5.1 Aromatic Hydrocarbon (Ah) Induction

Polycyclic aromatic hydrocarbons, such as the carcinogens 3-methylcholanthrene (3-MC) and benzo(a)pyrene, are prototypical inducers of several P450 proteins, most notably CYP1A1/2 and CYP1B1 (Kimura *et al.*, 1986; Sagami *et al.*, 1986). Nuclear "run-on" experiments have shown that the induction is due to an increase in the rate of transcription of these genes. (Whitlock, 1993; Sutter *et al.*, 1994). A much more potent inducer of CYP1A1 than 3-MC is the environmental contaminant 2,3,7,8, tetrachloridibenzo-*p*-dioxin (TCDD, dioxin) (Poland *et al.*, 1982). Later, Poland and his coworkers (1986) showed the existence of a mouse liver protein called "Ah receptor" which bound TCDD (and the aromatic hydrocarbons) saturably, reversibly and with high affinity. Hankinson (1995) identified and designated a protein termed Arnt that heterodimerizes with liganded AhR to generate an active, enhancer-binding transcription factor. Elegant experiments have indicated the presence of a TCDD-inducible, AhR-dependent and Arnt-dependent transcriptional enhancer and a transcriptional promoter upstream of the CYP1A1 gene (Whitlock, 1993; Hankinson, 1995). Thus, binding of the AhR-Arnt to the enhancer bends the DNA and facilitates the binding of transcription factors to the promoter.

Section 1.1.5.2 Phenobarbital Induction

Phenobarbital (PB) induces CYP2B1/2 genes in rat liver (Waxman and Azaroff, 1992) and the structurally related CYP2H1/2 in chick embryo liver (Gonzalez, 1989). Although PB induces transcription of various P450 genes in bacteria, birds and mammals, the mechanisms of induction are different. For example, ongoing protein synthesis is required for CYP2B1/2 induction but not for CYP3A (Waxman and Azaroff, 1992). In *Bacillus megaterium*, a 17-base pair regulatory element (called the "Barbie Box") is located upstream of the barbiturate-inducible CYP102 (BM-3) gene (Fulco, 1991). PB response elements upstream of the CYP2B1/2 and CYP2H1

genes (Waxman and Azaroff, 1992) have also been identified. Relatively little is understood, however, about the mechanism by which PB and PB-related chemicals induce P450 gene expression. Ligand binding studies have failed to detect a specific PB-binding protein. It seems possible that PB-related compounds act indirectly causing an accumulation of an endogenous steroid, as suggested by Shaw and his coworkers (1993), who demonstrated that PB responsiveness was influenced by glucocorticoids.

Section 1.1.5.3 Steroid Induction

The "catatoxic" steroid, pregnenolone-16a-carbonitrile (PCN) has also been shown to induce P450 enzymes (Okey, 1990), now designated the CYP3A family, which metabolize a variety of substrates, including testosterone and various drugs (Gonzalez *et al.*, 1986; Wrighton *et al.*, 1985). Exposure of rats to either PCN or dexamethasone leads to an increased rate of CYP3A gene transcription (Simmons *et al.*, 1987). A dexamethasone/PCN-response element has been identified upstream of the CYP3A1 gene (Burger *et al.*, 1992). Regulation of genes of the CYP3 family also occurs at the post-transcriptional level. The rat CYP3A1/2 proteins are stabilized by the macrolide antibiotic triacetyloleandomycin (TAO) (Watkins *et al.*, 1986). However, steroid receptors have not been found to be directly responsible for the induction of PB-responsive genes (Burger *et al.*, 1992)

Section 1.1.5.4 Peroxisome Proliferator Induction

Peroxisomal proliferators (PPs) cause an increase in transcription of CYP4 family of enzymes (Hardwick *et al.*, 1987). This induction has been demonstrated to be an early and crucial event, prior to the induction of enzymes involved in the β-oxidation of fatty acids like acyl coA oxidase, during the process of peroxisome proliferation in rat (Bell and Elcombe, 1991; Milton *et al.*, 1990). Chronic administration of peroxisome proliferators leads to liver carcinogenesis. The transcriptional nature of the induction response led to the hypothesis that PPs act via a receptor protein related to steroid hormone receptors. A receptor, now termed the PPAR, has since been

identified which could be activated by a variety of peroxisome proliferators including hypolipidaemic drugs and a plasticiser (Isseman and Green, 1990). In addition, response elements upstream of the clofibrate-responsive genes like the rat CYP4A1 (Aldridge *et al.*, 1995), rabbit CYP4A6 (Muerhoff *et al.*, 1992a) and acyl CoA oxidase (Osumi *et al.*, 1990) have been identified and shown to bind PPARα *in vitro*. The PPARα has been formally demonstrated to be an important factor for peroxisome proliferation. Lee and coworkers (1995) utilized PPARα-null mice to demonstrate that these 'knockout' mice were refractory to the effects of peroxisome proliferating chemicals.

Section 1.2 The Cytochrome P450 4 Family (CYP4)

Section 1.2.1 Discovery

In 1982, Orton and Parker noted an approximately 28-fold increase in lauric acid ω-hydroxylation activity, increase in liver size and cytochrome P450 in rat after administration of clofibrate (Orton and Parker, 1982). A clofibrate inducible lauric acid ω-hydroxylase, named CYP4A1, was first purifed from rats (Tamburini *et al.*, 1984). ω-hydroxylases have also been isolated from the lungs of pregnant rabbbits (Williams *et al.*, 1984), large colon (Kaku *et al.*, 1984) and placenta (Yamamoto *et al.*, 1986) of the rabbit and human and rat (Kawashima *et al.*, 1992) kidney.

The common nomenclature of Nelson *et al.*(1996) designated this subfamily as CYP4A, typically a group of cytochrome P450s with fatty acid ω and ω-1 hydroxylase activity. This family to date contains 14 genes, including genes from the rat, mouse, guinea pig, rabbit and human. These enzymes are thought to have a range of differing specificities for fatty acids and prostaglandins. Many of these genes are found at high levels in the liver and kidney, while others are expressed in sex- and tissue-specific patterns. The CYP4 family is further categorized into subfamilies and contains 66 genes from the rat, rabbit, mouse, human, guinea pig, cockroach, mosquito, *Drosophila*, housefly, beetle roach tobacco hornworm (Nelson *et al.*, 1996), including a newly cloned

mouse *Cyp4b1* gene (Imaoka *et al.*, 1995).

Section 1.2.2 The CYP4A Subfamily

The CYP4A subfamily in particular, contains 13 genes (Table 1.1). CYP4A1 (Hardwick *et al.*,

Gene	Species	cDNA/ Genomic	References
CYP4A1	Rat	cDNA Genomic	Hardwick *et al.*, 1987 Kimura *et al.*, 1989a
CYP4A2	Rat	Genomic	Kimura *et al.*, 1989a
CYP4A3	Rat	cDNA	Kimura *et al.*, 1989b
CYP4A4	Rabbit	cDNA Genomic	Matsubara *et al.*, 1987 Palmer *et al.*, 1993a
CYP4A5	Rabbit	cDNA	Johnson *et al.*, 1990
CYP4A6	Rabbit	cDNA Genomic	Yokotani *et al.*, 1989 Muerhoff *et al.*, 1992
CYP4A7	Rabbit	cDNA	Yokotani *et al.*, 1989
CYP4A8	Rat	cDNA	Stromstedt *et al.*, 1990
CYP4A9	Human	cDNA	(listed in Nelson *et al.*, 1991)
Cyp4a10	Mouse	cDNA	Henderson *et al.*, 1990 Bell *et al.*, 1993
CYP4A11	Human	cDNA	Palmer *et al.*, 1993b Imaoka *et al.*, 1993
Cyp4a12	Mouse	cDNA	Bell *et al.*, 1993
CYP4A13	Guinea Pig	cDNA	Bell *et al.*, 1993

Table 1.1 Genes in the CYP4A subfamily. The species from which the cDNA or genomic sequence was cloned are indicated. The laboratories of identification of the genes are listed as references.

1987; Kimura *et al.*, 1989a), CYP4A2 (Kimura *et al.*, 1989b), CYP4A3 (Kimura *et al.*, 1989b) and CYP4A8 (Stromstedt *et al.*, 1990) are rat genes; CYP4A4 (Matsubara *et al.*, 1987; Palmer *et al.*, 1993), CYP4A5 (Johnson *et al.*, 1990), CYP4A6 (Yokotani *et al.*, 1989; Muerhoff *et al.*, 1990) and CYP4A7 (Yokotani *et al.*, 1989) rabbit genes; CYP4A9 and CYP4A11 (Palmer *et al.*, 1993; Imaoka *et al.*, 1993) human genes; CYP4A13 a guinea pig gene (Bell *et al.*, 1993); and *Cyp4a10* (Bell *et al.*, 1993; Henderson *et al.*, 1994) and *Cyp4a12* (Bell *et al.*, 1993) are mouse genes.

Section 1.2.2.1 Rat CYP4A genes

In the rat, only 4 CYP4A genes have been identified, two of which are genomic (CYP4A1 and CYP4A2) and two are cDNA (CYP4A3 and CYP4A8). CYP4A1 has been isolated from a rat genomic library and was found to span 14.14 kb and contains 13 exons, but only the first 12 are coding regions (Kimura *et al.*, 1989a). No TATA box was found within 40bp upstream of the transcription start site. However, the gene possesses a GATTTA motif located between -20 and -30 positions. Also, a 19bp sequence was found to be highly conserved between CYP4A1 and CYP4A2 genes. CYP4A1 is deduced to encode for 509 amino acids.

CYP4A2 was isolated from a rat genomic library and was found to span 10.58 kb (Kimura *et al.*, 1989). It contains 12 exons and like CYP4A1, it possesses a GATTTA sequence motif instead of a TATA box. An interesting feature of CYP4A2 is the presence of a large direct repeat of 378 bp located at -440 to -817 and -1045 to -1422 upstream of the start site. CYP4A2 is deduced to encode for 504 amino acids.

CYP4A3 was isolated from a cDNA library from clofibrate-treated rat liver by screening with a CYP4A1 cDNA probe (Kimura *et al.*, 1989b). The cDNA encodes a protein of 507 amino acids and a calculated molecular weight of 58 kDa. The CYP4A3 cDNA shared 65% and 97% nucleotide and 72% and 96 % deduced amino acid sequence similarities with CYP4A1 and CYP4A2 respectively.

A fourth rat CYP4A cDNA, termed CYP4A8, was isolated from rat prostate (Stromstedt *et al.*, 1990). It is deduced to encode for 506 amino acids and the deduced amino acid sequence shows 76 % identity with CYP4A1.

The identities of each protein based on their deduced amino acid sequences for CYP4A1, CYP4A2, CYP4A3 and CYP4A8 are compared and shown in Table 1.2 below.

Section 1.2.2.2 Rat CYP4A enzyme activity

CYP4A1 is one of the most active fatty acid metabolizing P450s, and hydroxylates laurate, palmitate and arachidonate at the ω and, to a lesser extent, ω-1 position (Tamburini *et al.*, 1984; Hardwick *et al.*, 1987). The ω / ω-1 ratio is about nine-fold higher for laurate hydroxylation than palmitate hydroxylation by CYP4A1 (Aoyama *et al.*, 1990). CYP4A3 however has a preference for hydroxylation of palmitate at the ω position compared to CYP4A1 (Aoyama *et al.*, 1990). The same group also found that CYP4A1 and CYP4A3 were also able to hydroxylate prostaglandins E_1 and $F_{2\alpha}$, albeit only at their ω positions and at considerably lower rates than those for fatty acid oxidations (Aoyama *et al.*, 1990). CYP4A2, however, does not appear to metabolize prostaglandins A1 and A2 by oxidation (Yoshimoto *et al.*, 1986). It is unknown whether fatty acid hydroxylation by CYP4A proteins play a significant physiological role, although their role in the metabolism of arachidonic acid has been shown to be important in regulating water and electrolyte balance in the kidney (*vide infra*).

Section 1.2.2.3 Rat CYP4A gene expression

CYP4A1 is found at low constitutive levels in liver and kidney but is highly induced by clofibrate, especially in the liver. In the kidney, expression is localized to the proximal convoluted tubule and in the liver, expression was found to be localized to the centrolobular region (Bell *et al.*, 1991). This gene has also been found to be expressed at very low levels in the lung (Stromstedt *et al.*, 1994).

CYP4A2 was found to be expressed at low constitutive levels in male liver but was highly induced by clofibrate. In the kidney, the gene was found to be at high levels in untreated animals. Clofibrate treatment only resulted in a slight induction (Kimura *et al.*, 1989b). CYP4A2 was also detected in seminal vesicle, brain cortex, prostate, cerebellum and brainstem (Stromstedt *et al.*, 1994). No protein was detected in female rat liver or kidney (Sundseth and Waxman, 1992). The sex specificity of CYP4A2 is regulated by the different growth hormone profiles between the

sexes (Waxman *et al.*, 1991). In the male liver, 'continuous' infusion with growth hormone, a female pattern, results in feminization and consequently a repression of this gene (Sundseth and Waxman, 1992). In female rats, 'pulsatile' growth hormone treatment caused an increase in the level of CYP4A2, suggesting that it is the female profile of continuous GH secretion which is responsible for the suppression of CYP4A2 in female liver.

CYP4A3, like CYP4A1, was found in liver and kidney and inducible by peroxisome proliferators. But unlike CYP4A1, CYP4A3 was found to be expressed in a sex-independent manner in both tissues (Sundseth and Waxman, 1992). CYP4A3 was found to be expressed in lung, prostate, brain cortex, cerebellum, hypothalamus-preoptic area (HPOA), and brainstem in a pattern similar to CYP4A2. But unlike this gene, CYP4A3 was not found in seminal vesicle (Stromstedt *et al.*, 1994).

The CYP4A8 gene was found expressed mainly in the prostate and kidney in Northern blot experiments with a faint signal detected in the retina (Stromstedt *et al.*, 1994). None was detectable in liver RNA. In situ hybridization of kidney sections showed strongest expression in the proximal tubules. In the same study, castrated rats and androgen releasing implants were used to examine the expression of this gene. In castrated rats, CYP4A8 in the prostate was completely abolished, but was restored by testosterone treatment.

Section 1.2.2.4 Mouse *Cyp4a* genes

At present, 2 mouse genes, *Cyp4a10*, *Cyp4a12* are listed (Nelson *et al.*, 1996). *Cyp4a10*, a murine cDNA has been cloned in two different laboratories (Bell *et al.*, 1993; Henderson *et al.*, 1994). Bell and coworkers (1993) used a PCR-based method to clone a murine *Cyp4a10* cDNA fragment using primers derived from predicted highly conserved sequences in exon 8 and exon 12 of the gene. Henderson and coworkers (1994), on the other hand, isolated an identical but full length *Cyp4a10* cDNA clone of about 1.6 kb by screening a mouse liver cDNA library using a

full-length rat CYP4A1 cDNA as probe. The protein is deduced to comprise of 509 amino acids and its deduced sequence was found to be most identical (92%) to the rat CYP4A1 protein.

The second mouse *Cyp4a* gene to be identified is the *Cyp4a12* (Bell *et al.*, 1993). The cDNA fragment of *Cyp4a12* was cloned by PCR, also based on primers derived from the highly conserved exons 8 and 12. The *Cyp4a12* is the homologue of the rat CYP4A8, showing approximately 93% amino acid identity with it (Bell *et al.*, 1993). Table 1.2 is a summary of the percentage similarities of all full length rat and mouse CYP4A proteins compared. To date, no murine genomic *Cyp4a* gene has been cloned. Also, it is not known if other genes exist which belong to the *Cyp4a* subfamily in the mouse.

Protein	CYP4A1	CYP4A2	CYP4A3	CYP4A8	CYP4A10
CYP4A1	100	72	72	76	92
CYP4A2		100	96	71	72
CYP4A3			100	72	72
CYP4A8				100	76
CYP4A10					100

Table 1.2 Percent similarities of rat and mouse CYP4A proteins. Deduced amino acid sequences of rat and mouse CYP4A proteins are compared and similarities with one another presented as a percentage value. Only full length protein sequences were compared.

Section 1.2.2.5 Mouse *Cyp4a* enzyme activity

Little work has been done to investigate the metabolism of fatty acids by murine CYP4A proteins. Henderson and coworkers (1990), showed the presence of ω-hydroxylase activity in untreated mice, and that male mice exhibited roughly 3-fold higher activity than female mice. They also showed the presence of CYP4A proteins in Western blots; however the individual members of the *Cyp4a* subfamily were undetectable with the anti-CYP4A1 antibody used. There has been work, to show the constitutive levels of CYP4A-related proteins metabolizing lauric acid at the ω and ω-1 positions, with a slight preference for ω-1 hydroxylation (Hiratsuka *et al.*, 1996). This work, however, makes no distinction between the individual CYP4A proteins, or indeed other

Section 1.2.2.6 Mouse *Cyp4a* gene expression

Treatment of mice with peroxisome proliferators induced *Cyp4a10* RNA in liver and to a lesser extent in the kidney (Bell *et al.*, 1993). The *Cyp4a12* gene exhibited distinct regulation, however, showing constitutively high levels in male liver and kidney, and did not show any induction after treatment of male mice with a peroxisome proliferator (Bell *et al.*, 1993). By contrast, in untreated female kidney and liver, the expression of this gene was found to be at very low levels but highly inducible by peroxisome proliferators. The regulation of mouse CYP4A proteins in mouse liver and kidney has been studied previously. Work by McLeod and Shapiro (1989) has shown that mice exhibit a sexually differentiated pattern of GH secretion, which is more frequently pulsed in female than in male mice. A series of experiments were carried out on hypophysectomized mice and growth hormone (GH)-deficient 'little' mice to establish the role of pituitary hormones in the regulation of hepatic and renal proteins (Henderson *et al.*, 1994). In their study, Henderson and workers used polyclonal CYP4A antibodies to demonstrate the suppression of a supposedly male-specific protein under pulsatile GH treatment, which is female specific. Conversely, a protein was detected in 'little' mice when treated with pulsatile GH, implying the female-specific expression of this protein. A major flaw with these experiments is the use of a polyclonal CYP4A1 antibody which does not discriminate between the various isoenzymes, thereby adding to the confusion. It is therefore unclear which expressed protein was indeed being detected at all. Their paper also does not mention the effects of testosterone, a fundamental difference between the male and female animal. Testicular feminized mice (tfm) have been used in experiments to study mouse renal sexual dimorphism and testosterone responsiveness (Henderson and Wolf, 1991). This strain of mice carry an androgen receptor defect, making them phenotypically female but genotypically male. The gonads, hypothalamus and pituitary form a hormonal axis that regulate the expression of sex-specific enzymes, with the gonadal steroid act-

ing through the hypothalamus and pituitary. Male 'tfm' mice possess female genitalia and would not produce testosterone, would therefore differ fundamentally from normal male mice. As a result of the impaired gonadal-hypothalamic-pituitary axis of these 'male' mice, it is thus not possible to interpret any data obtained for these mice.

Section 1.2.2.7 Rabbit CYP4A genes

CYP4A4, CYP4A5, CYP4A6 and CYP4A7 are the four genes that have been identified in the rabbit (Nelson *et al.*, 1996). CYP4A4 was isolated by Matsubara and coworkers (1987) in rabbit lung. The deduced amino acid sequence shows it contains 506 amino acids with a calculated molecular weight of 58 515. It shows 74% and 77% similarity in the amino acid and the nucleotide sequence respectively compared with CYP4A1. Palmer and coworkers (1993a) have reported the gene structure of the CYP4A4 to be similar to the cofibrate-inducible genes CYP4A1 and CYP4A2 in terms of intron/exon structure. However, the 5' flanking sequences of these genes show little similarity. Like the CYP4A genes, CYP4A4 promoter does not contain a TATA consensus sequence or recognition elements for Sp1.

CYP4A6 and CYP4A7 are cDNA clones isolated from a rabbit kidney cDNA library (Yokotani *et al.*, 1989). Genomic DNA encoding the CYP4A6 has been also cloned (Muerhoff *et al.*, 1992b). The promoter region of this gene was found to contain an Sp1 consensus recognition sequence at -46 to -37 upstream of the start site. Also, a sequence matching the first 11 of 12 nucleotides that constitute a consensus HNF-4 binding site occurs just upstream of the transcription start site. There are also two CCAAT boxes located at -464 and - 448 in the 5' flanking region. A peroxisome proliferator response element (PPRE) was also found in a segment of the 5' flanking region at -740 to -723, which has been shown to regulate the induction of this gene by peroxisome proliferators *in vitro* (Muerhoff *et al.*, 1992a; Palmer *et al.*, 1994).

Section 1.2.2.8 Rabbit CYP4A enzyme activity

Unlike the rodent CYP4A proteins, CYP4A4 efficiently hydroxylates various prostaglandins (PGs) such as PGE, PGE_2, $PGF_{2\alpha}$, PGD_2, PGA, or PGA_2 (Kusunose *et al.*, 1985; Williams *et al.*, 1984; Roman *et al.*, 1993). None of the expressed CYP4A5, CYP4A6 or CYP4A7 hydroxylate PGE_1. Expressed CYP4A4 however, is unable to ω-hydroxylate lauric acid but palmitate metabolism has been shown for both expressed and purified enzyme (Roman *et al.*, 1993; Johnson *et al.*, 1990). CYP4A5 is active in the ω and ω-1 hydroxylation of laurate as are CYP4A6 and CYP4A7 (Johnson *et al.*, 1990) and to a lesser extent palmitic acid (Roman *et al.*, 1993). Towards arachidonic acid, all show significant ω-hydroxylase activity except CYP4A5 (Roman *et al.*, 1993). These data show that each enzyme in the subfamily exhibits a unique and marked substrate specificity.

Section 1.2.2.9 Rabbit CYP4A gene expression

CYP4A4 mRNA has not been shown to be detectable in the liver or kidney of untreated or clofibrated- treated rabbits but was shown to be induced in liver, lung, small intestine and uterus in pregnant animals (Palmer *et al.*, 1993a) or by treatment with progesterone (Matsubara *et al.*, 1987). Treatment of male rabbits with dexamethasone increased the levels of CYP4A4 mRNA in the lung and liver, but the levels were eightfold less than those seen for pregnant rabbits (Palmer *et al.*, 1993a). CYP4A5 was found to be expressed constitutively in the rabbit kidney, liver and small intestine (Johnson *et al.*, 1990; Roman *et al.*, 1993) but was only induced two-fold in the liver by clofibrate (Roman *et al.*, 1993). The level of CYP4A5 mRNA in the kidney did not appear to be affected by either clofibrate treatment or pregnancy (Roman *et al.*, 1993). CYP4A6 expression, however, is distinctly different. The mRNA for CYP4A6 was barely detectable in untreated animals but was increased by the administration of clofibrate in the liver (12-fold) and kidney (6-fold) (Yokotani *et al.*, 1989; Roman *et al.*, 1993). In contrast, CYP4A7 was expressed highly constitutively in the liver, kidney and small intestine, but its expression was enhanced only

in the liver by clofibrate treatment (Yokotani *et al.*, 1989; Roman *et al.*, 1993), with no induction in the kidney. Data on substrate specificity and expression patterns of the rabbit CYP4A proteins combine to show the peculiarity of the rat CYP4A genes, which suggests a distinct metabolic role for every isoform. This property makes the rabbit genes markedly different to the rodent genes.

Section 1.2.2.10CYP4A in Other Species

CYP4A9, CYP4A11 and CYP4A13 are the remaining CYP4A subfamily members listed (Nelson *et al.*, 1996). Only their cDNAs have been cloned. CYP4A9 and CYP4A11 are cloned from human (Palmer *et al.*, 1993b; Imaoka *et al.*, 1993). CYP4A11 cDNA was isolated from a human kidney using the rat CYP4A3 cDNA as a probe (Imaoka *et al.*, 1993). It was deduced to encode a protein of 519 amino acids and shows 76 %, 72% and 80 % similarities to the rat CYP4A1, CYP4A3 and rabbit CYP4A6 respectively. CYP4A13 was isolated by PCR as a partial cDNA fragment from the guinea pig (Bell *et al.*, 1993).

Section 1.2.2.11CYP4A enzyme activity in other species

The human CYP4A11 was shown to catalyze the ω-hydroxylation of lauric acid, palmitic acid and arachidonic acid in order of decreasing efficiency, with little or no activity towards prostaglandins A_1 and E_1 (Palmer *et al.*, 1993b). No data is available for CYP4A9 and CYP4A13.

Section 1.2.2.12CYP4A gene expression in other species

RNAse protection assays showed the presence of CYP4A11 mRNA in human kidney and liver but none in lung, skin, colon, small bowel, ileum, stomach, uterus among others (Palmer *et al.*, 1993b). In contrast to the great inducibility of CYP4A genes in the rodent, even high doses of methylclofenapate did not elicit an induction of CYP4A13 RNA in either male or female guinea pig liver (Bell *et al.*, 1993).

Section 1.2.3 Importance of CYP4 Family in Peroxisome Proliferation

The CYP4A family of enzymes, which are fatty acid hydroxylases, are one of the many enzymes

induced during peroxisome proliferation (Sharma *et al.*, 1988). The fact that peroxisome proliferation and the CYP4 family of enzymes are linked to lipid metabolism prompted much research to be directed towards the understanding of the function of the CYP4A gene subfamily. There is considerable evidence that peroxisome proliferators alter hepatic lipid metabolism which supports the idea that this could lead to peroxisome proliferation. It has been shown that CYP4A1 is induced by medium-chain (C_6-C_8) fatty acids in cultured hepatocytes and that it metabolizes preferentially long-chain fatty acids, which in turn can be converted to long-chain dicarboxylic acids. Long chain dicarboxylic acids (C_{14}-C_{18}) induce peroxisomal β-oxidation but not microsomal ω-oxidation in hepatocytes and stimulate [^{3}H] thymidine incorporation into DNA in hepatocyte culture (Lock *et al.*, 1989), thus contributing to the hyperplasia seen following peroxisome proliferator administration. This hypothesis requires the induction of CYP4A proteins prior to the induction of peroxisomal enzymes like acyl coA oxidase, which indeed is the case (Milton *et al.*, 1990; Bell *et al.*, 1991; Bell and Elcombe, 1991). The dual administration of clofibrate and the protein synthesis inhibitor cycloheximide did not prevent the transcription of CYP4A1 mRNA but inhibited both CYP4A1 and acyl coA oxidase activity inductions (Milton *et al.*, 1991). However, CYP4A1 mRNA induction by clofibrate was unimpeded whereas the corresponding mRNA for acyl coA oxidase was not induced, suggesting the requirement of a protein, possibly CYP4A, for the subsequent induction of acyl coA oxidase. Further evidence was provided by work which found that the specific inactivator of CYP4A1, 10-undecynoic acid (UDYA), inhibited clofibrate-mediated induction of acyl coA oxidase mRNA but not the induction of CYP4A1 mRNA in hepatocytes (Kaikaus *et al.*, 1993). This study indicates that the CYP4A1 ω-oxidation pathway, or even the CYP4A1 protein itself, plays an important part in the peroxisome proliferator induction of peroxisomal β-oxidation. This work also demonstrates that only long-chain dicarboxylic acids, the products of the ω-oxidation pathway, but not medium-chain dicarboxylic acids or long-chain monocarboxylic fatty acids, also enhanced peroxi-

somal β-oxidation activities (Kaikaus *et al.*, 1993). However, they also show that the long-chain dicarboxylic acids are poorly metabolized by hepatocytes and the sustained presence of these compounds due to their slow utilization may explain their greater potency as peroxisome proliferators compared to monocarboxylic fatty acids. Lee and coworkers (1995) have demonstrated that "knockout" mice lacking PPARα, which is a key transcription factor regulating induction of CYP4A, acyl coA oxidase and other enzymes, are not responsive to peroxisome proliferators. It would certainly further our understanding of the phenemenon to examine if 'knockout' mice lacking CYP4A genes undergo peroxisome proliferation. However, no mouse *Cyp4a* genes have been cloned to date for such key experiments to be undertaken. The available experimental evidence suggests that only rats and mice undergo peroxisome proliferation, with elevated levels of CYP4A expression. However, attention has only focussed on the study of the expression of the rat genes in an artificial cell line, using cotransfected PPARs from another species (eg. Aldridge *et al.*, 1994) when the mouse PPARs are the most well-characterized. It is clearly desirable to study the true response of a promoter region of a murine gene in hepatocytes with its endogenous PPAR as only this can demonstrate the real effects of peroxisome proliferators. It remains to be seen if a murine CYP4A gene possesses regulatory elements that are associated with its transcriptional activation by peroxisome proliferators.

Section 1.2.4 Importance of CYP4 in Arachidonic Acid Metabolism

Arachidonic acid, like other fatty acids and their derivatives, is a substrate for CYP-dependent oxidation reactions. The functional significance of arachidonic acid as the precursor of prostanoids and leukotrienes is also well established (Nelson and Strobel, 1987). Arachidonic acid is metabolized by CYP4A family proteins to ω and ω-1 hydroxyeicosatetraenoic acids (HETEs), epoxyeicosatrienoic acids (epoxides/EETs) and dihydroxyeicosatetraenoic acids (diols/DHTs), all which play important biological roles. In response to salt, urinary excretion of EETs was enhanced, reflecting increased renal synthesis of EETs (Capdevila *et al.*, 1992). Other studies have

described the capacity of exogenously added EETs to modify cell membrane permeability to peptide hormones and to alter cellular concentrations of ions such as H^+, Na^+, K^+ and Ca^{2+} (Fitzpatrick and Murphy, 1989). HETEs have been found on the human skin and been increased during psoriatic inflammation (Woollard, 1986). ω and ω-1 hydroxylated HETEs have been identified at several sites intrarenally that are pivotal to the regulation of renal function (Imig *et al*, 1993; Omata *et al*., 1992). 20-HETE was also shown to be a potent vasoconstrictor of isolated perfused renal arteries (Ma *et al*., 1993). An inhibitor of microsomal 20-HETE production, 17-octadecynoic acid (17-ODYA), has been shown to increase urine flow and sodium excretion whereas renal blood flow and glomerular filtration rate were not significantly altered (Zou *et al*., 1994). This suggests that endogenous cytochrome P450 derived metabolites of arachidonic acid are important in regulating renal function. 20-HETE production is concentrated in the proximal tubule with the highest specific activity in the S3 segment, although much can be produced by other tubular segments and non-tubular structures like the thick ascending limb of Henle's loop (Carroll *et al*., 1991). In the spontaneously hypertensive rat (SHR), which displays higher renal P450 arachidonate metabolism, specifically 20-HETE production (Omata *et al*., 1992; Sacerdoti *et al*., 1987), all three CYP4A proteins, CYP4A1, CYP4A2 and CYP4A3 have been shown to be expressed when 20-HETE production is at its maximum (Schwartzmann *et al*., 1996). This suggests that the activation of one or more of the CYP4A isoforms is involved in increased 20-HETE synthesis. Northern blot analysis used a CYP4A1 cDNA probe which could make no distinction between the isoforms due to the high sequence homology. The use of the highly sensitive RNAse protection to distinguish between the 3 genes would therefore be preferable. The cloning and understanding of the murine *Cyp4a* gene subfamily would enable the study of their function in "knockout" animals in relation to renal function and disease.

Section 1.3 The Peroxisome Proliferator Activated Receptor

Section 1.3.1 Discovery

The identification of a cellular protein, known as the aryl hydrocarbon (Ah) receptor (Poland and Knutson, 1982) enabled the understanding of the 'structure activity' relationship for dioxin and its congeners. Although dioxin, a liver carcinogen, and peroxisome proliferators both evoke a pleiotropic response, it is not clearly understood at a molecular level how peroxisome proliferators of varying chemical structures cause peroxisome proliferation. Lalwani and coworkers (1987) were the among the first to isolate a binding protein for nafenopin, a peroxisome proliferator. This protein has since been shown to be a member of the heat shock protein family HSP70. This discovery then focussed efforts in isolating a receptor that is the equivalent of the Ah receptor. A breakthrough was made in 1990 by Isseman and Green who isolated cDNAs for members of the steroid hormone receptor superfamily from mouse liver. The first peroxisome proliferator activated receptor was thus isolated, termed mPPARα (Isseman and Green, 1990). The PPAR is a member of the nuclear hormone receptor superfamily which is characterized by a DNA-binding domain composed of 2 highly conserved zinc fingers (Klug and Schwabe, 1995).

Chimeras of the PPAR were used in trans-activation assays to show its involvement in the transcription of reporter genes in response to the peroxisome proliferator nafenopin. A direct involvement of mPPARα in peroxisome proliferation was demonstrated recently by a 'knockout' of this gene : mice which were null for PPARα did not undergo peroxisome proliferation (Lee *et al.*, 1995).

Section 1.3.2 PPAR Subtypes

The cloning of further PPAR genes soon followed after the identification of the first PPAR. To date, PPARs have been cloned in the rat (rPPARα and rPPARβ) (Gottlicher *et al.*, 1992), *Xenopus* sp. (xPPARα, xPPARβ and xPPARγ) (Gottlicher *et al.*, 1992; Dreyer *et al.*, 1992; Dreyer *et al.*, 1993), human (hPPARα, hNUC1 or hPPARβ and hPPARγ1) (Sher *et al.*,1993; Schmidt

et al., 1992; Greene *et al.*, 1995), hamster (hamPPARγ) (Aperlo *et al.*, 1995) and 2 other PPAR genes in the mouse (mPPARβ or mPPARδ and mPPARγ1) (Chen *et al.*, 1993; Zhu *et al.*, 1993; Kliewer *et al.*, 1994; Tontonoz *et al.*, 1994).

Section 1.3.3 Tissue Specificity, Relative Abundance and Regulation

The physiological functions of the PPARs are very poorly understood. The tissue localization of various PPAR subtypes, however, has been studied by various research groups in an attempt to elucidate their physiological roles. By in situ hybridization using specific probes for PPARα, -β and -γ, the expression patterns have been examined in the adult rat (Braissant *et al.*, 1996). In the rat kidney and liver, PPARα and -β were found to be expressed at most prominent levels compared to the low -γ levels. PPARγ was expressed at very low levels in most tissues compared to the -α and -β, except in the white adipose tissue and the immune system where the -γ was at the highest level of the 3 subtypes. PPARβ was expressed at the highest level in the kidney, the CNS and the genital system. In this work, rat PPARβ and PPARγ specific probes were obtained by RT-PCR, using primers derived from the mouse PPARβ and -γ cDNA sequence, generating a subtype specific oligonucleotide of 135 and 403 nucleotides respectively. Each of these oligonucleotides is 96% homologous to its mouse PPAR counterpart. A similar study on the PPAR tissue distribution in mouse was undertaken (Jones *et al.*, 1995) using highly specific RNAse protection assays. In the liver, PPARα was the predominant PPAR, with 20- fold less PPARβ and virtually negligible levels of PPARγ. PPARα was also present at appreciable levels in kidney, brown adipose tissue (IBAT), brain and heart and at low levels elsewhere. PPARβ was present at highest levels in liver, kidney, IBAT, WAT and spleen. IBAT was the principal site of expression of PPARγ, suggesting that this receptor subtype plays an important role in lipid metabolism in brown adipose tissue.

The regulation of PPAR expression is not well understood and has only been studied in the PPA-

Rα subtype. There have been reports that PPAR RNAs are inducible in liver after treatment with peroxisome proliferators (Gebel *et al.*, 1992; Zhu *et al.*, 1993), however this was not observed in work done by other groups (eg. Jones *et al.*, 1995; Sorenson *et al.*, 1993). Dexamethasone, a synthetic glucocorticoid, however, has been shown to stimulate expression of PPAR mRNA up to 15 times over control (Steineger *et al.*, 1994). In the same study, insulin blocked this stimulatory effect totally and minor inductions of PPAR mRNA were observed when different fatty acids were administered. PPAR expression is also stimulated by stress and follows a diurnal rhythm, as demonstrated by Lemberger and co-workers (1996).

Section 1.3.4 PPAR Activation

Exogenous compounds, known to cause peroxisome proliferation, such as hypolipidaemic fibrate drugs, phthalate ester plasticizers and herbicides, have been shown to activate the PPAR (Isseman and Green, 1990; Dreyer *et al.*, 1992). However, these are not the natural ligands for PPAR. Fatty acids have been proposed as the endogenous activators of PPAR as high fat diets stimulate peroxisomal β-oxidation and peroxisome proliferation. Indeed, this has been shown to be true as a range of fatty acids has been demonstrated to activate PPARα (Keller *et al.*, 1993; Gottlicher *et al.*, 1992). Despite this, none of these PPAR-activating chemicals has been shown to bind it as a ligand would bind a receptor. Only recently, has it been shown that the arachidonic acid-derived prostaglandin J metabolite, PGJ_2, is a PPARγ ligand (Kliewer *et al.*, 1995; Forman *et al.*, 1995). This is the first ever demonstration of an endogenous ligand of PPAR of any kind. This follows the observation that an exogenous compound, the antidiabetic thiazolinediones, activate and competitively bind PPARγ with high affinity (Lehmann *et al.*, 1995). Interestingly, PPARγ mediates adipocyte differentiation, a process which is shown to be enhanced by thiazolinediones (Kliewer *et al.*, 1995; Forman *et al.*, 1995) and prostaglandins (Russell and Ho, 1976). These are key discoveries as this shows that PPARs are not only transcriptional mediators of the induction of various genes during peroxisome proliferation, they possibly play an important

physiological role in fatty acid and glucose metabolism as well.

Section 1.3.5 Peroxisome Proliferator Response Elements (PPREs)

Several nuclear hormone receptors possess the same P-box sequence as the PPARs and most of the response elements of this class consist of direct repeats AGGTCA (or TGACCT) motifs with a variable number of intervening nucleotides. These 'spacer elements' play an essential role in selective hormonal response. The response elements that have been characterized in PPAR target genes, termed peroxisome proliferator response elements (PPREs) are DR1, or direct repeats with one spacer nucleotide. It is also well established that RXRs (retinoid X receptors) act as a preferential partner in a functional PPAR-RXR heterodimeric receptor complex. The importance of RXR in this complex was shown in studies demonstrating that PPAR and RXR act synergistically to confer peroxisome proliferator responsiveness on a DR1 driven promoter (Kliewer *et al.*, 1992), and since then, in the promoter regions of currently identified PPAR target genes. PPREs have been identified in the genes for rat acyl coA oxidase (Osumi *et al.*, 1991), rabbit CYP4A6 (Muerhoff *et al.*, 1992a), the liver fatty acid binding protein , L-FABP (Isseman *et al.*, 1992), rat CYP4A1 (Aldridge *et al.*, 1995), the multifunctional enzyme (Zhang *et al.*, 1992; Bardot *et al.*, 1993) and malic enzyme (Castelein *et al.*, 1994). Mitochondrial enzymes like the rat HMG coA synthase (Rodriguez *et al.*, 1994), which is involved in ketogenesis, and the medium-chain acyl coA dehydrogenase (MCAD) (Gulick *et al.*, 1994) are also found to possess PPREs in the proximal promoter region. Another extrahepatic gene in which PPREs have been identified is the adipose tissue-specific aP2 gene (Tontonoz *et al.*, 1994).

Peroxisome proliferators induce the expression of CYP4A proteins (Gibson *et al.*, 1982; Sharma *et al.*, 1988), mitochondrial b-oxidation enzymes and the rate of ketogenesis (Mannaerts *et al.*, 1978) during peroxisomal proliferation, together with cytosolic proteins such as the liver FABP (Kaikaus *et al.*, 1993; Isseman *et al.*, 1992), aP2 (Tontonoz *et al.*, 1994) and the malic enzyme

(Hertz *et al.*, 1991). All these genes possess a response element which is recognized by PPARs. The rabbit CYP4A PPRE has been fairly well studied. Two distinct elements, named Z and X have been identified as functional elements in the promoter of the gene (Muerhoff *et al.*, 1992a). The Z element has been shown to mediate the major portion of the response to peroxisome proliferators by CYP4A6 reporter constructs in transient expression assays, binding PPAR-RXR much more efficiently than the other CYP4A6 PPREs. However, further work demonstrated that the typical PPRE motifs may not be sufficient to ensure a peroxisome proliferative response (Palmer *et al.*, 1994). Deletions or mutations within 6 nucleotides 5' of the DR1 dramatically diminished PPAR-RXR binding. PPAR-RXR binding, therefore, requires sequences immediately 5' of the DR1, which are conserved in natural PPREs and promote binding of PPAR-RXR heterodimers in preference to potential competitors such as ARP1 and RXRα homodimers. The identification of PPRE as recognition elements of PPARs is an important step in understanding the mechanism of transcriptional activation during peroxisome prolifereation. However, there are major limitations in these studies that complicate our understanding of the toxicological aspect of this phenomenon. For example, in the study of the CYP4A1 PPRE (Aldridge *et al.*, 1995), the Hepa1c1c7 cell line was used in transient transfection assays. These cells are known not to undergo peroxisome proliferation when treated with peroxisome proliferators. As such, it is not possible to extend in vitro results to actual events in vivo. Cells which undergo peroxisome proliferation, like liver hepatocytes, will give us a better understanding of the phenomenon. Secondly, the PPRE caused a 3-5 fold induction, but failed to cause any induction with the endogenous promoter of the gene, demonstrating that the in vivo response seen in the liver is more complex than that of the model system. Therefore, it is difficult to say that PPAR interacts with a DNA element in an artificial system and thus regulates peroxisome proliferation-induced transcription in the liver. This clearly remains to be shown unequivocally in a hepatocyte system with an endogenous promoter upstream of a reporter gene.

might therefore have promoted spontaneously initiated lesions. However, it is possible that older animals might be more susceptible to peroxisome proliferator-induced liver tumours due to reduced protective enzymes anyway. Nevertheless, the conflict of much of the data regarding the carcinogenicity of peroxisome proliferators means that it is difficult to classify carcinogenic chemicals into 'initiators' or 'promoters' simply by observations made with a single dose of a chemical. It is thus necessary to explore a whole range of issues from the cellular to the biochemical and molecular, regarding carcinogenicity so that a better understanding of the responses of cells to these chemicals can be achieved.

Section 1.3.6 Peroxisome Proliferation : Its Significance for Humans

Many attempts have been made to reproduce in other species the phenomenon of peroxisome proliferation that is readily observed in rats and mice. However, the available experimental data show that this phenemenon is restricted to these animals and that humans are clearly less responsive to peroxisome proliferators, as are dogs, marmosets, rhesus monkeys, guinea pigs and hamsters at dose levels that produce a marked response in rats and mice (reviewed in Bentley *et al.*, 1993). Currently, what little is known about the mechanisms of carcinogenicity of peroxisome proliferators suggests that the risk of these chemicals to humans is very low.

Section 1.4 Conclusion

The CYP4A subfamily of genes is believed to play a role in lipid metabolism and the regulation of renal function. Additionally, genes in this subfamily are induced at an early stage, a crucial event, during the process of peroxisome proliferation in mice and rats. While much effort has been directed towards the study of these genes in the rat, relatively very little is understood about them in the mouse. To date, only two cDNAs, *Cyp4a10* and *Cyp4a12*, of the *Cyp4a* subfamily of genes have been cloned from the mouse. It is not known if other members of the murine *Cyp4a* subfamily exist.

This is an attempt to isolate genomic clones of murine members of the *Cyp4a* subfamily. This study will be important for several reasons. The genetics of the mouse are well understood and powerful techniques like transgenic and 'knockout' technology are most well characterized using a mouse model. A detailed understanding of the murine *Cyp4a* subfamily would be crucial for understanding their evolution and regulation in mice and as a prelude to designing a strategy for gene-targetting experiments. This would be a powerful way of understanding the function of the *Cyp4a* gene subfamily in lipid and eicosanoid metabolism and their role in peroxisome proliferation.

Chapter 2 MATERIALS AND METHODS

Section 2.1 Materials

Section 2.1.1 General Materials

All chemicals were obtained at the highest grade possible. Polyethylene Glycol was obtained from Sigma and was Molecular Biology grade and had molecular weight of 8000. Hybond-N+ membranes were obtained from Amersham. X-ray film was manufactured by Fuji and obtained from Amersham. Methylclofenapate (synthesized by Lancaster Synthesis Ltd., Lancaster, UK) was a kind gift of Dr CR Elcombe (Zeneca CTL, Macclesfield, Cheshire). Oligonucleotides were synthesized by John Keyte of the Biopolymer Synthesis and Analysis Unit, Department of Biochemistry, University of Nottingham. UHP grade water (>13MΩ/cm) was produced by using a Purite Select Bio system. DEPC water was produced by adding 1ml of Diethyl Pyrocarbonate (dissolved in 5ml ethanol) to 1litre of UHP water. This was then stood for 1 hour and then autoclaved for 1 hour. PCR PromoterFinder Kit was manufactured by Clontech Laboratories and obtained from Cambridge Bioscience. Sequagel concentrate and diluent were manufactured by National Diagnostics and obtained from Flowgen. Sequenase Version 2.0 T7 Polymerase and reagents were manufactured by United States Biochemicals and obtained from Amersham. The following radioactive isotopes were obtained from ICN Flow at the stated specific activities:

(i) transcription grade [α-^{32}P]-CTP (3000Ci/mmol; 0.5mCi), (ii) endlabelling grade [γ-^{32}P]-ATP (7000Ci/mmol; 22.9mM) and (iii) sequencing grade [α-^{35}S]-dATP (6000Ci/mmol; 0.25mCi).

Section 2.1.2 Plasmids

λ genomic clones 4 and 16 are gifts of Dr David Bell (Department of Life Science, Nottingham University), which were initially isolated from a genomic library using riboprobes transcribed

from *Cyp4a10* and *Cyp4a12* partial cDNA clones, originally described in Bell *et al.* (1993). The rat CYP4A1 cDNA was a kind gift of Professor Gordon Gibson (Surrey University; Earnshaw *et al.*, 1988) from which was derived a *Cla* I/*Eco* RI fragment termed pIV2 (D.R. Bell). This was the 3' most cDNA fragment, corresponding to nucleotide position 900 to the end (exon 8 to the end). pZErO Zero Background/Kan Cloning Kit was obtained from Invitrogen. pGEM-T t-tailed vector kit was obtained from Promega.

Section 2.1.3 Animals

Male C57Bl6 mice, 25-30g each, were obtained from the Queens Medical Centre, University of Nottingham. Mice were fed a standard laboratory diet (Harlan Teklad) ad libitum. Groups of two mice were treated by intraperitoneal injection with either corn oil vehicle control or methylclofenapate (MCP) (25 mg/kg) daily for 3 days.

Section 2.1.4 Bacterial Genotypes

The E.coli strains below had the following genotypic traits and were obtained from Stratagene:

XL1-Blue {recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIqZDM15 Tn10 (Tetr) ; XL1-Blue MRA {D(mcrA) 183, D(mcrCB-hsdSMR-mrr)173, endA1, supE44, thi-1, gyrA96, relA1, lac};

XL1-Blue SURE {e14- (McrA-) D (mcrCB-hsdSMR-mrr) 171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC :: Tn5 (Kanr) uvrC [F' proAB lacIqZ DM15 Tn10 (Tet^r)};

XL1-Blue Top10F' had the following trait and was obtained from Invitrogen's Zero Background Cloning Kit:

XL1- Blue Top 10F' {F' [lacIq Tet^r] *mcr*A D(mrr-hsdRMS-mcrBC) f80lacZDM15 DlacX74 deoR recA1 araD139 D(ara-leu)7697 galU galK rpsL endA1 nupG}

ElectroMAX DH10B cells had the following trait and were obtained from Gibco BRL:

DH10B {F- mcrA D (mrr -hsdRMS-mcrBC), 80d lacZDM15, D lacX74, deoR, recA1, endA1 araD139, D(ara leu)7697, galU, galK, l-,rpsL, nupG}

Section 2.2 Methods

Section 2.2.1 General DNA Techniques

Section 2.2.1.1 Restriction Digest

DNA (5µg/µl) was digested with enzymes according to Sambrook *et al* (1988) in a 50µl reaction volume. The following were placed in a clean 1.5ml tube in this order :

10X restriction enzyme digestion buffer	X µl	⋆
DNA	5 µl	
dH_20	Y µl	
Restriction enzyme (2U)	2 µl, to make a total of 50µl.	

The mixture was then mixed by vortexing, and centrifuged in a Heraeus Benchtop Centrifuge for 10 seconds, then incubated at 37 ^oC for 45 minutes. The reaction was then stopped by adding a quarter volume of 7.5 M ammonium acetate (12.5µl) and then 2 volumes of ethanol (125µl) and placed on ice for 30 minutes. This precipitates the DNA. The tube was then centrifuged at 15000rpm for 5 minutes and the supernatant removed by gentle pipetting. The pellet was then washed with 1ml of cold 70% ethanol and centrifuged again for 2 minutes. The ethanol was then removed and the pelleted DNA was then air-dried and resuspended in 5µl of clean UHP water.

⋆

Restriction Enzyme Buffer System	**Working Concentration**
Bam HI 10X Universal Buffer	1 X
Eco RI 10X Universal Buffer	1.5X
Not I 10X Buffer 6	1 X
Sac I 10X Buffer 2	1 X

Components of Buffer Systems

10X Universal Buffer (Stratagene)

1M KAc
250mM Tris-Acetate pH 7.6
100mM MgAc
5mM β-mercaptoethanol

10X Buffer System 2 (NBL)

33mM Tris-Acetate pH 8.2
66mM KAc
10mM MgAc
0.5mM DTT 100mg/ml BSA

10X Buffer System 6 (NBL)

50mM Tris-HCl pH 7.8
100mM NaCl
10mM $MgCl_2$
1mM DTT

Section 2.2.1.2 Ligation

Ligation is the process of the formation of phosphodiester bonds in the presence of ATP between double- stranded DNAs with 3'-hydroxyl and 5' phosphate termini. Vector DNA and insert DNA were both appropriately restriction digested, and gel-purified and resuspended in an appropriate volume of clean UHP water. The following were added in a 1.5ml reaction tube:

1μg of vector	1 μl	
3μg of insert DNA	X μl (Where X can be varied)	
clean UHP H_2O	Y μl	To give a total of 8.5 μl

5X T4 Ligase Buffer	1 µl
T4 ligase (Gibco BRL)	0.5 µl, making a total reaction volume of 10µl.

The tube was then gently vortexed and touch-centrifuged in a Heraeus Benchtop Centrifuge to collect the contents. The tube was then incubated at 16°C overnight for the ligation to proceed.

Components of Buffer System

5X T4 Ligase Buffer (Gibco BRL)

250mM Tris-HCl, pH 7.6

50mM $MgCl_2$

5mM ATP

5mM DTT

25% (w/v) polyethylene glycol-8000

Section 2.2.1.3 GeneClean

The GENECLEAN II kit (BIO 101 Inc.) contains a specially formulated silica matrix called 'glassmilk' which binds single and double stranded DNA without binding DNA contaminants. This kit was used to purify DNA larger than 500bp from agarose gels made with 1X TAE. The reagents NaI solution (6.0M) and NEW wash (NaCl, ethanol, water) are provided in the kit. DNA was visualized on an ethidium bromide (0.5mg/ml) stained agarose gel made with 1X TAE, and gel- excised with a razor blade. The agarose slice was then weighed in a 1.5ml tube if the slice weighed less than 0.4g. 3 volumes of NaI stock solution were then added to the tube containing the agarose slice, where 1g is approximately 1ml. The tube was then incubated in a waterbath at a temperature of 45 °C-55 °C for about 5 minutes, with mixing after every minute. The gel slice should be completely dissolved. 5µl of 'glassmilk' was then added to approximately 5µg of DNA in solution and the whole suspension was then incubated at room temperature for

about 10 minutes to allow binding of the DNA to the silica matrix, with mixing every 1-2 minutes. The 'glassmilk' was then pelleted by centrifugation in a Heraeus benchtop centrifuge at 15000rpm for approximately 5 seconds. The NaI supernatant was then removed with a pipette. The pellet was then washed 3 times with NEW wash. To do this, 500µl of ice-cold NEW wash was pipetted into the tube, resuspending the 'glassmilk'. After resuspension, the matrix was pelleted by centrifugation for 5 seconds and the wash repeated twice more. The washed pellet was then resuspended with 10µl UHP water to elute the bound DNA from the 'glassmilk' by incubation of the tube between 45 °C-55 °C for 2-3 minutes. The tube was then centrifuged again for about 30 seconds to make a solid pellet. The supernatant, which contains the DNA, was then removed and placed in a clean tube for further manipulations.

Section 2.2.1.4 Southern Blotting

DNA transfer to Hybond N+ membranes (Amersham) was performed according to the Amersham Handbook. DNA was run on a 0.6% agarose/TBE gel and photographed against a fluorescent ruler. The gel was then placed in a container of denaturation buffer for 30 minutes at room temperature (or until 15 minutes after the dye colours have returned to normal, whichever is longer). The gel was then rinsed in dH_20 and transferred to a container of neutralization buffer for 15 minutes and then to another container of fresh solution for another 15 minutes. To set up a Southern Blot, a Tupperware box was filled with blotting buffer [10X SSC] to about 1.5cm deep. A platform (inverted Eppendorf racks) was then prepared by covering it with a wick made from 1 sheet of Whatman 3MM filter paper, cut to size, and saturated with 2X SSC. The gel was then placed on the wick, wells facing down, making sure no air bubbles were trapped underneath. The gel was then surrounded with Nescofilm to prevent blotting buffer being directly absorbed into the paper towels above. A sheet of Hybond-N+ membrane was then cut to exact size, wet with 2X SSC and placed on top of the gel, again making sure no air bubbles are trapped. 3 sheets of 3MM paper were then cut to size, saturated with 2X SSC and placed on top of the

membrane. A layer of nappy (Boots), was then placed on top of the 3MM paper. A glass plate was then placed on top of the nappy and a 1kg weight (1 litre of water) was then placed on top. The transfer was then allowed to proceed overnight. After blotting, the membrane was then marked where the wells were with pencil to allow later identification of the tracks. The membrane was then washed briefly and carefully in 2X SSC to remove any adhering agarose. The DNA was then air-dried and fixed by baking, DNA-side up, in an 80°C oven for 2 hours.

Components of Buffers

Denaturation Buffer

1.5M (w/v) NaCl

0.5M (w/v) NaOH

Blotting Buffer (10X SSC)

1.5M (w/v) NaCl

0.15M (w/v) Na_3Citrate pH 7.2

Neutralization Buffer

1.5M (w/v) NaCl

0.5M Tris-HCl pH 7.2

0.001M (w/v) EDTA

Section 2.2.1.5 DE-81 test

DE-81 filters (Whatman) are positively charged and strongly bind and retain nucleic acids, including oligonucleotides. Unincorporated nucleotides stick less tightly to the filters and are selectively removed by washing the filter extensively in 0.5M Na_2HPO_4. This property is used to calculate the efficiency of incorporation of a radioactive precursor into a desired product. For every DE-81 test, 2 pieces of DE- 81 filters were prepared by labelling using a soft-lead pencil, one to measure the total radioactivity in the reaction (ie, the 'unwashed' filter) and one to measure the incorporated activity (ie, the 'washed' filter). 1µl of the sample was spotted on each of the 2 filters and left to dry. Using forceps, the 'washed' filter was then swirled in 5ml of 0.5M

Na_2HPO4 for 5 minutes. This was then repeated 4 more times with fresh buffer. The filter was then washed in 5ml of DEPC water for 1 minute and then in 5ml 70% ethanol for another 1 minute. The filter was then left to dry at room temperature. Both the 'washed' and 'unwashed' filters were then measured for radioactivity with a geiger counter or in a scintillation counter. The amount of radioactivity on both filters were then compared and the proportion of precursor that has been incorporated can be calculated :

$$\frac{\text{cps in washed filter}}{\text{cps in unwashed filter}} = \text{\% incorporation}$$

Section 2.2.1.6 DiethylAminoEthyl Cellulose (DEAE) Column

Radiolabelled oligonucleotides were separated from unincorporated radiolabel on a DE52 column. Before preparing an ion-exchange column, pre-swollen DE52 Anion Exchange cellulose (Whatman) was first prepared in slurry form. 5g of dry resin was weighed out and suspended in 100ml of 1M NaCl/TE buffer [1.0M NaCl, 10mM Tris-HCl, pH 8.0, 1mM EDTA], and stirred for 2-3 minutes. The slurry was then allowed to settle and the supernatant was then decanted off. The resin was then redispersed in 100ml 0.1M NaCl/TE buffer [0.1M NaCl, 10mM Tris-HCl, pH 8.0, 1mM EDTA] and the slurry allowed to settle again. After decanting off the supernatant again, the slurry was then redispersed in 30ml of 0.1M NaCl/TE buffer. To prepare a column, 2.5ml-3ml of slurry was poured in a 1ml gilson pipette-tip using siliconized glass wool as support. The column was not allowed to dry up and was prepared just before use.

Section 2.2.1.7 Autoradiography

Gels containing ^{35}S were dried on 3MM paper (Whatman) using a Bio-Rad 583 Gel Drier. Gels containing ^{32}P were Southern transferred to Hybond N+ membrane and wrapped in Saran Wrap to prevent contamination of intensifying screens and film cassettes. The film was exposed at

varying lengths of time at -80°C to maximize the signal. Just before developing the film, the cassette was removed from the -80°C freezer and placed at room temperature for 5 minutes to thaw. The film was then developed in a dark room by the following procedure. The film was washed in 1x X-ray developer (Kodak LX24) for 5 minutes. The film was then washed in a water bath for 2 minutes, then washed in 1X Rapid Fixer (Ilford Hypam Fixer) for 5 minutes before returning it to the water bath for a further minute before the lights were switched on. The film was then rinsed under a running tap for a further 5 minutes before being dried and aligned.

Section 2.2.2 Library Screening

Section 2.2.2.1 Phage Titreing

A mouse B6/CBA [C57Blck6 x CBA] genomic library constructed in λ FIX II vector (Stratagene) was used. 10ml LB broth supplemented with 0.2% maltose and 10mM $MgSO_4$ in a sterile flask was inoculated with a single colony of host strain XL1-Blue MRA and grown overnight with shaking at 30°C. The cells were then spun down in a sterile tube for 10 minutes at 2000rpm in a Centaur Centrifuge. The media was then decanted off and the cell pellet gently resuspended in 10ml of 10mM $MgSO_4$ without vortexing. The cells were then diluted down to OD_{600} = 0.5 with 10mM $MgSO_4$. 1µl of phage was then added to 200µl of diluted host cells and incubated at 37°C for 15 minutes to allow for phage adsorption in a sterilin tube. 1µl of a 1:10, 1:20, 1:100 and 1:200 dilution of phage was also used for titering. 3.5 ml of top agar (48°C) was then added to the mixture, quickly swirled and poured evenly on a 100mm LB plate. The plates were then incubated overnight at 37°C. The plaques were then counted and the plaque-forming units per millilitre (pfu/ml) concentration of the library determined by the following formula:

$$\frac{\text{No of plaques x dilution factor}}{\text{Volume of extract plated}} = \text{pfu/ml}$$

Section 2.2.2.2 Library Plating

5ml LB broth supplemented with 0.2% maltose and 10mM $MgSO_4$ in a sterile flask was inoculated with a single colony of host strain XL1-Blue MRA and grown overnight with shaking at 30°C. The cells were then spun down in a sterile tube for 10 minutes at 2000rpm in a Beckman Centrifuge. The media was then decanted off and the cell pellet gently resuspended in 10ml of 10mM $MgSO_4$ without vortexing. The cells were then diluted down to OD_{600} = 0.5 with 10mM $MgSO_4$. The cells may be stored for 2-3 days at 4°C. The library was plated on large 150mm LB plates to 50 000 pfu/plate. 20 plates were used to screen 1 X 10^6 pfu. The appropriately diluted phage and 600µl of bacteria were incubated for 15 minutes at 37°C to allow phage adsorption in a sterilin tube. To the mixture was then added 7ml top agar (48°C), quickly swirled and poured onto a plate. The top agar was then allowed to set and the plate incubated at 37°C overnight without inverting and no more than 2 high. Agar plates were then chilled at 4°C the next day for 2 hours to prevent the agar sticking to the membranes during phage lifts.

Section 2.2.2.3 Phage Lifts

Phage lifts were done on Amersham Hybond-N+ membranes as described in the Amersham Hybond N+ Handbook. Each membrane was first labelled with a pencil before placing it carefully on the agar surface. The membrane and agar were then marked asymmetrically using a sterile needle to ensure correct orientation of plaques. After 1 minute, the membrane was then removed with a pair of forceps and placed, plaque side up, on a pad of 3MM paper (Whatman) soaked in Denaturing Solution and left for 7 minutes. A duplicate membrane was prepared in the same way, and marked through the holes used to align the first membrane. After denaturing, the membrane was then placed, plaque side up, on a pad of 3MM paper (Whatman) soaked in Neutralizing Solution and left for 3 minutes. The neutralization was repeated with a fresh pad soaked in the same solution. The membrane was then washed in 2X SSC/0.1% SDS and placed on dry paper towels to air dry, plaque side up. The DNA was then fixed on the membranes by baking

at 80°C for 2 hours.

Composition of Solutions

Denaturing Solution	*Neutralizing Solution*
1.5M NaCl	1.5M NaCl,
0.5M NaOH	0.5M Tris-HCl pH 7.2
	0.001M EDTA

Section 2.2.2.4 Hybridization

Under stringent conditions, prehybridization was carried out for about 2 hours and hybridization was performed overnight at 65°C, in Hybridization Buffer in a Techne HB-1D Hybridiser, colony-side towards the inside of the hybridization tubes. Thereafter, membranes were rinsed in 2X SSC/0.1% SDS at room temperature and then washed 3 times in Wash Buffer at 65°C for 15 minutes each. Membranes were then exposed to autoradiographic film at -80°C for varying amounts of time to optimize the signal.

Components of Buffers

Hybridization Buffer	*Wash Buffer*
*0.5M Sodium Phosphate buffer pH7.2	40mM sodium phosphate
1% bovine serum albumin	0.1% SDS
7% SDS	
1mM EDTA	

*NB. To prepare 0.1M Sodium Phosphate buffer pH7.2, 68.4ml of 1M Na_2HPO_4 and 31.6ml of 1M NaH_2PO_4 were mixed together and diluted to 1000ml with UHP water. To prepare 0.5M solution, the phosphate mix was diluted to 200ml with UHP water.

Section 2.2.2.5 Secondary Screening

To orient the filters, the film was lined up and where the needle poked through the membrane. The 'putative' clones were determined with help of the duplicate membranes. The 'plug' of agar was then removed with an inverted yellow Gilson pipette tip where the clone lined up with the film spot. The plug was then placed in 500µl of SM buffer and 50µl of chloroform and left at room temperature for 1 hour. The candidate phage was then diluted appropriately and titred with host cells on a 100mm LB plate so that one plate had approximately 50 plaques. Plaque lifts were then performed and hybridization performed as before. Isolated plaques were then picked from a secondary screen. Where positive plaques were too close to the background plaques to be distinguishable, a tertiary screen was performed until single, isolated plaques were isolated.

Components of Buffer

SM Buffer

5.8g/*l* NaCl

2.0g/*l* $MgSO_4.H_20$

50.0ml/*l* of 1M Tris-HCl pH 7.5

5.0ml/*l* 2% (w/v) gelatin

Section 2.2.3 Oligonucleotide Hybridization

The synthetic oligonucleotide was first end-labelled and purified as described (see 2.4.2). Prehybridization was carried out for about 2 hours and hybridization was performed overnight at 60°C, in Hybridization Buffer in a Techne HB-1D Hybridiser. Hybridization was carried out according to the method of Dilella and Woo (1987). The membrane was then rinsed in 2X SSC at room temperature, then washed in TMACl Wash Buffer at 60°C. The wash was then repeated twice more before autoradiography.

Components of Buffers

Hybridization Buffer

0.5M Sodium Phosphate Buffer pH 7.2

1% BSA

7% SDS

1mM EDTA

2X SSC

0.3M NaCl

30mM NaCitrate, pH 7.2

TMACl Wash Buffer

3M Tetramethylammonium Chloride (TMACl)

2mM EDTA pH 8.0

50mM Tris-HCl

Oligonucleotides used

5'-CCA AAG GCT TCT GGA ATT TAT-3', spanning a 3 residue deletion in CYP4A2 after amino acid 113 in exon 3, but not CYP4A3, was used as a probe to screen clones containing exon 3.

Section 2.2.4 Probe Preparation

Section 2.2.4.1 Transcribed RNA Probe

The appropriate subclone was first linearized by restriction digest and checked on an agarose gel. The DNA was then ethanol precipitated and resuspended in DEPC water to 1μg/μl concentration. The probe was then transcribed using either T7 or SP6 RNA polymerase and [α-32P] CTP according to the protocol (Lambda FIX II Library, Stratagene). The following were incubated in a 1.5ml tube at 37°C for 45 minutes:

Linearized DNA	1 μl
5X TCS Transciption Buffer	4 μl

Cold ATP/GTP/TTP mix [10mM each]	1 µl
RNAse Inhibitor	1 µl [40 Units]
DTT [0.75M]	1 µl
[α^{32}P] CTP [400-800 Ci/mmol, 10mCi/ml]	5 µl
T7 or SP6 Polymerase	1 µl [10 units]
DEPC water	6 µl

A DE-81 test was then performed to determine the efficiency of incorporation. A quarter volume of 7.5M ammonium acetate was then added to the reaction, followed by 2 volumes of ethanol to precipitate the probe for 45 minutes at -20°C. The contents were then centrifuged at 15000rpm in a Heraeus Benchtop Centrifuge for 15 minutes. 500µl of cold 70% ethanol was then added to the tube and centrifuged again for 5 minutes to wash and pellet the probe. The pellet of RNA probe was then air-dried briefly and resuspended in 500µl of DEPC water.

Components of Buffer

10X TCS

200mM Tris-HCl, pH 8.0,

40mM $MgCl_2$

10mM spermidine

250mM NaCl

Section 2.2.4.2 Oligonucleotide Probe Endlabelling

T4 Polynucleotide Kinase catalyzes the transfer of the γ-phosphate from ATP to the 5' terminus of polynucleotides or mononucleotides. The labelling of oligonucleotide was carried out in a 1.5ml tube with the following contents:

Oligonucleotide (30pmol/µl)	1 µl

10X T_4 Kinase Buffer	2 µl
[γ-^{32}P]ATP (7000mCi/mmol; 0.17mCi/ml)	2 µl
H_20	14 µl
T_4 Polynucleotide Kinase	1 µl

The contents were mixed and incubated at 37°C for 45 minutes. The efficiency of transfer of ^{32}P was then determined by DE-81 test. To separate the radiolabelled oligonucleotide from the unincorporated radioactivity, a DEAE ion-exchange column was used (see section 2.1.6). 100µl of 0.1M NaCl/TE [0.1M NaCl, 10mM Tris-HCl, pH 8.0, 1mM EDTA] was added to the end-labelling reaction, and then pipetted back up and loaded onto a DE52 column. 200µl aliquots of 0.1M NaCl/TE were passed through until the radioactivity flowing out through the column was less than 100 cps per 200µl fraction. The radiolabelled oligonucleotide bound to the column was then eluted in 100µl fractions by passing 1M NaCl/TE [1.0M NaCl, 10mM Tris-HCl, pH 8.0, 1mM EDTA] through the column, with the eluate monitored constantly with a geiger counter. The fractions containing the radiolabelled oligonucleotide were then pooled and tested for incorporation by DE-81 test.

Components of Buffer

10X T4 Kinase Buffer

700mM Tris-HCl, pH 7.6

100mM $MgCl_2$

50mM DTT

Section 2.2.5 Isolation of Phage DNA

Section 2.2.5.1 Maxi-Prep By Liquid Culture

5ml of fresh overnight *E. coli* XL1-Blue MRA host cells grown in LB containing 10mM $MgSO_4$

and 0.2% maltose, were mixed with various amounts of phage and incubated at room temperature for 5 minutes. They were then added to prewarmed 250ml of LB containing 10mM $MgSO_4$ and 0.2% maltose and incubated in an orbital shaker at 37°C. Lysis was achieved after approximately 8 hours, and sodium chloride was then added to a concentration of 0.5M and together with 1ml of chloroform and returned to the shaker for another 5 minutes. The culture was then centrifuged in a Beckman centrifuge using a JA20 rotor at 6500rpm for 10 minutes at 4°C. The supernatant was then treated with 250µl of DNAse I (1mg/ml stock) and 175µl RNAseA (10mg/ml stock) and incubated for 1 hour at 37°C. Polyethylene glycol (PEG) was then added to a concentration of 10% (w/v) and stirred to dissolve and incubated in a bucket of ice water for 1 hour. The culture was then centrifuged as above to precipitate phage. The pellet was then resuspended in 5ml SM buffer and 500µl chloroform. The suspension was then extracted with an equal volume of chloroform using a Centaur centrifuge at 2500rpm for 10 minutes. 50µl of 10% SDS and 50µl 0.5M EDTA (pH 8.0) were then added and incubated at 68°C for 15 minutes. The suspension was then phenol extracted, phenol/chloroform extracted and then chloroform extracted. An equal volume of isopropanol was then added to precipitate the DNA at room temperature for 30 minutes. The suspension was then microfuged for 5 minutes at 15000rpm to pellet the DNA. The pellet was then washed with 70% ethanol and air dried. The DNA was then resuspended in 300µl of water.

Composition of Buffer

SM Buffer

5.8g/*l* NaCl

2.0g/*l* $MgSO_4.H_20$

50.0ml/*l* of 1M Tris-HCl, pH 7.5

5.0ml/*l* 2% (w/v) gelatin

Section 2.2.5.2 Mini-Prep By Plate Lysate

E.coli MRA host cells were grown overnight at 37°C in 5ml of LB broth containing 0.2% maltose and 10mM $MgSO_4$. 1.5ml of the fresh cells were then pelleted by centrifugation at 15000rpm in a benchtop centrifuge for 30 seconds. The supernatant was then removed and the pellet gently resuspended in 0.5ml of 10mM $MgSO_4$. The cells were then diluted to OD_{600} = 0.5 with 10mM $MgSO_4$. Meantime, the phage were serially diluted in SM buffer. 1µl of various dilutions of phage were incubated with 200µl of cells in clean sterilin tubes at 37°C for 15 minutes to allow for phage attachment. To the phage-cell mixture, 3.5ml of top agar (cooled to 48°C) were then added, quickly swirled to mix and poured evenly on 10mm LB agar plates. The top agar was then allowed to dry and the plates incubated, uninverted and no more than 2 high, overnight at 37°C. A cells-only control was poured and incubated in the same way. After lysis, the plates were then flooded with 5ml of SM buffer and stored at 4°C for several hours with gentle shaking. The phage were then harvested from the plate with a pipette and placed in a Beckman JA-21 centrifuge tube. An extra 1ml of fresh SM buffer was added to the plate and stored an extra 15 minutes in a tilted position to allow the fluid to drain. This was then drained and combined with the first harvest. Sodium chloride was then added to a concentration of 0.5M and together with 0.1ml of chloroform and incubated in a shaker for 5 minutes. The culture was then centrifuged in a Beckman centrifuge using a JA- 21 rotor at 6500rpm for 10 minutes at 4°C. The supernatant was then treated with 5µl of DNAse I (1mg/ml stock) and 5µl RNAseA (10mg/ml stock) and incubated for 1 hour at 37°C. 0.6g of polyethylene glycol (PEG) was then added and stirred to dissolve and incubated in a bucket of ice water for 1 hour. The culture was then centrifuged as above to precipitate phage. The pellet was then resuspended in 0.5ml SM buffer and 50µl chloroform. The suspension was then extracted with an equal volume of chloroform using a benchtop centrifuge at 15000rpm for 5 minutes. 5µl of 10% SDS and 5µl 0.5M EDTA (pH 8.0) were then added and incubated at 68°C for 15 minutes. The suspension was then phe-

nol extracted, phenol/chloroform extracted and then chloroform extracted. An equal volume of isopropanol was then added to precipitate the DNA at room temperature for 30 minutes. The suspension was then microfuged for 5 minutes at 15000rpm to pellet the DNA. The pellet was then washed with 70% ethanol and air dried. The DNA was then resuspended in 50µl of water and stored at 4°C.

Section 2.2.6 Subcloning

Section 2.2.6.1 Electrocompetent Cell Preparation

E. coli strains XL1-Blue, SURE and TOP10F' were cultured in 500ml of LB broth containing 10mg/ml of tetracycline, to an OD of 0.6. DH10B were cultured without any antibiotics. The cells were prepared for electroporation by centrifugation in 250ml Beckman buckets using a JA-14 rotor in a Beckman JA2-21 centrifuge at 6000 rpm for 15 minutes. The pellets were then washed and resuspended 4 times in ice-cold, sterile UHP water and 4 times in ice-cold 10% glycerol, before being resuspended in 1 ml 10% glycerol. The cells can then be used immediately or stored at -80°C for future usage in 120µl aliquots.

Composition of Media

LB Broth

10 g/*l* bacto-tryptone

5 g/*l* bacto-yeast extract

10 g/*l* NaCl

Section 2.2.6.2 Electrotransformation

40µl of electrocompetent cells were mixed with approximately 10ng of DNA ligated into either pGEM7Zf(+), pGEM-T or pZErO and placed in a chilled Biorad Genepulser cuvette (1mm path length) and electroporated at 1.8kV. The cells were then immediately suspended in 1ml of

LB broth and incubated at 37°C for 1 hour before plating out on selective media.

Section 2.2.6.3 Vector Selection

pGEM7Zf(+) and pGEM-T Selection

Cells were selected on LB agar plates containing 10 μg/ml tetracycline (where necessary) and 50 μg/ml ampicilllin to select for the vector. To allow blue/white selection, the plates were spread with 30μl X-gal (20mg/ml stock) and 40 μl IPTG (20 mg/ml stock) and allowed to dry. The plates were then spread with 50 μl of the cells and incubated at 37°C overnight.

pZErO Selection

By zero background selection, recombinant colonies grow on the selective media due to an insertional disruption of a lethal gene *ccd*B. The CcdB protein acts by poisoning bacterial DNA-gyrase (topoisomerase II), an essential enzyme that catalyzes the ATP-dependent supercoiling of DNA. This causes DNA breakage, activation of the SOS response, and cell death. The pZErO-2 vector features the *lac*Zα-*ccd*B fusion gene. Cells were selected on LB containing 50μg/ml kanamycin to select for the vector and 1mM IPTG to select for recombinants. The plates were spread with 50μl of the transformed TOP10F' cells and incubated at 37°C overnight.

Section 2.2.6.4 DNA Extraction

Alkaline Lysis

Candidate clones, white colonies as picked by the use of blue/white selection, or just colonies by zero background selection, were used to inoculate 5ml cultures of LB broth containing the appropriate antibiotics and incubated at 37°C overnight. DNA was extracted using the alkaline lysis miniprep method. 1.5ml samples were taken from corresponding 5ml cultures and spun at 12 000g for 30 seconds to pellet the bacteria. Pellets were then resuspended in 100 μl of Solution

I. 200µl of freshly prepared Solution II was added and mixed by inversion at room temperature. 150µl of ice-cold Solution III was then added. The mixtures were then inverted and mixed thoroughly, and the bacterial lysates allowed to precipitate on ice for 10 minutes. The insoluble fractions were then separated by centrifugation at 12 000g for 5 min and the supernatant transferred to a fresh tube. Samples were phenol/chloroform extracted twice and ethanol precipitated. Pellets were washed in 70% ethanol, allowed to dry and resuspended in 50µl of UHP water.

Components of Solutions

Solution I	*Solution II*	*Solution III*
50mM glucose	0.2M NaOH	600ml/*l* 5M potassium acetate
25mM Tris-Cl pH 8.0	1% SDS	115ml/*l* glacial acetic acid
10mM EDTA pH 8.0		

Qiagen Mini Prep Kit

DNA was prepared by a Qiagen column for high purity for automated sequencing on an ABI 373A Fluorescent Sequencer. 3ml of an overnight culture were pelleted at 12 000g for 30 seconds and then resuspended in 0.3ml of Buffer P1. 0.3ml of Buffer P2 was then added, mixed by inversion and then incubated at room temperature for 5 minutes. 0.3ml of Buffer P3 was then added, the tubed mixed by inversion and incubated on ice for 10 minutes. The bacterial lysate was pelleted by centrifugation for 15 minutes at 12 000g, keeping the supernatant for purifying through a Qiagen tip. Meantime, a Qiagen tip was equilibrated by allowing 1ml of buffer QBT to drain through the column. The supernatant from the centrifugation was then added and allowed to enter the column. The supernatant from the centrifugation was then put through the column, followed by washing of the column with 1ml Buffer QC for 4 times. The column was then eluted with 0.8ml of Buffer QF. The collected fraction was then precipitated in 0.7 vol-

umes (0.56ml) of propan-2-ol for 15 minutes at room temperature. The DNA was pelleted by centrifugation at 12 000g for 30 minutes, washed in 1ml cold 70% ethanol and then air-dried. The pure DNA sample was then resuspended in 20µl of clean UHP water.

Components of Qiagen Buffers

Buffer P1	*Buffer P2*	*Buffer P3*
50mM Tris-Cl	200mM NaOH	3 M potassium acetate (pH 5.5)
10mM EDTA (pH 8.0)	1% SDS 100 mg/ml RNAse A	

Buffer QBT	*Buffer QC*
750mM NaCl	1.0M NaCl
50mM MOPS	50mM MOPS
15% ethanol	15% ethanol, pH 7.0
0.15% Triton X-100, pH 7.0	

Buffer QF

1.25M NaCl

50mM Tris-Cl

15% ethanol, pH 8.5

Section 2.2.6.5 PCR Cloning

A PromoterFinder DNA Walking kit (Clontech Laboratories) was used to walk upstream in murine genomic DNA. The kit contains 5 libraries of uncloned, adaptor-ligated genomic DNA fragments prepared from mice. Primary PCR reaction was performed using a gene-specific primer (GSP1) and an adaptor- specific primer (AP1) in 0.5ml tubes, containing the following master mix: 37.7µl H_20, 5µl 10X *Tth* PCR reaction buffer, 1 µl dNTP [10mM each of dATP, dTTP, dGTP and dCTP], 2.2µl $Mg(OAc)_2$ [25mM], 1µl AP1 [10mM], 1µl GSP1 [10mM], 1µl

50X Advantage *Tth* Polymerase Mix (Clontech) and 1µl of each of the 5 DNA libraries to an appropriately labelled tube. The tube was then gently tapped to mix and briefly centrifuged to collect the contents. The reaction mix was then overlaid with 2 drops of mineral oil and capped firmly. The reaction was then cycled in a Perkin Elmer DNA Thermal Cycler 480 using the following parameters :

1) 7 cycles - 94°C for 25 seconds, 72°C for 4 minutes
2) 32 cycles - 94°C for 25 seconds, 67°C for 4 minutes
3) 67 °C for an additional 4 minutes.

8µl of the primary PCR reaction products were then run on a 1.5% agarose/TAE gel. For the secondary PCR reaction, DNA bands were gel-excised, 'geneclean'-ed and re-amplified using a second gene-specific primer (GSP2) and a second adaptor-specific primer (AP2) in 0.5µl tubes containing the same mix except to use AP2 and GSP2 primers and 1µl of gel-excised DNA. PCR was performed using the following parameters :

1) 5 cycles - 94°C for 25 seconds, 72°C for 4 minutes
2) 18-22 cycles - 94°C for 25 seconds, 67°C for 4 minutes
3) 67°C for an additional 4 minutes

5µl of the secondary PCR reaction products were then analyzed on a 1.2% (w/v) agarose/TBE gel.

Gene Specific Primer Sequences

Cyp4a14 Walk 1 GSP 1 [5'-GAG ATA CCA TCC AAG TAC CTT GTA G-3']

GSP 2 [5'-CCC ATG GTT AGT AGT TTC TGG ATC GAG-3']

Cyp4a14 Walk 2 GSP 1 [5'-GGA TAG CTA GGA AGG CCC TGA GGA TC-3']

GSP 2 [5'- GAG GTT GGA GGA AGG AAG GAA TAT G-3']

Cyp4a10 Walk 1 GSP 1 [5'-GCA GCA GGA GCA GAC CGA GCA CAG A-3']

GSP 2 [5'-GGC TTG GGC TTA GAG CAG AGA CAC TG-3']

Adaptor Primer Sequences AP 1 [5'-GTA ATA CGA CTC ACT ATA GGG C-3']

AP 2 [5'-ACT ATA GGG CAC GCG TGG T-3']

Section 2.2.7 RNA Extractions

RNA was extracted from liver tissues using the method of Cathala and coworkers (1983). 0.5-1.0g of liver, flash-frozen in liquid nitrogen, was placed in a clean flat-bottomed glass universal and homogenized for 30 seconds in 5ml of lysis buffer using a Silverstone homogenizer. The homogenate was then quickly added to 35ml of ice cold 4M LiCl and mixed by inversion and allowed to precipitate overnight at 4°C. The sample was then spun at 11 000g (9500rpm) for 90 minutes at 4°C in a JA-21 rotor. The supernatant was removed and the pellet resuspended in 5ml of SDS/TE and 5ml phenol/chloroform (1:1). The sample was then placed on ice for 45 minutes and vortexed for 20 seconds every 5 seconds. Samples were frozen at -80°C for 30 minutes, allowed to defrost, and then spun at 11 000g for 15 minutes at 4°C in a JA-21 rotor. The aqueous phase was removed and extracted with phenol/chloroform as above. The aqeous phase was removed and ethanol precipitated overnight at -20°C. This was spun at 11 000g for 15 minutes at 4°C in a JA-21 rotor, the pellet washed in 90% ethanol and air-dried and resuspended in 500µl of DEPC water. The concentration of the RNA in solution was then determined by reading the optical density at 260nm, and worked out according to the following parameter:

$$1\ OD_{260nm}\ \text{Unit} = 40\mu g/ml\ \text{RNA}$$

To check for integrity, RNA was run out on a 0.7% agarose gel containing 0.1% SDS and 0.5μg/ml of ethidium bromide for prestain.

Components of Buffers

Lysis Buffer	*SDS/TE*
5M guanidine thiocyanate	0.1% SDS
10mM EDTA	1mM EDTA
50mM Tris-Cl	10mM Tris-Cl
8% (v/v) 2-mercaptoethanol	

Section 2.2.8 Primer Extension

Radiolabelled synthetic oligonucleotides Primer 1 [5'-GAG ATA CCA TCC AAG TAC CTT GTA G-3'] and Primer 2 [5' -CTG AGC AAG AAG GCC CAT TGG AAG AAC-3'] were used in First Strand cDNA synthesis using SUPERSCRIPT II Reverse Transcriptase (Gibco BRL). The following were added into a 1.5ml tube:

Oligonucleotide (approximately 2 pmoles/μl)	1 μl	
RNA (3μg/μl)	10 μl	(= 30μg)
Sterile DEPC water	1 μl	

making a total reaction volume of 12μl. The mixture was heated to 70°C for 10 minutes and placed on ice briefly, then touch-centrifuged to collect the contents of the tube. To anneal, the tube was then incubated at 55°C for 20 minutes. For the reverse transcription reaction, the following were then added to the tube:

5X First Strand Buffer	4 μl
0.1M DTT	2 μl

10mM dNTP mix	1 µl	[10mM each dATP, dGTP, dCTP and dTTP]

The contents were then mixed gently and incubated at 42°C for 2 minutes. 1µl SUPER-SCRIPT II [200 units] was then added and the reaction mixed by gently pipetting up and down. The reaction was then incubated at 42°C for 50 minutes. The reaction was then inactivated by heating at 70°C for 15 minutes. 0.1 volume of 3M sodium acetate (pH 5.2)(2µl) and 2.5 volumes of ethanol (55µl) were then added to the tube to precipitate the cDNA at -20°C for 1 hour. The tube was then centrifuged in a Heraeus Benchtop Centrifuge for 15000 rpm for 10 minutes to pellet the cDNA. The supernatant was then removed and the pellet allowed to air-dry behind a perspex screen. The pellet was then resuspended in 6µl of clean water and 4µl of Load Buffer.

Components of Buffer

5X First Strand Buffer (Gibco BRL)

250mM Tris-HCl, pH 8.3

375mM KCl 15mM $MgCl_2$

Load Buffer

95% formamide

20mM EDTA

0.05% bromophenol blue

0.05% xylene cyanol FF

Section 2.2.9 Sequencing

Section 2.2.9.1 Chain Termination Sequencing

DNA was prepared by Qiagen column and adjusted to about 1µg/µl in concentration. 5µl of template DNA was first denatured in a tube containing 1µl GSP1 primer [5'-GAG ATA CCA TCC

AAG TAC CTT GTA G-3'] (0.2pmol/µl) and 1µl NaOH (1M) and incubated at 42°C for 10 minutes. To neutralize, 1µl HCl (1M) was added to the sample, together with 2µl of 5X Sequenase Reaction Buffer and 0.5µl DMSO. The sample was then incubated at 42°C for 5 minutes. For labelling, the following were added to the sample :

DTT	1 µl
1X Labelling Mix	2 µl
[α-^{35}S] dATP	0.5 µl
Sequenase Version 2.0 T7 Polymerase (freshly diluted 1 in 8 in enzyme dilution buffer)	2µl
DMSO	0.61µl

The sample was then incubated at 20°C for 5 minutes. Meantime, 4 tubes were prepared and properly labelled, containing 2.5µl of each dideoxynucleotide (8mM each dideoxynucleotide) and 0.28µl DMSO and preincubated at 42°C. 3.5µl of labelling reaction was then added to each tube containing ddNTP, and the incubation was continued for a further 5 minutes. 4µl of STOP solution were then added to each tube which can then be stored at -80°C. Before loading 4.5µl on a sequencing gel, tubes were boiled for 2-3 minutes.

Components of Buffers

5X Sequenase Reaction Buffer

200mM Tris-HCl, pH 7.5

100mM $MgCl_2$ 5mM DTT 250mM NaCl

Enzyme Dilution Buffer

10mM Tris-HCl, pH 7.5

0.5mg/ml BSA

5X Labelling Mix (For 1X ,dilute 1 in 5 in UHP water)

7.5mM each of dGTP, dCTP and dTTP

95% formamide,20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF

Section 2.2.9.2 Automated Fluorescent Sequencing

DNA was purified by a Qiagen column before sequencing. Automated cycle sequencing was performed by J. Keyte, Biochemistry Dept, Nottingham University, using a Perkin Elmer ABI 373A Fluorescent Sequencer and an ABI Cycle Sequencing Kit. Sequencing was performed using the universal M13 forward and reverse primers, T7 and SP6 primers and various custom-made primers (see table below) for oligonucleotide-directed sequencing.

Primer Name	Sequence (5' - 3')
F1	GCA CCA TCT GGA AGG TTA GG
F2	CTA ATC TAG GTG GTC CTC TG
F3	GCT TCT CCC ATT TAT TAT CCT GTG
F4	ATG TGG TGT CAG CTA CAG AG
F5	GCT ATG CTG TTG AGT GGG TG
F6	CCA TTA GCC AAA GAG CAG TG
F7	CAG CAG CTG AAG TCT CTG CC
F8	CCA GAA CCA CCT TCA CAT AGT C
F9	GCC TTC CTA GCT ATC CAT CT
F10	TGC TGG CAC ATC CCA CCC TGG
F11	GCC ATT CCC ATC CTC AAG TG
F12	CAT AAC CTG TGC CAG AGA CTG A
F13	GAG TTG GTT CAT TCA CCC CA
R1	GCA TAG TCA GCC ATT CCT AC
R2	GCT CCA CAT GAC TAA GGG GCA AAC
R3	GAC CAT TGC TGT GTC CCA GG
R4	GTG ACA CTT CAG CCA TCA ATG
R5	AGC CCT GAG AAC CCA CAT TGA
R6	CTT CTG TCT TAG TCC CAG CAG
R7	ATG GAG AGG GAC TCC AGA ATG A
R8	CCC TTG CTG ATG AGG TTA G
R9	AGC CTG TGA CAG ACT CTT TC

Table 2.1 List of primers used in automated cycle sequencing of *Cyp4a14* gene. Primer synthesis and automated sequencing were performed by J. Keyte, Biochemistry Dept, Nottingham University.

Section 2.2.10 Gel Electrophoresis

Primer Name	Sequence (5' - 3')
10f1	CGT TGT GCG AGT CAA GAG CAT TG
10f2	CTG TAC TTG GGT GCC CAT GTA
10f3	CAT CTA GGC ACT GCT GCT C
10r1	TAC ATG GGC ACC CAA GTA CAG
10r2	CCC ATG AGA CTG GCC TAC CT
10r3	GCG TGC AGT ACC CAG GTG TTG

Table 2.2 List of primers used in automated cycle sequencing of the 5' flanking region of *Cyp4a10*. Primer synthesis and automated fluorescent sequencing were performed by J. Keyte, Biochemistry Dept, Nottingham University.

Section 2.2.10.1 Agarose Gel Electrophoresis

DNA agarose electrophoresis was performed using a Pharmacia GNA-100 Gel Electrophoresis Kit or a Stratagene Gel Kit. The appropriate amount of agarose was dissolved in the appropriate amount of 1X TBE or 1X TAE buffer by boiling and allowed to cool to around 50°C, before the addition of ethidium bromide (10mg/ml stock) to a concentration of 0.5μg/ml. The gel mix was then poured into a gel casting mould containing the appropriate well-forming comb and allowed to set. After setting, the gel was then placed in a gel tank and topped up with the appropriate buffer until just covering the gel. Samples were then loaded in a 10% solution of 10X loading buffer. Electrophoresis was carried out at either 30V (overnight) or 160V (2hours) in a Stratagene gel kit, or 80V (2 hours) in a Pharmacia Mini-gel Kit.

Components of Buffers

10X TBE (1 litre)	*50X TAE (1 litre)*
107.8g Tris base	242g Tris base
55g Boric Acid to pH 8.3	57.1ml glacial acetic acid
7.44g Disodium EDTA, dihydrate	100ml 0.5M EDTA pH 8.0, and adjust to pH 7.2

10X Loading Buffer

30% glycerol

0.25% bromophenol blue

0.25% xylene cyanol

Section 2.2.10.2Sequencing Gel Electrophoresis

Sequencing was performed on a Bio-Rad Sequi-gen Cell Electrophoresis Kit. To make sequencing gel, spacers and electrophoresis glass plates were cleaned with 0.1% SDS and then with ethanol and dried with tissue paper. The heavier plate was then coated with 1.5ml of Sigmacote (Sigma) and wiped clean with tissue. The glass plates were then placed together with the spacers (0.4mm thick) in place and clamped with the side pieces with the electrodes on the same side of the plates as the lip. The following chemicals were used: Sequagel 19:1 Acrylamide/Bis-acrylamide Concentrate, Sequagel Diluent and 10X TBE Buffer. The percentage of acrylamide in the gel can be calculated from the following formulae:

1. $\text{Volume}_{conc} = (\text{Volume}_{total} \times (\%\ \text{acrylamide})) / 25$

2. $\text{Volume}_{10X\ TBE} = 0.1 (\text{Volume}_{total})$

3. $\text{Volume}_{diluent} = \text{Volume}_{total} - (\text{Volume}_{conc} + \text{Volume}_{10X\ TBE})$

where Volume_{conc} = Volume of sequagel concentrate, Volume_{total} = total volume of gel, $\text{Volume}_{10X\ TBE}$ = volume of 10X TBE and $\text{Volume}_{diluent}$ = volume of sequagel diluent.

To make a 6% gel, 24ml sequagel concentrate, 66ml sequagel diluent and 10ml 10X TBE buffer were mixed in a beaker. This is good for 2 gels. Distilled water was poured through the lip to just above the electrode level. Then an 'acrylamide plug' for the bottom of the gel was prepared.

5ml of the above Sequagel mix was measured out in a Falcon tube, and 10µl TEMED and 50µl ammonium persulphate (25% stock concentration) were then added, and quickly mixed, to start the gel polymerizing. With the gel cast flat, the mixture was then pipetted into the base of the glass plates to about 4-5cm deep. The plug was then allowed to set for 45-60 minutes. To cast the remaining gel, 43ml of the Sequagel mix was then mixed with 25µl TEMED and 250µl ammonium persulphate (25% stock). With the top of the gel cast propped up at an angle, the mix was then pippetted into the space between the glass plates, constantly shifting so no bubbles were allowed to form within the gel mix. When full, a shark tooth comb was then inserted in an inverted position, making a gel with a flat horizontal edge. More gel mix was then pipetted onto the comb so as to cover most of it to prevent the gel drying out. The top of the gel cast was draped over with a wad of paper towel soaked in 10X TBE, then covered with Saran Wrap and allowed to set overnight propped up at an angle. After the gel had set, the shark tooth comb was inverted and inserted so the teeth just entered the gel surface. The trough adjacent to the gel cast was then emptied of water and replaced with 1X TBE and connected up. Before running samples, the sequencing gel apparatus was pre- run by passing 50W of power through it for about 90 minutes to warm it up to 50^{o}C, indicated by the temperature strip on the gel kit, at which sequencing gels run best.

Composition of Solutions

Sequagel Concentrate (Flowgen)	*Sequagel Diluent (Flowgen)*
237.5g/*l* acrylamide	500g/*l* 8.3M urea
12.5g/*l* methylene bis-acrylamide	
500g/*l* 8.3M urea	

10X TBE Buffer

0.89M Tris base ,0.89M boric acid pH 8.0, 20mM EDTA

Section 2.2.11 Sequence Analysis

DNA and protein sequences were analyzed using GCG (University of Wisconsin), Staden (Bonfield *et al.*, 1995), CLUSTALW, PHYLIP and WWW-based software packages at their respective default values unless otherwise stated.

Chapter 3 RESULTS

Section 3.1 Cloning of *Cyp4a14*

Section 3.1.1 λ clone 4

There are currently 4 members of the CYP4A subfamily of genes known in the rat : CYP4A1, CYP4A2, CYP4A3 and CYP4A8. The CYP4A subfamily of genes is poorly understood in the mouse. To date, only cDNAs belonging to two murine *Cyp4a* genes have so far been cloned : *Cyp4a10* and *Cyp4a12*, which are most similar to the rat CYP4A1 and CYP4A8 respectively. No murine Cyp4a gene has been cloned thus far. Furthermore, it is not clear if other genes in the subfamily exist in the mouse. To further our understanding of the murine gene subfamily, a genomic cloning project was undertaken to isolate clones belonging to the *Cyp4a* subfamily.

The murine genomic library used was constructed in the λ FIX II vector which accepts inserts from 9 to 23kb (Statagene). In this library, the manufacturers prepared genomic inserts partially digested with *Sau* 3A and cloned into the compatible *Xho* I site of the vector without re-creating the *Xho* I site. Wild type λ phage contain active *red* and *gam* genes in the stuffer fragment; thus, only recombinant phages with these genes replaced by a genomic fragment are able to grow on strains lysogenic for P2 phage lysogens, such as XL1-Blue MRA (P2), taking advantage of the *spi* (sensitive to P2 inhibition) selection.

λ genomic clone 4 was originally isolated (D.R. Bell) when the library of 5 $X10^5$ plaques was screened with riboprobes transcribed from murine *Cyp4a10* and *Cyp4a12* cDNA sequences of Bell and coworkers (1993). The clone was mapped, particularly with restriction enzymes which excise the phage arms, subcloning and sequencing. On this occasion, *Xba* I (which cuts in the polylinker thus releasing the genomic insert) and *Not* I restriction enzymes were used, both singly

and in combination with other enzymes like *Bam* HI and *Eco* RI.

Section 3.1.2 Restriction digestion map of λ 4 clone

Single restriction digestion of λ 4 with *Xba*I yielded 2.2 kb, 2.6 kb, 4 kb and 4.5 kb fragments together with the phage arms of 9kb (right) and 20kb (left) (Figure 3.1). Double restriction digest with *Xba*I and *Bam*HI enzymes gave 1.7kb, 2.5kb and 5kb fragments together with the phage arms. To identify which fragments would be of most interest, a 1.2 kb partial rat CYP4A1 cDNA was used in hybrization experiments. This clone, named pIV2, was derived from the full length 2.1 kb cDNA of Earnshaw and coworkers (1988). pIV2 spans from nucleotide position 900 to the 3' end of the cDNA, and corresponding to the 3' region of CYP4A1 commencing from exon 8 of the gene. The gel was blotted onto Hybond N+ membrane by Southern transfer and hybridized with the riboprobe as described in the Materials and Methods. From Figure 3.1, the 2.2 kb *Xba* I fragment hybridized most strongly to the riboprobe pIV2. In *Xba* I/*Bam* HI double digests, the riboprobe hybridized very strongly to the 1.7 kb fragment, indicating that *Bam* HI cuts within the 2.2 kb *Xba* I fragment, thus releasing the 1.7 kb fragment. This was further proved when sequencing revealed a *Bam* HI site within this *Xba* I fragment.

Section 3.1.3 Subcloning and sequencing of λ 4 clone

All *Xba* I fragments obtained from λ 4 clone were subcloned into pGEM7 vector and sequenced by oligonucleotide-directed automated sequencing. Preliminary sequence data obtained for λ 4:6 showed that λ 4 was distinct from the known murine *Cyp4a* cDNA sequences, *Cyp4a10* and *Cyp4a12*.

Instead, comparisons with the rat CYP4A genes indicated that λ 4:6 subclone which contains a 2.2kb *Xba* I fragment and which hybridized strongly to the pIV2 riboprobe, was similar to the rat CYP4A2 / CYP4A3 sequences, corresponding to nucleotide positions 6333 to 8351 of the rat CYP4A2 gene. λ 4:6 was further deduced to be most 5' of the clone because sequencing

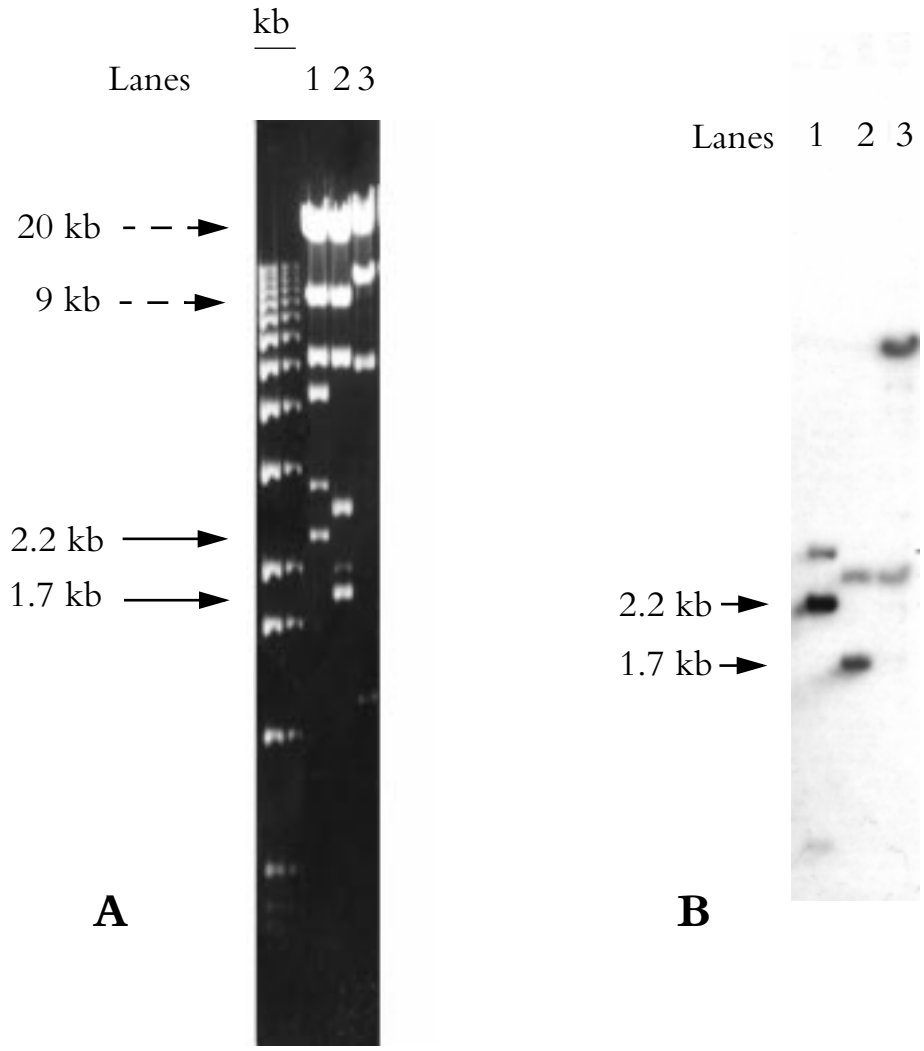


Figure 3.1 λ **4 clone mapped by restriction digest (A) and Southern hybridization (B).** Large-scale DNA was extracted from λ 4 phage plaque as described in Materials and Methods, restriction digested and run on a 0.7% agarose gel to separate the fragments. The gel was then blotted onto Hybond N+ membrane by Southern transfer and hybridized with the rat CYP4A1 partial cDNA pIV2 riboprobe (nucleotide positions 900 - end, corresponding to exon 8 - end) at 65 ^{o}C in 0.5M sodium phosphate buffer (pH7.2) containing 7 % SDS, 1 % BSA and 1mM EDTA. The membrane was then rinsed twice in 40mM sodium phosphate (pH 7.2) containing 0.1% SDS at room temperature and twice again at 65 ^{o}C for 30 minutes each and exposed using the BioRad GS-250 Molecular Imager Solid arrows indicate the size of *Xba* I fragments, while dashed arrows indicate the positions of the λ arms at 9 and 20kb. Lane 1 : *Xba*I digest, Lane 2 : *Xba* I + *Bam* HI digest, Lane 3 : *Bam* HI + *Eco* RI digest.

using Universal primers revealed a sequence which corresponded to the polylinker region of the λ FIX II vector. Since the exon sequences were distinct from both *Cyp4a10* and *Cyp4a12*, this proves that λ 4 clone was part of a novel murine gene. This gene has since been designated *Cyp4a14* by consultation with the cytochrome P450 nomenclature committee (D.R. Nelson, personal communication).

The 2.9 kb and 5.4 kb *Xba* I subclones, designated λ 4:22 and λ 4:5 were sequenced and contained exons 11 and 12 between them. However, no *Xba* I fragment was subcloned which contained exon 10. To confirm the deduced positions of these subclones and subclone a fragment that would cover exon 10, λ 4 was restriction digested with both *Not* I (which, like *Xba* I, releases the vector arms) and *Eco* RI (Figure 3.2). All *Not* I/*Eco* RI fragments were subcloned into

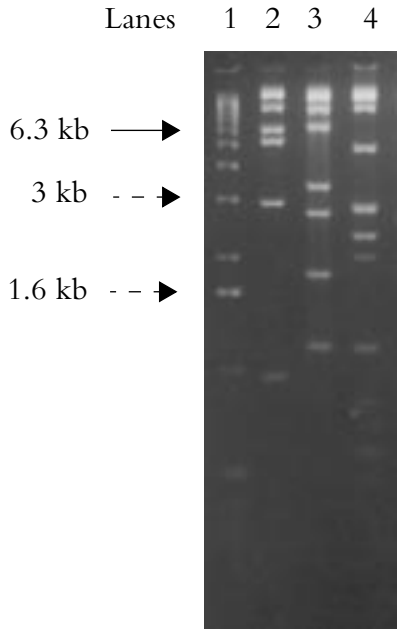


Figure 3.2 Double restriction digest analysis of λ 4 clone. DNA was prepared by a large-scale extraction method from λ 4 phage as described and double restriction digested with *Not* I + *Eco* RI (Lane 2), *Not* I + *Bam* HI (Lane 3) and *Xba* I + *Eco* RI (Lane 4). Products were then run on a 0.6 % agarose gel prestained with ethidium bromide to visualize the fragments. A 1 kb ladder was run in Lane 1. *Not* I/*Eco* RI fragments from Lane 2 were subcloned into pBlueScript vector and sequenced. Solid arrow points to a 6.3 kb *Not* I/*Eco* RI fragment. Dashed arrows point to two fragments of the kb ladder.

pBlueScript vector. The 6.3 kb *Not* I/*Eco* RI fragment, when sequenced, was shown to overlap sequences from subclones λ 4:6 and λ 4:22; additionally, exon 10 was also found to be contained in this 6.3 kb *Not* I/*Eco* RI fragment, between these two fragments (see Fig 3.3). Therefore, this 6.3 kb fragment spanned exons 6 through to 12 of this novel *Cyp4a14* gene.

Section 3.1.4 Map of λ4 clone

The deduced map of λ 4 is shown in Figure 3.3.

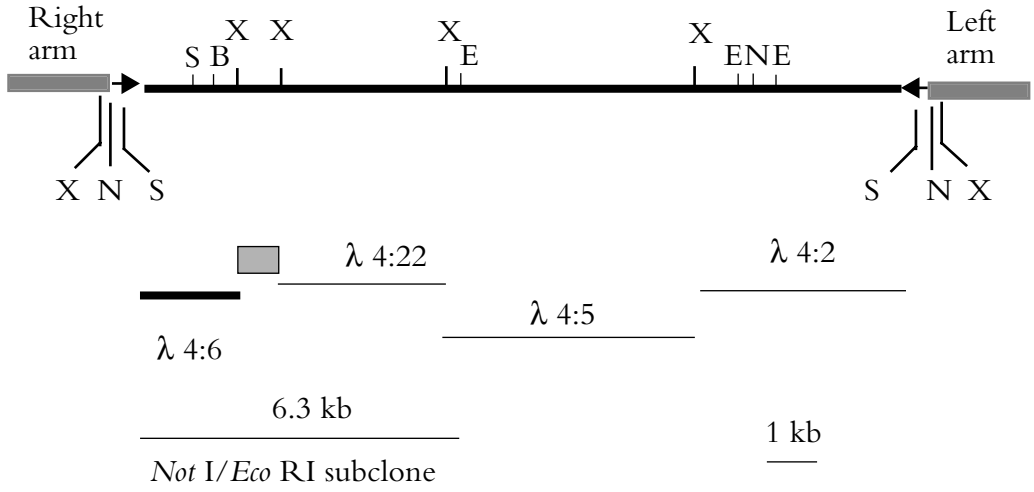


Figure 3.3 Map of λ 4 clone. Genomic DNA, indicated by a thick, bold line, was mapped with *Xba* I, *Sac* I, *Not* I and *Eco* RI. *Xba* I and *Not* I/*Eco* RI fragments were subcloned into pGEM7 vector and sequenced by automated fluorescent sequencing. λ 4 *Xba*I subclones and a *Not* I/*Eco* RI subclone are indicated at the bottom of the diagram, with λ 4:6 indicated by a bold line. Within λ 4:6, *Sac* I and *Bam* HI restriction sites mapped about 1kb and 1.7kb away from the right arm respectively. *Xba* I, *Not* I and *Sac* I sites are located within the polylinker of the vector. Arrowheads indicate the positions of the T3 (right arm) and T7 (left arm) bacteriophage promoters. X : *Xba* I, N : *Not* I, S : *Sac* I, B : *Bam* HI, E : *Eco* RI. The grey box denotes an *Xba* I fragment which was not successfully subcloned but which was contained within the *Not* I/*Eco* RI subclone of 6.3 kb. Only key restriction sites are shown.

To obtain the whole of the gene, a riboprobe was transcribed, as described in Materials and Methods, from λ 4:6, which is the 5' most *Xba* I subclone of λ 4 to re-screen the library. The genomic library of 5 x 10^5 phage plaques was plated onto 150mm agar plates and DNA was then transferred in duplicate onto circular Hybond N+ membranes according to Amersham's Hybond protocol, orientated using a needle, and denatured for seven minutes and then neutralized twice at three minutes each. Membranes were then baked at 80°C to fix the DNA onto the membranes and prehybridized at 65°C in a solution of 40mM sodium phosphate buffer (pH 7.2) containing 7% SDS, 1% BSA and 1mM EDTA for 2 hours in a Techne Hybridizer. λ 4:6 riboprobe

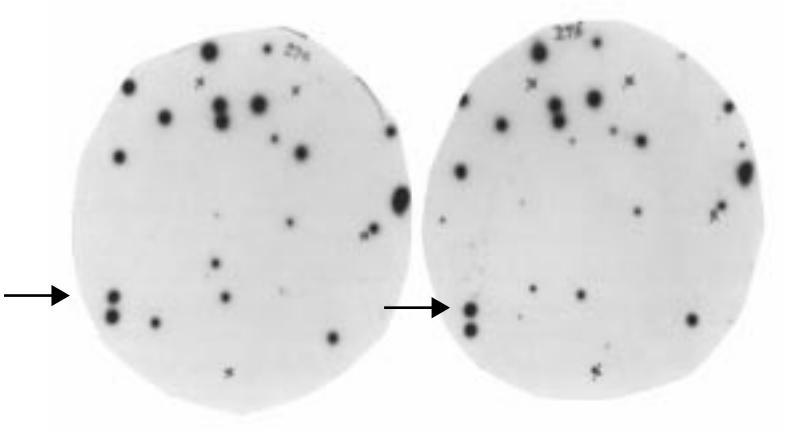


Figure 3.4 Autoradiograph showing a tertiary screen. A small area surrounding a positive candidate clone was picked from a primary screen of a genomic library, appropriately screened in a secondary screen at lower density, and the process repeated at an even lower density so as to obtain single distinct plaques in a tertiary screen as shown above. Hybond N+ membranes were used to transfer DNA from the phage genomic library in duplicate for the tertiary screen, orientated using a needle and re-hybridized with a riboprobe. Membranes were hybridized at 65 ^{o}C in 0.5M sodium phosphate buffer (pH7.2) containing 7 % SDS, 1 % BSA and 1mM EDTA. The membranes were then rinsed twice in 40mM sodium phosphate (pH 7.2) containing 0.1% SDS at room temperature and twice again at 65 ^{o}C for 30 minutes each and exposed to autoradiographic film for 72 hours at -80^{o}C. Arrows point to a candidate plaque which has hybridized to the probe in duplicate. Crosses within the autoradiograph denote the positions of assymmetrical needle-marks which were used to orientate the membranes relative to each other and to the library plate.

was then added to the tubes at the same temperature and left overnight. After hybridization, the membranes were then rinsed twice in 40mM sodium phosphate (pH 7.2) containing 0.1% SDS at room temperature and twice again in fresh wash buffer at 65 ^{o}C for 30 minutes each. The membranes were then exposed to autoradiographic film at -80 ^{o}C. Because the library was screened at very high density, secondary and tertiary screens were performed in which a small area surrounding a candidate plaque were picked and replated, thereby repeatedly reducing the density of the library until a pure candidate clone could be isolated. An example of a tertiary screen is shown in Figure 3.4.

Section 3.1.5 λ clone 12

Eleven independent candidate clones were identified by screening the library with the λ 4:6 ri-

boprobe. To identify which fragments of the clones hybridized to the λ 4:6 riboprobe, the candidate clones were restriction digested with *Xba* I (to excise the vector arms) and blotted onto Hybond N+ membrane and hybridized with the λ 4:6 riboprobe under stringent conditions as described. From the resulting autoradiograph in Figure 3.5, the riboprobe identified a higher molecular weight *Xba* I fragment (*ca* 5.5kb) in all but two candidate clones compared with the λ 4 2.2 kb *Xba* I fragment, the positive control. This indicated that clones have been isolated which contained sequences immediately upstream of the 5' end of and overlapping λ 4, an extension of approximately 3.3 kb. Of the eleven new clones, candidate clone number 1 had the least 5' sequences due to 2.9 kb *Xba* I fragment binding to the probe - an extension of just 0.7 kb.

Section 3.1.6 Detailed analysis of λ 9 and λ 12 clones

From Figure 3.5, nine λ clones out of the eleven were shown to have a 5.5 kb *Xba* I fragment which hybridized to the λ 4:6 riboprobe, thus having a similar 3.3 kb extension. Two clones belonging to this class, λ 9 and λ 12, were picked for detailed analysis by restriction digest and Southern hybridization, prior to establishing a strategy for subcloning and sequencing of the clones. DNA was prepared from the phage clones and restriction digested with *Not* I (which excises the phage vector arms) and various other restriction enzymes (not shown) (Figure 3.6). 1.7 kb and 1.3 kb *Not* I fragments were observed for λ 9 and λ 12 respectively. It is known that λ4 has a *Not* I site at its 3' end (Figure 3.3). Therefore, λ 9 and λ 12 have similar 5' extensions as observed in Southern hybridizations in Figure 3.5, but have different 3' ends. Thus, restriction digest with *Not* I, excised the vector arms of λ 9 and λ 12 and revealed that λ 12 is approximately 0.4 kb shorter than λ 9 on the 3' end.

As it was not possible to deduce from restriction mapping which of the two clones had a greater 5' extension, clones λ 9 and λ 12 were then further analyzed by directly sequencing the phage. Very pure DNA from the clones was extracted using the protocol according to the manufacturers

Qiagen and sequenced by priming from the vector arms. The partial sequencing data revealed that λ 12 was in fact longer than λ 9 by about 200 nucleotides on the 5' end (data not shown). Therefore, λ 12 was the clone chosen for further manipulations due it being the longer clone on the 5' end.

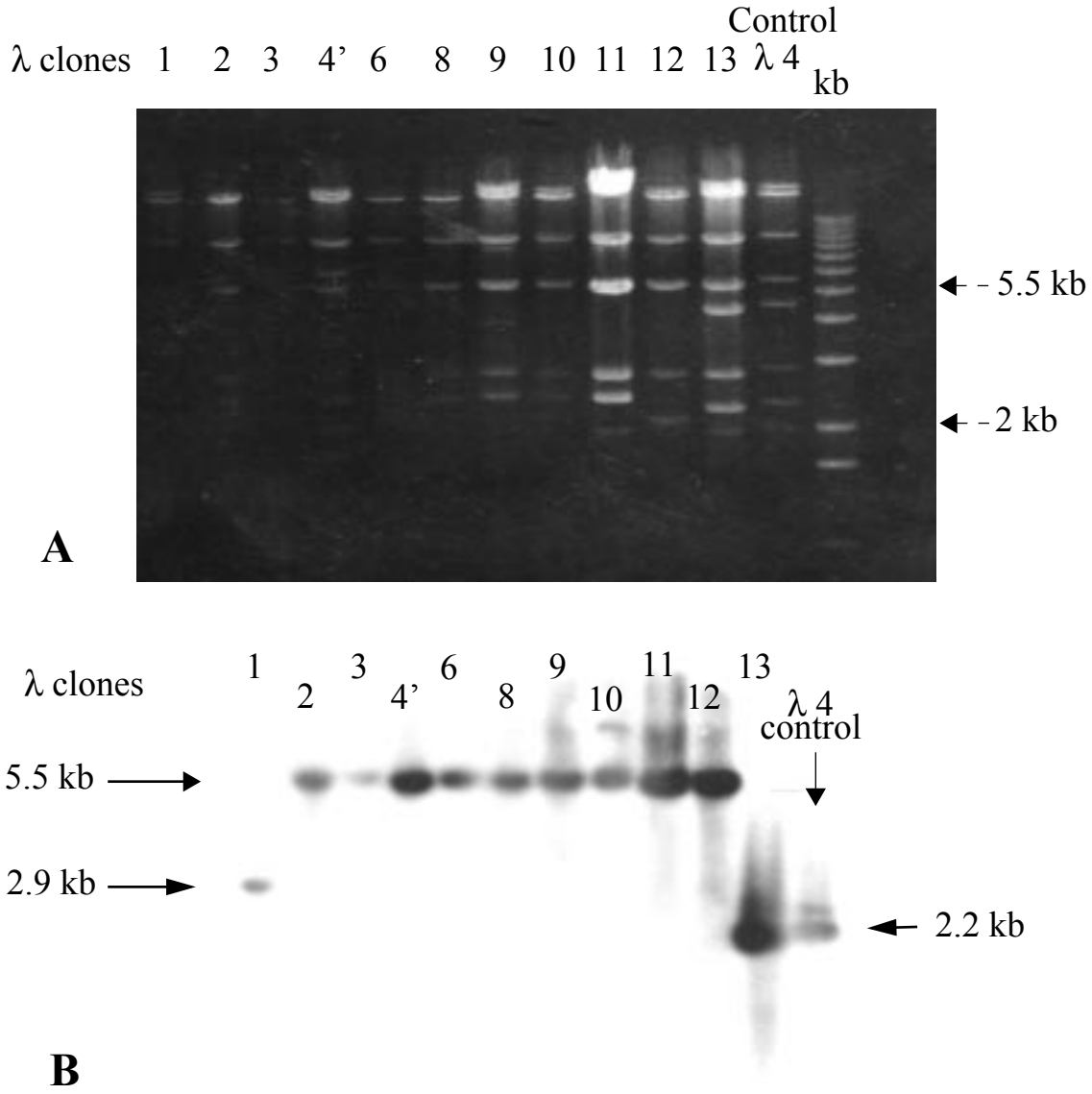


Figure 3.5 Mapping of λ clones by restriction digest (A) and Southern hybridization (B). DNA was prepared from the eleven genomic clones isolated from the genomic library when screened with λ 4:6 riboprobe. Clones were then restriction digested with *Xba* I and size-fractionated on a 0.7 % agarose gel. The gel was then blotted onto Hybond N+ membrane by Southern transfer and hybridized with λ 4:6 riboprobe at 65°C in 0.5M sodium phosphate buffer (pH 7.2) containing 7 % SDS, 1 % BSA and 1mM EDTA. The membrane was then rinsed twice in 40mM sodium phosphate (pH 7.2) containing 0.1% SDS at room temperature and twice again at 65°C for 30 minutes each and exposed to autoradiographic film at -80 °C. DNA from λ 4, the original clone isolated, was run alongside as a positive control, designated "control". The phage clones are indicated above as λ clones 1 to 13 while the sizes of the hybridized fragments are indicated on the side of the diagram with a bold arrow. Dashed arrows indicate the size of fragments from the 1 kb ladder. 'kb' denotes the 1kb ladder in diagram A.

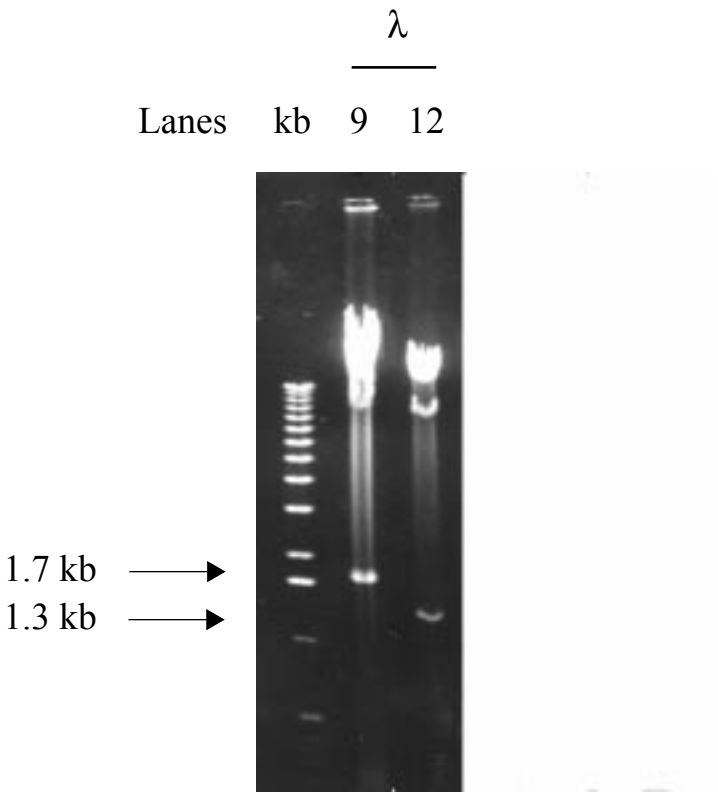


Figure 3.6 *Not* I digest of λ 9 and λ 12. Large scale DNA was extracted from λ phages 9 and 12 as described and were digested with *Not* I restriction enzyme . Products were then run on a 0.7 % agarose gel prestained with ethidium bromide to visualize the fragments. Solid arrows point to a 1.7 kb and a 1.3 kb *Not* I fragment for λ 9 and λ 12 respectively. 'kb' refers to the 1 kb ladder.

Section 3.1.7 Subcloning and sequencing of λ 12

Since *Xba* I produced a 5.5 kb fragment in λ 12 which hybridized to the λ 4:6 riboprobe, this enzyme was chosen to digest λ 12 to produce fragments for subcloning. In particular, the goal was to subclone a 5' extension of the gene. λ 12 DNA was extracted by the large-scale method as described, and restriction digested with *Xba* I and an aliquot run on a 0.7% agarose gel to ensure the digest was to completion. DNA was then ethanol-precipitated and ligated with appropriately digested pGEM7 vector overnight. The ligation mixture was then ethanol precipitated

and electroporated into *E. coli* XL1-Blue host cells and selected on appropriate media containing IPTG and X-Gal as described in Materials and Methods. White colonies were picked and plasmid-extracted as described, and re-digested with *Xba* I to release any insert cloned into the vector. Initially, 48 candidate clones were shown to be *Xba* I subclones but no 5.5 kb *Xba* I fragment was cloned, although other *Xba* I fragments were successfully ligated into the vector.

An alternative cloning strategy was employed in which *Not* I (which also excises the vector arms) and *Bam* HI (which cuts about 0.4kb inside the 3' end of the 5.5 kb *Xba* I fragment) were used for restriction digest. *Not* I and *Bam* HI double digests should therefore produce a 5.1 kb fragment covering an identical upstream extended region, in addition to two other fragments (3 kb and 6.2 kb, data not shown). Similarly, after plasmid extractions from 120 white candidate clones, and *Not* I/*Bam* HI double digests, no 5.1 kb fragment was observed, but both the 3 kb and 6.2 kb *Not* 1/*Bam* HI fragments have been successfully subcloned. Thus, all *Xba* I and *Not* I/*Bam* HI fragments have been cloned into pGEM7 except the 5' extended region.

A method, using 'colony lifts', was then devised to screen much larger numbers of bacterial colonies for the 5' extended fragment. XL1-Blue transformed with *Xba* I or *Not* I/*Bam* HI ligation mixtures were spread on sterile Hybond N+ membranes laid on agar plates and incubated overnight. Membranes were then orientated using four needle-marks, then peeled off, denatured and neutralized as described in the Amersham Hybond Handbook and hybridized with the λ 4:6 riboprobe. Initially, good positives were obtained in duplicate from the high-density primary screen, but with every secondary and tertiary screen to enable isolating a single colony, the radioactive signals were lost. This, together with the failure to clone the fragment after repeated attempts, suggested that the 5' extended region had a DNA structure which might be unstable due to a repetitive sequence.

Another attempt was made to clone the 5' extended region with a change of host cells from XL1-Blue to *E. coli* MRA cells, which were the original host cells of the phage library. Since the Cyp4a clones in the phage library apparently propagated stably in *E. coli* MRA cells, the unstable DNA fragment could be accommodated more efficiently in these original host cells. The original host cells are deficient in the *E. coli* restriction enzymes McrA and McrCB, thus DNA is expected to be more stable in these cells. After plasmid extractions from 96 white candidate colonies, again, all *Xba* I and both the 3 kb and 6.2 kb *Not* I/*Bam* HI fragments from λ 12 were cloned except the 5.5 *Xba* I or 5.1 kb *Not* I/*Bam* HI fragment. Another change to recombination-deficient host cells *E. coli* SURE (Stratagene) and DH10B (Gibco) also produced no positive result after analysis of 48 subclones.

λ 12 was also digested with *Sac* I (which also excises the phage vector arms) and *Bam* HI in combination to map it as an aid to subcloning smaller fragments of the DNA (Figure 3.7). The gel was blotted onto Hybond N+ membrane to identify specific extended fragments in a hybridization with λ 4:6 riboprobe. The resulting autoradiographic image (Figure 3.7) indicated the identification of a 3.6kb fragment generated from the double digest. The latter was deduced to be a *Sac* I-only 3.6kb fragment as a *Sac* I restriction map of λ 12 also demonstrated the presence of a fragment of the same size (Figure 3.8), and this *Sac* I fragment also hybridised to the riboprobe (not shown).

In an attempt to solve the problem of subcloning a whole DNA fragment which was unstable, attention was then focussed on subcloning the 5' extended region in parts; a *Sac* I fragment of 3.6 kb was particularly of interest because it hybridized to the λ 4:6 riboprobe, and was therefore immediately upstream of it. This indicates *Sac* I cuts within the 5.5 kb *Xba* I fragment. Since *Sac* I cuts within λ 4:6 about 1 kb inside the 5' end, it can be deduced that *Sac* I cuts again within the extended region, about 3.6kb away from the *Sac* I site in λ 4:6. This would explain the 3.6kb

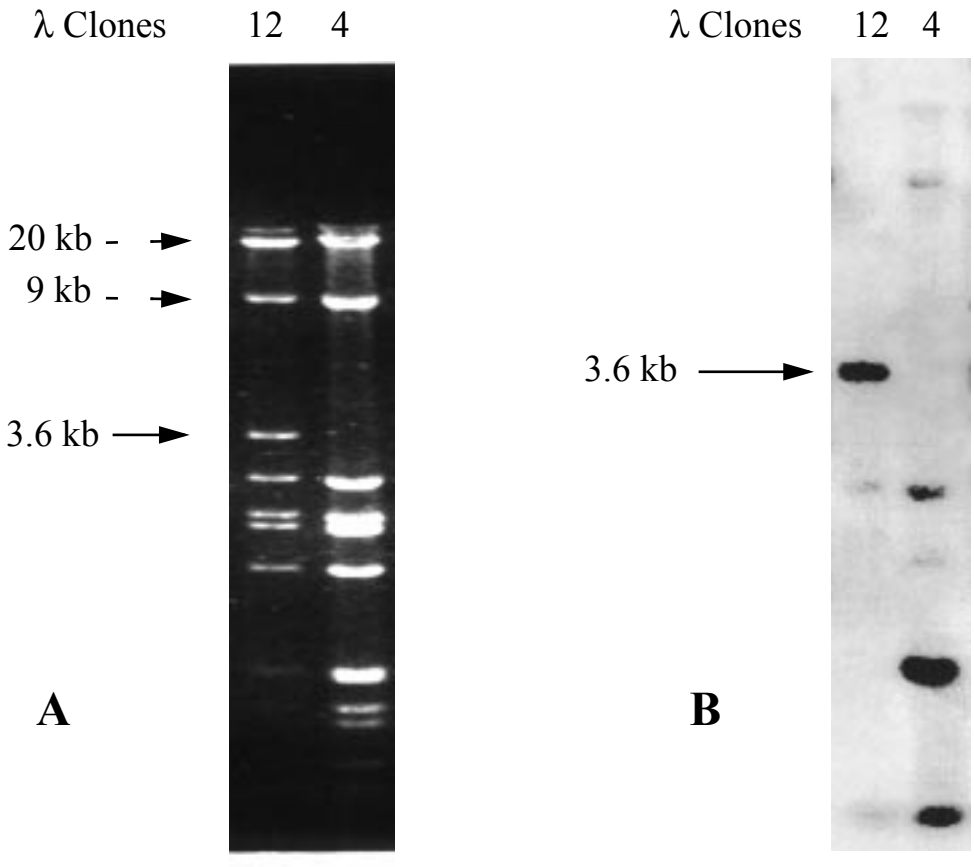


Figure 3.7 Mapping of *Cyp4a14* clones λ 4 and λ 12 by double restriction digest with *Bam* HI and *Sac* I (A) and Southern hybridization (B). Large-scale DNA was extracted from both phage clones as described in Materials and Methods, restriction digested and run on a 0.7% agarose gel to separate the fragments. The gel was then blotted onto Hybond N+ membrane by Southern transfer and hybridized with the rat CYP4A1 cDNA-derived pIV2 riboprobe (nucleotide 900 to end) at 65 °C in 0.5M sodium phosphate buffer (pH7.2) containing 7 % SDS, 1 % BSA and 1mM EDTA. The membrane was then rinsed twice in 40mM sodium phosphate (pH 7.2) containing 0.1% SDS at room temperature and twice again at 65 °C for 30 minutes each and exposed using the BioRad GS-250 Molecular Imager

Sac I fragment hybridizing to the λ 4:6 riboprobe due to a common region of about 1 kb in length. This hypothesis is confirmed by sequence data of λ 4:6. A deduced map of λ 12 is shown in Figure 3.9.

λ 12 was digested with *Sac* I and ligated into pGEM7 vector. After electroporation into both SURE and DH10B cells, and DNA extracted from 168 and 84 candidate clones each, again, the 3.6kb *Sac* I fragment was resistant to cloning. Six other *Sac* I fragments were successfully subcloned into pGEM7 vector (see Table 3.1).

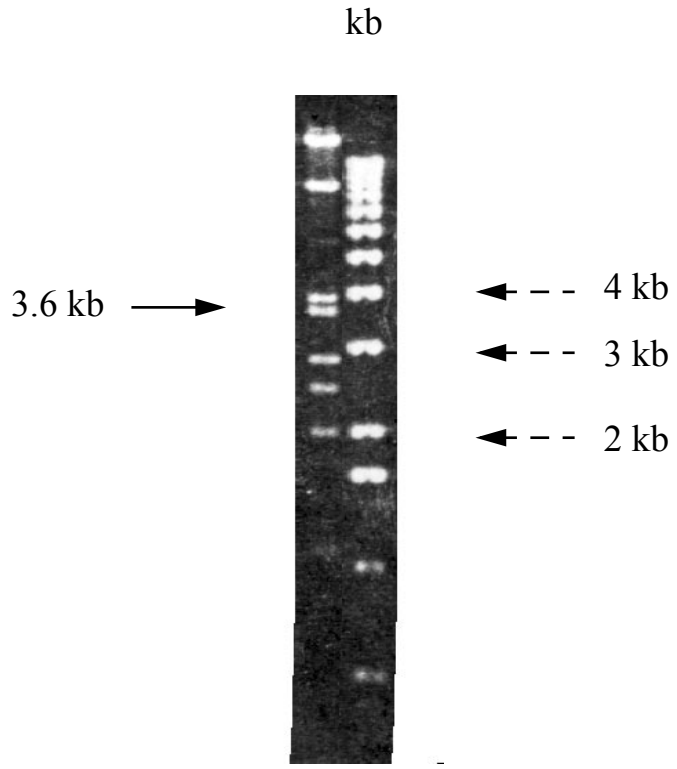


Figure 3.8 Restriction digest of *Cyp4a14* λ clone 12. DNA was restriction digested with Sac I enzyme and run on a 0.7% agarose gel prestained with ethidium bromide to separate the fragments as described in Materials and Methods. A solid arrow on the left side of the picture points to a 3.6kb *Sac* I fragment while broken arrows on the right side of the picture indicate the sizes of fragments in the ladder. 'kb' denotes the 1 kb adder.

A new vector, pZErO (Invitrogen), was then used for subcloning the 3.6 kb *Sac* I fragment. This vector is a component of Invitrogen's Zero Background/Kan Cloning Kit which is based on the vector containing the lethal *E. coli* gene, *ccd*B. Insertion of a DNA fragment disrupts expression of the *ccd*B gene, thereby permitting growth of only positive recombinants upon induction with IPTG. The recommended host strain for pZErO is *E. coli* TOP10F'. λ 12 was digested with *Sac* I and products run on a 0.7% agarose/TAE gel. After visualization of the fragments, the 3.6kb fragment was excised from the gel, purified by the geneclean method and ligated into the appropriately digested pZErO vector, and electroporated into electrocompetent TOP10F' host cells, and plated out on media containing kanamycin (which selects for the vector) and IPTG to select for the positive recombinant colonies. Using this cloning kit, the unstable 3.6 kb *Sac* I

fragment, that was so far resistant to subcloning after numerous attempts, was remarkably stable in pZErO and was thus finally cloned after the first attempt with this vector.

Sac I generated fragments of λ 12 were subcloned into a plasmid vector as summarized in Table 3.1. Subclone pG711 (which contains a 1.1 kb *Sac* I fragment) was sequenced by automated fluorescent sequencing and was found to be most 5' of λ 12, containing exon 1 of *Cyp4a14* by analogy with CYP4A2. Subclone pZ36 (which contains a 3.6 kb *Sac* I fragment), was found to contain exons 2, 3, 4 and 5. This *Cyp4a14* subclone also contains a region of CT dinucleotide repeats spanning over 70 bp in intron 2, which might have contributed to its instability. The sequence at the 3' most end of pZ36 (from λ 12) was identical to the 5' end of λ 4. This demonstrates that both λ 4 (the first λ clone isolated) and λ 12 are contiguous on the genome and that both clones belong to *Cyp4a14*.

Name of Subclone	Plasmid	Insert size
pG705	pGEM7	0.5 kb
pG711	pGEM7	1.1 kb
pG720	pGEM7	2.0 kb
pG723	pGEM7	2.3 kb
pG728	pGEM7	2.8 kb
pG738	pGEM7	3.8 kb
pZ36	pZErO	3.6kb

Table 3.1 Table of subclones of *Cyp4a14*. Large scale DNA was prepared from λ 12 phage as described and was digested with *Sac*I and subcloned into either pGEM7 or pZErO vectors. The subclones were designated according to the plasmid vector they were cloned into and their fragment size.

Therefore, the first clone isolated, λ 4, was the 3' half of *Cyp4a14*, and covers exons 6 through to 12. λ 12, the second and overlapping clone isolated, contains the 5' half of *Cyp4a14*, and contains the remaining exons 1 to 5 of the gene (See Section 3.2). Thus, the whole *Cyp4a14* gene was cloned and is contained in these two overlapping phage clones.

A deduced map of λ clone 12 is shown in Figure 3.9. The extended region was deduced to be adjacent to the left arm of the vector, by a hybridization experiment in which a *Bam* HI fragment was identified by the λ 4:6 riboprobe, which was over 20kb (left arm) in size (not shown). Direct sequencing of the phage clone using T7 primer (which primes off the left arm of the vector) confirmed that pG711 was adjacent to the arm.

Section 3.1.8 Cloning the 5' upstream region of *Cyp4a14*

To elucidate the regulation of the *Cyp4a14* gene, the promoter region must be cloned to examine which regulatory motifs are present. In order to isolate an overlapping phage clone that contains the promoter region of the gene, the genomic library of 5 x 10^5 phage plaques was screened again, using a riboprobe transcribed from pG711 subclone (containing the 1.1 kb *Sac* I fragment and exon 1), the 5' most subclone of λ 12. After a tertiary screen, eight independent phage candidate clones were isolated. DNA was extracted from these eight phage clones using a large-scale DNA extraction method as described in Methods and restriction digested with *Sac* I again (since *Sac* I excises the phage arms to release the genomic DNA). All phage clones possessed the same *Sac* I restriction map (not shown). After Southern blotting, the Hybond N+ membrane was hybridized with the pG711 riboprobe to identify any extended regions. On this occasion, the probe hybridized to 1.1 kb *Sac* I fragments obtained from all eight positive phages (not shown). This, together with the identical *Sac* I maps for all eight phages, indicated an identical 5' end to λ 12. This could mean that the *Cyp4a14* promoter region might be poorly represented in this particular genomic library. To solve the problem, a PCR-based approach was used to clone a small portion of the promoter region initially, with the option to then re-screen the genomic library using a probe derived from this PCR product if required.

Clontech's PromoterFinder DNA walking kit was used to clone regions upstream of the *Cyp4a14* start site after extensive library screening failed to identify any phage clones apart from

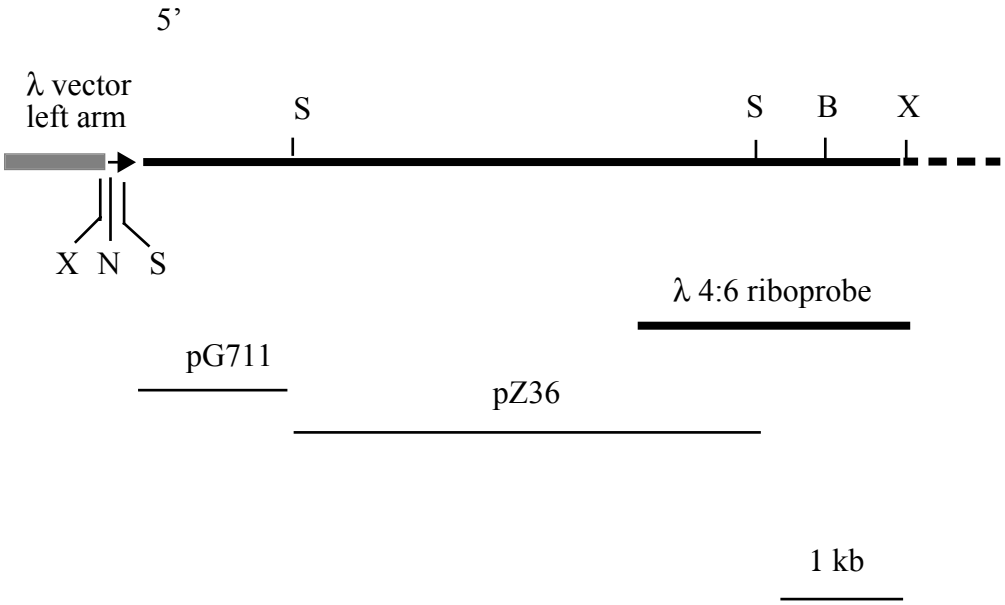


Figure 3.9 Map of the 5' end of λ 12 clone. Genomic DNA, indicated by a thick, bold line, was mapped with *Sac* I and subcloned into pGEM7 and pZErO vectors. λ 12 *Sac* I subclones pG711 and pZ36 are indicated beneath the phage map, with λ 4:6 riboprobe indicated by a bold line. Only the *Xba* I, *Not* I and *Sac* I sites within the polylinker of the left arm (20kb) are indicated. The arrowhead indicates the position of the T7 bacteriophage promoter. The dashed line refers to 3' region of the gene mapped in less detail. X : *Xba* I, N : *Not* I, S : *Sac* I, B : *Bam* HI, E : *Eco* RI. The 5' end of the gene is on the left hand side of the diagram. The vector arm is not drawn to scale.

clones identical to λ 12. This kit contains five libraries of uncloned genomic DNA fragments, all ligated to an adaptor of known sequence. In the construction of the libraries by Clontech, genomic DNA was isolated from mice, and five separate aliquots of DNA were then thoroughly digested with five different restriction enzymes : *Eco*R V, *Sca* I, *Dra* I, *Pvu* II and *Ssp* I. Each pool of DNA fragments was then ligated to the PromoterFinder adaptor. Each walk consisted of two PCR reactions : The primary PCR reaction used the outer Adaptor Primer AP1 provided in the kit and an outer, gene-specific primer (GSP1) derived from the gene. The primary PCR reaction mixture was then diluted and used as a template for a secondary PCR using the nested Adaptor Primer AP2 and a nested gene-specific primer (GSP2), upstream of GSP1 of the gene. Where appropriate, DNA bands of interest obtained from the primary PCR were also excised from the gel and purified by the gene-clean method as described in Materials and Methods. These DNA samples were then resuspended in a suitable volume of water and subjected to sec-

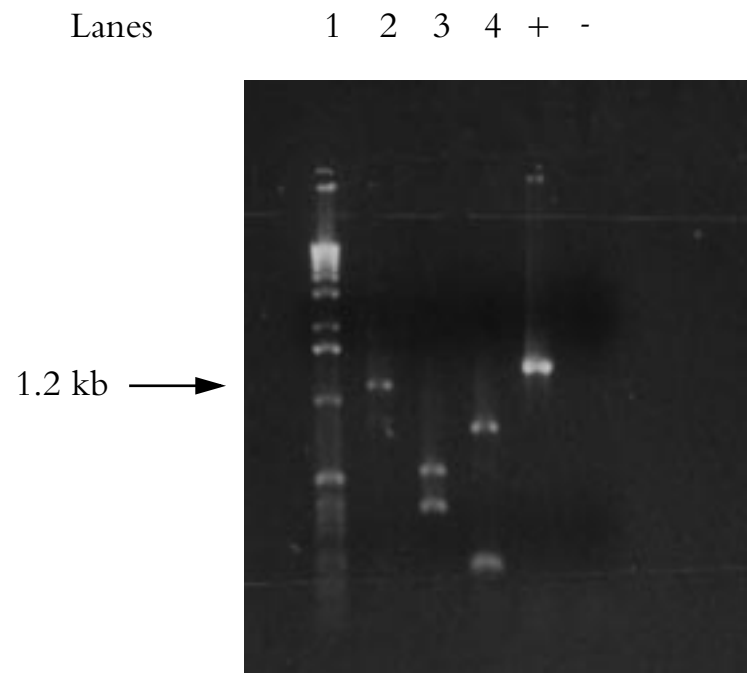


Figure 3.10 Identification of *Cyp4a14* secondary PCR products. Primary PCR reaction from the three libraries was diluted 1 in 20 and subjected to secondary PCR confirmation using *Cyp4a14* gene specific primers and adaptor specific primers as described in Materials and Methods. 5µl of the 50µl secondary PCR reaction were run on a 1.2% agarose gel. The arrow points to a 1.2 kb PCR fragment identified in Lane 2. This fragment was then gel-excised and cloned into pGEM-T vector and named pGT12. Lane 1: kb ladder; Lane 2 : *Dra I* library; Lane 3 : *Pvu* II library; Lane 4 : *Ssp* I library. '+' denotes positive control performed with *Dra* I library using positive control primers as described in the Clontech manual. '-' denotes negative control performed with no genomic DNA. Primary PCR used GSP1 (bottom strand, 5'-**GAG ATA CCA TCC AAG TAC CTT GTA G**-3', corresponding to nucleotide positions 47 to 23 of the *Cyp4a14* cDNA) while secondary PCR used GSP2 (bottom strand, 5'-**CCC ATG GTT AGT AGT TTC TGG ATC GAG**-3') located upstream of GSP 1 within the 5' untranslated region of exon 1 of *Cyp4a14*.

ondary PCR for confirmation.

In the primary PCR reaction, a gene-specific primer GSP1 (bottom strand, 5'-GAG ATA CCA TCC AAG TAC CTT GTA G-3', corresponding to nucleotide positions 47 to 23 of the *Cyp4a14* cDNA) derived from exon 1 of *Cyp4a14* and Adaptor Primer AP1 (top strand, 5'-GTA ATA CGA CTC ACT ATA GGG C-3') were primers used for primary PCR at the recommended annealing and extension temperature of 67 °C as specified in the Clontech protocol. From the resulting primary PCR, distinct bands of various molecular weights were observed when PCR products from the *Dra* I, *Pvu* II and *Ssp* I libraries in particular were run on an agarose gel (not shown). The products were then diluted 1 in 20 with clean UHP water and subjected to secondary PCR

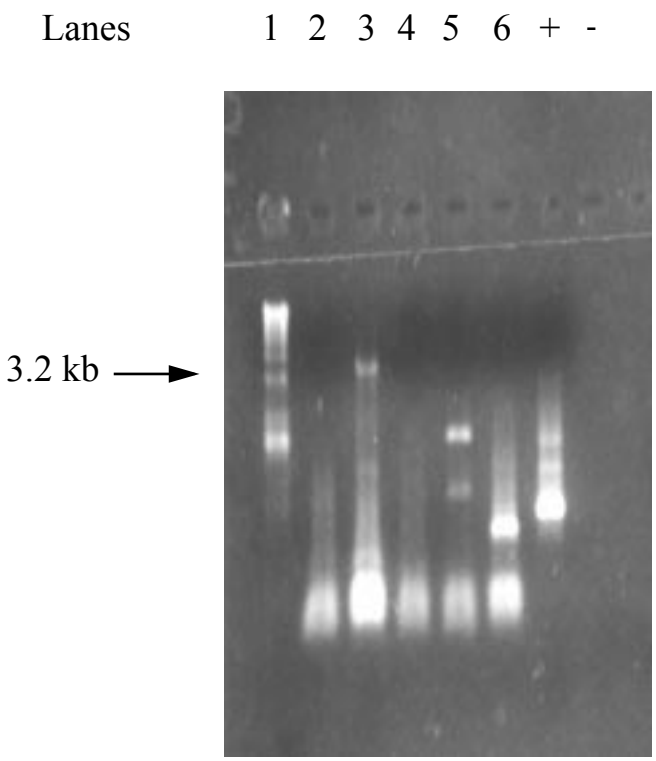


Figure 3.11 Identification of *Cyp4a14* secondary PCR reaction products. Primary PCR reaction was diluted 1 in 20 and subjected to secondary PCR confirmation using *Cyp4a14* gene specific primers derived from pGT12 subclone and adaptor specific primers as described. 5µl of the 50µl secondary PCR reaction were run on a 1.2% agarose gel. The arrow points to a 3.2 kb PCR fragment identified in Lane 3 (*Sca* I library). This fragment was then excised from the gel, purified by the gene-clean method, and cloned into pGEM-T vector and named pGT32. Lane 1: kb ladder; Lane 2 : *Eco*R V library; Lane 3 : *Sca* I library; Lane 4 : *Dra* I library; Lane 5 : *Pvu* II library; Lane 6 : *Ssp* I library. '+' denotes positive control performed with *Pvu* II library using positive control primers as described and produced a fragment with the expected molecular weight for this particular library. '-' denotes negative control performed with no genomic DNA. Primary PCR reaction used GSP 1 (bottom strand, 5'-**GGA TAG CTA GGA AGG CCC TGA GGA TC**-3') and Adaptor Primer AP1 (top strand, 5'-**GTA ATA CGA CTC ACT ATA GGG C**-3') and secondary PCR used a nested secondary primer GSP 2 (bottom strand, 5'-**CTA GGA GAG GTT GGA GGA AGG AAG GA**-3') upstream of GSP1 and Adaptor Primer AP2 (top strand, 5'-**ACT ATA GGG CAC GCG TGG T**-3').

confirmation using a second, internal gene-specific primer GSP2 (bottom strand, 5'-CCC ATG GTT AGT AGT TTC TGG ATC GAG-3') located upstream of GSP1 and within the untranslated region of exon 1 of *Cyp4a14* and Adaptor Primer AP2 (top strand, 5'-ACT ATA GGG CAC GCG TGG T-3'*)*. 5 µl of the 50 µl PCR products from each library were run on an agarose gel (Figure 3.10). The longest secondary PCR product, a 1.2 kb fragment from the *Dra* I library was then chosen for further analysis. This PCR fragment was excised from the gel, purified by the gene-clean method, and cloned into pGEM-T vector and named pGT12. This clone was then puri-

fied by a Qiagen DNA extraction protocol and fully sequenced on both strands by automated fluorescent sequencing.

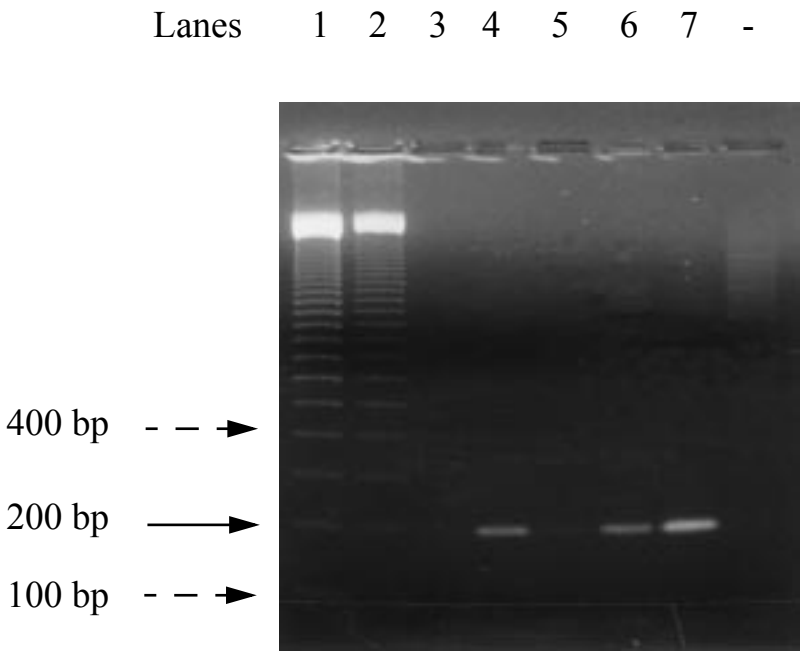


Figure 3.12 Analysis of junction sequence between *Cyp4a14* subclones pGT12 and pGT32. The two subclones were analyzed for contiguity on the *Cyp4a14* gene by PCR analysis. PCR was performed on mouse genomic DNA, represented by all five Clontech libraries, according to the secondary PCR protocol as detailed in Materials and Methods. 5µl of the 50µl PCR reaction of each library was analyzed on a 1.2 % agarose gel. Lanes 1 & 2 : 100 bp ladder, Lane 3 : *Eco* RV library; Lane 4 : *Sca* I library; Lane 5 : *Dra* I library; Lane 6 : *Pvu* II library; Lane 7 : *Ssp* I library. '-' denotes a negative control performed with no genomic DNA. The PCR reaction used primers derived from pGT12 (bottom strand, 5'-**GGA TAG CTA GGA AGG CCC TGA GGA**-3') and pGT32 (top strand, 5'-**AAG ATA ATC ATG ATC CAG GTG AGG ACA**-3'). The solid arrow indicates 200 bp fragments identified which correspond to the calculated distance between the PCR primers across the junction while broken arrows indicate the sizes of fragments of the ladder.

To walk further upstream of the gene, a separate set of PCR reactions was performed using primers derived from the 5' most sequences of the pGT12 clone. Primary PCR reaction using GSP 1 (bottom strand, 5'-GGA TAG CTA GGA AGG CCC TGA GGA TC-3') derived from pGT12 and Adaptor Primer AP1 (top strand, 5'-GTA ATA CGA CTC ACT ATA GGG C-3') was performed again at the annealing and extension temperature of 67°C according to the Clontech protocol as described in Materials and Methods. Primary PCR performed on all five Clontech libraries yielded bands of various molecular weights when reaction products were run on an agarose gel (not

shown). PCR reaction products from all five libraries were then diluted 1 in 20 and secondary PCR was then carried out using a nested secondary primer GSP 2 (bottom strand, 5'-CTA GGA GAG GTT GGA GGA AGG AAG GA-3') upstream of GSP1 and Adaptor Primer AP2 (top strand, 5'-ACT ATA GGG CAC GCG TGG T-3'*)*. Secondary PCR products were then run out on a gel to visualize the fragments. The longest fragment, a 3.2 kb band obtained in the *Sca* I library was chosen for further analysis, excised from the gel, purified by the gene-clean method and cloned into pGEM-T vector and named pGT32 (Figure 3.11).

To confirm that clones pGT12 and pGT32 which were obtained by PCR are contiguous on the genome, PCR was performed across the junction using primers derived from pGT12 [bottom strand, 5'-GGA TAG CTA GGA AGG CCC TGA GGA TC-3'] and derived from pGT32 [top strand, 5'-AAG ATA ATC ATG ATC CAG GTG AGG ACA-3']. This will produce a calculated 200 bp fragment (the distance between the primers) if the 2 PCR clones are contiguous on the chromosome. As shown in Figure 3.12, 200bp fragments were observed thus demonstrating that pGT32 is indeed located upstream of pGT12.

Section 3.1.9 Transcription start site of *Cyp4a14*

In order to determine the start site of transcription, primer extension analysis experiments were performed. Six male C57Bl6 mice, weighing 25-30 g each, were treated, in groups of three, with corn oil control (300µl volume) or methylclofenapate (MCP; 25 mg/kg, in 300µl corn oil vehicle volume) by *i.p.* injection for 3 days daily (A. Bell, N. Horley and D. Brady). Liver RNA was then extracted by the methodology of Cathala and coworkers (1983) as described in Materials and Methods and was pooled.

Two primers were used for primer extension. Primer 1 used was a 25-mer oligonucleotide (bottom strand, 5'-GAG ATA CCA TCC AAG TAC CTT GTA G-3') derived from exon 1 of *Cyp4a14* and corresponding to nucleotide positions 47 to 23 of the deduced cDNA and primer 2 was a

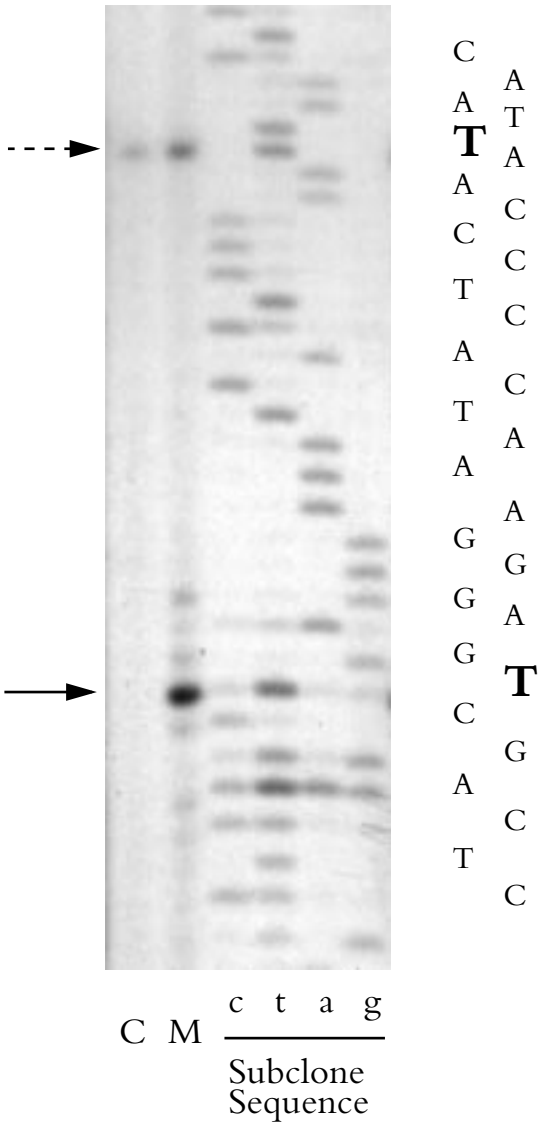


Figure 3.13 Primer extension analysis of the *Cyp4a14* gene. Male C57Bl6 mice were treated with 25 mg/kg methylclofenapate (M) or corn oil control (C) for 4 days, and 30μg of liver RNA, extracted by the method of Cathala and coworkers (1983), were used for primer extension analysis, at an annealing temperature of 55 ^{o}C, with primer 1 (bottom strand, 5'-**GAG ATA CCA TCC AAG TAC CTT GTA G**-3') corresponding to nucleotide positions 23 to 47 of the *Cyp4a14* cDNA, endlabelled with [γ-32P]-ATP as described. A subclone of λ 12, pG711, was used for sequencing with the same oligonucleotide primer and run in parallel on a denaturing 10 % acrylamide sequencing gel. The sequencing tracks are labelled c, t, a and g. The nucleotide sequence is indicated on the side of the gel, and the position of the major and minor primer extension products are indicated by a solid and dashed arrow respectively. The nucleotides indicated in bold correspond to nucleotides upstream of the primer where the extension has terminated.

27-mer oligonucleotide (bottom strand, 5'-CTG AGC AAG AAG GCC CAT TGG AAG AAC-3') and corresponding to nucleotide positions 77 to 51 of the *Cyp4a14* cDNA. Primers were end-labelled with [γ-^{32}P] ATP using T4 kinase. 2pmoles of the radiolabelled oligonucleotide were mixed with 30μg of liver RNA at an annealing temperature of 55^{o}C for 20 minutes and extend-

ed using the Superscript II Reverse Transcriptase supplied by Gibco BRL as detailed in Materials and Methods. The extension products were then run on a 10% denaturing acrylamide gel alongside a DNA sequencing ladder of a λ 12 subclone pG711 as a size marker, sequenced using primer 1. In the first instance, using primer 1, the principal extension product from RNA extracted from livers of MCP-treated mice was markedly more abundant than that from mice treated with the corn oil control (Figure 3.13).

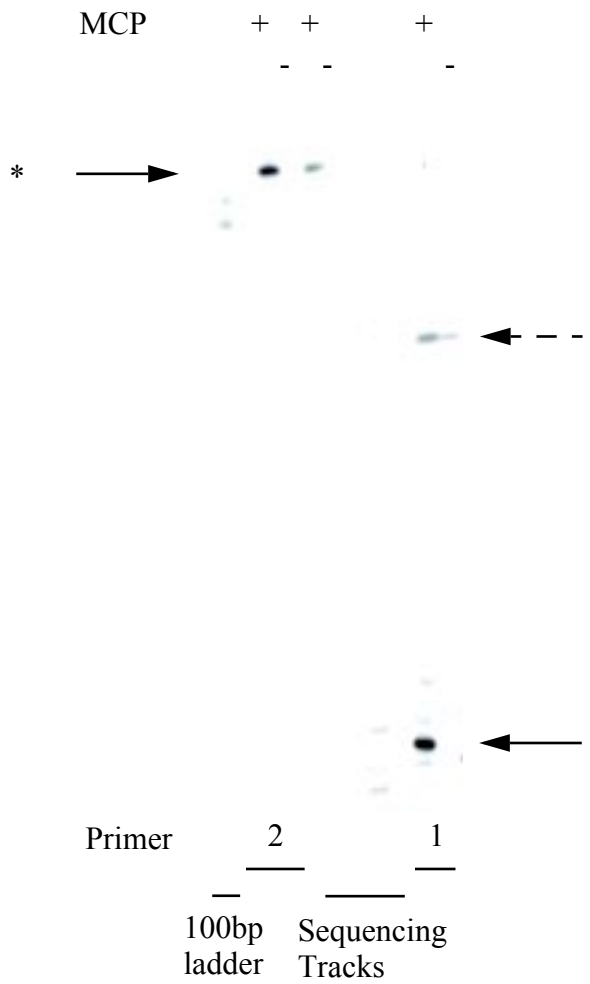


Figure 3.14 Primer extension analysis of the *Cyp4a14* gene with two primers. Male C57Bl6 mice were treated with 25mg/kg methylclofenapate (+) or corn oil control (-) for 3 days, and 30μg of liver RNA were used for primer extension analysis as described. A subclone of λ 12, pG711, was used for sequencing with the original primer 1(bottom strand, 5'-**GAG ATA CCA TCC AAG TAC CTT GTA G**-3', corresponding to nucleotide positions 23 to 47 of the *Cyp4a14* cDNA) and run in parallel on denaturing 10% acrylamide sequencing gel. The sequencing tracks were not visible at this exposure. The positions of the major and minor primer extension products are indicated by a solid and dashed arrow respectively. The asterisk (*) denotes an extended product (only in MCP-treated livers) with primer 2 (bottom strand, 5'-**CTG AGC AAG AAG GCC CAT TGG AAG AAC**-3'),corresponding to nucleotide positions 51 to 77 of the *Cyp4a14* cDNA). Sequencing tracks, primed with primer 1, are not visible on this exposure, but were clearly legible on extended exposure.

However, a slightly longer extension product of less abundance was also detected in corn oil vehicle-treated liver and only marginally induced by MCP. In order to rule out the possibility of the gene having an alternate transcription start site, primer extension analysis was carried out with an additional primer, primer 2 (corresponding to nucleotide positions 51 to 77 of the cDNA). This primer only detected a single, major transcription start site (Figure 3.14). No bands were observed that were of a higher molecular weight that would correspond to the minor band seen in Figure 3.13. Thus, the minor band seen with primer 1 is an artefact and not reproducible with primer 2. Here, primer extension assays using two gene specific primers confirmed that only one transcription start site exists for *Cyp4a14*; this was mapped to a T nucleotide, 26 nucleotides upstream of the putative start of protein translation.

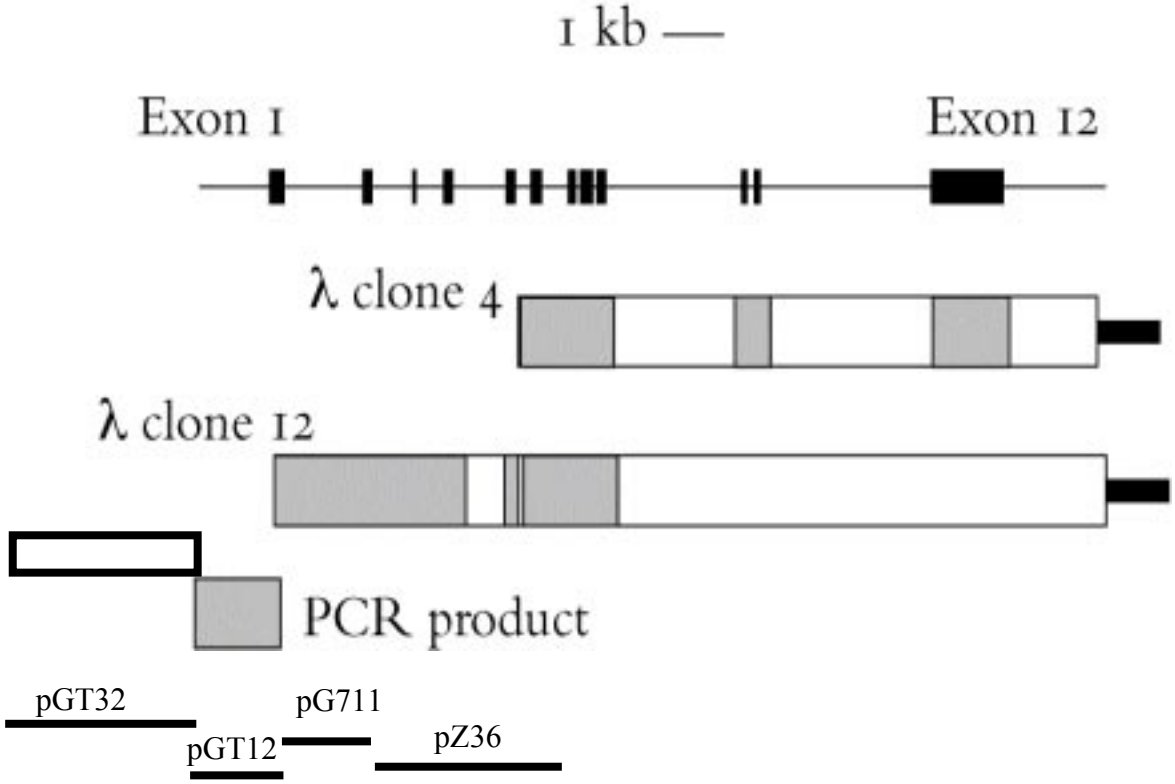


Figure 3.15 Cartoon summarizing the λ clones obtained for the murine *Cyp4a14* gene. *Cyp4a14* clones λ 4, λ 12 were isolated from the genomic library, and the clones pGT12 and pGT32 were obtained by PCR. . Black boxes indicate positions of exons, grey areas indicate regions where sequencing was performed in full while unshaded areas indicate regions where only partial sequencing was performed.

Section 3.2 Sequence analysis of λ clones

In this genomic cloning project, λ clones of the murine *Cyp4a* genes, *Cyp4a10*, *Cyp4a12* and an

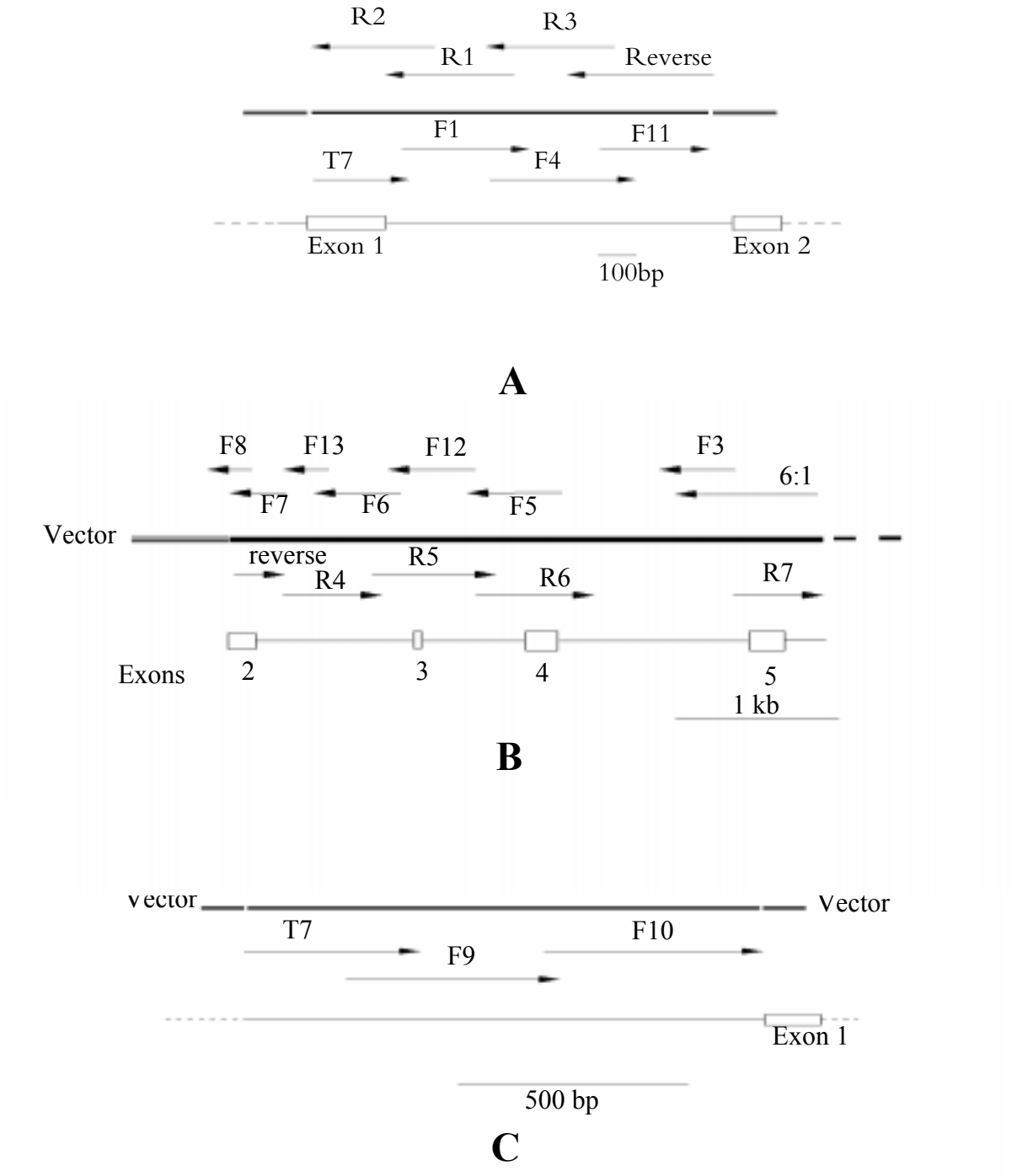


Figure 3.16 Sequencing scheme of subclones (A) pG711, (B) pZ36, (C) pGT12. Phage DNA was extracted as described, restriction digested with either *Xba* I (for pG711) or *Sac* I (for pZ36) cloned into the respective sites of pGEM7 vector or pZErO respectively. For pGT12, a 1.2kb PCR product obtained after confirmation with secondary gene-specific primers was gel-excised and subcloned into pGEM-T vector. All clones were sequenced by oligonucleotide-directed automated sequencing. Arrows denote direction and approximate length of sequencing runs. The names of the primers are indicated above the primers. The positions of the exons (empty boxes) of the gene are indicated beneath the diagram.

entirely novel *Cyp4a14*, have been isolated and characterized. In Figure 3.15, a cartoon summarizes the λ clones and genomic clones produced by PCR obtained for the murine *Cyp4a14* gene, also showing the positions of the exons and the regions of full or partial DNA sequencing. Subclones from λ clones 4 were sequenced according to primers designed by S. Kuo and D.R. Bell. All other subclones were sequenced by automated fluorescent sequencing by oligonucleotide-walking and the detailed scheme and direction of sequencing runs for *Cyp4a14* subclones are presented in detail in Figure 3.16.

Section 3.2.1 Genomic Sequence of *Cyp4a14*

The Cyp4a14 gene spans about 13 kb and contains 12 exons (a translational stop codon, followed by a poly-A addition site is found located within exon 12; S.Kuo and D.R. Bell). The genomic sequence, from the 1.2 kb 5' flanking region up to and including exon 4 of *Cyp4a14,* was thus assembled from the sequencing data separately obtained from the subclones pGT12, pG711 and pZ36. This complete sequence is shown in Figure 3.17.

```
                          -1100
TTAACTATAG GGCACGCGTG GTCGACGGCC GGGGCTGGTA AAATTTATTT
TTTATTTAAT TTTTTTCATG CAATATAATT TGATCATATT CCTTCCTTCC
                          -1000
TCCAACCTCT CCTAGATCCT CAGGGCCTTC CTAGCTATCC ATCTTCATGT
TAATGGATAG ACTGACAACC AAAACATTCT TTCTCTGCTT AAATAATATC
                          -900
TCCATAAAAT CTATAAATAA ATGAGGTAGT TGGAAACTAT CTCAGCACTT
TTCAATTGAT TGGCTAGTAA TCCTTCAATA TCTCATTTTT TTAACTTTCG
                          -800
CTTTATCTAT TCTGTGTGTA CATTAATTTT TTTCAGGCAA GGCATAATAT
ATATATAATT GGACTGATTT CTTTATTAGA GTTTGCCCTA TGTGAGGTCA
                          -700
AGAAATATTC TTAAATTAAT GAGTGACTGA ATAAGTGATG GGCAATTTAA
GTTTTTAGAA AAGAAAGGTT TTATTATTCC ATTCAGTCAA GATAGTGAGA
                          -600
CAGAGAAAGA GTCTGTCACA GGCTGTGTAT GTGGTGAGGC TGATTGAGTC
TTGAGCCACC TGAATGCAAC TGCACTGTTC CACCTGCTGG CACATCCATC
                          -500
CTGGATCAAT CTGGAGTGTG ACTGTGACAA GTCTCAGATA AAATGGAAGA
AACAGCTGGA TTTGGAGTCC AGATGCAAAG ATGACTATAG GTAGAAACTT
                          -400
TCAGCAATTA CATTCATCTG AACACACCAA CTACTGTTGT CATCATTTCA
CCCTGAAATT AGGAAAATAG TACAAGCAGC TACACCTATT ACATGTTTGG
                          -300
TAAATTAGAA TGTGAATTTC TTAATATCCA GGTTAATGTC TAGTCCATGA
```

```
CTTTACCTCA TCAGCAAGGG ATATACATAA CATGCAATAT GTGCTCAATA
                       -200
AATAGTTGTG AGTTAGTTCA GAGAAATGGG AATTGGTATA CATATAGATG
TTACCAAGAC TAGATACTAG AGATTTGTTT TTACTGTTTA CCAAAGCTGA
                       -100
TGTTGCAGAT TAATAAACTT TGGATTCTGA GGTCAGTCTC TGTCTGTCTT
CTCCATTCCC CCCTCCCACA AGTAGGTGTG TCTACCTTCT CATGACTTAA
                       +1
ATGAATTAAC CCTCACTAAA GGGAGTCGAC TCGATCCAGA AACTACTAAC
CATGGGTTTT TTTTTATTTA GCCCTACAAG GTACTTGGAT GGTATCTCTG
  M  G  F   F  L  F   S  P  T  R   Y  L  D   G  I  S
                        100
GGTTCTTCCA ATGGGCCTTC TTGCTCAGTC TATTTCTGGT GCTGTTCAAG
G  F  F  Q   W  A  F   L  L  S   L  F  L  V   L  F  K
GCAGTCCAAT TCTACTTACG AAGGCAGTGG CTGCTCAAGA CCCTCCAGCA
 A  V  Q   F  Y  L  R   R  Q  W   L  L  K   T  L  Q  H
                       200
TTTCCCATGC ATGCCTTCCC ACTGGCTTTG GGGGCACCAT CTGAAGGTAG
  F  P  C   M  P  S   H  W  L  W   G  H  H   L  K
GAAGAAGGGA GAGGGTGGAA GAGTGTGAAG AGCTGAATGA AGCTATGAGA
                       300
GTTTAGTATT AGTATACAGG GCAGTCTGCT GAGTGGTAAG CATGTAGTTT
CTGAAGAACG CTGGGTTTAC TTCTAAGAGT GGGTTTGCCC TTAGTCATGT
                       400
GGAGCATGAT TTCTTTTTAA AGGACACATT ATGACAATGG ACTCAAATCT
CAGGCTGTAG AAGGGCATGG GAAAATGTGG TGTCAGCTAC AGAGAAACAA
                       500
CTAATCTAGG TTGTTCTCTG AGAGAAATTG TAGAAGTGAT GATTTCTATG
TTGTATATCA ATTTCCTCTG AGATTGAAAC TGCCCATTTT GGAGAGAAAT
                       600
GTCTGATTAC AATACAGGGC TTTCTAGAGG TAAAATATTC CATCTAGAAA
GAAAGCCTCT TACAATAAAG AAAGTTCTTT GGGCAGGTGA AACTGGCTGG
                       700
ATTCATTTGC CATTCCCATC CTCAAGTGTA TGTAAGCAGT AAGAATGGCT
GACTATGCTA CTAACTCTCA GTATAATGTA ATTTTGTAAG GGTCTGCATG
                       800
AAGAGTCACT CACTGAGGGC TCACACGCTA GCCACATTTA AGAAGAGCAA
AATTCAGAAT TACTACAGAA CACAGAAGTT CTTACCACAG TAGCTTCTCT
                       900
AAGGAGCTTT AGTTTCATCT AACAAGACTC CAGGATCTGG CCAGGACAAG
TCCTGGGACA CAGCAATGGT CTGAATAGAA GAAAAAAAAA AGACAGATTA
                       1000
TCAAACCAAC CCAAATTTTT ACCCTTTCTT GAACATGCCA GAGTTATGCC
CCTGTGGTGG TCCTGACTCT TCCATACCAC ACAAAGTACT TGATGTCAAA
                       1100
GAACATCTTC GTGAATTCAA CTTTTCTTTG ATGTTCTGTT CCCAAAGGAC
                                                      D
AAGGAGCTCC AGCAGATTCT TATATGGGTA GAGAAATTCC CAAGTGCCTG
 K  E  L   Q  Q  I  L   I  W  V   E  K  F   P  S  A  C
                       1200
CTTACAGTGT CTCTCGGGGA GCAATATACG AGTCCTGCTT TATGATCCTG
  L  Q  C   L  S  G   S  N  I  R   V  L  L   Y  D  P
ACTATGTGAA GGTGGTTCTG GGGAGATCAG GTGAGAGTCC AAACTCATTA
D  Y  V  K   V  V  L   G  R  S
                       1300
CAGCTGTAGC ATATCTGCCA CCCAACATTC AAGACCCAAG TCTGTGTAAT
TTTAGAAAAG AGTGACACTT CAGCCATCAA TGTCTAAATA ATACCTACCA
                       1400
```

```
TTCTTTCTGT CCAGAGACCA ACATATGGTT AGCAAATACT TACATAGTTC
TCATCACTTA AAATACTTGC ACAAGGCTTA AATTCTCTCT CTCTCTCTCT
                       1500
CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTTC
TGTTTTTTCT TTTGCTCTTA AAACATATTG AGTGGATATG TTTTACAAAT
                       1600
TTCCCAAACG TTATAATGGC AGAGACTTCA GCTGCTGGTA CATTGAAGTT
CTCTTTCCTA AGTCCAGCTA ACAGGCAACA CAAAGACAGA AAGGAATGGG
                       1700
AGGTAGTTTG AGTGAAAAGG CAGCAGCATG AAATCGACAG AAAACTAATG
TTGTGGTCGC TGAAATGGAT GGACTGGTTT TGTTGCTCAA GCCCAATGAC
                       1800
CTAATTTTAA GCCCTGAGAC CCACATTGAA AATGGGGTGA ATGAACCAAC
TCCACAAATT TACCTTCTGA ACTCCACACA GACACTGTAG TAGGTGCAAT
                       1900
TCTGAAAAAC TAAAGGAAAG CTTGCAGTTT GAACACCATC CCCCATCTGG
TCCTCAAGAC CTGTCTTACA TTTCAGATCC AAAGGCTTCT GGAATTTATC
                            D  P   K  A  S   G  I  Y
                       2000
AATTCTTTGC TCCCTGGATT GGTAAGTATA TTAAAACAAG AAGTGTGTTT
Q  F  F  A   P  W  I
TCTTCACTGC TCTTTGGCTA ATGGAGAACA ACAGGTAACT AGTCACAAGT
                       2100
ATCACAATCC AGTCCTAAAT CAAGCTTTAT TTCTCTTCTT CCAATCAGTC
CTTAGTCCTT TTTAATGCTG GTGTGAGTGT GCATGATACC TGGATTCTGC
                       2200
AATATTATTT CTTCAGACAC TCATTGCCAG ACATTTTTAT ATTTCCCTCT
TGACATGACC CCAACACATT TTCATCCCTC ATTTCACTTC TGTCTTAGTC
                       2300
ACAGCAGATT TATCAGACAC ACTGGAAATC TCACTTTCTG TCTTTGTCTA
AAAGAAAAAA AAAACCTCAC TATTACTGAA CTTTGATTTC CCCTTTTTTA
                       2400
AATCCCAGTG AATGTAACCA AATGATTCCA GGGGTACTCT TAATTCCACT
TCTGTCAGAC ACATTCACAT ATCAGTCTCT GGCACAGGTT ATGGTTTGCT
                                      G   Y  G  L  L
                       2500
CCTGTTGAAT GGGAAGAAGT GGTTCCAGCA TCGGCGGATG TTGACTCCAG
  L  L  N   G  K  K   W  F  Q  H   R  R  M   L  T  P
CCTTCCACTA TGACATCCTC AAACCCTATG TCAAAATCAT GGCGGACTCT
A  F  H  Y   D  I  L   K  P  Y   V  K  I  M   A  D  S
                       2600
GTCAATATAA TGCTAGTGAG TTGCTCTGTC CCTTTCTCAT TTTATTACAC
 V  N  I   M  L
CCACTCAACA GCATAGCCGC
```

Figure 3.17 Genomic sequence of *Cyp4a14* spanning only the 5' flanking region and exon 4. The *Cyp4a14* gene begins at the transcription start site position +1 and the start of protein translation is indicated in a bold ATG codon at position +26. The amino acids, in their single letter abbreviations, are indicated beneath the four exon sequences from which they are deduced. The (CT) dinucleotide repeats in intron 2 are indicated in bold.

Exons 6 to 12 were sequenced by S. Kuo and D.R. Bell. Within pZ36 subclone, intron 4 was not sequenced to completion due to its length (850 bp). Therefore, exon 5 was sequenced separately using oligo-directed automated sequencing from the 3' end of pZ36. Exon 5 sequence is presented in Figure 3.18.

```
  1    GATAAATGGG AGAAGCTTGA TGGCCAGGAC CACCCTCTGG AGATCTTCCA
        D  K  W   E  K  L  D   G  Q  D   H  P  L   E  I  F  H
 51    CTGTGTTTCA TTGATGACAC TGGACACTGT TATGAAGTGT GCTTTCAGCT
         C  V  S   L  M  T   L  D  T  V   M  K  C   A  F  S
101    ACCAAGGCAG TGTTCAGTTG GATGAG
       Y  Q  G  S   V  Q  L   D  E
```

Figure 3.18 Sequence of *Cyp4a14* exon 5. Exon 5, in pZ36 subclone, was sequenced by automated fluorescent sequencing. The deduced amino acid residues are indicated beneath the nucleotide sequence by their single letter abbreviations. Numbers are arbitrary and refer to the nucleotide positions in the exon. Deduced cDNA sequence of *Cyp4a14*

Section 3.2.2 Analysis of *Cyp4a14* promoter

To further the understanding of the regulation of the *Cyp4a14* gene expression, the promoter region of the gene was analyzed for any transcription factor binding sites and other putative regulatory motifs (Figure 3.19). In the region upstream of the transcription start site, no consensus TATA box was found. However, a 19 bp element and a GAATTAA box were present, located at a similar distance (49 bp and 17 bp respectively) from the transcription start site as in CYP4A1 and CYP4A2. Both motifs are conserved in the rat CYP4A1 and CYP4A2 genes (Kimura *et al.*, 1989a). Interestingly, the 19bp element has been shown to be crucial for the basal transcription of the rat CYP4A1 (Bell and Elcombe, unpublished results) and is thus likely to have a similar role for this gene. It is also a putative binding site for Sp1, an ubiquitous transcription factor. By utilizing the Transcription Element Search Software (TESS) using the *Transfac* database, two separate PPRE half sites were identified approximately 90 bp and 725 bp of the start of transcription. No recognizable PPRE, a directly-repeated AGGTCA motif spaced by one nucleotide, was detected. Other putative transcription factor binding sites include those for SP1 and HNF-1. DNA sequence similarity with CYP4A2 in the 5' non-coding region is relatively high (84 %) for approximately 350bp upstream of the transcription start site but is reduced sharply thereafter (data

not shown). Interestingly, this region of reduced similarity corresponds to the position of a repeat in the CYP4A2 gene.

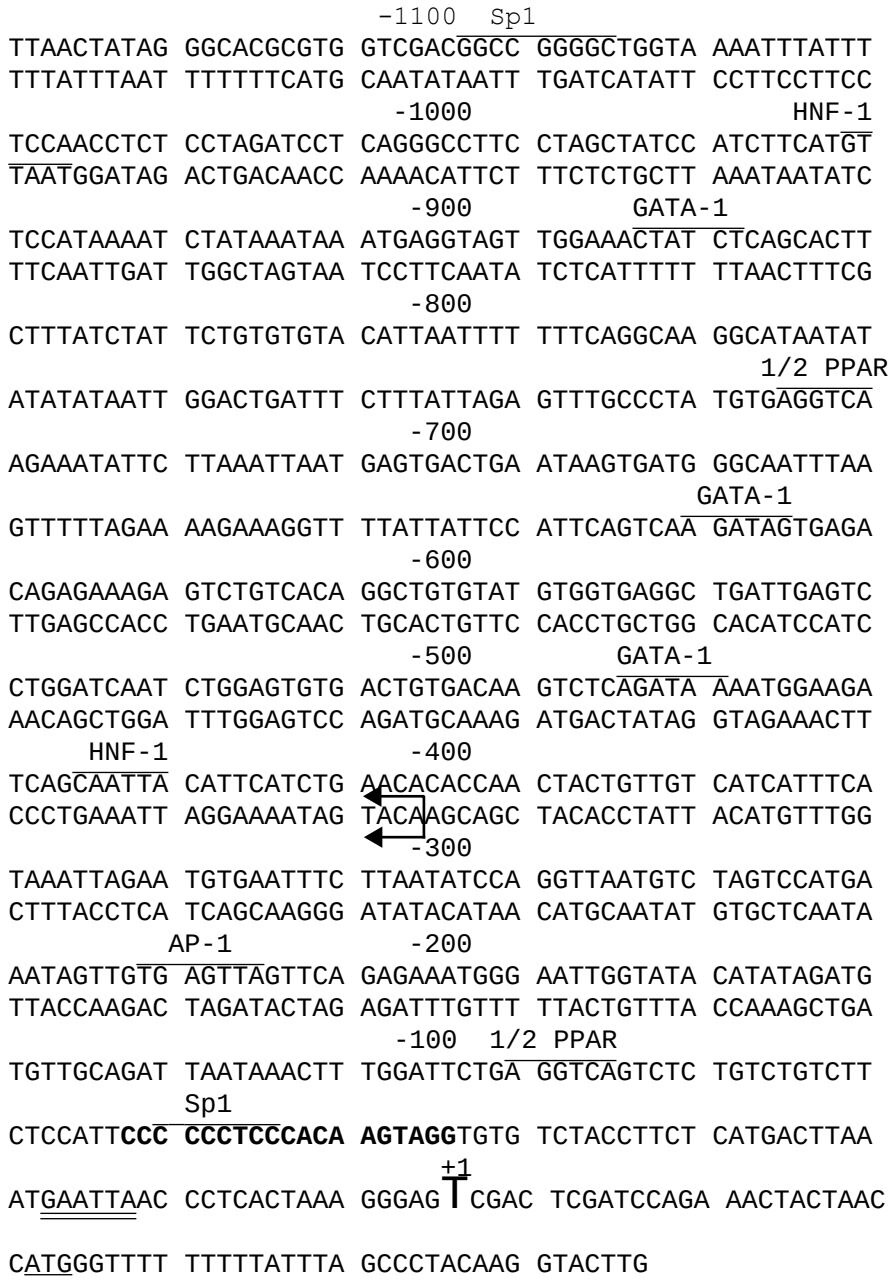


Figure 3.19 Analysis of 5' flanking sequence of *Cyp4a14*. Nucleotide sequence from the promoter region of *Cyp4a14* was analyzed for putative transcription factor binding sites by utilizing the Transcription Element Search Software (TESS) using the *Transfac* database on the WWW. Putative transcription factor binding sites are overlined. The start of gene transcription, denoted by +1, as determined by primer extension analyses is an enlarged T nucleotide, located 26 nucleotides upstream of the start of translation. The conserved 19bp element (indicated in bold) and the GAATTA box (double-underlined) are located 49 and 17 nucleotides upstream of the start of transcription respectively. Double arrows 350 bp upstream of the transcription start site indicate the region where similarity with CYP4A2 is reduced sharply.

Section 3.2.3 Analysis of exon/intron junctions of *Cyp4a14*

The exon/intron junctions of *Cyp4a14* are analogous to the previously reported positions in the

rat CYP4A2. By analysis of the *Cyp4a14* gene, the sequence of exon 1 contains a putative ATG codon, the start of protein translation. 26 nucleotides upstream of this triplet is located the start of transcription, a T nucleotide, that has already been determined by primer extension using primers derived from exon 1 of the *Cyp4a14* gene. Additionally, the sequence of exon 12 contains a stop codon followed by a consensus poly-A addition site downstream from the stop codon, indicating that exon 12 is the final exon of this gene (not shown, S. Kuo and D.R. Bell). All exons of the murine and rat CYP4A genes are compared in Table 3.2.

From the table, *Cyp4a14* showed highest similarity to the rat CYP4A2 and CYP4A3, and marginally more similar to CYP4A3. Specifically, exons 4, 8, 11 and 12 were very highly conserved, with up to 95 % or more nucleotide similarity between *Cyp4a14* and CYP4A2 / CYP4A3, as indicated by a single asterisk. Compared with the other CYP4A genes, these exons are also most highly conserved. Additionally, exon 3 was also highly similar between *Cyp4a14* and CYP4A3, but not CYP4A2. The exon sequences demonstrate that *Cyp4a14* is slightly more similar to CYP4A3 than to CYP4A2.

Similarities of the introns or regions within introns may also provide information about evolution in the CYP4A subfamily. Table 3.3 shows the similarities of several introns compared between *Cyp4a14* and CYP4A2, by halves.

Exon/ Gene	CYP4A1	CYP4A2	CYP4A3	CYP4A8	*Cyp4a10*	*Cyp4a12*
1	70	91	90	69	67	-
2	69	87	89	77	72	-
3	70	91	** 98	70	73	-
4	84	* 98	* 98	81	84	-
5	76	93	92	76	75	-
6	64	87	88	69	63	-
7	83	92	94	78	83	-
8	86	* 96	* 97	79	88	80
9	79	92	92	86	78	85
10	75	88	91	79	73	75
11	78	* 93	* 94	82	78	81
12	85	* 95	* 95	91	85	-
Average	77	92	93	78	77	-

Table 3.2 Comparison of *Cyp4a14* exon nucleotide sequences with rat and mouse CYP4A genes. Individual exons of *Cyp4a14* are compared by BESTFIT for similarity with respective exons in rat and mouse CYP4A genes. Single asterisks '*' denote key exons of *Cyp4a14* and CYP4A2 / CYP4A3 which share high sequence similarity (93% and over) while double asterisks '**' denote an especially high sequence similarity in exon 3 between *Cyp4a14* and CYP4A3 only. Only exons 8 - 11 are sequenced in *Cyp4a12* (Bell et al., 1993).

Intron	% similarity 5' half	% similarity 3'half
1	86	82
2	84	82
3	89	88
6	81	84
7	88	**92**
8	85	86
10	70	**95**

Table 3.3 Comparison of intron sequences *Cyp4a14* with CYP4A2. Only introns 1, 2 and 3 and 6, 7, 8 and 10 are compared, by halves, with the corresponding regions of CYP4A2. The value in bold refers to the 3' half of introns 7 and 10, adjacent to the highly conserved exons 8 and 11, which show very high nucleotide conservation between *Cyp4a14* and CYP4A2. Sequences for *Cyp4a14* introns 6, 7 8 and 10 are obtained from S. Kuo and D.R. Bell.

The full lengths of the introns of *Cyp4a14* may be compared by BESTFIT with the correspond-

ing introns of CYP4A2, giving similarities of up to 90 % (intron 7). However, the similarities of the seven introns between *Cyp4a14* and CYP4A1 were consistently very poor. An example of this is intron 2 (total of 696 nucleotides) of *Cyp4a14*, where only 166 nucleotides at 66 % identity with a corresponding region in CYP4A1 (not shown) while the whole intron in *Cyp4a14* is 83% similar to that in CYP4A2. Thus, from the relatively high conservation of intron sequences, it is possible to deduce that the mouse *Cyp4a14* gene is more similar to the rat CYP4A2 than CYP4A1.

Additionally, the analysis of the introns in detail, as summarized in the table above, revealed that, specifically, the 3' halves of intron 7, which is adjacent to the highly conserved exon 8 (96% nucleotide similarity), are more conserved between *Cyp4a14* and CYP4A2 than the 5' halves. Exon 8 in fact encodes for a highly conserved 13 amino acid motif, which is unique only to CYP4 proteins (Bradfield *et al.*, 1991; see later sections). Similarly, the 3' half of intron 10, which is adjacent to exon 11, is 95% identical to the corresponding region in CYP4A2. Exon 11 is highly conserved, and together with exon 12, code for the PFSGG(A/S)RNCIG amino acid motif that is highly conserved in all CYP4 proteins (see later sections). Thus, in these two examples, the intron sequences immediately upstream of the highly conserved exons 8 and 11, are highly conserved together with the exons themselves.

Within intron 2 of *Cyp4a14* is located a region of (CT)n dinucleotide repeats, where n=32. Such a feature was also found in CYP4A1 (3' UTR, n=25) but not in CYP4A2 (although there are twenty-four (CA) repeats in intron 2). These (CT)n repeats have previously been reported to be located upstream of a Drosophila heat shock protein hsp26 gene, which appeared to play a role in formation of the wild-type chromatin structure and the DNAse I hypersensitive sites of the gene, thus regulating its gene expression (Lu *et al.*, 1992). It would be interesting to see if these repeats have a role in organizing the chromatin structure of *Cyp4a14* (or indeed the rat

CYP4A genes) and contributing to its expression.

Section 3.2.4 Deduced amino acid sequence of *Cyp4a14*

The protein coding sequence of *Cyp4a14* was deduced from its predicted exon regions. The start of transcription of the *Cyp4a14* RNA was determined by primer extension and mapped to a T nucleotide 26 bases upstream of the putative start of protein translation. Thus, the cDNA sequence of the gene was deduced and the amino acid sequence of the protein predicted; these are presented in Figure 3.20.

```
  1   TCGACTCGAT CCAGAAACTA CTAACCATGG GTTTTTTTTT ATTTAGCCCT
                                 M  G   F  F  L   F  S  P          8
 51   ACAAGGTACT TGGATGGTAT CTCTGGGTTC TTCCAATGGG CCTTCTTGCT
      T  R  Y   L  D  G  I   S  G  F   F  Q  W   A  F  L  L       25
101   CAGTCTATTT CTGGTGCTGT TCAAGGCAGT CCAATTCTAC TTACGAAGGC
        S  L  F   L  V  L   F  K  A  V   Q  F  Y   L  R  R       41
151   AGTGGCTGCT CAAGACCCTC CAGCATTTCC CATGCATGCC TTCCCACTGG
        Q  W  L  L   K  T  L   Q  H  F   P  C  M  P   S  H  W     58
201   CTTTGGGGGC ACCATCTGAA GGACAAGGAG CTCCAGCAGA TTCTTATATG
        L  W  G   H  H  L  K   D  K  E   L  Q  Q   I  L  I  W     75
251   GGTAGAGAAA TTCCCAAGTG CCTGCTTACA GTGTCTCTCG GGGAGCAATA
        V  E  K   F  P  S   A  C  L  Q   C  L  S   G  S  N       91
301   TACGAGTCCT GCTTTATGAT CCTGACTATG TGAAGGTGGT TCTGGGGAGA
      I  R  V  L   L  Y  D   P  D  Y   V  K  V  V   L  G  R      108
351   TCAGATCCAA AGGCTTCTGG AATTTATCAA TTCTTTGCTC CCTGGATTGG
        S  D  P   K  A  S  G   I  Y  Q   F  F  A   P  W  I  G     125
401   TTATGGTTTG CTCCTGTTGA ATGGGAAGAA GTGGTTCCAG CATCGGCGGA
        Y  G  L   L  L  L   N  G  K  K   W  F  Q   H  R  R       141
451   TGTTGACTCC AGCCTTCCAC TATGACATCC TCAAACCCTA TGTCAAAATC
      M  L  T  P   A  F  H   Y  D  I   L  K  P  Y   V  K  I      158
501   ATGGCGGACT CTGTCAATAT AATGCTAGAT AAATGGGAGA AGCTTGATGG
        M  A  D   S  V  N  I   M  L  D   K  W  E   K  L  D  G     175
551   CCAGGACCAC CCTCTGGAGA TCTTCCACTG TGTTTCATTG ATGACACTGG
        Q  D  H   P  L  E   I  F  H  C   V  S  L   M  T  L       191
601   ACACTGTTAT GAAGTGTGCT TTCAGCTACC AAGGCAGTGT TCAGTTGGAT
      D  T  V  M   K  C  A   F  S  Y   Q  G  S  V   Q  L  D      208
651   GAAAATTCCA AGTTGTATAC TAAGGCTGTC GAGGATCTAA ACAACCTGAC
        E  N  S   K  L  Y  T   K  A  V   E  D  L   N  N  L  T     225
701   TTTCTTTCGC CTGCGGAATG CCTTTTATAA GTACAACATC ATCTACAATA
        F  F  R   L  R  N   A  F  Y  K   Y  N  I   I  Y  N       241
751   TGTCCTCTGA TGGACGTTTG TCCCACCATG CCTGCCAGAT TGCTCACGAG
      M  S  S  D   G  R  L   S  H  H   A  C  Q  I   A  H  E      258
801   CACACAGATG GAGTGATCAA GATGAGGAAG TCTCAGCTGC AGAATGAGGA
        H  T  D   G  V  I  K   M  R  K   S  Q  L   Q  N  E  E     275
851   AGAGCTGCAG AAGGCCAGGA AGAAGAGACA CTTGGATTTC TTGGACATCC
        E  L  Q   K  A  R   K  K  R  H   L  D  F   L  D  I       291
901   TCTTGTTTGC CAGAATGGAG GATAGGAACA GCTTGTCTGA TGAGGACCTG
        L  L  F  A   R  M  E   D  R  N   S  L  S  D   E  D  L     308
951   CGTGCAGAGG TGGACACATT CATGTTTGAG GGTCATGACA CTACAGCCAG
        R  A  E   V  D  T  F   M  F  E   G  H  D   T  T  A  S     325
```

```
1001   TGGAATTTCC TGGATTTTCT ATGCTCTGGC CACCCACCCT GAGCACCAAC
         G  I  S   W  I  F   Y  A  L  A   T  H  P   E  H  Q          341
1051   AGAGATGCAG AGAGGAGGTG CAGAGCATTC TGGGTGATGG AACCTCTGTC
       Q  R  C  R   E  E  V   Q  S  I   L  G  D  G   T  S  V          358
1101   ACATGGGACC ATCTGGGCCA GATGCCCCTA CACCACCATG TGCATCAAGG
        T  W  D   H  L  G  Q   M  P  Y   T  T  M   C  I  K  E         375
1151   AGGCCCTGAG GCTCTATCCA CCAGTAATAT CTGTGAGTGG AGAGCTCAGC
         A  L  R   L  Y  P   P  V  I  S   V  S  R   E  L  S          391
1201   TCACCTGTCA CCTTCCCAGA TGGACGCTCC ATACCCAAAG GTATCACAGC
       S  P  V  T   F  P  D   G  R  S   I  P  K  G   I  T  A          408
1251   CACAATTTCC ATTTATGGCC TACATCATAA CCCACGTTTC TGGCCAAACC
        T  I  S   I  Y  G  L   H  H  N   P  R  F   W  P  N  P         425
1301   CAAAGGTGTT TGACCCCTCT AGATTTGCAC CAGATTCTTC TCACCATAGC
         K  V  F   D  P  S   R  F  A  P   D  S  S   H  H  S          441
1351   CATGCTTATC TGCCATTCTC AGGAGGATCA AGGAACTGCA TTGGGAAACA
       H  A  Y  L   P  F  S   G  G  S   R  N  C  I   G  K  Q          458
1401   GTTTGCTATG AACGAGCTGA AGGTGGCTGT GGCCCTGACC CTGCTCCGCT
        F  A  M   N  E  L  K   V  A  V   A  L  T   L  L  R  F         475
1451   TTGAATTGCT GCCAGATCCC ACCAGGATCC CAGTCCCCAT TGCAAGACTT
         E  L  L   P  D  P   T  R  I  P   V  P  I   A  R  L          491
1501   GTGTTGAAGT CCAAGAATGG GATCCACCTG TGTCTCAAGA AGCTAAGATA
       V  L  K  S   K  N  G   I  H  L   C  L  K  K   L  R  *          508
1551   ATA
```

Figure 3.20 Deduced cDNA and amino acid sequence of *Cyp4a14*. An ATG triplet (underlined), at position +26, denotes the putative start of protein translation. An enlarged T nucleotide at position 1 denotes the start of transcription as determined by primer extension. Nucleotide numbering is given at the left hand side while peptide numbering is given at the right hand side of the figure. Sequence from exons 6 to 12 are from S. Kuo and D.R. Bell. Amino acids are indicated by their single letter codes. '*' refers to the stop codon.

Section 3.2.4.1 Comparison of amino acid sequence of *Cyp4a14*

When the genomic sequence of *Cyp4a14* was analyzed and compared with the rat CYP4A2 gene structure, it was found that the *Cyp4a14* exon positions were analogous to those in CYP4A2. The predicted open reading frame for *Cyp4a14* was found to encode a protein of 508 amino acids. The predicted amino acid sequences encoded by each exon were compared with corresponding amino acid sequences between CYP4A14 and the rat and mouse CYP4A proteins (Table 3.4).

Exon / Protein	CYP4A1	CYP4A2	CYP4A3	CYP4A8	CYP4A10	CYP4A12
1	58	87	85	54	61	-
2	66	85	87	69	59	-
3	73	75	** 93	73	67	-
4	91	* 98	* 98	88	91	-
5	80	95	95	80	80	-
6	57	82	84	67	63	-
7	83	94	97	81	86	-
8	90	* 97	* 97	86	92	95
9	84	93	93	84	84	89
10	82	82	86	73	82	68
11	85	85	85	85	85	85
12	93	* 95	* 96	95	93	-
Average	79	89	91	78	79	-

Table 3.4 Percent identity of peptide sequences by exon between *Cyp4a14* and the rat and mouse CYP4A genes. Amino acid sequences were deduced from individual exons of Cyp4a14 and compared with the corresponding amino acid sequences in rat and mouse proteins. "*' denotes peptide sequence of extremely high similarity (95% and over) between CYP4A14 and CYP4A2 and CYP4A3 proteins. '**' denotes an especially high peptide sequence similarity deduced from exon 3, between CYP4A14 and CYP4A3 only. Only exons 8 to 11 are sequenced in Cyp4a12 (Bell *et al.*, 1993).

A comparison of the amino acid sequence with other CYP4A genes shows that *Cyp4a14* is, on average, more similar to the rat CYP4A2 / CYP4A3 proteins (92 / 93 %), than it is to the murine *Cyp4a10* protein (77%; Table 3.4). Additionally, the amino acid sequences, deduced exon by exon, were compared with corresponding peptide sequences of rat and mouse CYP4A proteins. From the comparison above, amino acid sequences corresponding to exons 4, 8 and 12 were very highly conserved between CYP4A14 and CYP4A2 and CYP4A3 proteins. These same exons were also very highly conserved at the nucleotide level, as shown previously in Table 3.2. Additionally, deduced amino acid sequence corresponding to the translated exon 3 had a uniquely high similarity only between CYP4A14 and CYP4A3 proteins. Thus, in exon 3, CYP4A14 was significantly more similar to CYP4A3 protein than to the CYP4A2 protein, es-

pecially at the protein level. This was due to a distinctive 3 amino acid deletion in CYP4A2, but not in any other CYP4A protein.

For CYP4A1, CYP4A8 and CYP4A10 proteins, all twelve exons showed relatively low deduced amino acid sequence similarity with the corresponding exons of CYP4A14, except for exon 12. Exon 12 was the most highly conserved exon, encoding for a highly conserved cysteine residue and several others around it, all of which comprise the haem-binding domain which is an essential feature of all cytochrome P450 proteins.

Section 3.2.5 Distinction between murine CYP4A2 and CYP4A3 homologues

Comparison of the deduced amino acid sequences of *Cyp4a14* with the closely related CYP4A2 and CYP4A3 proteins revealed that in exon 3, CYP4A2 can be distinguished from CYP4A3 and CYP4A14 proteins by a 3 amino acid deletion after amino acid 113 of CYP4A2 but not CYP4A3 or CYP4A14. This corresponded to a 9-nucleotide deletion in the cDNA of CYP4A2, but not in CYP4A3 and CYP4A14 (Figure 3.21).

To examine if any of the λ phage clones isolated were similar to CYP4A2 or CYP4A3 in this region, DNA was extracted from phage clones which contain exon 3 and spotted onto Hybond N+ membrane, and after fixing, were hybridized with an oligonucleotide probe spanning a

```
CYP4A2        389  CCA AAG CCT             TAT
                   Pro Lys Pro             Tyr
CYP4A3        343  CCA AAG GCT TCT GGA ATT TAT
                   Pro Lys Ala Ser Gly Ile Tyr
Cyp4a14       356  CCA AAG GCT TCT GGA ATT TAT
                   Pro Lys Ala Ser Gly Ile Tyr
```

Figure 3.21 Scheme showing the region in exon 3 where CYP4A2 differs from CYP4A3 and *Cyp4a14*. Numbers denote the nucleotide positions of the respective sequences in the cDNA. The deduced amino acid for each protein is indicated beneath each codon.

corresponding region of deletion in the CYP4A2 homologue. The oligonucleotide, (5'-CCA AAG GCT **TCT GGA ATT** TAT-3'), which corresponds to nucleotides 356 to 376 in the *Cyp4a14*

cDNA, was used (with the nucleotides in bold the deleted region in CYP4A2 homologues). CYP4A2 equivalents would lack the nucleotides indicated in bold and would therefore not hybridize to the probe. As shown in Figure 3.22, all clones hybridized to the oligonucleotide, indicating an identical sequence in this region to the radiolabelled oligonucleotide. Thus, all the genomic clones containing exon 3 were identical to CYP4A3 in this region

Section 3.2.6 Conserved regions in *Cyp4a14* and non-CYP4 proteins

To better understand the structure/function relationship of P450 proteins, the three-dimensional structures of CYP101, CYP102 and CYP108 have previously been compared (Ravichandran *et al.*, 1993). The authors compared this alignment with other P450 proteins and reported that only 3 residues were absolutely conserved in the P450 superfamily of proteins. These were the Cys residue in the haem-binding region and the Glu and Arg residues in the K helix.

As shown in Figure 3.23, which shows an alignment of primary structures of P450 proteins, several common residues are present in all P450 proteins regardless of family. This agrees with the observations of Ravichandran and coworkers (1993) in the following ways : A Cys residue at position 454 of CYP4A14 is absolutely conserved. Surrounding this cysteine residue in CYP4A14 in the haem-binding region, Pro446, Phe447 residues and a Gly450 are also common. Additionally, Glu375 and Arg378 residues are conserved; these correspond to the same residues in the same helix as assigned by Ravichandran and coworkers (1993) for CYP102 and assigned by Edwards and coworkers (1989) for CYP2 and CYP17 families.

From Figures 3.23 and 3.24, regions 1 to 6 have previously been proposed to be key substrate binding areas (substrate recognition regions, SRS) in CYP2 proteins (Gotoh, 1992). Since then, extensive work on various CYP2 subfamilies have shown that virtually all residues identified as critical for the substrate specificity of CYP2 proteins fall within or near the putative SRSs proposed (eg, Straub et al., 1993). Here, the sites of substrate binding of CYP4A14 were inferred

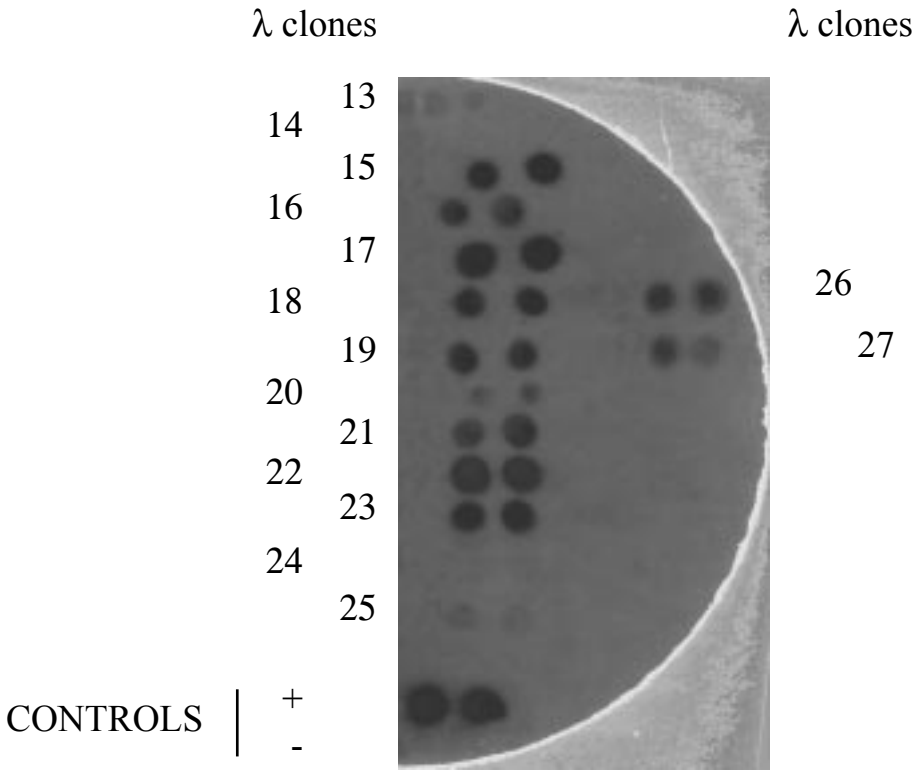


Figure 3.22 Dot hybridization of λ clones to detect a 3 amino acid deletion. λ clones were dot-blotted, in duplicate, onto Hybond N+ membrane, and hybridized with a radiolabelled oligonucleotide, **5′-CCA AAG GCT TCT GGA ATT TAT-3**' (which corresponds to nucleotides 356 to 376 in the *Cyp4a14* cDNA), overnight at 60°C and was then rinsed in 2X SSC at room temperature. The membrane was then washed twice in TMACl Washing Buffer at 60°C according to the method of Dilella and Woo (1987) before exposing to autoradiographic film at -80°C for 24 hours. The middle 9 bp of this oligonucleotide (in bold) are absent in the corresponding region in exon 3 of CYP4A2 but not in CYP4A3 and *Cyp4a14*. λ clones are indicated on the sides of the figure. Positive control used λ 12 while negative control used λ 4 which lacks exon 3. λ clones 14, 24 and 25 were false positives obtained during the library screen while λ 13 was previously shown to be similar to λ 4.

from the primary sequence alignment with the CYP2C4 protein. In site 4, two threonine residues located at positions 322 and 323 appear to be essential to all P450s with an Asp or Glu just N-terminal to it, as is a Gly319 residue.

It has been found preferable to compare secondary structure of proteins as this is often sacrificed in alignment of primary sequence (Edwards *et al.*, 1989) in favour of primary sequence homology. Thus, hydropathy plots of CYP2C4 and CYP4A14 were compared with each other to assess the accuracy of the assignments of the SRSs in CYP4A14 (Figure 3.24). CYP4A14 possesses a slightly longer N-terminal tail, but the hydropathy profile matched well generally with that of

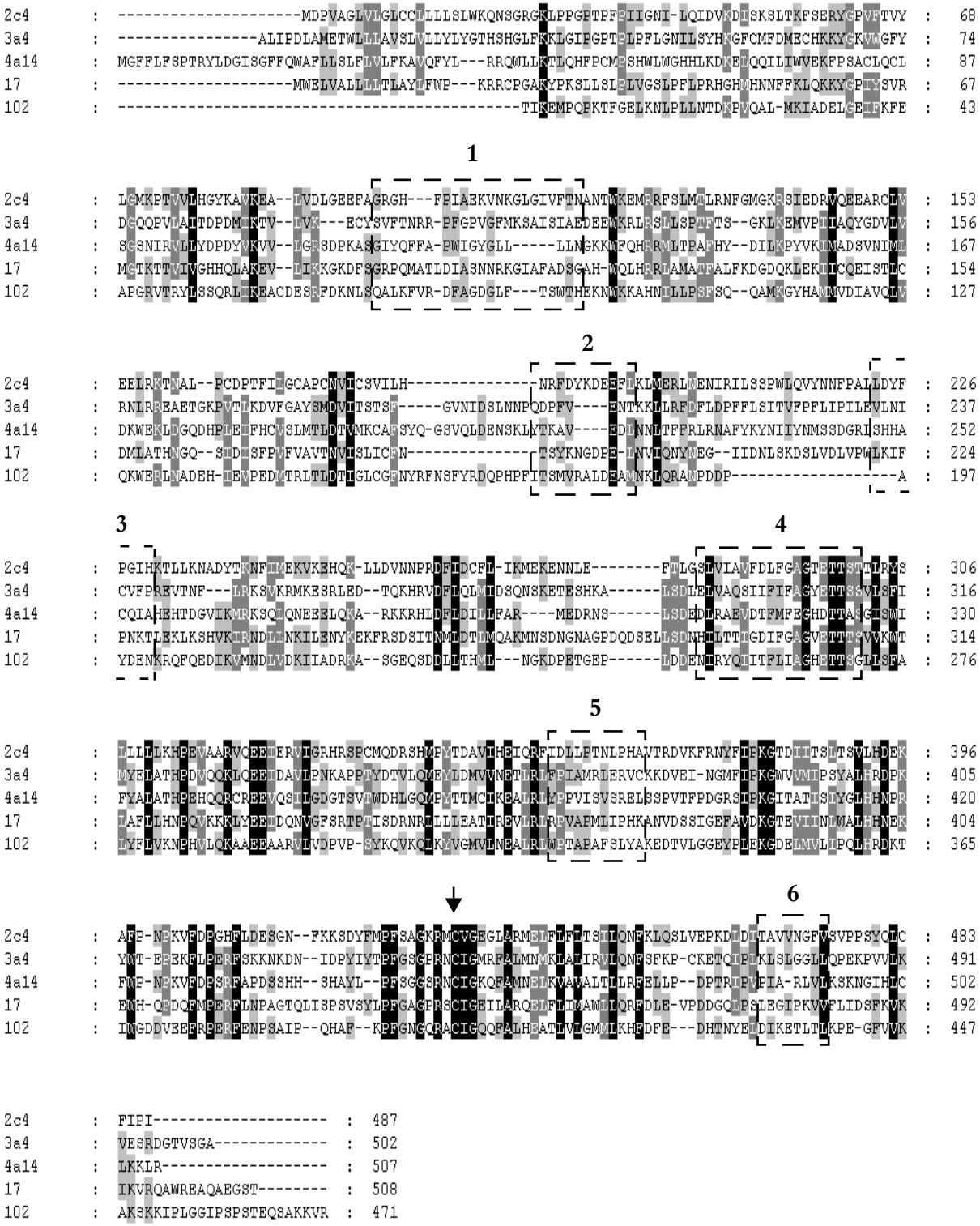


Figure 3.23 Alignment of cytochrome P450 proteins by primary sequence similarity. Deduced amino acid sequences were aligned using PILEUP and highlighted using BOXSHADE programmes. Identical residues have a black background and white text, conserved residues have a dark grey background, and similar residues have a light grey background; a dot represents a space introduced into the alignment. Substrate recognition sites (SRS) have been proposed (Gotoh, 1992) and similar regions are deduced in CYP4A14 by reference to CYP2C4. These regions are enclosed in a dashed box labelled (1 to 6). A single arrowhead points to the conserved Cysteine residue in the haem-binding domain at position 454 in CYP4A14. CYP2C4 (2c4), CYP3A4 (3a4), CYP4A14 (4a14), CYP17A (17) and CYP102 (102) residue positions are indicated to the right of the alignment.

CYP2C4. It has been proposed that the α-helical domains of the cytochrome P450 families are

well conserved, including the CYP2 family (Edwards *et al.*, 1989). This comparison suggests that this is the case, and extends previous data suggesting that the three-dimensional structure of cytochrome P450 proteins is fairly well conserved and this is applicable to CYP4A14.

Section 3.2.7 Putative substrate binding regions of *Cyp4a14*

As shown in Figure 3.23, putative substrate recognition sites (SRS) of CYP4A14 were inferred by first aligning its deduced primary amino acid sequence with that of CYP2C4. Hydropathy plots of CYP2C4 and CYP4A14 matched each other well (Figure 3.24), and the SRS positions tentatively assigned in CYP4A14 appear to be in a similar three-dimensional environment to those in CYP2C4. Thus, taken together with the primary sequence alignments, it is likely that these regions in CYP4A14 could have a similar role in substrate recognition. But, differences between the two proteins also exist - most notably, in site 4, where CYP2C4 is much more hydrophobic than CYP4A14, and the N-terminal end of site 1 in CYP2C4 is distinctly more hydrophobic than in CYP4A14. These differences in hydrophobicity are not surprising since CYP2 and CYP4 proteins have different substrate requirements. Thus, sites 1 and 4 could play a role in the substrate specificities of the two proteins.

Section 3.2.8 Sequence alignment of *Cyp4a14* and other CYP4 proteins

A comparison of the peptide sequence of CYP4A14 with all vertebrate members of the CYP4 family is shown in Figure 3.25. This alignment indicates there are many stretches of amino acids which are very similar in CYP4 proteins, especially in and around the six tentatively assigned sites 1-6. In site 1, motif PWIGYGLLLL is absolutely conserved in the CYP4A proteins and remained highly similar in the CYP4B, CYP4F, CYP4C and CYP4D subfamilies. From Figures 3.23 and 3.24, this region was highly divergent between the families. This specific region might play a role in the different substrate specificities of different families of P450 enzymes. Just C-terminal to site 1 is another motif that is highly conserved in CYP4 proteins and previously unreported. This motif, HR RMLLTPGFHYDIL, is located between residue positions 139 to 152

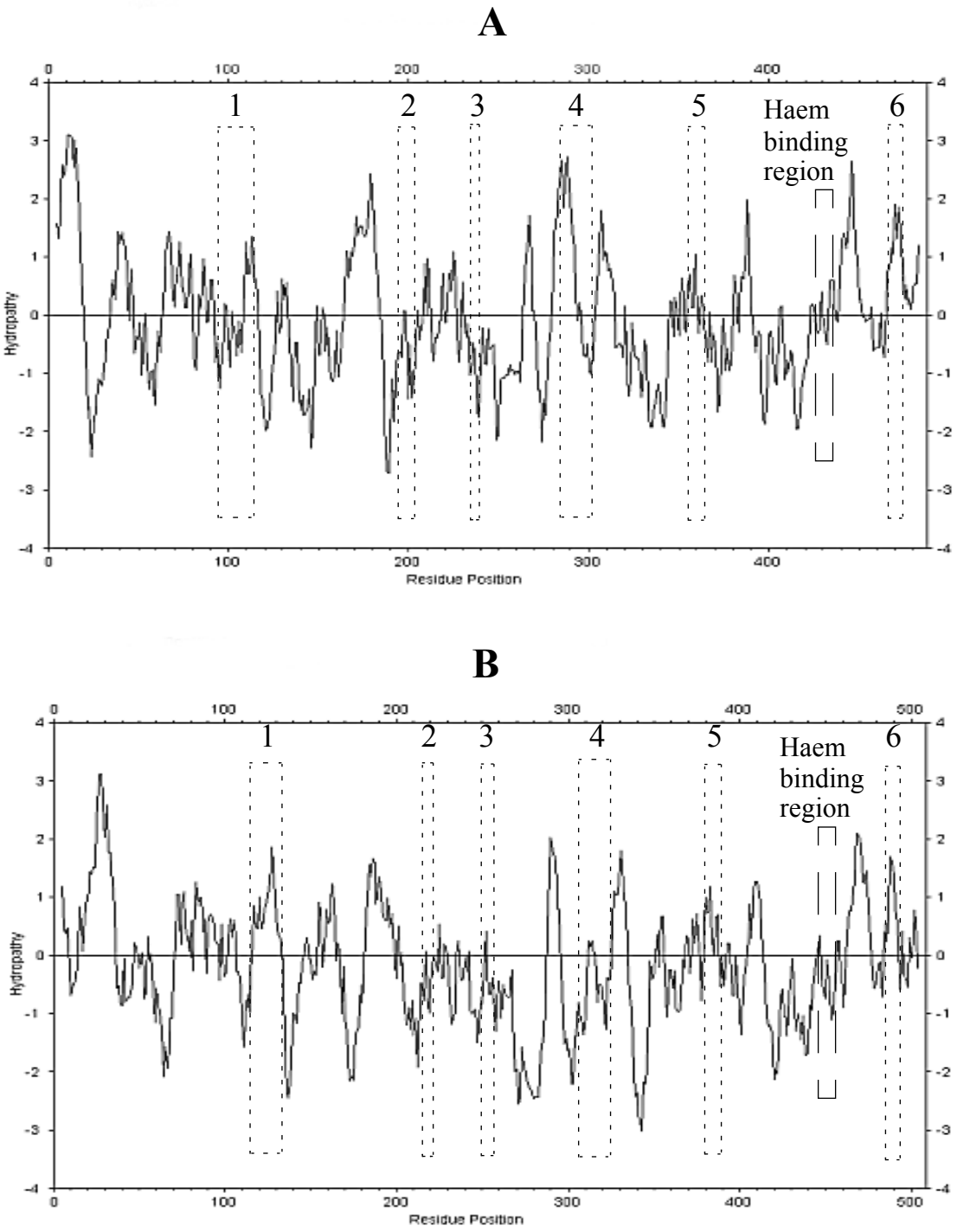
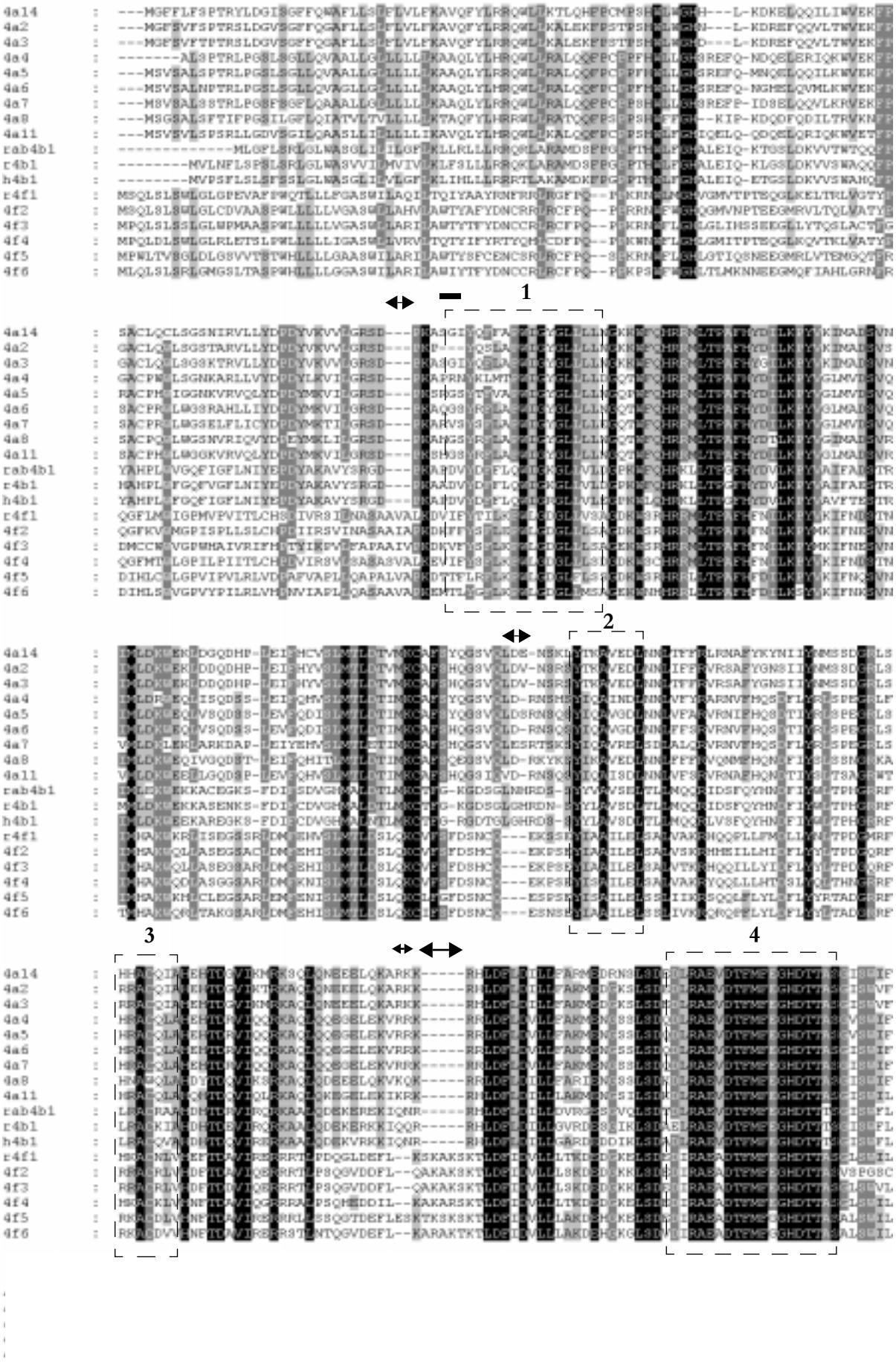


Figure 3.24 Hydropathy profiles of CYP2C4 (A) and *Cyp4a14* (B) proteins obtained with the parameters of Kyte and Doolittle (1982). (A) Regions 1 to 6 are putative substrate recognition sites (SRS) identified by Gotoh (1992) in CYP2C4. (B) Regions 1 to 6 are putative substrate binding sites inferred from primary sequence homology with CYP2C4 obtained in Figure 31. SRS 1 to 6 are boxed with dashed lines, while the haem binding domain, which contains a conserved cysteine residue common to all P450 proteins is boxed with longdash lines.

and encoded by exon 4 in CYP4A14. As noted, exon 4 is one of several exons highly conserved at the nucleotide level in CYP4A proteins. In substrate recognition site 4, a 16 amino acid motif,


1
2
3
4

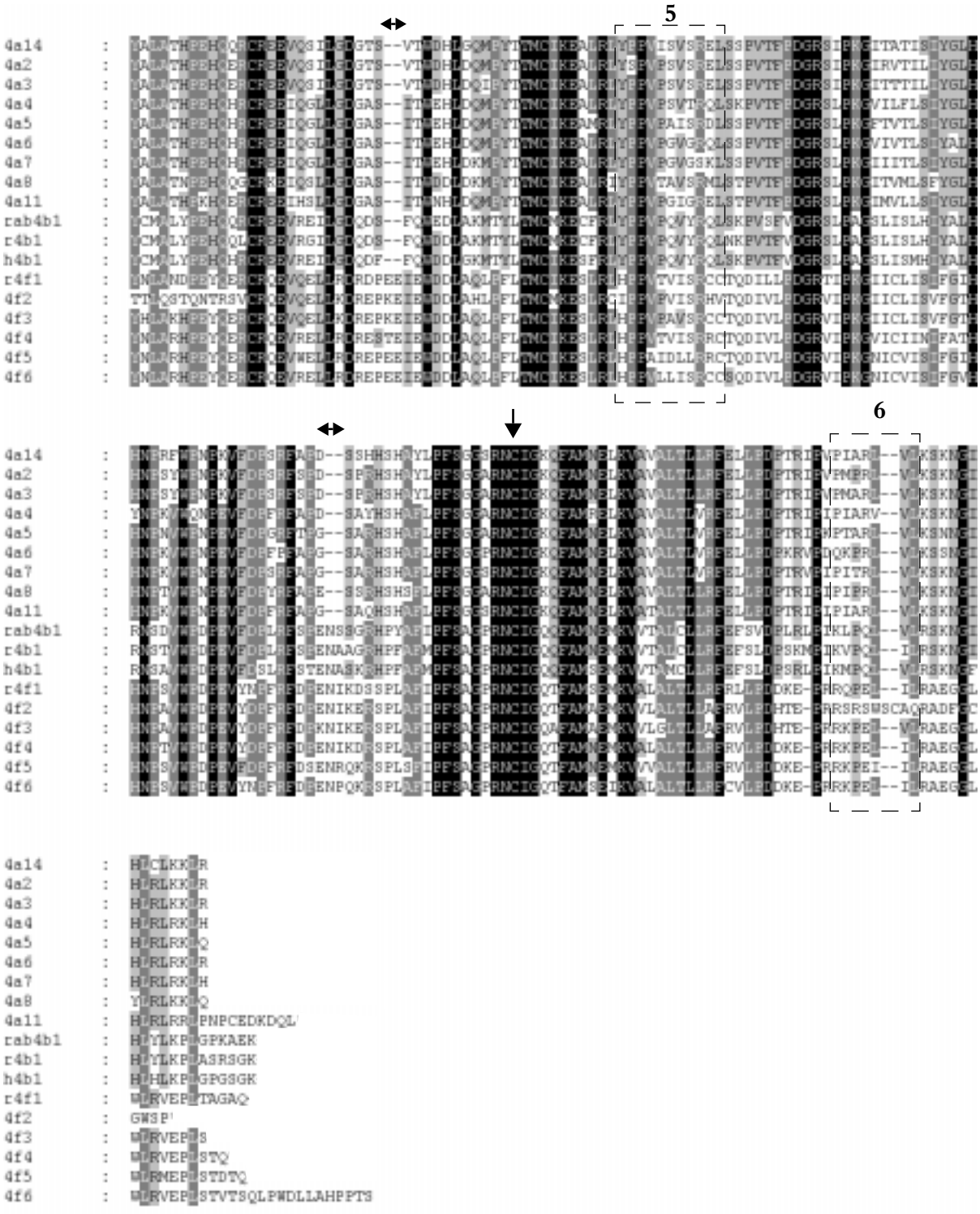


Figure 3.26 Alignment of CYP4 proteins by primary sequence similarity. Deduced amino acid sequences were aligned using PILEUP and highlighted using BOXSHADE. Regions enclosed in a dashed box denote putative substrate binding sites (1 to 6) and were inferred by aligning CYP4A14 with the published substrate recognition sites (SRS) of CYP2C4 as deduced by Gotoh (1992). Double arrow heads indicate regions where CYP4A and CYP4B primary protein sequences differed from CYP4F. A single arrowhead points to the conserved Cysteine residue at position 454 in CYP4A14. A bold bar indicates the region of a 3 amino acid deletion in CYP4A2 but not in CYP4A3. Protein sequences of YP4A14 (indicated by "4a14") , CYP4B1 (4b1), CYP4F1 (4f1) to CYP4F6 (4f6), CYP4C1 (4c1) and CYP4D1 (4d1) to CYP4D2 (4d2) are indicated to the left of the alignment. Prefixes m, r, h, c and dm denote rabbit, human, cockroach and *Drosophila* proteins respectively.

LRAEVDTFMFEGHDTT, was most notably non-divergent between CYP4A and CYP4B proteins. From Figure 3.23, the motif (E/D)TT within this sequence was conserved in all P450 proteins. This region is thought to play a role in oxygen activation and would thus be a key requirement in all P450 enzymes (Raag *et al.*, 1989; Raag *et al.*, 1991). In the haem-binding region, the motif RNCIG appears to be highly conserved in the CYP4 protein and in non-CYP4 proteins (in Figure 3.23). Indeed, this region has been proposed to be a signature for the cytochrome P450 superfamily (Gonzalez, 1989).

CYP4A2 and CYP4A3 are very similar to each other but can be distinguished between each other by a 3 amino acid deletion in CYP4A2 but not in CYP4A3, indicated by a bold bar.This deletion, indeed was not present in any other CYP4 proteins and is thus unique to CYP4A2. There are several regions in the protein sequences, as indicated, where CYP4A and CYP4B proteins differ from CYP4F proteins, either through insertions or deletions of 2 to 5 residues. Indeed, phylogenetic analysis of the CYP4 proteins suggest that CYP4A proteins are more similar to the CYP4B proteins than to the CYP4F proteins and vice versa (Figure 3.27).

Section 3.2.9 Phylogenetic analysis of *Cyp4a14*

Human, mouse and rat proteins of CYP4A, CYP4B and CYP4F subfamilies and the invertebrate CYP4C and CYP4D proteins were aligned and analyzed using CLUSTALW, which utilizes the Neighbour-Joining Method of Saitou and Nei (1987). The rat CYP2C4 was used as the outgroup, a sequence that was distinct from all the other proteins and should branch to the outside of the tree. From Figure 3.27, it appears that CYP4A, CYP4B and CYP4F proteins form three distinct clusters. Within the CYP4A cluster, CYP4A14 is most related to CYP4A2 / CYP4A3 proteins while CYP4A1 is most similar to the CYP4A10. In the CYP4B cluster, mouse CYP4B1 and rat CYP4B1 are more similar to each other than to the human CYP4B1. In the CYP4F cluster, the rat proteins are also quite distinct from the human proteins. Again, as already

noted from the primary sequence comparisons, the CYP4A and CYP4B subfamilies appear to be more similar to each other than to the CYP4F subfamily.

The CYP4A, CYP4B and CYP4F families consistently form 3 separate clusters in other methods of analysis like PHYLIP and with CLUSTALW these clusters were shown to be absolutely consistent by BOOTSTRAP analysis (not shown), of Felstenstein (1985). However, the invertebrate CYP4 proteins, CYP4C1, CYP4D1 and CYP4D2 were less confidently placed here and in various other methods of phylogenetic analyses (not shown). In Figure 3.27, these proteins have been placed more distant to the mammalian CYP4 proteins than the designated 'outgroup' CYP2C4, which was expected to be the most distant member. Using PHYLIP and PROTPARS, these invertebrate CYP4 proteins were placed more closely with the mammalian CYP4

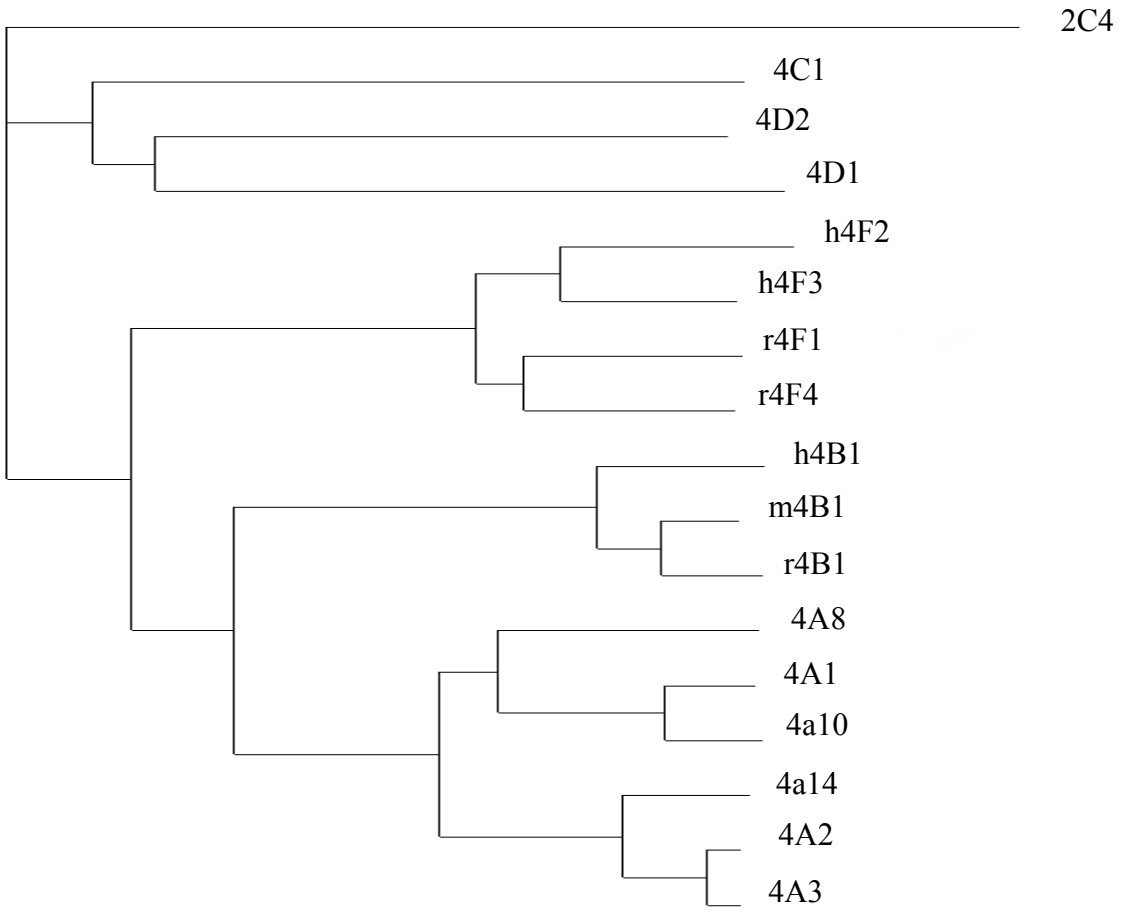


Figure 3.27 Phylogenetic analysis of CYP4 proteins. CYP4 deduced amino sequences were aligned and analyzed by CLUSTALW using the Neighbour-Joining Method (Saitou and Nei, 1987) with CYP2C4 as the outgroup.. 'h', 'm' and 'r' denote human, mouse and rat CYP4B1 proteins respectively.

proteins, leaving the outgroup most distant in the trees which is the 'expected' outcome (not shown). Therefore, these analyses show that there is no single method of analysis that is 'correct', and that distantly related CYP4 proteins like the mammalian and invertebrate CYP4 proteins may be less confidently placed relatively to each other in a phylogenetic tree.

Section 3.3 Cloning of *Cyp4a12*

Section 3.3.1 λ clone 16

λ 16, a gift from D.R. Bell, was the second genomic clone isolated when the genomic library was first screened with riboprobes transcribed from *Cyp4a10* and *Cyp4a12* cDNA clones (D.R. Bell). In order to map the phage clone, large scale DNA was extracted from λ 16 and was restriction digested with *Xba* I (which releases the genomic DNA from the vector arms), singly and in a double digest with *Bam* HI, and blotted onto Hybond N+ membrane and hybridized to the rat partial CYP4A1 cDNA riboprobe pIV2. (Fig 3.28).

Section 3.3.2 Subcloning and sequencing of λ 16

DNA was prepared by the large scale method from phage λ 16 and restriction digested with *Xba* I enzyme. Before ligation, an aliquot of the restriction digest was run out on a 0.6 % agarose gel to make sure the restriction digest was to completion. All *Xba* I fragments were subcloned into pGEM7 by ligation overnight at 4 ^{o}C. Selection for clones was as described in Materials and Methods. To prepare for automated sequencing, the 6.5 kb and 3 kb *Xba* I subclones were purified by the protocol supplied by Qiagen and sequenced by oligonucleotide-directed automated sequencing.

Sequence analysis of the 6.5 kb and 3 kb *Xba* I fragments showed exons 10, 11 and 12 were contained in these two subclones (done by Sharon Kuo). To identify the gene represented by this λ clone, the sequences of the exon regions were analyzed in detail : the high sequence similarity with the published rat CYP4A8 cDNA sequence (Stromstedt *et al.*, 1990) indicated that λ 16 was

a genomic clone of the murine *Cyp4a12* gene, the mouse homologue of the CYP4A8. A map of λ 16 is shown in Figure 3.29).

Section 3.3.3 Strategy for the rescreening of the genomic library

Cyp4a12 gene expression has been shown in this laboratory to be regulated in a male-specific fashion, with high level expression in male liver and kidney and low levels in females (Bell *et al.*, 1993; Heng *et al.*, 1997). It would thus be interesting to examine the regulation of this gene by

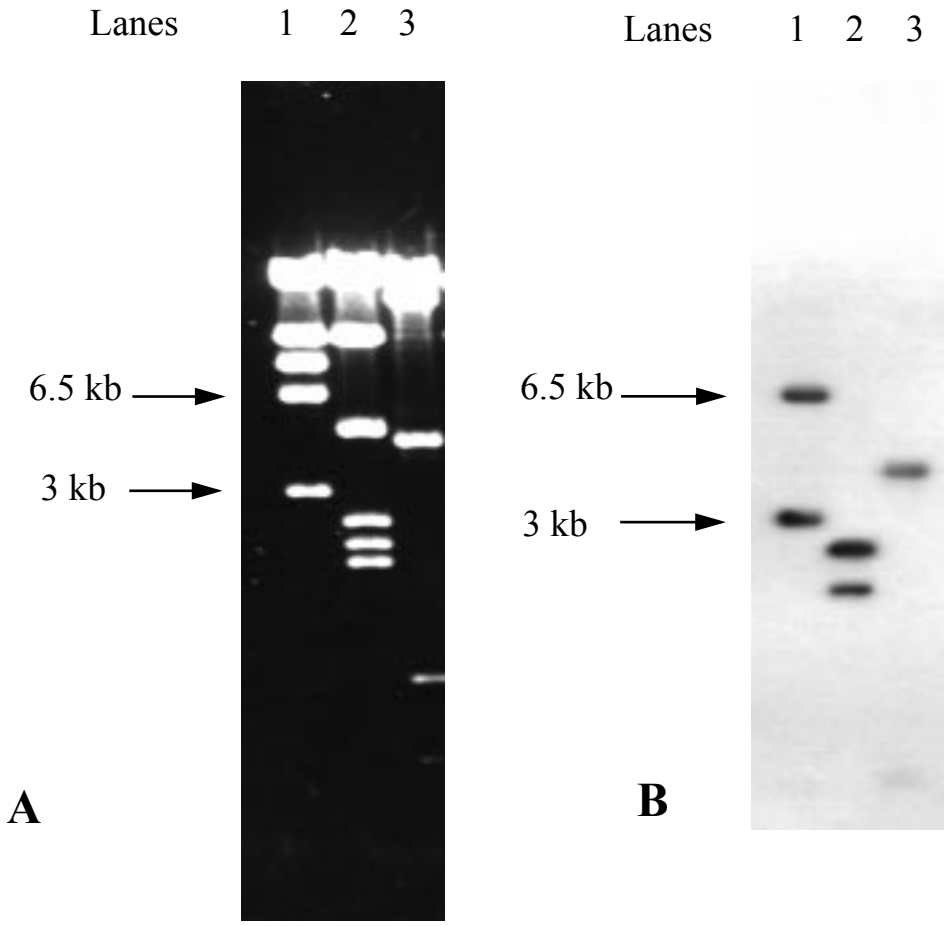


Figure 3.28 Mapping of λ 16 clone by restriction digest (A) and Southern hybridization (B). Large scale phage DNA was extracted as described in Materials and Methods, restriction digested and run on a 0.7% agarose gel to separate the fragments. To identify the fragments which hybridize to the rat CYP4A1 partial cDNA riboprobe pIV2 (corresponding to nucleotide 900 to end), gel was transferred onto Hybond N+ membrane as described and hybridized overnight with the riboprobe at 65°C in 40mM sodium phosphate (pH 7.2) containing 1% BSA, 7% SDS and 1mM EDTA. Membrane was then washed three times in 40mM sodium phosphate buffer (pH 7.2) containing 0.1% SDS at 65oC before exposing the membrane on X-ray fim at -80°C for 48 hours. Lane 1 : *Xba* I digest, Lane 2 : *Xba* I + *Bam* HI digest, Lane 3 : *Eco* RI + *Bam* HI digest. Arrows indicate the two *Xba* I fragments which hybridized to the riboprobe.

studying the promoter region of the gene in detail. The intention was then to screen the genomic library to obtain an overlapping phage clone that would contain the 5' flanking region of *Cyp4a12*.

λ 16:4 was chosen as the *Cyp4a12* subclone to be used for riboprobe synthesis for the following reason : from the sequence analysis, this subclone was found to be the most 5' clone of λ 16 as it contained the T3 promoter sequence from the left λ vector arm, followed by exon 10 and then exon 11 further downstream. The intron preceding exon 10 was only sequenced on the top strand so as to 'walk' to exon 10 (about 2kb away). However, for the purposes of designing a riboprobe for library screening, only 500 bp of *Cyp4a12* intron sequence immediately 3' of the T3 sequence in λ 16:4 were fully sequenced.

Sequence analysis of this 500bp region of λ 16:4 against the databases at EMBL by the FASTA programme of Pearson and Lipman (1988), revealed this stretch of nucleotides contained repetitive sequences, with up to 90 % identity in a 460 bp overlap with the mouse L1Md-9 repetitive

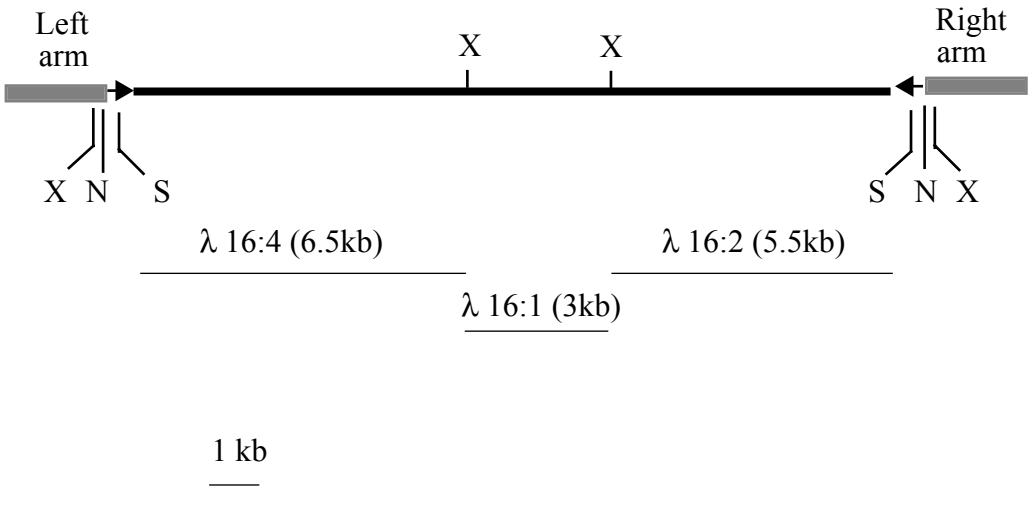


Figure 3.29 Map of the λ 16 clone. Genomic DNA, indicated by a thick, bold line, was mapped with *Xba* I and subcloned into pGEM7 vector. λ 16 *Xba* I subclones λ 16:4, λ 16:1 and λ 16:2 are indicated beneath the phage map. Only the *Xba* I, *Not* I and *Sac* I sites within the polylinker of the left arm (20kb) and right arm (9kb) are indicated. The arrowheads indicate the position of the T7 (right arm) and T3 (left arm) bacteriophage promoters. X : *Xba* I, N : *Not* I, S : *Sac* I. The 5' end of the gene is on the left hand side of the diagram. The vector arms are not drawn to scale.

sequence, for example. This meant that a 500bp riboprobe derived from this region of λ 16:4 would hybridize with many other unrelated genes which might have such a repetitive sequence. In fact, the genomic library was screened with a riboprobe transcribed from this region and an inappropriately high number of clones (at least 800) was identified in a primary screen. This probe was thus not suitable for screening the library to isolate *Cyp4a12* clones.

Section 3.4 Cloning of *Cyp4a10*

Section 3.4.1 PCR Cloning

Cloning of the promoter region of *Cyp4a10* was undertaken using Clontech's PromoterFinder DNA Walking Kit. Primary PCR was performed using an adaptor specific primer AP1 (top strand, 5'-GTA ATA CGA CTC ACT ATA GGG C-3'), supplied by Clontech, and a gene-specific primer GSP1 (bottom strand, 5'-GCA GCA GGA GCA GAC CGA GCA CAG A-3') derived from exon 1 of the published *Cyp4a10* sequence (Henderson *et al.*, 1990), corresponding to nucleotide positions 100 to 75. After running out a tenth of the total volume on a 2% agarose gel to visualize the amplified fragments (not shown), all five PCR reactions and the positive and negative control reactions were then diluted 1 in 20 in UHP water and subjected to secondary PCR confirmation using a Clontech-supplied, nested adaptor specific primer AP2 (top strand, 5'-ACT ATA GGG CAC GCG TGG T-3') and gene-specific primer GSP2 (bottom strand, 5'-GGG TTG GGC TTA GAG CAG AGA CAC TC-3') which corresponds to nucleotide positions 37 to 12 of the *Cyp4a10* cDNA. Both primary and secondary PCR reactions were performed at the specified annealing and extension temperature of 67 ^{o}C. The secondary PCR reaction, which confirmed products obtained from the primary PCR with a second, upstream PCR primer belonging to *Cyp4a10*, were run on a 1.2% agarose gel (Fig 3.30) and the longest PCR product, a 1.4 kb fragment was chosen for further analysis. The fragment was then excised from the gel, purified using the gene-clean method and cloned into pGEM-T, designated pGT14 and sequenced.

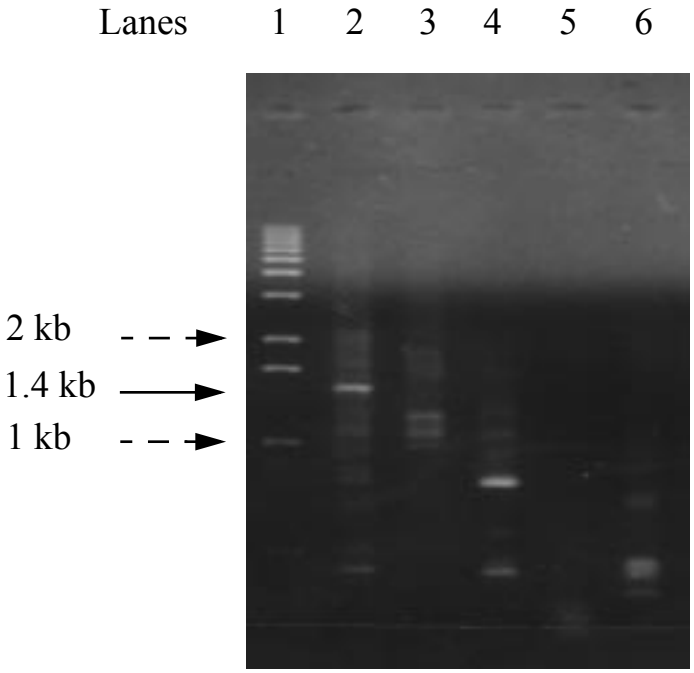


Figure 3.30 Identification of *Cyp4a10* secondary PCR products. Primary PCR reactions from all five libraries were diluted 1 in 20 and subjected to secondary PCR confirmation of bands as detailed in Materials and Methods. 5µl of the 50µl secondary PCR reaction volume were run on a 1.2% agarose gel. The arrow points to a 1.4 kb PCR fragment identified in Lane 2. This fragment was then gel-excised and cloned into pGEM-T vector and named pGT14. Secondary PCR reaction used gene-specific primer GSP 2 (bottom strand, 5'**GGC TTG GGC TTA GAG CAG AGA CAC TC**-3'), corresponding to positions 37 to 12 of the *Cyp4a10* cDNA, and Adaptor primer AP2 (top strand, 5'-**ACT ATA GGG CAC GCG TGG T**-3') according to the protocol at the recommended annealing/extension temperature of 67°C. Broken arrows point to two sized fragments in the ladder. Lane 1: kb ladder; Lane 2 : *Eco*R V library; Lane 3 : *Sca* I library; Lane 4 : *Dra* I library; Lane 5 : *Pvu* II library; Lane 6 : *Ssp* I library. Positive control (performed with the *Dra* I library with positive control primers and produced the expected size) and negative control reactions are not shown on this diagram.

Section 3.4.2 Sequence analysis

The PCR clone, named pGT14, containing 1.4 kb of the 5' flanking region of *Cyp4a10*, was sequenced completely by automated fluorescent sequencing. BESTFIT analysis of the DNA sequence of pGT14 revealed highest sequence identity with the rat CYP4A1 and was entirely distinct from the other CYP4A genes in this region, including *Cyp4a14*, consistent with the fact that *Cyp4a10* is the murine homologue of CYP4A1. The pGT14 and the corresponding region in CYP4A1 show 86 % identity with each other. Sequence analysis of pGT14 also demonstrated the presence of a 19 bp element and a GATTTA box upstream of exon 1 (Fig 3.31). Both motifs are highly conserved in CYP4A1 and CYP4A2 genes (Kimura *et al.*, 1989a). PPREs comprise

a AGGTCA direct repeat spaced by one nucleotide, which have been implicated in the transcriptional activity of a CYP4A gene (eg. Aldridge *et al.*, 1995). No such binding site exists for the *Cyp4a10* in this part of the 5' flanking region. However, other putative binding sites are located for transcription factors like Sp1 and C/EBP.

```
                                                        -1400
ACTATAGGGC ACGCGTGGTC GTTTGAATAT GCTTGGCCCA GGGAGTGGCA
            Sp1         C/EBP
CTATTTGGAG GCGTGGCCTT GTTGCAAGAA CTGTGCCACT GTGGGCATGG
                                                 -1300
TCTTTATGAC CCTCATCCTA CCTTCCCTAG AAGCCAGTGT TCTCCTATTT
                                                      Sp1
GCCTTCAGAA CAAGATGTTG AACTCTCAGC TCCTCCTGTA CCATGCCTGC
                                                 -1200
CCAGATGCTG CCATGTTCCC TCCTTGATGA TAATGGACTG AACCTATTAA
ATTTTGTTCT TTTACAAGTT GCCTTGGTCA TGGTATTTGT TCACAGCAGT
                                                   -1100
AAAACCCTAA ATAAGAAAAG CTATGTAGCT GAATAATAGC TGAAGCTTTC
TAGTAGTTAT TGAGAAATAA CATGGGCAAC CCAAGTACAG TAAATTAAAG
 Sp1                                               -1000
GGCCCCATTT AGATTCAATT GATTTGGTCC TGTTGGAGAC ACACTATATA
AAAGCTCTCA CAAGTCCAAG ACAATAATTC TAAAACCAAG AGAAGTATGT
                                                   -900
CAAGTCCTAT ATAAGCAAGG CCTTGGTAGA CTAACAGTAG ATTCCACAGC
AGAAACTTTT CAAACATGGA TTAAACTCTC CAAGTAAAAG AGTTAAATTG
                          c-Myc                    -800
GCTGAATAGA TTTTTTTAAT GAAATTCAAC TGTAGGCTTC TACAGAAAAA
AAAAAACCCA CAAGAAGGTG CCGAGGAATG GAAAGAAACC CATTAGCTCC
                                                   -700
CCCAAAGCAA AACAAGCACA AGAACCTTTG GCACCCCCAG GCATGGACTG
CCCATGAGAC TGGCCTACCT AGATGGATAA GAGCAGCAGT GCCTAGATGT
                                                   -600
TCTCCAGTCT GAATCGGCAA CCACTCACCA TCCATTTGGC ATGGACAATT
                                                    AP1
CAAAAAACTA TTAGATACTG GATCTGCCAT GGAAACCCAA GTAACTCACA
                     Sp 1                          -500
AAAACAACAA ATAAAGCTAT GGGGCCTGAG AAAGAGCCAG TTTGTGAGGA
                                  TFIID
TCCCCATATC ACACAGGACT TCAACAAGAT ATTTATTTTT ATATGACAAC
                                                   -400
CTTGTTTGAT GTTTAATAAT TTAATAAAAC TTTGCAAGAA AAGATTAAAA
ATAAGTGGAT GAATATTTAA GGAGATTATA ATTGTTCTGC TTAATTTATG
                                                   -300
CGTGCAGTAC CCAGGTGTTG AAATATCACA CTACATCGAT AAGTGTAGAC
AGCGAATACA ATGAAAAGGA AAATAAAAGG GAAACAACTC AATGCTCTTG
                          Sp1                      -200
ACTCGCACAA CGAAAAGGAA AGTGGGCAGA GGTAGGCACT GACAGCTGGA
AAAAAAAATG TTGTCACTGA TTTTTCAAGG GTTAATGATT CCTGTTTATT
               1/2 GR                              -100
TATATTTAAC CAACTGAACT ATGAGATCAA GATTGGTCAG CCTCACCCTT
   Sp 1
CTCCCTCCCA CAGGTAGGCG GGCAATCCTT ACAGGGATTT AGGCAGGAGG
```

+1

TAGAAAGGAT TATTGAAGAG TGAGGACCTC AACAAAGGTG AACTAT(G)

Figure 3.31 Sequence of the *Cyp4a10 5'* flanking region. DNA was cloned by PCR using Nucleotide sequence from the Cyp4a10 clone, pGT14, was analyzed for putative transcription factor binding sites by utilizing the Transcription Element Search Software (TESS) using the *Transfac* database on the WWW. Putative transcription factor binding sites are overlined. The putative start of protein translation begins at the single-underlined ATG triplet (Henderson*et al*., 1990), , denoted by +1. The conserved 19bp element (indicated in bold) and the GATTTA box (double underlined) are located 94 and 59 nucleotides upstream of the start of translation respectively

Chapter 4 DISCUSSION

The cytochrome P450 4A subfamily (CYP4A) was originally discovered as fatty acid ω-hydroxylating enzymes which are induced by peroxisome proliferating chemicals in the rat (Gibson *et al.*, 1982; Hardwick *et al.*, 1987). These proteins are rapidly induced and are now believed to be crucial for peroxisome proliferation in rat liver (eg Kaikaus *et al.*, 1993). Besides being involved in peroxisome proliferator-induced hepatocarcinogenesis, CYP4A proteins are also believed to be important in normal lipid and eicosanoid metabolism and the regulation of renal function.

To date, thirteen CYP4A members are listed (Nelson *et al.*, 1996). Four genes have already been identified in the rat : CYP4A1 and CYP4A2 are cloned as genomic sequences (Kimura *et al.*,1989a), while CYP4A3 (Kimura *et al.*, 1989b) and CYP4A8 (Stromstedt *et al.*, 1990) are cDNA clones identified. While a great deal is known about the rat and rabbit *CYP4A* genes, relatively little is understood about the *Cyp4a* genes in the mouse. To date, only two cDNA clones have been isolated : *Cyp4a10* (Bell *et al.*, 1993; Henderson *et al.*, 1994) and *Cyp4a12* (Bell *et al.*,1993). It is also not known if other members of this subfamily exist in the mouse. Thus, this genomic cloning project was undertaken to study the *Cyp4a* gene subfamily in detail in the mouse and, as a prelude to future gene-targetting studies, to take advantage of the widely utilized techniques in murine genetic manipulation.

For cloning the genomic *Cyp4a10*, two separate gene-specific primers derived from exon 1 of the published cDNA sequence of *Cyp4a10* (Henderson *et al.*, 1994) were used in a genomic PCR walking strategy to identify and confirm *Cyp4a10* genomic fragments. After secondary PCR reaction, various genomic fragments were isolated (Figure 3.30). However, only the longest DNA fragment, a 1.4 kb fragment was chosen for cloning and analysis. This fragment was amplified in the first PCR and it was then important, for it to be confirmed by re-amplification in secondary PCR using another primer derived from the sequence of *Cyp4a10*. The 3' end of

this PCR clone, designated pGT14, shared an identical nucleotide sequence with the 5' end of the cDNA sequence of *Cyp4a10*. Additionally, pGT14 showed very high identity with the corresponding region in its homologue, the rat CYP4A1 (86%) but not with any other CYP4A gene. These observations were essential to demonstrate that pGT14 is part of *Cyp4a10* and contains the 5' flanking region of the gene.

When a mouse genomic library was screened using riboprobes transcribed from the *Cyp4a10* and *Cyp4a12* cDNA of Bell and coworkers (1993), two independent and distinct phage clones were initially isolated. One, λ clone 16, was found to be a genomic clone representative of the 3' half of *Cyp4a12* (Figure 3.28) : only regions corresponding to exons 10, 11 and 12 were sequenced and these coding regions were identical in sequence to the corresponding regions of the reportedcDNA sequence of *Cyp4a12* (Bell *et al.*, 1993).

In order to obtain genomic clones spanning the whole of *Cyp4a12*, the genomic library had to berescreened for a clone that covered the 5' region of the gene. However, this was unsuccessful due to the following reason. Initially, only 500bp of genomic sequence on the 5' most end of λ 16 was sequenced to be used as a probe for the hybridization. Unfortunately, this region was found to contain repeated sequences and when compared against the databases, was highly homologous to many other genes containing such a repetitive sequence. This meant that a probe derived from this sequence would hybridize with many other irrelevant genomic sequences. Indeed, a hybridization experiment using this sequence as a riboprobe confirmed this suspicion; aninappropriately high number of positive clones was identified in the genomic library. Therefore, under the circumstances, the 5' half of *Cyp4a12* was not isolated from the library. Instead, attention was turned to the other clone isolated from the genomic library.

The second phage clone isolated from the original library screen, λ clone 4, was distinct from the existing murine *Cyp4a* genes and thus represented the 3' half of an entirely novel murine*Cyp4a*

member (Figures 3.3 & 3.15). A riboprobe was derived from the 5' most subclone of λ 4, and used to re-screen the library in order to obtain the whole of the gene. λ clone 12 was subsequently isolated which was found to contain exon 1 of this novel *Cyp4a* gene (Figures 3.9& 3.15). This novel gene has since been designated *Cyp4a14*. This genomic cloning project has therefore identified a third and previously unknown member of the murine *Cyp4a* subfamily.

Cyp4a14 expression was found to be very highly induced by peroxisome proliferating chemicals (Heng *et al.*, 1997). In order to determine if any unique protein binding sites are located in the promoter region of *Cyp4a14* which could play a role in the high inducibility of this gene, an attempt was made to "walk" further upstream of the gene by rescreening the library. The re-screening of the library was performed using a riboprobe derived from λ clone 12 and containing exon 1 of *Cyp4a14*. Unfortunately, all clones isolated and characterized were identical to λ clone 12 and no new clones were isolated. This could be due to poor representation of the promoter region of *Cyp4a14* in the genomic library.

In order to get around this problem, a PCR-based approach was then undertaken : a primer was derived from the exon 1 sequence of *Cyp4a14*, and was used for primary PCR, together with an adaptor-specific primer supplied. After confirmation by secondary PCR reaction, this time using a second gene-specific primer upstream of the first primer and within the untranslated region of exon 1 of *Cyp4a14*, and a second, nested adaptor-specific primer. The longest PCR product, a 1.2 kb genomic fragment was then cloned, designated pGT12 and analyzed (Figure 3.10). This fragment was confirmed to be the 5' flanking region of *Cyp4a14* as it was amplifiable using two different primers which were specific for the gene. Also, analysis of the nucleotide sequences of the junction sequence between the pGT12 clone and λ clone 12 revealed that the sequence was identical; this sequence analysis was essential to prove that the two clones belonged to and were contiguous on the *Cyp4a14* gene. A further 3.2kb of sequence, immediately 5' of

the 1.2kb PCRfragment has been amplified and cloned (Figure 3.11). This was done using two separate gene-specific primers derived from the the 1.2kb PCR fragment, the product of the first PCR "walk". This clone has not been completely sequenced, but preliminary sequence data of this clone showed that it shares the same sequence junction with the 1.2 kb fragment. As further proof that this clone is contiguous and further upstream of the 1.2kb clone, PCR has been performed across this junction and a fragment of expected size between the PCR primers has been detected on a gel(Figure 3.12).

Section 4.1 *Cyp4a* genomic structure

The *Cyp4a14* gene contains 12 exons and has a similar exon/intron structure as the rat CYP4A2 gene. To map the transcription start site of *Cyp4a14*, primer extension experiments were performed. Although initial data showed two primer-extended products (Figure 3.13), a separate mapping experiment using a second primer was key in establishing that *Cyp4a14*, in fact, only had a single transcription start site (Figure 3.14); this was mapped to a T nucleotide, located at a similar distance from the putative start of protein translation as the rat CYP4A2 gene (Figure 3.17).

Cyp4a14 is highly similar to the rat CYP4A2 and CYP4A3 genes. It shares 92 % and 93% identity at the cDNA level (Table 3.2) and 89% and 91% identity at the peptide level (Table 3.4) to CYP4A2 and CYP4A3 respectively. Thus, *Cyp4a14* is slightly more similar to CYP4A3 than to CYP4A2. This is further suggested by the presence of a 3-amino acid deletion in exon 3 in CYP4A2 but not in CYP4A3 or CYP4A14 proteins (Figure 3.21). As this corresponds to a 9 nucleotide deletion in the cDNA, an attempt was made to examine if all the λ clones isolated that contained exon 3 lacked this deletion or not; in other words, is there a murine clone that would be more similar to CYP4A2 than to CYP4A3? An oligonucleotide spanning this 9bp deletion was radiolabelled and used in a hybridization experiment with all λ clones containing exon

3, such that clones that did not hybridize to the probe would contain this 9bp deletion and would thus be more CYP4A2-like. The result of this experiment demonstrated that all clones hybridized to the probe and that they were all more similar to CYP4A3, that is, they were all *Cyp4a14* clones (Figure3.22).

The implications of this simple hybridization experiment were that no murine homologue for the rat CYP4A2 gene was identified : there are only three *Cyp4a* genes in the mouse but the rat possesses four members in the CYP4A subfamily. The cloning of *Cyp4a14*, the third and last *Cyp4a* member in the mouse is crucial to enable in-depth analyses of the entire *Cyp4a* gene subfamily, including aspects of its evolution, enzyme activity, transcriptional regulation and physical structure.

The rat CYP4A2 and CYP4A3 are themselves very similar to each other (97% amino acid identity), it is therefore likely that the gene duplication event in the rat giving rise to CYP4A2 must have occurred after the formation of the rat and mouse lineages. In addition, the three mouse genes (*Cyp4a10, Cyp4a12* and *Cyp4a14*) are highly similar to the rat genes (CYP4A1, CYP4A8 and CYP4A3) respectively, these three genes must have existed before the speciation of the rat and mouse.

The promoter regions of *Cyp4a10* and *Cyp4a14* have been cloned and analyzed in detail. Like other CYP4A genes, both *Cyp4a10* and *Cyp4a14* do not contain a TATA consensus sequence in their promoters. However, a 19bp element is present in both genes. This motif is highly conserved also in the rat CYP4A1 and CYP4A2 genes (Kimura *et al.*, 1989a). The promoter region of *Cyp4a10* is highly similar to the rat CYP4A1 gene but not to the other *Cyp4a* genes. *Cyp4a14* is highly similar to CYP4A2, in the intron region and even in the 5' flanking region, showing up to 84% similarity, but not with CYP4A1. However, this high similarity withCYP4A2 is dramatically reduced about 350bp upstream of the start of transcription (Figure 3.19). Interestingly, this

region is close to the position of two 358bp direct repeats which run from -440to -817 in the CYP4A2 gene (Kimura *et al.*, 1989a).

The presence of repetitive sequences in mammalian genomes is well known, although their role-sare not completely understood. The *Alu* family of repeats, for example, is a major family of repeats in human DNA (Schmid and Jelinek, 1982). These sequences are interspersed throughout the human genome by insertion into unoccupied DNA sequences (Van Arsdell *et al.*, 1981). Itis currently believed the large number of *Alu* repeats inserted in gene regions could have been major sources of variation during primate evolution. Indeed, an example of this is the human adult α-globin-like gene region (Hess *et al.*, 1983). These researchers studied the extensive sequence homology between the flanking regions of the two human adult α-globin-like genes, α2 and α1, which are interrupted by regions of nonhomology; one 1kb region 5' to the α2-globin gene and a 0.5kb and 0.2kb region 5' to the α1-globin gene. Interestingly, these regions and sequences around them contain *Alu* family repeats. Thus, this insertion of repetitive sequences may contribute to the diversity between duplicate genes and their flanking regions. In the mouse,insertion of repeated DNA elements is also not unknown. An example of this is the insertion of the LINE (long interspersed) repeated DNA (which are also present in the rat CYP4A2 gene; Kimura *et al.*, 1989a) in the intron of a member of the mouse Proline-rich Protein family of genes (Ann *et al.*, 1988).

The role of repeated sequences present in CYP4A2 (but not in *Cyp4a14*) remains to be elucidated. Characterizing the genomic structure of the rat CYP4A3 would be of interest since it is more similar to the mouse *Cyp4a14*, to see if repeated elements also exist. Both the rat CYP4A3 and the mouse *Cyp4a14* are not sexually differentiated, (nor are the murine *Cyp4a10* and the rat CYP4A1 which also lack these direct repeats), but the CYP4A2 is male-specific (Sundseth and Waxman, 1992; Heng *et al.*, 1997). It would thus be interesting to see if the insertion of this re-

peated sequence after duplication from CYP4A3, could have altered the CYP4A2 locus, and thus alter the regulation of this gene.

The individual exon sequences of *Cyp4a14* share consistently high similarity with corresponding exon regions in CYP4A2 and CYP4A3. Exons 4, 8, 11 and 12, in particular, show outstandingly high nucleotide similarity of up to 98 % (Table 3.2). Correspondingly, the amino acids encoded by these exon sequences are up to 98% identical between CYP4A14 and CYP4A2/CYP4A3 (Table 3.4). Due to their unusually high conservation, these exons might encode functionally-important regions of the protein.

Introns, on the other hand, are subject to less functional constraint and consequently less preservation of the nucleotide sequences is expected to occur during evolution (Van Den Berg *et al.*, 1978). Nevertheless, it has previously been noted that certain introns can be relatively more conserved than others within a gene. Nakajima-Iijima and coworkers (1985) observed that the nucleotide sequences of intron 3 of the β-actin genes were unexpectedly homologous between human and rat, with up to 89% similarity while other introns only showed 40-59% similarity. This observation was not repeated between other species. It was thus speculated that this homology might contribute to the expression and regulation of the gene specific to mammalian cells, because between human and chicken, this intron was interestingly the most divergent intron in the β-actin gene (less than 39% similarity; Nakajima-Iijima *et al.*,1985). When the intron sequences surrounding the conserved exons were analyzed for any possible conservation between *Cyp4a14* and CYP4A2, the following observation is particularly noteworthy : intron halves closest to the conserved exons 8 and 11 were, interestingly, up to 95% similar between *Cyp4a14* and CYP4A2 (Table 3.3). However, by the same token, it is unclear why the 5' and 3' halves of intron 3, adjacent to the highly conserved exon 4, were similarly conserved between *Cyp4a14* and CYP4A2. Thus, not only were exons 8 and 11 highly conserved,the non-coding regions around

them were also highly conserved, indicating that there exists an important selective bias for the especial conservation of these exons and their peptide products. Thus, the conserved exons and their translated protein motifs were of interest. Exon 8 putatively encodes a LRAEVDTFM-FEGHDTT 16-amino acid motif. This motif contains a stretch of 13 amino acids previously proposed to be a signature for the CYP4 family due to its high conservation, even in the CYP4C1 enzyme isolated from the cockroach *Blaberus sp.* (Bradley *et al.*, 1991). In CYP101, mutation of a Thr252 residue (corresponding to the final Thr residue in this motif) to an Ala residue produced an enzyme in which substrate hydroxylation is uncoupled from electron transfer, producing hydrogen peroxide and excess water instead of the CYP101 product, 5-exo-hydroxycamphor (Raag *et al.*, 1991). Furthermore, X-ray crystallographic analysis of this region indicated severe distortion in the mutant molecule (Raag *et al.*, 1991). This suggests that the Thr252 is a key active site residue, forming part of the proton delivery pathway.The importance of the Thr and the residues around it, thus, cannot be overemphasized, and this motif is probably keenly involved in the ω-hydroxylation of specific CYP4 substrates.

Exons 11 and 12 encode the more well-known RNCIG motif. The cysteine residue is known to be involved in directly binding haem, as established by a series of experiments, culminating in conclusive work on the crystal structure of the CYP101 protein (Poulos *et al.*, 1986). This motif is thus a crucial feature in the P450 protein and has been proposed to be a signature for the cytochrome P450 superfamily of enzymes (Gonzalez, 1989).

Exon 4 encodes for a highly conserved motif HRRMLLTPGFHYDIL, located between residuepositions 139 and 152 in CYP4A14 (Figure 3.26). This observation marks the identification of yet another highly conserved amino acid motif in CYP4A proteins never before reported. The role of this motif is undetermined, but could play a role in the ω-hydroxylation of the enzyme as deduced by comparing the tertiary structures of CYP4A14 with CYP2C4 (*vide infra*), whose

substrate binding regions are more well-documented. Exon 3, additionally, was highly conserved but only between *Cyp4a14* and CYPA3 (Table 3.2). As mentioned, this exon contains a 9bp deleted region in *CYP4A*2 only but not in *Cyp4a14* or CYP4A3 (Figure 3.21). The consequences of this deletion with respect to CYP4A enzyme activity remain to be elucidated.

Section 4.2 CYP4A enzyme function

In 1992, Gotoh proposed the existence of six substrate recognition sites (SRS) in the CYP2 proteins based on analyses of amino acid and coding nucleotide sequences. Since then, several reports have reaffirmed the accuracy of his assignments of the SRS'. Recently, Szklarz and coworkers (1996), for example, used hybrid enzymes and site-directed mutagenesis at various residues within these proposed SRS'; they found that all mutations caused changes in activity of the CYP2B5 protein. To examine the feasibility of deducing the substrate binding sites of theCYP4A14 protein, its amino acid sequence has been compared with the CYP2C4 protein based on their primary sequence homology (Figure 3.23), and their hydropathy indices (Figure 3.24). From the latter, it is concluded that although the two proteins had poor identity at the amino acid level, the tertiary structures of the cytochrome P450 proteins are fairly well conserved, an observation similarly made by Edwards and coworkers (1989). Thus, six SRS' in CYP4A14 are proposed; the deduced positions of the SRS' of the CYP4A14 protein were located in similar hydrophobic environments as those of the CYP2 protein. However, as noted, differences in hydrophobicity also exist, especially in SRS 1 and 4 (Figure 3.24). Due to the CYP2 and CYP4 proteins having different substrate specificities, these sites could have a major role to play in regulating the enzyme activities of these proteins.

As previously mentioned, besides the RNCIG amino acid motif in the P450 haem-binding region, two other motifs were highly conserved only in the CYP4 enzyme : HRRMLLTPGFHYDIL andLRAEVDTFMFEGHDTT. Interestingly, these CYP4-specific motifs map very close to

SRS 1 and within SRS 4 respectively (Figure 3.26). Thus, since these motifs are unique to CYP4 and not CYP2 proteins, they (and SRS1 and SRS4) could play a major part in the enzyme activity of the CYP4 protein. All the CYP4A proteins were also compared with one another and all SRS' were highly conserved, which is not unexpected.

While the enzymatic activity of the protein encoded by the novel murine *Cyp4a14* has not been demonstrated, it is likely that it metabolizes fatty acid by ω and ω-1 oxidation due to its high similarity to CYP4A1, CYP4A2 and CYP4A3, proteins of known enzyme activity (Hardwick *et al.*, 1987; Aoyama *et al.*, 1990).

Section 4.3 CYP4 family evolution

From the primary sequence alignments of the CYP4 proteins, it appeared that there were several regions of amino acid deletions and insertions which indicated CYP4A, CYP4B and CYP4F proteins were rather distinct from one another (Figure 3.26). Moreover, CYP4A and CYP4B proteins appear to be more different to the CYP4F proteins than they do to each other. This suggested that CYP4F proteins were a more distant related member of the CYP4 family compared to CYP4A and CYP4B proteins. To study the relationship of these subfamilies in greater detail,phylogenetic analysis was performed on the CYP4 proteins using CLUSTALW and PHYLIP with CYP2 as an "outgroup", a sequence which is distinctly different from the proteins. These methods of analysis consistently showed that the CYP4A, CYP4B and CYP4F proteins were grouped into three distinct clusters (Figure 3.27). Furthermore, in agreement with observations made with primary sequence alignments, the phylogenetic tree demonstrated that CYP4A and CYP4B were, again, more closely related to each other than either of them to CYP4F.

It is widely accepted that mammalian gene families normally arise by duplication of an ancestral-gene and its flanking sequence; thus, family members are usually located close to one another-

when they first arise. Recently, Trusov and Dear (1996) argued that with subsequent nucleotide substitutions and continued divergence of these newly duplicated genes, physical distances would be least between more closely-related family members : the evolutionary tree of the gene familycould be a good indication of its physical map. Thus, based on the above observations on the relatedness of the *Cyp4a*, *Cyp4b* and *Cyp4f* subfamilies, the theory of Trusov and Dear would imply that *Cyp4a* would be physically closer to *Cyp4b* than to the apparently more distant *Cyp4f* subfamily.

Recent experimental work done by Sharon Kuo in this laboratory has provided formal evidence that this is in fact true : not only are all three murine *Cyp4a* genes physically linked and contained in a single YAC clone of 220kb, the *Cyp4b1* gene is also located in this clone (published in Heng *et al.*, 1997). Furthermore, while this linkage is replicated in human, the human CYP4F2 gene is located on a separate chromosome. Thus, the physical relatedness of *Cyp4a* and *Cyp4b* would suggest that the duplication events giving rise to *Cyp4a/Cyp4b* genes was a more recent occurrence, and that in the murine *Cyp4* family of proteins, the *Cyp4f* subfamily would be the more ancient compared to the *Cyp4a/Cyp4b* genes.

The mouse *Cyp4a* genes have previously been localized to mouse chromosome 4 (Bell *et al.*,1993; Henderson *et al.*, 1994), which also appears to contain the *Cyp4b* subfamily (Heng *et al.*,1997). This area of co-localization is syntenic to human chromosome 1, which was found by Nhamburo and coworkers (1989) to contain the human CYP4B1 gene. Thus, this conserved arrangement of the *Cyp4a* and *Cyp4b* subfamilies on the chromosome must have existed before the rodent diverged from humans in evolutionary history. However, each subfamily appears to have evolved at different rates from each other. The relatedness of the human, rabbit and guinea pig CYP4A proteins are compared in Table 4.1 while the CYP4B proteins are compared against one another in Table 4.2.

	4a10	4a14	4A1	4A2/3	4A8	4A11	4A13	4A4	4A5	4A6	4A7
Mouse CYP4A10	-	79	91	72	81	77	82	76	76	77	73
CYP4A14		-	71	90	78	73	78	72	73	72	72
Rat CYP4A1			-	72	76	77	82	75	76	77	75
CYP4A2/3				-	71	74	77	72	74	73	72
CYP4A8					-	76	77	74	76	76	75
Human CYP4A11						-	79	80	82	81	79
Guinea Pig CYP4A13							-	80	78	82	82
Rabbit CYP4A4								-	87	88	85
CYP4A5									-	90	85
CYP4A6										-	87

Table 4.1 Similarity of the CYP4A proteins in mouse, rat, human, guinea pig and the rabbit. The relatedness of the CYP4A proteins are expressed as percent identity by BESTFIT analyses.Only full length protein sequences are compared.

	mCYP4B1	rCYP4B2	hCYP4B1	rbCYP4B1
mCYP4B1	-	92	85	88
rCYP4B2		-	86	88
hCYP4B1			-	90
rbCYP4B1				-

Table 4.2 Similarity of the CYP4B proteins in mouse, rat, human and the rabbit. The relatedness of the CYP4B proteins are expressed as percent identity.The prefixes'm','r','h' and'rb' refer to the mouse, rat, human and rabbit sequences respectively.

As shown, members of the CYP4A subfamily are as low as 72 % identical, compared to members of the CYP4B subfamily which are no less than 85% identical to one another. Thus, the CYP4A subfamily members are considerably less conserved across species compared with the CYP4B gene subfamily despite being in tandem on the same chromosome. This implies that the evolution of the CYP4A genes are particularly rapid compared with the neighbouring CYP4B subfamily. Additionally, rabbit and human CYP4A genes are approximately equally similar to the

mouse and rat *Cyp4a* genes (Table 4.1); therefore, the duplication event giving rise to the *Cyp4a* genes must have happened after the human/rabbit/guinea pig lineage diverged from the mouse/rat. Furthermore, the CYP4A genes appear to have evolved significantly more than the neighbouring CYP4B subfamily in the mouse and rat. Ohta (1989) suggested that once an initial single duplication of an ancestral gene had occurred, in this case the CYP4A ancestral gene, the chance of subsequent gene duplication via unequal crossing-over would be increased. Thus, these molecular events would lead to an increased number of *Cyp4a* genes, which would then individually undergo further differentiation.

There is much debate surrounding whether generation time has a significant effect on the rate of molecular evolution (Wilson *et al.*, 1977; Kimura, 1983) : are the rates of evolution constant 'per generation' rather than 'per year' ? Wu and Li (1985) reported that, when eleven genes from the mouse and rat were compared with those from other species, the rodent genes seemed to have higher rates of nucleotide substitution. More recently, when the *eta* globin pseudogene DNA and protein sequence data were compared, the rate of evolution is higher in the rodent lineage than in the primate lineage (Li *et al.*, 1996).In further support of the hypothesis, the *Cyp4a* genes in the mouse and rat have evolved at a significantly higher rate compared to these genes in the other species. The simple explanation for the higher rates in rodents could be that rodents have shorter generation times and thus, higher mutation rates. Alternatively, it has been speculated that the difference in evolution rates could be due to differences in DNA repair mechanisms (Catzeflis *et al.*, 1987)

The evolution of the rodents themselves is not so straightforward either, and has been rather controversial among biologists. While traditional taxonomy, mainly based on comparative morphology, classifies the guinea pig into the order of the Rodents (Novacek, 1992), much recent data obtained from molecular analysis of mitochondrial DNA has argued in favour of the inclusion of

the guinea pig in a new mammalian order distinct from Rodentia (D'Erchia *et al.*, 1996; Cao *et al.*, 1997). Indeed, this confusion is reflected by the data of Table 4.1, which shows that the guinea pig CYP4A13 gene is approximately equally similar to all mouse, rat and human CYP4A genes as the rabbit CYP4A genes (more than 77% identity), although the mouse and rat genes themselves are relatively distant to the rabbit genes (less than 77% identity). Thus the guinea pig CYP4A13 certainly appears to be evolved in a distinct fashion from the CYP4A genes of the more established rodents, the mouse and rat.

Section 4.4 *Cyp4a* expression and significance

The transcriptional activity of several CYP4A genes, for example, CYP4A1, is believed to be-regulated by the interaction of the peroxisome proliferator activated receptor (PPAR) with aresponse element located in the 5' flanking region of the gene (Aldridge *et al.*, 1995). Like the CYP4A1, the *Cyp4a14* gene is also highly inducible by a peroxisome proliferator, as demonstrated both by primer extension experiments (Figures 3.13 & 3.14), and formally, by RNAse protection assays performed by Paul Jones in this laboratory (published in Heng *et al.*, 1997). The *Cyp4a10* has also previously been shown to be highly inducibleby a peroxisome proliferator (Bell *et al.*, 1993). As such, the 5' flanking sequences of both genes have been cloned by PCR and analyzed. However, no classical peroxisome proliferator response element (PPRE) was found in the sequence immediately upstream of the start of transcription of either *Cyp4a10* or *Cyp4a14*, although two separate half sites were found in *Cyp4a14* (Figures3.19 & 3.31). Recently, additional experiments provided evidence that not only is the full PPRE site required for transcriptional activation, deletions or mutations within six nucleotides (TGAACT) 5' of the PPRE motif dramatically diminished PPAR-RXR heterodimer binding in the rabbit *CYP4A*6 (Palmer *et al.*, 1995). Thus, the DR1 motif is not sufficient to constitute a PPRE. This six-nucleotide sequence is highly conserved in CYP4A1, acyl coA oxidase among others (Palmer *et al.*, 1995). However, this sequence was not seen 5' of the PPRE half sites of *Cyp4a14*,suggesting that these are probably

not legitimate binding sites for PPAR-RXR heterodimers. Nevertheless, it remains a possibility that, like CYP4A1, the response element would be located much farther upstream. Thus, pGT32, the unsequenced PCR clone containing 3.2 kb of 5' flanking region, could be analyzed further for any putative binding sites. Even if such a putative PPRE was located within the upstream sequences, reporter gene experiments would have to be undertaken to demonstrate its function in binding PPAR-RXR heterodimers and regulating the transcription of *Cyp4a10* and *Cyp4a14* by a peroxisome proliferator. It has recently been shown in this laboratory using RNAse protection assays that *Cyp4a10* RNA (35-fold) and *Cyp4a14* RNA(1000-fold) are highly induced in liver tissues treated with a peroxisome proliferator (Bell *et al.*,1993; Heng *et al.*, 1997). It is unknown precisely what, if any, additional factors might regulate the extraordinarily high induction of *Cyp4a14*.

As noted earlier, the 5' flanking regions of *Cyp4a10* and *Cyp4a14* do not contain a TATA box,but possess a 19bp element (5'- (T)CCCCCCTCCCACAAGTAGG-3') which is similarly conserved in CYP4A1 and CYP4A2 (Kimura *et al.*, 1989a). The integrity of the 19bp element has been shown to be important in regulating the basal transcriptional activity of CYP4A1 (Bell and Elcombe, unpublished results) and is therefore likely to have a similar regulatory role in the murine genes. Analysis of the sequence in and around this 19bp element indicated that its pyridimidine-rich sequence, specifically its 'CTCCC' sequence, might serve as a binding motif for the Sp1 transcription factor (Figures 3.19 & 3.31).

The Sp1 family of transcription factors are zinc-finger proteins, are abundant nuclear proteins in most cells (Saffer *et al.*, 1991) and have been implicated in the expression in many genes. Appropriate DNA methylation patterns are important for normal development (Li *et al.*, 1992). Recent experiments have linked Sp1, and its related proteins, to the prevention of methylation of CpG islands in mammals (Brandeis *et al.*, 1994; Macleod *et al.*, 1994). The Sp1 proteins typically bind

GC or GT boxes (eg. Hagen *et al.*, 1992). Recently, Jiang and coworkers (1996) showed that the Sp family of transcription factors are also able to upregulate promoter activity by binding to a CTCCC motif in the promoter region of the mouse hepatocyte growth factor (HGF) gene, a motif which is also conserved in the human HGF promoter. However, the regulation of the *Cyp4a* genes is likely to be more complicated. Ihn and coworkers (1996) recently identified a TCCCCC motif in the promoter region of the human α2(I) collagen promoter which appeared to be a binding site for an unidentified transcriptional repressor. Interestingly, this motif is highly conserved in the promoters of the novel murine *Cyp4a14* gene and of the previously known rodent *Cyp4a* genes. Furthermore, both these motifs are present in tandem within the 19bp element of the *Cyp4a* genes. Thus, it is entirely conceivable that the basal transcription of *Cyp4a* genes is dependent on the complex interaction of both positively- and negatively-acting proteins on the highly conserved 19bp element.

Sp1 is also thought to be critical for the human β-globin Locus Control Region (LCR, which contains Sp1 binding sites) activity by maintaining an open chromatin structure (eg. Philipsen *et al.*, 1993). An open chromatin structure maintained by the upstream LCR of a gene locus has been associated with the coordinate regulation and expression of a cluster of genes (eg. Ellis *et al.*, 1996). In *Drosophila*, CT dinucleotide repeats are believed to have such an effect in the *hsp26* gene (Lu *et al.*, 1993). In intron 2 of *Cyp4a14*, CT dinucleotide repeats are also found which perhaps might have a similar role in regulating the expression of this gene (Figure 3.17). However, it remains to be discovered if the *Cyp4a* locus contains DNAseI hypersensitive sites due to an open, transcriptionally active chromatin structure and a locus control region that will regulate the expression of all its genes in this locus.

The expression of the genes in the rat and mouse *Cyp4a* subfamily are remarkably different and complex. It is unclear why such closely related genes have such distinct patterns of expression.

Even highly identical isozymes within a species may differ in their expression patterns. For example, although the rat CYP4A2 and CYP4A3 are 97% identical, theCYP4A2 is male-specific while CYP4A3 and the highly similar murine *Cyp4a14*, show no sexual dimorphism (Sundseth and Waxman, 1992; Heng *et al.*, 1997). *Cyp4a14*, in RNAse protection experiments, was highly induced in the liver and kidney of both male and female animal by the peroxisome proliferator, MCP. Consitutive expression of the novel murine *Cyp4a14*, however, was extremely low in tissues examined like spleen, lung, brain, gut and gonads (Heng *et al.*, 1997). The homologous rat CYP4A1 and murine *Cyp4a10* are modestly expressed in untreated liver and kidney but are highly induced by peroxisome proliferators (Hardwick *et al.*, 1987; Bell *et al.*, 1993). As discussed, the three rat genes CYP4A1, CYP4A3 and CYP4A8 have very similar equivalents in the mouse designated *Cyp4a10*, *Cyp4a14* and *Cyp4a12* respectively and must have arisen before the mouse/rat divergence. The rat CYP4A8 was shown to be expressed in the prostate and not regulated in a sex-specific way (Stromstedt *et al.*, 1994), but its murine homologue *Cyp4a12* shows very low expression in the prostate but is constitutively expressed in liver and kidney of male rats (Bell *et al.*, 1993; Heng *et al.*, 1997). Since it is unknown whether the three *Cyp4a* genes in the rat/mouse ancestor previously showed any sex-specificity, it is difficult to predict if the sex-specificity traits of the present *Cyp4a* isoforms had been gained or lost. However, since the male-specific CYP4A2 was deemed to have arisen, probably by a recent duplication from the highly similar CYP4A3 (but which is not sexually differentiated) in the rat, its sexually dimorphic trait must have been gained only after the rat species had radiated from the mouse.

A search of the database revealed several murine EST (Expressed Sequence Tag) sequences,which are randomly cloned and sequenced cDNAs, and are a reflection of the expression of murine *Cyp4a* genes. These cDNAs (accession numbers Aa123331, Aa268984, Aa060595,W41435, W13618 and W18764), corresponding to *Cyp4a14*, have been isolated from liver and extrahepatic sources like kidney and total foetus. ESTs corresponding to *Cyp4a10*

were also identified in a separate homology search, with the accession numbers Aa273591, Aa10968,Aa473036, Aa073857 and W13605. These RNAs were isolated also from extrahepatic sources like the kidney, T-cells and the heart in the mouse. These data for *Cyp4a10* and *Cyp4a14* demonstrate that these genes are expressed at moderate levels in many tissues, and that CYP4A proteins could well serve a function in organs other than the liver. Oddly, no ESTs corresponding to the *Cyp4a12* gene were found from the search of the sequence database, although this gene is known to be expressed constitutively in the male liver and kidney.

Interestingly, the enzyme activities and patterns of expression of the *Cyp4a* genes are distinctly different in the rabbit. While *Cyp4a* enzymes in the rat and mouse favour the metabolism of lauric acid (Hardwick *et al.*, 1987), the four rabbit CYP4A proteins, CYP4A4, CYP4A5, CYP4A6 and CYP4A7 were found to exhibit a distinct substrate specificity profile across a panel of substrates, which included lauric acid, palmitic acids and arachidonic acid (Roman *et al.*, 1993). For example, CYP4A5 showed little activity towards arachidonic or palmitic acids as compared to that toward lauric acid while CYP4A6 and CYP4A7 were relatively active metabolizers of arachidonic acid. The regulation of these genes is also quite distinct from the rodent *Cyp4a* genes : One CYP4A enzyme, CYP4A4, was found to be highly induced in the lungs of progesterone-treated or pregnant rabbits (Matsubara *et al.*, 1987) while CYP4A6 was the most inducible by clofibrate (Roman *et al.*, 1997). Since the enzyme activities and regulation of the rabbit CYP4A genes are clearly different from the rodents, it must be that the primordial CYP4A locus only evolved to give rise to these four rabbit isoforms, but only after the rabbit species had formed in evolutionary history. It is unclear why the rabbit and the rodents, should have *Cyp4a* subfamily members which are so closely related but that have evolved to possess such distinct properties.

Section 4.5 CYP4A proteins in peroxisome proliferation

In the liver, the elevated expression of *Cyp4a* genes caused by peroxisome proliferating chemi-

calsis believed to be a key event prior to peroxisome proliferation in the mouse and rat (eg Kaikaus*et al.*, 1993). In male liver, *Cyp4a10* and *Cyp4a14* are highly induced by MCP (Bell *et al.*, 1993;Heng *et al.*, 1997). *Cyp4a14* RNA, especially, was induced 1000-fold above control levels (Heng *et al.*, 1997) and could thus be a major mediator in murine peroxisome proliferator-induced carcinogenesis. On the other hand, *Cyp4a12*, due to its high constitutive expression in liver,might have more of a housekeeping role, maintaining normal lipid metabolism.

Section 4.6 CYP4A proteins in renal function

Cyp4a proteins also have been implicated in the metabolism of eicosanoids such as arachidonic acid in the kidney, which generates 20- and 19-HETEs (Sacerdoti *et al.*, 1987). Significant amounts of this eicosanoid can be produced by the thick ascending limb of Henle's loop (TALH). 20-HETE is a potent vasoconstrictor of blood vessels (Ma *et al.*, 1993), inhibits K^+ channel activity in renal tubules and microvessels and produces a sustained increase in intracellular Ca^{2+}concentration (Wang and Lu, 1995). Thus, the synthesis of 20-HETE by *Cyp4a* proteins may contribute to the tubular ion transport mechanisms, while in the renal vasculature, *Cyp4a* metabolites may exert their effects on renal blood flow and tubuloglomerular feedback (Zou *et al.*, 1994a). The importance of *Cyp4a* proteins are further emphasized in experiments involving 17-ODYA, a suicide-subtstrate inhibitor of ω-hydroxylation of arachidonic acid (Zou *et al.*,1994b). Zou and coworkers found that infusion of 17-ODYA into renal artery or renal cortical interstitium , increases urine flow and sodium excretion, associated with the increase in papillary blood flow and an increase in renal interstitial hydrostatic pressure. These observations suggest that *Cyp4a* metabolites of arachidonic acid exert a vasoconstrictor influence on the renal arteries and affect tubular sodium and water reabsorption. In this respect, due to the very low levels of expression in the control male kidney of *Cyp4a10* or *Cyp4a14*, these gene products may have little or no significance in maintaining renal function. However, *Cyp4a12* is constitutively highly expressed in the kidney suggesting there might be a distinct physiological role for the

Cyp4a12 enzyme. *Cyp4a12* could well be the more important *Cyp4a* gene in the regulation of renal metabolism and function.

Section 4.7 CYP4A proteins in inflammation control

LTB_4, a dihydroxy fatty acid leukotriene B_4, has an important role during inflammation as this signalling molecule induces a cascade of molecular and cellular events that triggers an immune-response and the gradual congregation of cells from the immune system to the site of injury. The duration and extent of the inflamed state thus depends on the turnover of the LTB_4 molecule in the body. Interestingly, LTB_4 is inactivated through metabolism by the microsomal ω- andper-oxisomal β-oxidation pathways (Samuelsson *et al.*, 1987; Jedlitschky *et al.*, 1991). Since CYP4A proteins are ω-hydroxylating enzymes which could thus control the levels of LTB_4, there could be a link between CYP4A enzymes and the control of inflammation. However, *Cyp4a* genes are in turn, dependent on transcriptional activation by PPAR binding to response elements in the5' flanking regions of these genes (eg. Aldridge *et al.*, 1995). Therefore, the increase or decrease in LTB_4 levels, and hence inflammation, could be dependent on the transcriptional activation of the *Cyp4a* genes by PPAR.

Recently, evidence for the possible role of PPAR in inflammation control was published in which Devchand and coworkers (1996) identified LTB_4 as a natural ligand for the PPARα. They also demonstrated that LTB_4 had the ability to induce the transcription of PPAR-regulated genes, an example of which is the *Cyp4a* subfamily. Further proof of the role PPAR might play in inflammation control came about with experiments with PPARα-deficient mice, in which inflammation was prolonged. However, it is entirely possible that due to the lack of PPARα, the *Cyp4a* genes are poorly expressed in these mutant mice, thus leading to the delayed degradation of LTB_4. Unfortunately, no data for the expression of *Cyp4a* genes was reported in their work. Nevertheless, this is the first report of a binding ligand of the orphan receptor PPARα. It is thus

entirely possible that fatty acids and eicosanoids also can regulate their own metabolism by CYP4A proteins through direct activation of PPAR.

Section 4.8 Summary and future aims

This project set out to study the murine *Cyp4a* gene subfamily in detail. Genomic sequences of the murine *Cyp4a* genes, *Cyp4a10* and *Cyp4a12* were cloned. In addition, a novel murine *Cyp4a* member, designated *Cyp4a14*, has been identified, which is crucial for understanding the murine *Cyp4a* subfamily in its entirety, in terms of its evolution, regulation and expression and enzyme activity. The whole of *Cyp4a14* has been cloned, including its 5' flanking region. *Cyp4a14* is most similar to the rat CYP4A3 gene and is thus its homologue.

The cloning of *Cyp4a14* demonstrates the presence of a third member of the *Cyp4a* subfamily in mouse while there are four CYP4A members in rat. The rat CYP4A2 does not have a homologue in the mouse and must have arisen from CYP4A3 after rat/mouse speciation. In light of the possible key involvement of the *Cyp4a* genes in murine physiology, it would be interesting to examine if mice deficient in *Cyp4a* genes are susceptible to peroxisome proliferator-induced carcinogenesis or have impaired renal function and inflammation control. It will then be possible in the not too distant future to target the entire *Cyp4a* locus for disruption in embryonic stem cells and the subsequent generation of mutant "knockout" mice.

Chapter 5 REFERENCES

Aldridge, T.C., Tugwood, J.D. and Green, S. (1995). Identification and Characterization of DNA Elements Implicated in the Regulation of CYP4A1 Transcription. *Biochem. J.* **306** : 473 - 479.

Ann, D.K., Smith, M.K. and Carlson, D.M. (1988). Molecular evolution of the mouse proline-rich protein multigene family. *J. Biol. Chem.* **263 (22)** : 10887 - 10893.

Aoyama, T., Hardwick, J.P., Imaoka, S., Funae, Y., Gelboin, H.V. and Gonzalez, F.J. (1990). Clofibrate-inducible rat hepatic P450s IVA1 and IVA3 catalyze the ω- and (ω-1)-hydroxylation of fatty acids and the ω-hydroxylation of prostaglandins E1 and F2a. *J. Lipid. Res.* **31** : 1477 - 1481.

Aperlo, C., Pognonec, P., Saladin, R., Auwerx, J. and Boulukos, K.E. (1995). cDNA cloning and characterization of the transcriptional activities of the hamster peroxisome proliferator-activated receptor haPPAR gamma. *Gene.* **162 (2)** : 297 - 302.

Ashby J, Brady, A., Elcombe, C.R., Elliott, B.M., Ishmael, J., Odum, J., Tugwood, J.D., Kettle, S. and Purchase, I.F.H. (1994). Mechanistically-based human hazard assessment of peroxisome proliferator-induced hepatocarcinogenesis. *Human. Exp. Toxicol.* **13 (S2)** : S1 - S117.

Bains, S.K., Gardiner, S.M., Manniveiler, K., Gillett, D. and Gibson, G.G. (1985). Immunochemical study on the contribution of hypolipidaemic-induced cytochrome P-452 on the metabolism of lauric acid and arachidonic acid. *Biochem. Pharmacol.* **34** : 3221 - 3229.

Baltimore, D. (1981). Gene conversion - some implications for immunoglobulin genes. *Cell.* **24 (3)** : 592 - 594.

Bardot, O., Aldridge, T.C., Latruffe, N. and Green, S. (1993). PPAR-RXR heterodimer acti-

vates a peroxisome proliferator response element upstream of the bifunctional enzyme gene. *Biochem. Biophys. Res. Commun.* **192 (1)** : 37 - 45.

Bell, D.R. and Elcombe, C.R. (1991). Induction of acyl coA oxidase and cytochrome P450IVA1 RNA in rat primary hepatocyte culture by peroxisome proliferators. *Biochem. J.* **280 (1)** : 249 - 253.

Bell, D.R., Plant, N.J., Rider, C.G., Na, L., Brown, S., Ateitalla, I., Acharya, S.K., Davies, M.H., Elias, E., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Elcombe, C.R. (1993). Species-specific induction of cytochrome P450 4A RNAs : PCR cloning of partial guinea-pig, human and mouse CYP4A cDNAs. *Biochem. J.* **294** : 173-180

Bentley, P., Calder, I., Elcombe, C.R., Grasso, P., Stringer, D and Wiegand, H.J. (1993). Hepatic Peroxisome proliferation in rodents and its significance for humans. *Food. Chem. Toxicol.* **31 (11)** : 857 - 907.

Berge, R.K., Aarsland, A., Kryvi, H., Bremer, J. and Aarsaether, N. (1989). Alkylthioacetic acid (3-thia fatty-acids) - a new group of non-beta-oxidizable, peroxisome-inducing fatty-acid analogS .1. a study on the structural requirements for proliferation of peroxisomes and mitochondria in rat-liver. *Biochim. Biophys. Acta.* **1004 (3)** : 345 - 356.

Black, S. and Coon, M. (1986). Comparative Structures of P-450 Cytochromes. *In* Cytochrome P450 : Structure, Mechanisms, and Biochemistry, ed. by P.R. Ortiz de Montellano, pp.161-216, Plenum Publishing Corp., New York.

Bonfield, J.K., Smith, K.E. and Staden, R. (1995). A new DNA sequence assembly program. *Nucleic Acids Research.* **23 (24)** : 4992 - 4999.

Bradfield, J.Y., Lee, Y.H. and Keeley, L.L. (1991). Cytochrome P450 family 4 in a cockroach :

Yee M. Heng
Molecular cloning and regulation by hypertrehalosemic hormone. *Proc. Natl. Acad. Sci. USA*. **88** : 4558-4562

Braissant, O., Foufelle, F., Scotto, C., Dauca, M. and Wahli, W. (1996). Differential expression of peroxisome proliferator-activated receptors (PPARs): tissue distribution of PPAR-alpha, -beta, and -gamma in the adult rat.PPAR distribution in rat. *Endocrinology*. **137 (1)** : 354 - 366.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelssohn, M., Nemes, A., Temper, V., Razin, A. and Cedar, H. (1994). Sp1 elements protect a CpG island from *de novo* methylation. *Nature*. **371** : 435 - 438.

Bronfmann, M., Amigo, L. and Morales, M.N. (1986). Activation of hypolipemic drugs to acyl-coenzyme A thioesters. *Biochem. J.* **239 (3)** : 781 - 784.

Bronfmann, M., Orellana, A., Morales, M.N., Bieri, F., Waechter, F., Staubli, W. and Bentley, P. (1989). Potentiation of diacylglycerol-activated protein kinase C by acyl coenzyme A thioesters of hypolipemic drugs. *Biochem. Biophys. Res. Comm.* **159 (3)** : 1026 - 1031.

Burger, H.J., Schuetz, J.D., Schuetz, E.G. and Guzelian, P.S. (1992). Paradoxical transcriptional activation of rat liver cytochrome P-450 3A1 by dexamethasone and the antiglucocorticoid pregnenolone 16 alpha-carbonitrile: analysis by transient transfection into primary monolayer cultures of adult rat hepatocytes. *Proc. Natl. Acad. Sci. (USA).* **89 (6)** : 2145 - 2149.

Cao, Y., Okada, N. and Hasegawa, M. (1997). Phylogenetic position of guinea pig revisited. *Mol. Biol. Evol.* **14 (4)** : 461 - 464.

Capdevila, J.H., Shouzou, W., Yan, J., Karara, A., Jacobson, H.R., Falck, J.R., Guengerich, F.P. and Dubois, R.N. (1992). Cytochrome P450 arachidonic acid epoxygenase : Regulatory control of the renal epoxygenase by dietary salt loading. *J. Biol. Chem.* **267** : 21720 - 21726.

Carroll, M.A., Sala, A., Dunn, C.E., McGiff, J.C. and Murphy, R.C. (1991). Structural identification of cytochrome P450-dependent arachidonate metabolites formed by rabbit medullary thick ascending limb cell. *J. Biol. Chem.* **266** : 12306 - 12312.

Castelein, H., Gulick, T., Declercq, P.E., Mannaerts, G.P., Moore, D.D. and Baes, M.I. (1994). The peroxisome proliferator activated receptor regulates malic enzyme gene expression. *J. Biol. Chem.* **269 (43)** : 26754 - 26758.

Cathala, G., Savouret, J.F., Mendez, B., West, B.L., Karin, M., Martial, J.A. and Baxter, J.D. (1983). A method for isolation of intact, translationally active ribonucleic acid. *DNA.* **2 (4)** : 329 - 335.

Cattley, R.C., Conway, J.G. and Popp, J.A. (1987). Association of persistent peroxisome proliferation and oxidative injury with hepatocarcinogenicity in female F344 rats fed di(2- ethylhexyl)phthalate for 2 years. *Cancer Letters.* **38 (1-2)** : 15 - 22.

Cattley, R.C., Marsman, D.S. and Popp, J.A. (1991). Age-related susceptibility to the carcinogenic effect of the peroxisome proliferator WY-14,643 in rat-liver. *Carcinogenesis.* **12 (3)** : 469 - 473.

Catzeflis, F.M., Sheldon, F.H., Ahlquist, J.E. and Sibley, C.G. (1987). DNA-DNA hybridization evidence of the rapid rate of muroid rodent DNA evolution. *Mol. Biol. Evol.* **4** : 242 - 253.

Chen, F., Law, S.W. and O'Malley, B.W. (1993). Identification of 2 messenger ppar related receptors and evidence for the existence of 5 subfamily members. *Biochem. Biophys. Res. Commun.* **196 (2)** : 671 - 677.

Christiansen, E.N., Gray, T.J.B. and Lake, B.G. (1985). Effect of very long chain fatty acids on peroxisomal b-oxidation in primary rat hepatocyte cultures. *Lipids.* **20 (12)** : 929 - 932.

Christy, R.J., Yang, V.W., Ntambi, J.M., Geiman, D.E., Landschulz, W.H., Friedman, A.D., Nakabeppu, Y., Kelly, T.J. and Lane, M.D. (1989). Differentiation-induced gene expression in 3T3-L1 preadipocytes: CCAAT/enhancer binding protein interacts with and activates the promoters of two adipocyte-specific genes. *Genes Dev.* **3 (9)** : 1323 - 1335.

Claysonn, D.B., Mehta, R. and Iverson, F. (1994). International Commission for Protection Against Environmental Mutagens and Carcinogens. Oxidative DNA damage - the effects of certain genotoxic and operationally non-genotoxic carcinogens. *Mutat. Res.* **317 (1)** : 25 - 42.

Cohen, A.J. and Grasso, P. (1981). Review of the hepatic response to hypolipidaemic drugs in rodents and assessment of its toxicological significance to man. *Food. Cosmet. Toxicol.* **19 (5)** : 585 - 605.

D'Erchia, A.M., Gissi, C., Pesole, G., Saccone, C. and Arnason, U. (1996). The guinea pig is not a rodent. *Nature.* **381** : 597 - 600.

Devchand, P.R., Keller, H., Peters, J.M., Vazques, M., Gonzalez, F.J. and Wahli, W. (1996). The PPARa-leukotriene B4 pathway to inflammation control. *Nature.* **384** : 39 - 43.

Diczfalusy, U., Alexson, S.E. and Pedersen, J.I. (1987). Chain-shortening of prostaglandin F2 alpha by rat liver peroxisomes. *Biochem. Biophys. Res. Commun.* **144** : 1206 - 1213.

DiLella, A.G. and Woo, S.L. (1987). Hybridization of genomic DNA to oligonucleotide probes in the presence of tetramethylammonium chloride. *Methods Enzymol.* **152** : 447 - 451.

Distlerath, L.M., Reilly, P.E.B., Martin, M.V., Davis, G.G. and Wilkinson, G.R. (1985). Purification and characterization of the human liver cytochrome P-450 involved in debrisoquine 4-hydroxylation and phenacetin O-deethylation, two prototypes for genetic polymorphism in oxidative drug metabolism. *J. Biol. Chem.* **260** : 9057 - 9067.

Dreyer, C., Krey, G., Keller, H., Givel, F., Helftenbein, G. and Wahli, W. (1992). Control of the peroxisomal b-oxidation pathway by a novel family of nuclear hormone receptors. *Cell.* **68** : 879 - 887.

Dreyer, C., Keller, H., Mahfoudi, A., Laudet, V., Krey, G. and Wahli, W. (1993). Positive regulation of the peroxisomal beta-oxidation pathway by fatty acids through activation of peroxisome proliferator-activated receptors (PPAR). *Biol Cell.* **77 (1)** : 67 - 76.

Earnshaw, D., Dale, J.W., Goldfarb, P.S. and Gibson, G.G. (1988). Differential splicing in the 3' non-coding region of rat cytochrome P-452 (P450IVA1) mRNA. *FEBS Letters.* **236 (2)** : 357 - 361.

Edwards, R.J., Murray, B.P., Boobis, A.R. and Davies, D.S. (1989). Identification and Location of a-helices in Mammalian Cytochromes P450. *Biochemistry.* **28 (9)** : 3762 - 3770.

Elcombe, C.R., Rose, M.S. and Pratt, I.S. (1985). Biochemical, histological and ultrastructural changes in rat and mouse liver following the administration of trichloroethylene; possible relevance to species differences in hepatocarcinogenicity. *Toxicol. Appl. Pharmacol.* **79** : 365 - 376.

Felsenstein, J. (1985). Confidence Limits on Phylogenies : An Approach Using the Bootstrap. *Evolution.* **39** : 783 - 791.

Fitzpatrick, F.A. and Murphy, R.C. (1989). Cytochrome P450 metabolism of arachidonic acid : formation and biological actions of epoxygenase-derived eicosanoid. *Pharmacol. Rev.* **40** : 229 - 241.

Forman, B.M., Tontonoz, P., Chen, J., Brun, R.P., Spiegelman, B.M. and Evans, R.M. (1995). 15-Deoxy-delta 12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR gamma. *Cell.* **83** : 803 - 812.

Freytag, S.O. and Geddes, T.J. (1992). Reciprocal regulation of adipogenesis by *Myc* and C/EBP alpha. *Science*. **256 (5055)** : 379 - 382.

Fulco, A.J. (1991). $P450_{BM-3}$ and other inducible bacterial P450 cytochromes: biochemistry and regulation. *Annu. Rev. Pharmacol. Toxicol.* **31** : 177 - 203.

Garfinkel,D. (1958). Studies on pig liver microsomes. I. Enzyme and pigment composition of different microsomal fractions. *Arch. Biochem. Biophys.* **77** : 493 - 509.

Gebel, T., Arand, M. and Oesch, F. (1992). Induction of the peroxisome proliferator activated receptor by fenofibrate in rat liver. *FEBS Lett.* **309 (1)** : 37 - 40.

Gibson, G.G., Orton, T.C. and Tamburini, P.P. (1982). Cytochrome P-450 induction by clofibrate. *Biochem. J.* **203** : 161 - 1678.

Gonzalez, F.J., Song, B.-J. and Hardwick, J.P. (1986). Pregnenolone 16α-carbonitrile-inducible P450 gene family : gene conversion and differential regulation. *Mol. Cell. Biol.* **6** : 2969 - 2976.

Gonzalez, F.J. (1989). The Molecular Biology of Cytochrome P450s. *Pharmacol. Reviews*. **40 (4)** : 243 - 288.

Gonzalez, F.J. and Nebert, D.W. (1990). Evolution of the P450 gene superfamily : Animal-plant "warfare", molecular drive, and human genetic differences in drug oxidation. *Trends Genet.* **6** : 182 - 186.

Gottlicher, M., Widmark, E., Li, Q. and Gustafsson, J.-A. (1992). Fatty acids activate a chimera of the clofibric acid-activated receptor and the glucocorticoid receptor. *Proc. Natl. Acad. Sci. USA*. **89** : 4653 - 4657.

Gotoh, O. (1992). Substrate Recognition Sites in Cytochrome P450 Family 2 (CYP2) Proteins

Yee M. Heng
Inferred from Comparative Analyses of Amino Acid and Coding Nucleotide Sequences. *J. Biol. Chem.* **267 (1)** : 83-90

Greene, M.E., Blumberg, B., McBride, O.W., Yi, H.F., Kronquist, K., Kwan, K., Hsieh, L., Greene, G. and Nimer, S.D. (1995). Isolation of the human peroxisome proliferator activated receptor gamma cDNA: expression in hematopoietic cells and chromosomal mapping. *Gene Expr.* **4 (4-5)** : 281 - 299.

Gulick, T., Crescis, Caira T., Moore, D.D. and Kelly, D.P. (1994). The peroxisome proliferator activated receptor regulates mitochrondrial fatty acid oxidative enzyme gene expression. *Proc. Natl. Acad. Sci. (USA).* **91 (23)** : 11012 - 11016.

Hagen, G., Muller, S., Beato, M. and Suske, G. (1992). Cloning by recognition site screening of two novel GT box binding proteins : a family of Sp1 related genes. *Nucleic Acids Research.* **20** : 5519 - 5525.

Hankinson, O. (1995). The aryl hydrocarbon receptor complex. *Annu. Rev. Pharmacol. Toxicol.* **35** : 307 - 340.

Hardwick, J.P, Song, B.-J., Huberman, E. and Gonzalez, F.J. (1987). Isolation, Complementary DNA Sequence, and Regulation of Rat Hepatic Lauric Acid ω-Hydroxylase (Cytochrome P-$450_{LA\omega}$). *J. Biol. Chem.* **262 (2)** : 801- 810.

He, Y.-A., Luo, Z., Klekotka, P.A., Burnett, V.L. and Halpert, J.R. (1994). Structural Determinants of Cytochrome P450- 2B1 Specificity : Evidence for Five Substrate Recognition Sites. *Biochemistry.* **33** : 4419 - 4424.

Henderson, C.J., Scott, A.S., Yang, C.S. and Wolf, C.R. (1990). Testosterone-mediated regulation of mouse renal cytochrome P-450 enzymes. *Biochem. J.* **206** : 675-681.

Henderson, C.J., Bammler, T. and Wolf, C.R. (1994). Deduced amino acid sequence of a murine cytochrome P450 *Cyp4a* protein : developmental and hormonal regulation in liver and kidney. *Biochim. Biophys. Acta.* **1200** : 182- 190.

Heng, Y.M, Kuo, S.C.-W., Jones, P.S., Savory, R., Schultz, R.M., Tomlinson, S. R., Gray, T.J.B. and Bell, D.R. (1997). A Novel Murine P450 Gene, *Cyp4a14*, is Part of a Cluster of *Cyp4a* and *Cyp4b*, but not CYP4F, Genes in Mouse and Man. *Biochem. J.* **325 (3)** : 741 - 749.

Hertz, R., Bar-Tana, J., Sujatta, M., Pill, J., Schmidt, F.H. and Fahimi, H.D. (1988). The induction of liver peroxisomal proliferation by beta,beta'-methyl-substituted hexadecanedioic acid (MEDICA 16). *Biochem. Pharmacol.* **37 (19)** : 3571 - 3577.

Hertz R, Nikodem V, Ben-Ishai A, Berman I, Bar-Tana J. (1996). Thyromimetic mode of action of peroxisome proliferators: activation of 'malic' enzyme gene transcription. *Biochem J.* **319(Pt 1)** : 241-248

Hess, R., Staubli, W. and Riess, W. (1965). Nature of the hepatomegalic effect produced by ethyl-chlorophenoxy-isobutyrate in the rat. *Nature* (London). **208 (13)** : 856 - 858.

Hess, J.F., Fox, M., Schmid, C. and Shen, C. -K. J. (1983). Molecular evolution of the human adult a-globin-like gene region : Insertion and deletion of the *Alu* family repeats and non-*Alu* DNA sequences. *Genetics*. **80** : 5970 - 5974.

High, S., Dobberstein, B. (1992). Mechanisms that determine the transmembrane disposition of proteins. *Curr. Opin. Cell. Biol.* **4 (4)** : 581 - 586.

Hinton, R.H., Mitchell, F.E., Mann, A., Chescoe, D., Price, S.C., Nunn, A., Grasso, P. and Bridges, J.W. (1986). Effects of phthalic acid esters on the liver and thyroid. *Environ. Health Perspect.* **70** : 195 - 210.

Hiratsuka, M., Matsuura, T., Watanabe, E., Sato, M. and Suzuki, Y. (1996). Sex and Strain Difference in Constitutive Expression of Fatty Acid ω-Hydroxylase (CYP4A-Related Proteins) in Mice. *J. Biochem.* **119** : 340 -345.

Hoffman, S.M.G., Fernandez-Salguero, P., Gonzalez, F.J. and Mohrenweiser, H.W. (1995). Organization and evolution of the cytochrome P450 CYP2A-2B-2F s ubfamily gene-cluster on human chromosome 19. *J. Mol. Evol.* **41 (6)** : 894 - 900.

Holmes, R.S. and Masters, C.J. (1972). Species specific features of the distribution and multiplicity of mammalian liver catalase. *Arch. Biochem. Biophys.* **148 (1)** : 217 - 223.

Ihn, H., Ohnishi, K., Tamaki, T., LeRoy, E.C. and Trojanowska, M. (1996). Transcriptional regulation of the human α2(I) collagen gene. *J. Biol. Chem.* **271 (43)** : 26717 - 26723.

Imaoka, S., Ogawa, H., Kimura, S. and Gonzalez, F.J. (1993). Complete cDNA Sequence and cDNA-Directed Expression of CYP4A11, a Fatty Acid ω-Hydroxylase Expressed in Human Kidney. *DNA and Cell Biology.* **12 (10)** : 893 - 899.

Imaoka, S., Hiroi, T., Tamura, Y., Yamazaki, H., Shimada, T., Komori, M., Degawa, M. and Funae, Y. (1995). Mutagenic Activation of 3-Methoxy-4-aminoazobenzene by Mouse Renal Cytochrome P450 CYP4B1 : Cloning and Characterization of Mouse CYP4B1. *Arch. Biochem. Biophys.* **321 (1)** : 255 - 262

Imig, J.D., Falck, J.R., Gebremedhin, D., Harder, D.R. and Roman, R.J. (1993). Elevated renovascular tone in young spontaneously hypertensive rats : Role of cytochrome P450. *Hypertension.* **22** : 357 - 364.

Isseman, I. and Green, S. (1990). Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature.* **347** : 645 - 650.

Isseman, I., Prince, R., Tugwood, J. and Green, S. (1992). A role for fatty acids and liver fatty acid binding protein in peroxisome proliferation ?. *Biochem. Soc. Trans.* **20** : 824 - 826.

Jahn, C.L., Hutchison, C.A., Phillips, S.J., Weaver, S., Haigwood, N.L., Voliva, C.F. and Edgell, M.H. (1980). DNA sequence organization of the beta-globin complex in the BALB/c mouse. *Cell.* **21 (1)** : 159 - 168.

Jansson, J.O., Ekberg, S., Isaksson, O., Mode, A. and Gustaffson, J.A. (1985). Imprinting of growth hormone secretion, body growth, and hepatic steroid metabolism by neonatal testosterone. *Endocrinology.* **117 (5)** : 1881 - 1889.

Jedlitschky, G., Huber, M., Volkl, A., Muller, M., Leier, I., Muller, J., Lehmann, W.-D., Fahimi, H.D. and Keppler, D. (1991). Peroxisomal degradation of leukotrienes by b-oxidation from the w-end. *J. Biol. Chem.* **266 (36)** : 24763 - 24772.

Jiang, J. -G., Chen, Q., Bell, A. and Zarnegar, R. (1997). Transcriptional regulation of the hepatocyte growth factor (HGF) gene by the Sp family of transcription facators. *Oncogene.* **14** : 3039 - 3049.

Johnson, E.F. , Walker, D.W. Griffin, K.J., Clark, J.E., Okita, R.T., Muerhoff, A.S. and Masters, B.S.S. (1990). Cloning and expression of three rabbit kidney cDNAs encoding lauric acid omega-hydroxylases. *Biochemistry.* **29 (4)** : 873 - 879.

Jones, P.S., Savory, R., Barratt, P., Bell, A.R., Gray, T.J.B., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Bell, D.R. (1995). Chromosomal localisation, inducibility, tissue-specific expression and strain differences in three murine peroxisome-proliferator-activated-receptor genes. *Eur. J. Biochem.* **233 (1)** : 219 - 226.

Kaikaus, N.M., Chan, W.K., Lysenko, N., Ray, R., Ortiz de Montellano, P.R. and Bass, N.M.

(1993). Induction of peroxisomal fatty acid β-oxidation and liver fatty acid-binding protein by peroxisome proliferators - mediation via the cytochrome P4504A1 w-hydroxylase pathway. *J. Biol. Chem.* **268** : 9593 - 9603.

Kaku, M., Ichihara, K., Kusunose, E., Ogita, K., Yamamoto, S., Yano, I. and Kusunose, M. (1984). Purification and Characterization of Cytochrome P-450 Specific for Prostaglandin and Fatty Acid Hydroxylase Activities from the Microsomes of Rabbit Small Intestinal Mucosa. *J. Biochem.* **96 (6)** : 1883 - 1891.

Kalb, V.F. and Loper, J.C. (1988). Proteins from eight eukaryotic cytochrome P-450 families share a segmented region of sequence similarity. *Proc. Natl. Acad. Sci. (USA).* **85 (19)** : 7221 - 7225.

Kase, B.F., Bjorkhem, I. and Pedersen, J.I. (1983). Formation of cholic acid from 3 alpha, 7 alpha, 12 alpha-trihydroxy-5 beta-cholestanoic acid by rat liver peroxisomes. *J. Lipid. Res.* **24 (12)** : 1560 - 1567.

Kawashima, Y., Narcisa, U.Y. and Kozuka, H. (1989). Sex-related difference in the inductions by perfluoro-octanoic acid of peroxisomal β-oxidation, microsomal 1-acylglycerphosphocoline acetyltransferase and cytosolic long-chain acyl coA hydrolase in rat liver. *Biochem. J.* **261** - 595 - 600.

Kawashima, H., Kusunose, E., Kubota, I., Maekawa, M. and Kusunose, M. (1992). Purification and NH_2-terminal amino acid sequences of human and rat kidney fatty acid ω-hydroxylases. *Biochim. Biophys. Acta.* **1123** : 156 - 162

Kvannes, J., Eikhom, T.S. and Flatmark, T. (1994). The peroxisomal beta-oxidation enzyme-system of rat-heart - basal level and effect of the peroxisome proliferator clofibrate . *Biochim. Bio-*

Yee M. Heng
phys. Acta. (General Subjects). **1201 (2)** : 203 - 216.

Keller, H., Dreyer, C., Medin, J., Mahfoudi, A., Ozato, K. and Wahli, W. (1993). Fatty acids and retinoids control lipid metabolism through activation of peroxisome proliferator-activated receptor-retinoid X receptor heterodimers. *Proc. Natl. Acad. Sci. (USA).* **90 (6)** : 2160 - 2164.

Kliewer, S.A., Umesono, K., Noonan, D.J., Heyman, R.A. and Evans, R.M. (1992). Convergence of 9cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. *Nature.* **358** : 771 - 774.

Kliewer, S.A., Forman, B.M., Blumberg, B., Ong, E.S., Borgmeyer, U., Mangelsdorf, D.J., Umesono, K. and Evans, R.M. (1994). Differential expression and activation of a family of murine peroxisome proliferator-activated receptors. *Proc. Natl. Acad. Sci. (USA*). **91(15)** : 7355-7359.

Kliewer, S.A., Lenhard, J.M., Wilson, T.M., Patel, I., Morris, D.C. and Lehman, J.M. (1995). A prostaglandin J2 metabolite binds peroxisome proliferator-activated receptor gamma and promotes adipocyte differentiation. *Cell.* **83** : 803 - 812.

Klug, A. and Schwabe, J.W.R. (1995). Protein motifs. 5. zinc fingers. *FASEB J.* **9 (8)** : 597 - 604.

Kimura, M. (1983). The neutral theory of molecular evolution. *p 367*, Cambridge University Press, London.

Kimura, S., Gonzalez, F.J. and Nebert, D.W. (1986). Tissue-specific expression of the mouse dioxin-inducible P1450 and P3450 genes : differential transcriptional activation and mRNA stability in liver and extrahepatic tissues. *Mol. Cell. Biol.* **6** : 1471 - 1477.

Kimura, S., Hanioka, N., Matsunaga, E. and Gonzalez, F.J. (1989a). The Rat Clofibrate-Inducible CYP4A Gene Subfamily I. Complete Intron and Exon Sequence of the CYP4A1 and CYP4A2 Genes, Unique Exon Organization, and Identification of a Conserved 19-bp Upstream Element. *DNA*. **8 (7)** : 503-516.

Kimura, S., Hardwick, J.P., Kozak, C.A. and Gonzalez, F.J. (1989b). The Rat Clofibrate-Inducible CYP4A subfamily II. cDNA Sequence of IVA3, Mapping of the *Cyp4a* Locus to Mouse Chromosome 4, and Coordinate and Tissue-Specific Regulation of the CYP4A genes. *DNA*. **8 (7)** : 517 - 525.

Klingenberg, M. (1958). Pigments of rat liver microsomes. *Arch. Biochem. Biophys*. **75** : 376 - 386.

Kolvraa, S. and Gregarsen, N. (1986). In vitro studies on the oxidation of medium-chain dicarboxylic acids in rat liver. *Biochim. Biophys. Acta*. **876 (3)** : 515 - 525.

Kraulis, S. (1991). MOLSCRIPT : A program to produce both detailed and schematic plots of protein structure. *J. Appl. Crystalog*. **24** : 946-950

Kraupp-Grasl, B., Huber, W., Taper, H. and Schulte-Hermann, R. (1991). Increased susceptibility of aged rats to hepatocarcinogenesis by the peroxisome proliferator nafenopin and the possible involvement of altered liver foci occurring spontaneously. *Cancer Res*. **51 (2)** : 666 - 671.

Kronbach, T., Kemper, B. and Johnson, E.F. (1991). A Hypervariable Region of P450IIC5 Confers Progesterone 21-Hydroxylase Activity of P450IIC1. *Biochemistry*. **30 (25)** : 6097 - 6102

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol*. **157 (1)** : 105 - 132.

Lake, B.G., Gangolli, S.D., Grasso, P. and Lloyd, A.G. (1975). Studies on the hepatic effects of

orally administered di-(2-ethylhexyl)phthalate in the rat. *Toxicol. Appl. Pharmacol.* **32** : 355 - 367.

Lake, B.G., Kozlen, S.L., Evans, J.G. and Gray, T.J.B. (1987). Effect of prolonged administration of clofibric acid and di-(2-ethylhexyl)-phthalate on hepatic enzyme activities and lipid peroxidation in the rat. *Toxicology*. **44** : 213 - 228.

Lake, B.G., Evans, J.G., Gray, T.J.B. and Korosi, S.A. (1989). Comparative studies on nafenopin-induced hepatic peroxisome proliferation in the rat, syrian hamster, guinea pig and marmoset. *Toxicol. Appl. Pharmacol.* **99** : 148 - 160.

Lalwani, N.D., Alvares, K., Reddy, M.K., Reddy, M.N., Parikh, I. and Reddy, J.K. (1987). Peroxisome proliferator-binding protein: identification and partial characterization of nafenopin-, clofibric acid-, and ciprofibrate-binding proteins from rat liver. *Proc. Natl. Acad. Sci. (USA).* **84 (15)** : 5242 - 5246.

Lazarow, P.B. and Fujiki, Y. (1985). Biogenesis of peroxisomes. *Annu. Rev. Cell. Biol.* **1** : 489 - 530.

Lee, S.-T., Pineau, J., Drago, E.K., Lee, J.W., Owens, D.L., Kroetz, P.M., Fernandez-Salguero, H., Westphal and Gonzalez, F.J. (1995). Target disruption of the a isoform of the peroxisome proliferator-activated receptor gene in mice results in abolishment of the pleiotropic effects of peroxisome proliferators. *Mol. Cell. Biol.* **15** : 3012 - 3022.

Lehmann, J.M., Moore, L.B., Smith-Oliver, T.A., Wilkison, W.O., Willson, T.M. and Kliewer, S.A. (1995). An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). *J. Biol. Chem.* **270 (22)** : 12953 - 12956.

Lemberger, T., Saladin, R., Vazquez, M., Assimacopoulos, F., Staels, B., Desvergne, B., Wahli, W. and Auwerx, J. (1996). Expression of the peroxisome proliferator-activated receptor alpha

gene is stimulated by stress and follows a diurnal rhythm. *J. Biol. Chem.* **271 (3)** : 1764 - 1769.

Li, E., Bestov, A.R. and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell.* **69** : 915 - 926.

Li, W.-H., Ellsworth, D.L., Krushkal, J., Chang, B.H. and Hewett-Emmett, D. (1996). Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Mol. Phylogenet. Evol.* **5 (1)** : 182 - 187.

Lock, E.A., Mitchell, A.M. and Elcombe, C.R. (1989). Biochemical mechanisms of induction of hepatic peroxisome proliferation. *Annu. Rev. Pharmacol. Toxicol.* **29** : 145 - 163.

Lu, Q., Wallrath, L.L., Allan, B.D., Glaser, R.L., Lis, J.T. and Elgin, S.C.R. (1992). Promoter sequence containing $(CT)_n.(GA)_n$ repeats is critical for the formation of the DNAse I hypersensitive sites in the *Drosophila hsp26* gene. *J. Mol. Biol.* **225 (4)** : 985 - 998.

Lundgren, B., Andersson, K. and Depierre, J.W. (1992) . Effects of dietary-treatment with 11 dicarboxylic-acids, diethylcarboxylic esters and fatty-acids on peroxisomal fatty-acid beta-oxidation, epoxide hydrolases and lauric acid omega- hydroxylation in mouse-liver. *Biochem. Pharmacol.* **43 (4)** : 785 - 792.

Ma, Y.-H., Gebremedhin, D., Schwartzman, M.L., Falck, J.R., Clark, J.E., Masters, B.S.S., Harder, D.R. and Roman, R.J. (1993). 20-hydroxyeicosatetraenoic acid is an endogenous vasoconstrictor of canine renal arcuate arteries. *Circ. Res.* **72** : 126 - 136.

MacLeod, J.N. and Shapiro, B. H. (1989). Growth Hormone Regulation of Hepatic Drug-Metabolizing Enzymes in Mouse. *Biochem. Pharmacol.* **38 (10)** : 1673 - 1677.

Macleod, D., Charlton, J., Mullins, J. and Bird, A.P. (1994). Sp1 sites in the mouse APRT gene

promoter are required to prevent methylation of the CpG island. *Genes Dev.* **8** : 2282 -2292.

Mannaerts, G.P., Thomas, J., Debeer, L.J., McGarry, J.D. and Foster, D.W. (1978). Hepatic fatty acid oxidation and ketogenesis after clofibrate treatment. *Biochim. Biophys. Acta.* **529 (2)** : 201 - 211.

Mannaerts, G.P., Van Veldhoven, P.P., Van Brocekhoven, A., Vandebroek, G. and Debeer, L.J. (1982). Evidence that peroxisomal acyl-coA synthetase is located at the cytoplasmic side of the peroxisomal membrane. *Biochem. J.* **204** : 17 - 23.

Martin , D.I.K., Fiering, S. and Groudine, M. (1996). Regulation of β-globin gene expression : straightening out the locus. *Current Opin. Gen. Dev.* **6** : 488 - 495.

Matsubara, S., Yamamoto, S., Sogawa, K., Yokotani, N., Fujii-Kuriyama, Y., Haniu, M., Shively, J.E., Gotoh, O., Kusunose, E. and Kusunose, M. (1987). cDNA Cloning and Inducible Expression during Pregnancy of the mRNA for Rabbit Pulmonary Prostaglandin ω-Hydroxylase (Cytochrome P-450$_{p-2}$). *J. Biol. Chem.* **262 (27)** : 13366 0 13371.

Matsunaga, E., Umeno, M. and Gonzalez, F.J. (1990). The rat P450-IID subfamily - complete sequences of 4 closely linked genes and evidence that gene conversions maintained sequence homogeneity at the heme-binding region of the cytochrome P450 active site. *J. Mol. Evol.* **30 (2)** : 155 - 169.

McGuire, E.J., Gray, R.H. and Iglesia, D.L. (1992). Chemical structure-activity relationships : peroxisome proliferation and lipid regulation in rats. *J. Am. College Toxicol.* **11 (3)** : 353 - 361.

Milton, M.N., Elcombe, C.R. and Gibson, G.G. (1990). On the mechanism of induction of microsomal cytochrome P450IVA1 and peroxisome proliferation in rat liver by clofibrate. *Biochem. Pharmacol.* **40 (12)** : 2727 - 2732.

Mitchell, A.M., Lhuguenot, J.C., Bridges, J.W. and Elcombe, C.R. (1985). Identification of the proximate peroxisome proliferator(s) derived from di(2-ethylhexyl) phthalate. *Toxicol. Appl. Pharmacol.* **80 (1)** : 23 - 32.

Morohashi, K., Fujii-Juriyama, Y., Okada, Y., Sogawa, K., Hirose, T., Inayama, S. and Omura, T. (1984). Molecular cloning and nucleotide sequence of cDNA for mRNA of mitochondrial cytochrome P-450 (SCC) of bovine adrenal cortex. *Proc. Natl. Acad. Sci. (USA).* **81** : 4647 - 4651.

Muerhoff, A.S., Griffin, K.J. and Johnson, E.F. (1992a). The peroxisome proliferator-activated receptor mediates the induction of CYP4A6, a cytochrome P450 fatty acid omega-hydroxylase, by clofibric acid. *J. Biol. Chem.* **267 (27)** : 19051 - 19053.

Muerhoff, A.S., Griffin, K.J. and Johnson, E.F. (1992b). Characterization of a Rabbit Gene Encoding a Clofibrate-Inducible Fatty Acid ω-Hydroxylase : CYP4A6. *Arch. Biochem. Biophys.* **296 (1)** : 66 - 72.

Mukherjee, R., Jow, L., Noonan, D. and McDonnell, D.P. (1994). Human and rat peroxisome proliferator activated receptors (PPARs) demonstrate similar tissue distribution but different responsiveness to PPAR activators. *J. Steroid. Biochem. Mol. Biol.* **51 (3-4)** : 157 - 166.

Neat, C.E., Thomassen, M.S. and Osmundsen, H. (1988). Induction of peroxisomal β-oxidation in rat liver by high-fat diets. *Biochem. J.* **186** : 369 - 371.

Nebert, D.W. and Gonzalez, F.J. (1985). P450 genes : Structure, evolution and regulation. *Annu. Rev. Biochem.* **56** : 943 - 945.

Nebert, D.W., Adesnik, M., Coon, M.J., Estabrook, R.W., Gonzalez, F.J., Guengerich, F.P., Gunsalus, I.C., Johnson, E.F., Kemper, B., Levin, W., Phillips, J.R., Sato, R. and Waterman,

M.R. (1987). The P-450 gene superfamily : Recommended nomenclature. *DNA* **5** : 1 - 11.

Nelson, D.R. and Strobel, H.W. (1987). Evolution of Cytochrome P-450 Proteins. *Mol. Biol. Evol.* **4 (6)** : 572-593.

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

Nemali, M.R., Usuda, N., Reddy, M.K., Oyasu, K., Hashimoto, T., Osumi, T., Rao, M.S. and Reddy, J.K. (1988). Comparison of constitutive and inducible levels of expression of peroxisomal beta-oxidation and catalase genes in liver and extrahepatic tissues of rat. *Cancer Res*. **48 (18)** : 5316 - 5324.

Nakajima-Iijima, S., Hamada, H., Reddy, P. and Kakunaga, T. (1985). Molecular structure of the human cytoplasmic β-actin gene : interspecies homology of sequences in the introns. *Proc. Natl. Acad. Sci. (USA)*. **82** : 6133 - 6137.

Nhamburo, P.T., Gonzalez, F.J., McBridem O.W., Gelboin, H.V. and Kimura, S. (1989). Identification of a new P450 expressed in human lung : complete cDNA sequence, cDNA-directed expression, and chromosome mappingl. *Biochemistry*. **28** : 8060 - 8066.

Novacek, M.J. (1992). Mammalian phylogeny : shaking the tree. *Nature*. **356** : 121 - 125.

Ohta, T. (1989). Role of gene duplication in evolution. *Genome*. **31** : 304 - 310.

Okey, A.B. (1990). Enzyme induction in the cytochrome P-450 system. *Pharmacol. Ther.* **45 (2)** : 241 - 298.

Omata, K., Abraham, N.G. and Laniado-Schwartzmann, M. (1992). Arachidonic acid w/w-1 hydroxylase along the nephron of the spontaneously hypertensive rat. *Am. J. Physiol.* **262** : F592 599.

Omura, T. and Ito, A. (1991). Biosynthesis and intracellular sorting of mitochondrial forms of cytochrome P450. *Methods Enzymol.* **206** : 75 - 81.

Orellana, A., Hidalgo, P.C., Morales, M.N., Mezzano, D. and Bronfman, M. (1990). Palmitoyl-coa and the acyl-coa thioester of the carcinogenic peroxisome-proliferator ciprofibrate potentiate diacylglycerol- activated protein kinase-c by decreasing the phosphatidylserine requirement of the enzyme. *Euro. J. Biochem.* **190 (1)** : 57 - 61.

Orton, T.C. and Parker, G.L. (1982). The effect of hypolipidaemic agents on the hepatic microsomal drug-metabolizing enzyme system of the rat. *Drug Metabolism and Disposition.* **10** : 110 - 115.

Osumi, T., Hashimoto, T. and Ui, N. (1980). Purification and properties of acyl CoA oxidase from rat liver. *J. Biochem.* **87** : 1735 - 1746.

Osumi, T., Wen, J.-K. and Hashimoto, T. (1991). 2 cis-acting regulatory sequences in the peroxisome proliferator-responsive enhancer region of rat acyl-coA oxidase gene. *Biochem. Biophys. Res. Commun.* **175** : 866- 871.

Palmer, C.N.A., Griffin, K.J. and Johnson, E.F. (1993a). Rabbit prostaglandin omega-hydroxylase (CYP4A4) : gene structure and expression. *Arch. Biochem. Biophys.* **300 (2)** : 670 - 676.

Palmer, C.N.A., Richardson, T.H., Griffin, K.J., Hsu, M.-H., Muerhoff, A.S., Clark, J.E. and Johnson, E.F. (1993b). Characterization of a cDNA encoding a human kidney, cytochrome P-450 4A fatty acid ω-hydroxylase and the cognate enzyme expressed in *Escherichia coli*. *Biochim.*

Yee M. Heng
Biophys. Acta. **1171** : 161 - 166

Palmer, C.N.A., Hsu, M.-H., Muerhoff, A.S., Griffin, K.J. and Johnson, E.F. (1994). Interaction of the Peroxisome Proliferator-activated Receptor α with the Retinoid X Receptor α Unmasks a Cryptic Peroxisome Proliferator Response Element That Overlaps an ARP-1-binding Site in the CYP4A6 Promoter. *J. Biol. Chem.* **269** : 18083 - 18089.

Palmer, C.N.A., Hsu, M.-H., Griffin, K.J. and Johnson, E.F. (1995). Novel sequencing determinants in peroxisome proliferator signalling. *J. Biol. Chem.* **270 (27)** : 16114 - 16121.

Palosaari, P.M. and Hiltunen, J.K. (1990). Peroxisomal bifunctional protein from rat liver is a trifunctional enzyme possessing 2-enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and Δ^3, Δ^2-enoyl-CoA isomerase activities. *J. Biol. Chem.* **265 (5)** : 2446 - 2449.

Pearson , W.R. and Lipman, D.J. (1988). Improved Tools for Biological Sequence Comparison. *Proc. Natl. Acad. Sci. (USA*). **85** : 2444 - 2448.

Philipsen, S., Pruzina, S. and Grosveld, F. (1993). The minimal requirements for activity in transgenic mice of hypersensitive site 3 of the beta globin locus control region. *EMBO J.* **12** : 1077 - 1085.

Poland, A. and Knutson, J.C. (1982). 2,3,7,8-tetrachlorodibenzo-p-dioxin and related halogenated aromatic hydrocarbons: examination of the mechanism of toxicity. *Annu. Rev. Pharmacol. Toxicol.* **22** : 517 - 554.

Poland, A., Glover, E., Ebetino, F.H. and Kende, A.S. (1986). Photoaffinity labeling of the *Ah* receptor. *J. Biol. Chem.* **261 (14)** : 6352 - 6365.

Porter, T.D. and Coon, M.J. (1991). Cytochrome P-450. Multiplicity of isoforms, substrates,

and catalytic and regulatory mechanisms. *J. Biol. Chem.* **266 (21)** : 13469 - 13472.

Poulos, T.L. Finzel, B.C., Gunsalus, I.C., Wagner, G.C. and Kraut, J. (1986). The 2.6-A crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.* **260** : 16122 - 16130.

Poulos, T.L., Finzel, B.C. and Howard, A.J. (1987). High-resolution crystal structure of cytochrome 450_{cam}. *J. Mol. Biol.* **195** : 687 - 700.

Price, S.C., Hinton, R.H., Mitchell, F.E., Hall, D.E., Grasso, P., Blane, G.F. and Bridges, J.W. (1986). Time and dose study on the response of rats to the hypolipidaemic drug fenofibrate. *Toxicology*. **41** : 169 - 191.

Raag, R. and Poulos, T.L. (1989). Crystal Structure of the Carbon Monoxide-substrate-cytochrome P-450_{cam} Ternary Complex. *Biochemistry*. **28** : 7586 - 7592.

Raag, R., Martinis, S.A., Sligar, S.G. and Poulos, T.L. (1991). Crystal Structure of the Cytochrome P-450_{cam} Active Site Mutant Thr252Ala. *Biochemistry*. **30** : 11420 - 11429.

Ravichandran, K.G., Boddupalli, S.S., Hasemann, C.A., Peterson, J.A. and Deisenhofer, J. (1993). Crystal structure of hemoprotein domain of $P450_{BM-3}$, a prototype for microsomal P450s. *Science*. **261** : 731 - 736.

Reddy, J.K. and Kumar, N.S. (1979). Stimulation of catalase synthesis and increase of carnitine acetyltransferase activity in the liver of intact female rats fed clofibrate. *J. Biochem. (Tokyo)* **85 (3)** : 847 - 856.

Reddy, J.K., Azarnoff, D.L. and Hignite, C.E. (1980). Hypolipidaemic hepatic peroxisome proliferators form a novel class of chemical carcinogens. *Nature*. **283 (5745)** : 397 - 398.

Reddy, J.K. and Lalwani, N.D. (1983). Carcinogenesis by hepatic peroxisome proliferators: eval-

uation of the risk of hypolipidemic drugs and industrial plasticizers to humans. *Crit. Rev. Toxicol.* **12 (1)** : 1 - 58.

Rodriguez, J.C., Gil-Gomez, G., Hegardt, F.G. and Haro, D. (1994). Peroxisome proliferator-activated receptor mediates induction of the mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase gene by fatty acids. *J. Biol. Chem.* **269** : 18787 - 18772.

Roman, L.J., Palmer, C.N.A., Clark, J.E., Muerhoff, A.S., Griffin, K.J., Johnson, E.F. and Masters, B.S.S. (1993). Expression of Rabbit Cytochromes P4504A Which Catalyze the ω-Hydrozylation of Arachidonic Acid, Fatty Acids, and Prostaglandins. *Arch. Biochem. Biophys.* **307 (1)** : 57-65.

Russell, T.R. and Ho, R. (1976). Conversion of 3T3 fibroblasts into adipose cells: triggering of differentiation by prostaglandin F2alpha and 1-methyl-3-isobutyl xanthine. *Proc. Natl. Acad. Sci. (USA).* **73 (12)** : 4516 - 4520.

Sabatini, D.A., Kreibich, G., Morimoto, T. and Adesnik, M. (1982). Mechanisms for the incorporation of proteins in membranes and organelles. *J. Cell. Biol.* **92** : 1 - 22.

Sacerdoti, D., Abraham, N.G., McGiff, J.C. and Schwartzman, M.L. (1987). Renal cytochrome P450-dependent metabolism of arachidonic acid in spontaneously hypertensive rats. *Biochem. Pharmacol.* **37** : 521 - 527.

Saffer, J.D., Jackson, S.P. amd Armarella, M.B. (1991). Developmental expression of Sp1 in the mouse. *Mol. Cell. Biol.* **11** : 2189 - 2199.

Sagami, I., Ohmachi, T., Fujii, H. and Watanabe, M. (1986). Pulmonary cytochrome P450 from 3-methylcholanthrene-treated hamsters. Purification, characterization, and metabolism of benzo(a)pyrene. *J. Biochem.* **100** : 449 - 457.

Saitou, N. and Nei, M. (1987). The Neighbour Joining Method : A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **4** : 406 - 425.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning. A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Samuelsson, B., Dahlen, S.-E., Lindgren, J.-A., Rouzer, C.A. and Serhan, C.N. (1987). Leukotrienes and lipoxins: structures, biosynthesis, and biological effects. *Science.* **237 (4819)** : 1171 - 1176.

Schepers, L., Van Veldohoven, P.P., Casteels, M., Eyssen, H.J. and Mannaerts, G.P. (1990). Presence of 3 acyl-coa oxidases in rat-liver peroxisomes - an inducible fatty acyl-coA oxidase, a noninducible fatty acyl-coa oxidase, and a noninducible trihydroxycoprostanoyl-coA oxidase. *J. Biol. Chem.* **265 (9)** : 5242 - 5246.

Schmid, C.W. and Jelinek, W.R. (1982). The *Alu* family of Dispersed Repetition Sequences. *Science.* **216 (4550)** : 1065 - 1070.

Schmidt, A., Endo, N., Rutledge, S.J., Vogel, R., Shinar, D. and Rodan, G.A. (1992). Identification of a new member of the steroid hormone receptor superfamily that is activated by a peroxisome proliferator and fatty acids. *Mol. Endocrinol.* **6 (10)** : 1634 - 1641.

Schwartzman, M.L., da Silva, J.-L., Lin, F., Nishimura, M. and Abraham, N.G. (1996). Cytochrome P450 4A Expression and Arachidonic Acid Omega-Hydroxylation in the Kidney of the Spontaneously Hypertensive Rat. *Nephron.* **73** : 652 - 663.

Sharma, R.K., Lake, B.G., Foster, J. and Gibson, G.G. (1988). Microsomal cytochrome P452 induction and peroxisome proliferation by hypolipidaemic agents in rat liver : a mechanistic interrelationship. *Biochem. Pharmacol.* **37** : 1193 - 1201.

Sher, T., Yi, H.F., McBride, O.W. and Gonzalez, F.J. (1993). cDNA cloning, chromosomal mapping and functional characterization of the human peroxisome proliferator activated receptor. *Biochemistry*. **32** : 5598 - 5604.

Simmons, D.L., McQuiddy, P. and Kasper, C.B. (1987). Induction of the hepatic mixed-function oxidase system by synthetic glucocorticoids : transcriptional and post-transcriptional regulation. *J. Biol. Chem*. **262** : 326 - 332.

Small, G.M ., Burdett, K. and Connock, M.J. (1982). Clofibrate-induced changes in enzyme activities in liver, kidney and small intestine of male mice. *Ann. New York Acad. Sci*. **386 (May)** : 460 - 463.

Sorenson, H.N., Hvattum, E., Paulssen, E.J., Gautvik, K.M., Bremer, J. and Spydevold, O. (1993). Induction of peroxisomal acyl-coA oxidase by 3-thia fatty acid, in hepatoma cells and hepatocytes in culture is modified by dexamethasone and insulin. *Biochim. Biophys. Acta*. **1171** : 263 - 271

Steineger, H.H., Sorensen, H.N., Tugwood, J.D., Skrede, S., Spydevold, O. and Gautvik, K.M. (1994). Dexamethasone and insulin demonstrate marked and opposite regulation of the steady-state mRNA level of the peroxisomal proliferator-activated receptor (PPAR) in hepatic cells. *Eur. J. Biochem*. **225** : 967 - 974.

Straub, P., Lloyd, M., Johnson, E.F. and Kemper, B. (1993). Cassette mutagenesis of a potential substrate recognition region of cytochrome P450 2C2. *J. Biol. Chem*. **268** : 21997 - 22003.

Stromstedt, M., Hayashi, S.-I., Zaphiropoulos, P.G. and Gustaffson, J.-A. (1990). Cloning and Characterization of a Novel Member of the Cytochrome P450 Subfamily IVA in Rat Prostate. *DNA and Cell Biology*. **9 (8)** : 569-577

Styles, J., Elliott, B.M., Lefevre, P.A., Robinson, M., Pritchard, N., Hart, D. and Ashby, J. (1985). Irreversible depression in the ratio of tetraploid:diploid liver nuclei in rats treated with 3'-methyl-4-dimethylaminoazobenzene (3'M). *Carcinogenesis*. **6 (1)** : 21 - 28.

Sundseth, S.S. and Waxman, D.J. (1992). Sex-dependent Expression and Clofibrate Inducibility of Cytochrome P450 4A Fatty Acid ω-Hydroxylases. *J. Biol. Chem*. **267 (6)** : 3915 - 3921.

Sutter, T.R., Tang, Y.M., Hayes, C.L., Wo, Y.Y.P., Jabs, E.W., Li, X., Yin, H., Cody, C.W. and Greenlee, W.F. (1994). Complete cdna sequence of a human dioxin-inducible messenger-rna identifies a new gene subfamily of cytochrome-P450 that maps to chromosome-2. *J. Biol. Chem*. **269 (18)** : 13092 - 13099.

Svoboda, D.J. and Azarnoff, D.L. (1979). Tumors in male rats fed ethyl chlorophenoxyisobutyrate, a hypolipidemic drug. *Cancer Res*. **39 (9)** : 3419 - 3428.

Szklarz, G.D., He, Y.Q., Kedzie, K.M., Halpert, J.R. and Burnett, V.L. (1996). Elucidation of amino acid residues critical for unique activities of rabbit cytochrome P450 2B5 using hybrid enzymes and reciprocal site directed mutagenesis with rabbit cytochrome P450 2B4. *Arch. Biochem. Biophys*. **327 (2)** : 308 - 318.

Tamburini, P.P., Masson, H.A., Bains, S.K., Makowski, R.J., Morris, B. and Gibson, G.G. (1984). Multiple forms of hepatic cytochrome P450 : purification, characterization, and comparison of a novel clofibrate-induced isozyme with other forms of cytochrome P-450. *Eur. J. Biochem*. **139** : 235 - 246.

Tomaszewski, K.E. and Melnick, R.L. (1994). In vitro evidence for involvement of CoA thioesters in peroxisome proliferation and hypolipidaemia. *Biochim. Biophys. Acta*. **1220 (2)** : 118 - 124.

Tontonoz, P., Hu, E., Graves, R.A., Budavari, A.I. and Spiegelman, B.M. (1994). mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. *Genes & Dev.* **8 (10)** : 1224 - 1234.

Trusov, Y.A. and Dear, P.H. (1996). A molecular clock based on the expansion of gene families. *Nucleic. Acids. Res.* **24 (6)** : 995 - 999.

Van Arsdell, S.W., Denison, R.A., Bernstein, L.B., Weiner, A.M., Manser, T. and Gesteland, R.F. (1981). Direct repeats flank 3 small nuclear RNA pseudogenes in the human genome. *Cell.* **26 (1)** : 11 - 17.

Van den Berg, J., van Ooyen, A., Mantei, N., Schambock, A., Grosveld, G., Flavell, R.A. and Weissmann, C. (1978). Comparison of cloned rabbit and mouse beta-globin genes showing strong evolutionary divergence of two homologous pairs of introns. *Nature.* **276 (5683)** : 37 - 44.

Van Veldhoven, P. P., Vanhove, G., Vanhoutte, F., Dacremount, G., Parmentier, G., Eyssenm H.J. and Mannaerts, G.P (1991). Identification and purification of peroxisomal branched fatty acyl-coA oxidase. *J. Biol. Chem.* **266** : 24676 - 24683.

Wang, W. and Lu, M. (1995). Effect of arachidonic acid on activity of the apical K^+ channel in the thick ascending limb of the rat kidney. *J. Gen. Physiol.* **106** : 727 - 743.

Warren, J.R., Simmon, V.F. and Reddy, J.K. (1980). Properties of hypolipidaemic peroxisome proliferators in the lymphocyte [^{3}H] thymidine and *Salmonella* mutagenesis assays. *Cancer Res.* **40** : 36 - 41.

Watkins, P..B., Wrighton, S.A., Maurel, P., Schuetz, E.G., Mendez-Picon, G., Parker, G.A. and Guzelian, P.S. (1986). Macrolide antibiotics inhibit the degradation of the glucocorticoid-responsive cytochrome P-450p in rat hepatocytes *in vivo* and in primary monolayer culture. *J. Biol. Chem.* **261** : 6264 - 6271.

Waxman D.J. (1984). Rat hepatic cytochrome P450 isoenzyme 2c : identification as a male-specific, developmentally induced steroid 16α-hydroxylase and comparison to a female-specific cytochrome P-450 isoenzyme. *J. Biol. Chem.* **259** : 15481 - 15490.

Waxman, D.J., Dannan, G.A. and Gunguerich, F.P. (1985). Regulation of rat hepatic cytochrome P450 : age-dependent expression, hormonal imprinting, and xenobiotic inducibility of sex specific isoenzymes. *Biochemistry.* **24** : 4409 - 4417.

Waxman, D.J., Pampori, N.A., Ram, P.A., Agrawal, A.K. and Shapiro, B.H. (1991). Interpulse interval in circulating growth-hormone patterns regulates sexually dimorphic expression of hepatic cytochrome-P450. *Proc. Natl. Acad. Sci. (USA).* **88 (15)** : 6868 - 6872.

Waxman, D.J. and Azaroff, L. (1992). Phenobarbital induction of cytochrome-p-450 gene-expression. *Biochem. J.* **281 (3)** : 577 - 592.

Weibel, E.R., Staubli, W., Gnagi, H.R. and Hess, F.A. (1969). Correlated morphometric and biochemical studies in the liver cell. I. Morphometric model, stereological methods and normal morphometric data for liver. *J. Cell Biol.* **42** : 68 - 112.

Wilson, A.C., Carlson, S.S. and White, T.J. (1977). Biochemical evolution. *Annu. Rev. Biochem.* **46** : 573 - 639.

Whitlock, J.P. (1993). Mechanistic aspects of dioxin action. *Chem. Res. Toxicol.* **6 (6)** : 754 - 763.

Williams, D.E., Hale, S.E., Okita, R.T. and Masters, B.S.S. (1984). A Prostaglandin ω-Hydroxylase Cytochrome P-450 (P-450$_{PG-\omega}$) Purified from Lungs of Pregnant Rabbits. *J. Biol. Chem.* **259 (23)** : 14600 - 14608

Woollard, P.M. (1986). Stereochemical difference between 12-hydroxy-5,8,10,14-eicosatet-

raenoic acid in platelets and psoriatic lesions. *Biochem. Biophys. Res. Commun.* **136 (1)** : 169 - 176.

Wrighton, S.A., Maurel, P., Schultz, E.G., Watkins, P.B., Young, B. and Guzelian, P.S. (1985). Identification of the cytochrome P450 induced by macrolide antibiotics in the rat liver as the glucocorticoid responsive cytochrome p-450p. *Biochemistry.* **24** : 2171 - 2178.

Wu, C.-I. and Li, W.-H. (1985). Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. (USA).* **82** : 1741 - 1745.

Yamada, J., Sakuma, M. and Suga, T. (1992). Induction of peroxisomal beta-oxidation enzymes by dehydroepiandrosterone and its sulfate in primary cultures of rat hepatocytes. *Biochim. Biophys. Acta.* **1137 (2)** : 231 - 236.

Yokotani, N., Bernhardt, R., Sogawa, K., Kusunose, E., Gotoh, O., Kusunose, M. and Fujii-Kuriyama, Y. (1989). Two Forms of ω-Hydroxylase toward Prostaglandin A and Laurate. *J. Biol. Chem.* **264** (36) : 21665 - 21669.

Yoshimoto, M., Kusunose, E., Yamamoto, S., Maekawa, M. and Kusunose, M. (1986). Purification and characterization of two forms of cytochrome P-450 from rat kidney cortex microsomes. *Biochem. Int.* **13 (5)** : 749 - 755.

Zhang, B., Marcus, S.L., Sajjadi, F.G., Alvares, K., Reddy, J.K., Subramani, S., Rachubinski, R.A. and Capone, J.P. (1992). Identification of a peroxisome proliferator-responsive element upstream of the gene encoding rat peroxisomal enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase. *Proc. Natl. Acad. Sci. (USA).* **89 (16)** : 7541 - 7545.

Zhu, Y., Alvares, K., Huang, Q., Rao, M.S. and Reddy, J.K. (1993). Cloning of a new member of the peroxisome proliferator-activated receptor gene family from mouse liver. *J. Biol. Chem.* **258** : 26817 - 26820.

Zou, A. -P., Imig, J.D., Ortiz de Montellano, P.R., Sui, Z., Falck, J.R. and Roman, R.J. (1994a). Effect of P450 ω-hydroxylase metabolites of arachidonic acid on tubuloglomerular feedback. *Am. J. Physiol. (Renal Fluid Electrolyte Physiol)*. **266** : F934 - F941.

Zou, A.-P., Imig, J.D., Kaldunski, M., Ortiz de Montellano, P.R., Sui, Z. and Roman, R.J. (1994b). Inhibition of renal vascular 20-HETE production impairs autoregulation of renal blood flow. *Am. J. Physiol*. **266** : F275 - F282.