TRANSCRIPTIOINAL CONTROL OF THE HUMAN CYP3A4 GENE

A THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
BY

Vincent Bombail

August 2003

Molecular Toxicology Group, Biochemistry Section, School of Biomedical and Molecular Sciences, University of Surrey, Guildford GU2 7XH

© Vincent Bombail 2003
ABSTRACT.

CYP3A4 is the most abundant P450 enzyme expressed in the human liver and it is responsible for the metabolism of approximately 50% of all clinically administrated drugs. The CYP3A4 gene is transcriptionally regulated by xenobiotics and previous work has demonstrated the first 300bp of proximal promoter to be the minimal requirement for such activation. Several nuclear receptors (CAR, SXR) have been shown to be involved in the induction of the CYP3A4 gene. The aim of this work was to further delineate the molecular basis of CYP3A4 gene expression.

In vitro DNase I footprinting was carried out using HepG2 nuclear extracts to map the sites of DNA-protein interaction within the -301/+7 region of the CYP3A4 gene. Putative protein assignment for these sites using in silico analysis revealed the potential binding of transcription factors previously shown to be involved in the regulation of other CYP genes (Sp1, HNF3 and C/EBPα) at the identified protein-DNA interaction sites. These regulatory regions were then disrupted by mutagenesis and their functional effect assessed by transient transfections of reporter gene plasmids into HuH7 hepatoma cells. Statistically significant reductions of reporter gene expression were observed when putative C/EBPα and HNF sites were altered, in both the basal and rifampicin (SXR ligand) induced states. This finding suggests the involvement of proteins binding at these sites in the regulation of the CYP3A4 gene expression.

An examination of the CYP3A4 promoter (−1056/+7) from 11 human DNA samples exhibiting a 14.3 fold variability in CYP3A mediated metabolism failed to show the presence of mutations. Protein-DNA interaction analysis were carried out within the newly identified CYP3A4*1E and CYP3A4*1F alleles as well as the CYP3A4*1B allele. The results implicate the Sp1 transcription factor in the regulation of the CYP3A4 gene, albeit at a more distal binding site.

The findings described in this thesis suggest a substantial involvement of transcription factors other than SXR/CAR in expression of CYP3A4. Because of the polymorphic expression of several liver- and hormone-dependent transcription factors their role in CYP3A4 regulation must be taken into account to understand drug-induction mechanisms and assess variability in inter-individual drug response.
ACKNOWLEDGMENTS

I would like to thank my supervisors, Dr Nick Plant and Prof Gordon Gibson, for their guidance and support throughout this project.

My thanks also go to the Molecular Toxicology lab members over the past 3 years especially Mr Peter Kentish, Anna, Patricia, Sihem, Michael, Jo, Lindz, Nasser, Taher, Sarah, Karen, Michelle and many visiting scientists and project students for providing moral support and daily entertainment.

Drs Wafaa El-Sankary and Hossein Hamzeiy both deserve to be mentioned in a separate paragraph for being great friends and collaborators besides being great lab mates.

Alexis and Eve, Tushna and Vas, you all know why you are on this page.

In those formative years it was not all about Science. So thank you to... all my Guildford and London friends for hours of fun in the sun, Les, the old teachers for getting me started, merci aux amis d’en France pour tous les bons moments (Eric, Chris A et B, Seb, Géraldine, Séverine, Cathy, Véro), John Peel, Boards of Canada, Mr Scruff and the Mancunian friends, who were there during the redaction of this thesis.

Finalement merci à Maman, Papa, Alexandre et Amandine. Cette thèse est dédiée à Mamie qui aurait été heureuse de voir ce travail finalement achevé.

Unless otherwise stated all the experimental work presented in this thesis was performed by myself.
I am grateful to the University of Surrey for providing financial support.
**ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS</td>
<td>Ammonium Persulfate</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine Triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CAR</td>
<td>Constitutive Androstane Receptor</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>CYP</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethyl Pyrocarbonate</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle Medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>DNase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide Triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EM</td>
<td>Extensive Metaboliser</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electro-Motility Shift Assay</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen Receptor</td>
</tr>
<tr>
<td>ER6</td>
<td>Everted Repeat 6 (6 bp spacer)</td>
</tr>
<tr>
<td>GR</td>
<td>Glucocorticoid Receptor</td>
</tr>
<tr>
<td>HNF</td>
<td>Hepatocyte Nuclear Factor</td>
</tr>
<tr>
<td>kb</td>
<td>kilobases</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani medium</td>
</tr>
<tr>
<td>LBD</td>
<td>Ligand Binding Domain</td>
</tr>
<tr>
<td>MEM</td>
<td>Minimum Essential Medium</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>mut</td>
<td>Mutant</td>
</tr>
<tr>
<td>PAR</td>
<td>Pregnane Activated Receptor</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>PB</td>
<td>Phenobarbital</td>
</tr>
<tr>
<td>PBRE</td>
<td>Phenobarbital Response Element</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PCN</td>
<td>5-pregnen-3β-ol-20-one-16α-carbonitrile</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene Glycol</td>
</tr>
<tr>
<td>PM</td>
<td>Poor Metaboliser</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonyl Fluoride</td>
</tr>
<tr>
<td>PXR</td>
<td>Pregnan X Receptor</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RXR</td>
<td>Retinoid X Receptor</td>
</tr>
<tr>
<td>SDM</td>
<td>Site Directed Mutagenesis</td>
</tr>
<tr>
<td>SEAP</td>
<td>Secretory Endothelial Alkaline Phosphatase</td>
</tr>
<tr>
<td>SSCP</td>
<td>Single Strand Conformation Polymorphism</td>
</tr>
<tr>
<td>SXR</td>
<td>Steroid and Xenobiotic Receptor</td>
</tr>
<tr>
<td>TCDD</td>
<td>2,3,7,8 tetrachlorodibenzo-p-dioxin</td>
</tr>
<tr>
<td>TCPOBOP</td>
<td>1,4-bis[2-(3,5-dichloropyridyloxy)] benzene</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylene diamine</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra-Violet</td>
</tr>
<tr>
<td>VDR</td>
<td>Vitamin D Receptor</td>
</tr>
<tr>
<td>WT</td>
<td>Wild Type</td>
</tr>
<tr>
<td>XRE</td>
<td>Xenobiotic Response Element</td>
</tr>
<tr>
<td>XREM</td>
<td>Xenobiotic Response Element Module</td>
</tr>
</tbody>
</table>
# Table of Contents

Abstract ........................................................................................................................ i
Acknowledgements ...................................................................................................... ii
Abbreviations ............................................................................................................... iii
List of figures ............................................................................................................... x
List of tables ................................................................................................................ xiii

Chapter 1 – Introduction .................................................................................................1
1.1. Cytochrome P450 enzymes .................................................................................... 1
1.1.1. Drug metabolism ........................................................................................... 1
1.1.2. The cytochrome P450 superfamily ................................................................ 1
1.1.3. Major CYP enzymes involved in drug metabolism .................................... 4
1.2. Human CYP3A subfamily .................................................................................... 5
1.2.1. CYP3A4 ......................................................................................................... 7
1.2.2. CYP3A5 ......................................................................................................... 7
1.2.3. CYP3A7 ......................................................................................................... 7
1.2.4. CYP3A43 ....................................................................................................... 8
1.2.5. Drug metabolism by CYP3A enzymes ......................................................... 8
1.3. The CYP3A4 enzyme ........................................................................................... 9
1.3.1. Substrates ..................................................................................................... 10
1.3.2. Regulation of CYP3A4 enzyme activity ..................................................... 10
1.3.3. Drug-drug interactions and first-pass metabolism ....................................... 11
1.3.4. CYP3A4 gene structure ............................................................................. 11
1.4. Gene regulation mechanisms in eukaryotic organisms ........................................ 12
1.4.1. Transcription activation ............................................................................... 14
1.4.2. Nuclear receptors ....................................................................................... 15
1.4.3 CYP gene regulation .................................................................................... 19
1.4.3.1. Potential levels of CYP gene regulation ................................................ 19
1.4.3.2. An example of CYP gene regulation: induction of the CYP1A1 gene ..... 19
1.5. Regulation of the CYP3A4 gene ........................................................................ 21
1.5.1. CYP3A4 is transcriptionally activated ....................................................... 21
1.5.2. CYP3A4 regulatory regions ........................................................................ 21
1.5.3. Regulation of the CYP3A4 transcriptional activation by nuclear receptors 23
1.5.3.1. The glucocorticoid receptor (GR, NR3C1) ................................................ 23
1.5.3.2. The steroid and xenobiotic receptor (SXR, NR1I2) .................................. 24
1.5.3.3. CAR (NR1I3) ...................................................................................... 26
1.5.3.4. Retinoic X Receptor α (RXRα, NR2B1) ................................................... 27
1.5.3.5. Hepatocyte nuclear factor 4 (HNF4, NR2A sub-family) ......................... 27
1.5.3.6. Nuclear receptor cross talk ................................................................... 28
1.6. Polymorphic expression of CYP3A4 and contribution to inter-individual
    variability in drug metabolism ........................................................................... 29
1.7. Project aims ............................................................................................................ 32
# Chapter 2 - Materials and Methods

## 2.1. Materials
- 2.1.1. General chemicals and plasticware ............................................................. 34
- 2.1.2. Plasmids ...................................................................................................... 34

## 2.2. General microbiology techniques
- 2.2.1. Transformation of bacteria .......................................................................... 43
- 2.2.2. Preparation of glycerol stocks ..................................................................... 43
- 2.2.3. Growth of bacterial cultures ....................................................................... 43

## 2.3. General molecular biology techniques
- 2.3.1. Agarose gel electrophoresis ........................................................................ 44
- 2.3.2. DNA extraction from agarose gel ............................................................... 44
- 2.3.3. DNA quantification ....................................................................................... 45
- 2.3.4. Restriction digest ........................................................................................ 45
- 2.3.5. Ethanol precipitation .................................................................................... 45

## 2.4. Cloning
- 2.4.1. PCR cloning .................................................................................................. 45
- 2.4.2. Ligation .......................................................................................................... 46
- 2.4.3. Selection ........................................................................................................ 46
- 2.4.4. Sequencing ..................................................................................................... 47

## 2.5. Plasmid preparation
- 2.5.1. Minipreps .................................................................................................... 47
- 2.5.2. Endofree maxiprep ...................................................................................... 48

## 2.6. General cell culture techniques
- 2.6.1. Culture conditions ....................................................................................... 49
- 2.6.2. Subculture of cells ....................................................................................... 49
- 2.6.3. Storage and recovery of cells in liquid nitrogen ......................................... 49

## 2.7. Nuclear protein extract preparation
- 2.7.1. HepG2 nuclear protein extraction ............................................................... 50
- 2.7.2. Protein electrophoresis and Coomassie staining ......................................... 51
- 2.7.3. Protein quantification .................................................................................. 51

## 2.8. Electro mobility shift assay (EMSA)
- 2.8.1. Electro mobility shift assay with the -301/+7 CYP3A4 proximal promoter fragment ................................................................. 52
  - 2.8.1.1. Probe preparation ................................................................................... 52
  - 2.8.1.2. Binding reaction ...................................................................................... 52
  - 2.8.1.3. Gel electrophoresis ............................................................................... 53
  - 2.8.2. Gel shift assay to determine the effect of DNA mutations on nuclear protein binding .............................................................. 53
  - 2.8.2.1. Probe preparation .................................................................................. 53
  - 2.8.2.2. Binding reaction ..................................................................................... 54
  - 2.8.2.3. Electrophoresis ...................................................................................... 55

## 2.9. In vitro footprinting assays
- 2.9.1. Overview of the method ............................................................................ 55
- 2.9.2. Probe preparation ....................................................................................... 57
- 2.9.3. In vitro footprinting ................................................................................... 58
- 2.9.4. Sequencing reactions .................................................................................. 58
2.9.5. Cerenkov counting ................................................................. 59
2.9.6. Denaturing gel electrophoresis ................................................ 59
2.9.7. Autoradiography .................................................................. 59
2.9.8. Binding site analysis ............................................................... 60

2.10. Mutation detection in 11 human liver samples ....................... 60
2.10.1. Liver samples used in this study ......................................... 60
2.10.2. Genomic DNA extraction from human liver ...................... 61
2.10.3. Polymerase Chain Reaction (PCR) .................................... 61
2.10.4. Sequencing ....................................................................... 63
2.10.4.1. CYP3A4 promoter region ............................................. 63
2.10.4.2. Exon 7 .......................................................................... 63
2.10.5. Sequence analysis .............................................................. 64

2.11. Site Directed Mutagenesis ..................................................... 64
2.11.1. Initial SDM protocol ........................................................... 66
2.11.2. Second SDM protocol ........................................................ 67
2.11.3. Generation of competent E.coli BMH 71-18 mutS bacterial cells 68

2.12. Transient transfection protocol for the reporter gene assay..... 68
2.12.1. Plasmids .......................................................................... 68
2.12.2. Cell seeding ....................................................................... 69
2.12.3. Transfection ...................................................................... 70
2.12.4. Dosing of the cells with rifampicin (day 1 and day 3) ........... 70
2.12.5. Chemiluminescent detection of SEAP protein activity ......... 71
2.12.6. Data analysis of the reporter gene assay ......................... 72

Chapter 3 - Characterisation of DNA:protein interactions in the CYP3A4 proximal promoter ........................................ 73
3.1. Introduction ............................................................................ 73
3.2. Results .................................................................................. 74
3.2.1. Electrophoretic motility shift assay for the proximal promoter (-307/+7 region) ........................................... 74
3.2.2. Footprinting analysis ............................................................. 74
3.2.2.1. Positive control using recombinant AP2 and SV40 enhancer element ..................................................... 75
3.2.2.2. Optimisation of in vitro DNase I footprinting of the −301/+7 fragment ...................................................... 77
3.2.2.3. Footprinting of 3 overlapping fragments and composite model of nuclear protein binding to the CYP3A4 −301/+7 proximal promoter fragment .............................................. 77
3.3. Discussion ............................................................................. 81
3.3.1. Electrophoretic motility shift assay for the proximal promoter (-307/+7 region) ........................................... 81
3.3.2. In vitro DNase I footprinting of the −301/+7 fragment: method optimisation and technical considerations .................................................................................. 81
3.3.3. General footprinting pattern analysis of the 3 overlapping fragments and composite model of nuclear protein binding to the CYP3A4 proximal promoter. 82
3.3.4. Transcription factors potentially binding to the CYP3A4 promoter region 84
3.3.4.1. Specificity protein 1 (Sp1) ............................................... 84
3.3.4.2. Hepatocyte Nuclear Factor 3 (HNF3) and members of the HNF family ...................................................... 85
3.3.4.3. CCAAT/enhancer binding protein alpha (C/EBPα) ....... 86
6.2.1. Liver specific transcription factors and inflammation............... 137
6.2.2. Liver specific transcription factors and inter-individual variability in drug metabolism.......................................................... 138
6.2.2.1. Mutations in liver specific transcription factors................................. 138
6.2.2.2. Variability in the parameters which regulate liver specific transcription factor expression......................................................... 139
6.3. Role of epigenetics on gene regulation.............................................. 141
6.4. Conclusion......................................................................................... 142
6.5. Future directions................................................................................ 143

7. Bibliography........................................................................................... 145
LIST OF FIGURES

Figure 1.1. The cytochrome P450 clans 3
Figure 1.2. Genomic structure of the human CYP3A locus 6
Figure 1.3. Stages in eukaryotic gene expression where regulation can potentially take place 13
Figure 1.4. Chromatin structure and potential location of the regulatory sequences 15
Figure 1.5. Schematic representation of a nuclear receptor 17
Figure 1.6. Phylogenetic tree connecting 65 known NR genes 18
Figure 1.7. Mechanisms of transcriptional activation by AhR and negative feedback regulation of AhR by AhRR 20
Figure 1.8. The CYP3A4 promoter structure and putative transcription factor binding sites, as predicted by bioinformatic analysis 22
Figure 1.9. The CYP3A4V mutation (CYP3A4*1B) located in the NFSE motif 30
Figure 2.1. Map of the human SXR expression plasmid (hPXR-pSG5) 37
Figure 2.2. Map of the pSEAP2-promoter plasmid, an alkaline phosphatase reporter gene plasmid 38
Figure 2.3. Map of the pSEAP2-promoter plasmid with the -301/+7 insert from the CYP3A4 proximal promoter 39
Figure 2.4. Map of the pSEAP2-basic reporter gene plasmid with the distal CYP3A4 gene enhancer XREM 40
Figure 2.5. Map of the pSEAP2-basic plasmid with the XREM insert and the -301/+7 insert from the CYP3A4 promoter (pWT) 41
Figure 2.6. Map of the pSEAP2-basic reporter gene plasmid with the distal enhancer XREM insert and the -1201/-61 insert from the CYP3A4 gene promoter 42
Figure 2.7. Outline of the various in vitro footprinting procedures used 56
Figure 3.1. Autoradiograms of EMSA experiments 76
Figure 3.2. Typical autoradiogram of an in vitro DNase I footprinting experiment on the -301/+7 region of the CYP3A4 gene

Figure 3.3. Composite model of the protein binding to the -301/+7 region of the CYP3A4 gene using in vitro DNase I footprinting on the -301/+7, -388/-83 and -240/+69 promoter fragments

Figure 4.1. Novel CYP3A4 alleles within the context of the putative response elements

Figure 4.2. Reporter gene assay experiments in HepG2 cells demonstrating the effect of the novel mutants on the transcriptional activation by xenobiotics

Figure 4.3. Effect of the CYP3A4*1B, *1E and *1F alleles on HepG2 nuclear protein binding

Figure 4.4. Alignment of the DNA oligomer sequences used to investigate protein binding to the NFSE region in this study and in two previous reports

Figure 5.1. A single base pair mutation alters responsiveness of the -301/+7 region of the CYP3A4 promoter to synthetic and endogenous steroids but not to rifampicin

Figure 5.2. The -190 T>C substitution does not alter the binding pattern of HepG2 nuclear proteins to -301/+7 fragment of the CYP3A4 promoter

Figure 5.3. Binding of HepG2 nuclear proteins to the WT oligomer is reduced by the T>C-190 transition

Figure 5.4. Comparison of the rifampicin induction profile for the pWT plasmid (with a -301/+7 insert) and a longer CYP3A4 promoter fragment (-1201/-61)

Figure 5.5. Comparison of the rifampicin induction profile for pWT and pWT* plasmids

Figure 5.6. Summary of the plasmids generated during the SDM experiments

Figure 5.7. Effect of the mutations in pWT and pWT* on the basal levels of reporter gene expression upon transfection in HuH7 cells

Figure 5.8. Rifampicin response of the ER6 mutant construct upon
transfection in HuH7 cells

**Figure 5.9.** Rifampicin response of the HNF mutant constructs upon transfection in HuH7 cells

**Figure 5.10.** Rifampicin response of the C/EBP mutant constructs upon transfection in HuH7 cells

**Figure 5.11.** Summary of the transfection experiments for the Sp1* and pSp1 plasmids

**Figure 5.12.** Representation of the locations where bases are modified in the pMUT, pHNF, and pHNF* plasmids

**Figure 6.1.** Summary of the mechanisms involved in the hepatic regulation of the CYP3A4 gene
LIST OF TABLES

Table 1.1. Mean IC₅₀ values for inhibitors of CYP3A4, 5 and 7 mediated metabolism of 7-benzyloxy-4-trifluoromethylcoumarin 9
Table 1.2. Examples of reactions performed by CYP3A4 10
Table 2.1. Suppliers of items used in this study 36
Table 2.2. Oligomer sequences for the EMSA studies 54
Table 2.3. PCR primers used to generate the DNA probes for DNase I footprinting 57
Table 2.4. Human liver samples used in this study 61
Table 2.5. PCR primers used in the mutation detection study 62
Table 2.6. Cycling conditions for the PCR reactions 63
Table 2.7. SDM mutagenic primers 65
Table 2.8. SDM selection primers 65
Table 2.9. Plasmids used in the transient transfection experiments 69
Table 4.1. Variation in 3 reported promoter sequences and comparison with the 11 liver samples from the current study 92
Table 4.2. Novel alleles identified following screening of a population of 101 subjects 94
CHAPTER 1 – INTRODUCTION

1.1. Cytochrome P450 enzymes.

1.1.1. Drug metabolism.

In order to maintain physiological homeostasis, a vast number of exogenous or endogenous compounds must be removed from the organism or transformed to modify their biological activity. Water-soluble molecules can be readily excreted whereas hydrophobic molecules must go through a biotransformation process to increase their water solubility. These reactions are called drug metabolism and are divided in two phases. Phase I consists of the addition or the unmasking of a functional group (e.g. hydroxyl) on the molecule to be excreted. This may be followed by a phase II reaction where a conjugate (e.g. glucuronic acid or sulphate) is added to the product (and possibly to the substrate) of the phase I metabolism (Gibson and Skett, 1994). The metabolites may be further processed, for instance compounds conjugated to glutathione may be involved in the biosynthesis of mercapturic acid conjugates (Gibson and Skett, 1994). The excretion of the biotransformed compounds may later take place through the kidneys (urine), the gastrointestinal tract (bile and faeces), the lungs (exhaled air) and body fluids such as sweat and breast milk (Taylor and Reide, 1998).

1.1.2. The cytochrome P450 superfamily.

A whole variety of phase I reactions are carried in the presence of NADPH and oxygen by the mixed function oxidase (MFO) system, which is located in the endoplasmic reticulum of the cell and is composed of the NADPH cytochrome P450 reductase, lipids
and cytochrome P450 enzymes (Gibson and Skett, 1994). Cytochrome P450 enzymes are the terminal elements of an electron transfer chain and they are haeme containing proteins (or haemoproteins) composed of a ferric protoporphyrin IX prosthetic group non covalently associated to an apoprotein which binds the substrate and the oxygen necessary to the MFO reaction (Gibson and Skett, 1994). They catalyse the biotransformation of compounds (RH) into hydroxylated metabolites (ROH) in the presence of NADPH and molecular oxygen according to the following stoichiometry:

$$NADPH + H^+ + O_2 + RH \rightarrow NADP^+ + H_2O + ROH$$

Cytochrome P450’s (or CYP) constitute a superfamily of drug metabolism enzymes which probably arose 3.5 billion years ago and are found in a vast number of taxa in the animal and plant kingdom, filamentous fungi, yeast and bacteria (Nelson et al., 1996). It is thought the first CYP enzymes were involved in endogenous compound metabolism and that they later were involved in xenobiotic detoxification processes (Lewis et al., 1998). Because of the increasing number of CYPs identified, a standard nomenclature has been agreed. The cytochrome P450 enzymes have been grouped in families and subfamilies according to the degree of amino acid sequence similarity (Nelson et al., 1996). A family contains enzymes with 40 % or greater similarity in the amino acid sequence of the apoprotein. The homology is greater within a sub family, where there is a protein similarity greater than 55 % (Nelson et al., 1993). For instance CYP3A4 represents the fourth gene in the A subfamily of the CYP3 family and CYP3A4 is the protein. More recently a higher order of nomenclature, the clans, has been proposed (http://dmelson.utmem.edu/Clans. June 2003). A clan contains several families on the basis of their structural similarity and for instance the CYP3 clan contains CYP3, 5, 6, 9, 13, 25, 28, 30 and 45 (Nelson, 1999). The diagram in figure 1.1 shows an example of CYP gene clans in animal species.

A depository of CYP data is frequently updated and is available on the Internet (Dr David Nelson, Cytochrome P450 homepage, http://dmelson.utmem.edu/cytochromeP450.html. June 2003).
Figure 1.1. The cytochrome P450 clans (Taken from Nelson, 1999). Only one member gene for each family is represented within a given clan. This diagram represents animal CYP genes.
1.1.3. Major CYP enzymes involved in drug metabolism.

The human CYP enzymes involved in xenobiotic metabolism mainly belong to the CYP1, 2 and 3 families (Wrighton and Stevens, 1992). The CYP1 family enzymes (CYP1A1, 1A2 and 1B1) transform and sometimes bioactivate several environmental pollutants especially the procarcinogen polyaromatic hydrocarbons (Nebert and Russell, 2002). The CYP1A2 enzymes appear to metabolise 10-20 drugs, whereas the other CYP1 family members may not be greatly involved in the metabolism of pharmaceutical compounds (Nebert and Russell, 2002). With 13 subfamily and 16 genes in human, the CYP2 family is one of the most important in drug metabolism (Nebert and Russell, 2002). It contains the CYP2C subfamily (CYP2C8, 9, 18 and 19) that is, at least partially, involved in the transformation of half of the commonly used drugs (Nebert and Russell, 2002). The CYP2D6 gene encodes an enzyme which catalyses the transformation of a large number of drugs (more than 75, Nebert and Russell, 2002), including the 4-hydroxylation of debriquisone. Population studies of debriquisone metabolism revealed the existence of a bimodal distribution, which consists in poor (PM) and extensive (EM) metabolisers (Gibson and Skett, 1994). Analyses of PM patient livers allowed the identification of a genetic polymorphism in PM subjects, resulting in the production of an incorrectly spliced CYP2D6 mRNA and consequently in the absence of the CYP2D6 protein (Gonzalez et al., 1988). The identification of such mutation provided a molecular basis for the observed inter-individual variability. The only member of the CYP2E subfamily, CYP2E1, is another important enzyme since it induced by a variety of small organic molecules (e.g. ethanol) and that many of the metabolites are more chemically reactive than the substrates, creating a possible toxicological risk such as during carcinogen activation (Wrighton and Stevens, 1992). Other CYP2 enzymes such as CYP2A6, CYP2A13, CYP2F1 and the phenobarbital inducible CYP2B6 are also capable of transforming drugs (Nebert and Russell, 2002). The CYP3 gene family has a very important role in drug metabolism since its members are known to metabolise at least 120 frequently prescribed pharmaceutical compounds (Nebert and Russell, 2002) together
with endogenous steroid hormones and bile acids. The four members of the CYP3A subfamily are described below (section 1.2).

The three CYP enzyme families described above are also capable of transforming endogenous substrates such as arachidonic acid and its eicosanoid metabolites, which are signalling molecules for a diversity of biological processes such as e.g. platelet aggregation or disaggregation (Nebert and Russell, 2002).

1.2. Human CYP3A subfamily.

Due to the high amino acid sequence similarity observed in the CYP3 family, only one subfamily (CYP3A) exists with four different genes having been described so far in human: 3A4, 3A5 and 3A7 (Nelson et al., 1996) and the newly identified CYP3A43 (Domanski et al., 2001). The CYP3A genes are located on the short arm of chromosome 7 (7q21-q22.1) (Nelson et al., 1996) and the locus has recently been mapped (see figure 1.2) (Finta and Zaphiropoulos, 2000; Gellner et al., 2001). Early reports described another CYP3A gene, CYP3A3 (Nelson et al., 1996), however, its existence has now been refuted since it appears that it was a pseudogene (CYP3P) (Finta and Zaphiropoulos, 2000; Gellner et al., 2001).

CYP3 is a very important family of drug metabolising enzymes because of their occurrence and their ability to metabolise a wide range of substrates (Wrighton and Stevens, 1992). It is estimated that 60% of all currently used drugs are metabolised by enzymes from this family which represents approximately 30% of the total hepatic CYP content in man (see Plant and Gibson, 2003 and references therein).
Figure 1.2. Genomic structure of the human CYP3A locus (drawn to scale and adapted from Finta and Zaphiropoulos, 2000; Gellner et al., 2001). The grey arrows represent the coding regions of the different CYP3A subfamily members and their relative orientations. The black boxes represent two pseudogenes (CYP3AP2 and CYP3AP1).
1.2.1. CYP3A4.

CYP3A4 is found in the majority of individuals and is the major hepatic P450 enzyme in healthy adult liver, however, it is not expressed in foetal liver (Lacroix et al., 1997). Additionally, CYP3A4 mRNA has been detected using in situ hybridisation in extrahepatic tissues such as the stomach, colon, oesophagus and in the mucosa of the small intestine segments (McKinnon et al., 1995). This enzyme will be further described in section 1.3.

1.2.2. CYP3A5.

CYP3A5 appears to be polymorphically expressed in adults and was found to be expressed only in a fraction of the population tested, ranging from 10-20% in Caucasians, 33% in Japanese to 55% in African Americans (Lamba et al., 2002). This polymorphic expression seems to be caused by two mutations in the CYP3A5 coding region, termed CYP3A5*3 and CYP3A5*6 (Kuehl et al., 2001). These two mutants encode mRNA molecules with disrupted splice sites which result in either a truncated protein at amino acid 102 (CYP3A5*3) or a missing exon 7 and a truncated protein at residue 184 (CYP3A5*6) (Kuehl et al., 2001). The enzyme seems to be able to metabolise a lower number of substrates than CYP3A4 (see section 1.2.5), therefore its contribution to hepatic drug metabolism may not be significant (Wrighton and Stevens, 1992).

1.2.3. CYP3A7.

CYP3A7 is only present in human foetus and is the major foetal hepatic enzyme (Wrighton and Stevens, 1992). Using a polyclonal antibody reacting against CYP3A4 and CYP3A7 it was found that during development (foetus, newborn, adult) the hepatic CYP3A levels remain fairly constant (Lacroix et al., 1997). However, the relative amounts of these enzymes vary, CYP3A7 being the dominant form in foetal liver and
after birth CYP3A4 level rises rapidly and reaches adult hepatic content after one week (Lacroix et al., 1997).

1.2.4. **CYP3A43**.

Using genetic analysis of the CYP3A locus (Gellner et al., 2001) and RT-PCR (Domanski et al., 2001), the CYP3A43 gene was identified. Messenger RNA from this gene was detected in the liver, kidney, pancreas and prostate, the latter organ being the site with the highest level of expression. Although the CYP3A43 gene is inducible by rifampicin (Gellner et al., 2001) and the enzyme is able to catalyse testosterone hydroxylation (Domanski et al., 2001), a classic CYP3A reaction, its role in drug metabolism has yet to be established.

1.2.5. **Drug metabolism by CYP3A enzymes**.

Several studies have identified a degree of variability in the expression level of the CYP3A enzymes present in the human liver (reviewed in Lamba et al., 2002). It is therefore important to characterise the relative metabolic capability of each CYP3A enzyme in order to determine which enzymes are likely to affect hepatic drug metabolism. In a comparative study of CYP3A4, 5 and 7 mediated metabolism, the transformation of 10 CYP3A substrates was assessed in vitro (Williams et al., 2002). Overall it was found that, for the substrates tested, metabolite formation was lower with CYP3A7 compared to CYP3A4 and CYP3A5 and that for several substrates (e.g. nifedipine) CYP3A4 was faster (and therefore more efficient at metabolising) than CYP3A5. In another study the substrate interaction with the CYP3A enzyme active site was assessed by measuring the inhibition of 7-benzyloxy-4-trifluoromethylcoumarin (BFC) O-dealkylation by diverse compounds (reviewed in Ekins et al., 2003). The results presented in table 1.1 demonstrate that for all chemicals used, the Inhibitory Concentration 50 (IC50) was the lowest for CYP3A4, which suggests a higher affinity of the compounds tested with the enzyme active site. Taken together these studies suggest
that among the CYP3A subfamily members, CYP3A4 is the most significant in terms of metabolic capability, making it the most important enzyme in hepatic drug metabolism.

Table 1.1. Mean IC$_{50}$ values for inhibitors of CYP3A4, 5 and 7 mediated metabolism of 7-benzyloxy-4-trifluoromethylcoumarin (adapted from Ekins et al., 2003).

<table>
<thead>
<tr>
<th>Test substance</th>
<th>CYP3A4 IC$_{50}$ (µM)</th>
<th>CYP3A5 IC$_{50}$ (µM)</th>
<th>CYP3A7 IC$_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astemizole</td>
<td>0.59</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td>Cisapride</td>
<td>0.038</td>
<td>5.4</td>
<td>16.5</td>
</tr>
<tr>
<td>Clotrimazole</td>
<td>0.002</td>
<td>0.042</td>
<td>0.046</td>
</tr>
<tr>
<td>Cyclosporin A</td>
<td>0.46</td>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>1.8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Ketoconazole</td>
<td>0.016</td>
<td>0.1</td>
<td>0.42</td>
</tr>
<tr>
<td>Mibefradil</td>
<td>0.006</td>
<td>0.23</td>
<td>0.44</td>
</tr>
<tr>
<td>Midazolam</td>
<td>0.46</td>
<td>5.6</td>
<td>24</td>
</tr>
<tr>
<td>Nicardipine</td>
<td>0.008</td>
<td>0.14</td>
<td>0.35</td>
</tr>
<tr>
<td>Nifedipine</td>
<td>6.6</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Nimodipine</td>
<td>0.62</td>
<td>3</td>
<td>3.6</td>
</tr>
<tr>
<td>Terfenadine</td>
<td>1.4</td>
<td>11</td>
<td>29</td>
</tr>
<tr>
<td>Troleandomycin</td>
<td>0.36</td>
<td>1.52</td>
<td>8.9</td>
</tr>
<tr>
<td>Verapamil</td>
<td>0.12</td>
<td>2.2</td>
<td>9.1</td>
</tr>
</tbody>
</table>

1.3. The CYP3A4 enzyme.

The first CYP3A protein was discovered in the female rat (Lu et al., 1972). This protein is inducible and has the same spectral properties as those induced by treatment with phenobarbital in rats but has a different substrate specificity. Watkins and co-workers (Watkins et al., 1985) reported the identification of a 52 kDa protein (called Human Liver P450 or HLP, now CYP3A4) induced by dexamethasone, phenobarbitone and PCN.
1.3.1. Substrates.

CYP3A4 has a very wide range of substrates and can metabolise environmental pollutants, pharmaceutical compounds or endogenous molecules such as testosterone (Li et al., 1995). About 60% of clinically used drugs are metabolised by this enzyme (Cholerton et al., 1992). The substrate size varies from acetaminophen (M=151) to cyclosporine A (M=1201) (Guengerich, 1999). The various types of reactions performed by CYP3A4, as well as examples of substrates are given in table 1.2. Since this large number of substrates may compete for the CYP3A4 active site (Shou et al., 1994) and may increase the protein levels (see below) there is therefore a substantial potential for drug-drug interactions in man (see section 1.3.3).

Table 1.2. Examples of reactions performed by CYP3A4 (adapted from Li et al., 1995).

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Use / Occurrence</th>
<th>Metabolic Pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaminophen</td>
<td>Analgesic</td>
<td>N-oxidation</td>
</tr>
<tr>
<td>Aflatoxin B1</td>
<td>Food hepatocarcinogen</td>
<td>C-oxidation</td>
</tr>
<tr>
<td>(Z)-Benzaldehyde</td>
<td>Insect growth regulator</td>
<td>Dehydration</td>
</tr>
<tr>
<td>Clonazepam</td>
<td>Anticonvulsant</td>
<td>Nitroreduction</td>
</tr>
<tr>
<td>Cocaine</td>
<td>Recreational drug</td>
<td>N-demethylation</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>Antibiotic, Immunosuppressant</td>
<td>O-demethylation</td>
</tr>
<tr>
<td>Testosterone, Oestradiol</td>
<td>Steroid hormones</td>
<td>C-hydroxylation</td>
</tr>
<tr>
<td>Zonisamide</td>
<td>Anticonvulsant</td>
<td>Reductive deamination</td>
</tr>
</tbody>
</table>

1.3.2. Regulation of CYP3A4 enzyme activity.

CYP3A4 substrates are also often inducers, the later being known to influence the production of enzyme in an active form. However, several compounds can alter the activity of this enzyme by directly interacting with the protein. Binding of molecules to the CYP3A4 enzyme and steric interactions between substrates (at least in vitro) can have
regulatory effects on CYP3A4 catalytic activity (Shou et al., 1994). In another study it was found that mifepristone (RU 486) inhibits CYP3A4 by covalent binding to the apoprotein (He et al., 1999) and is therefore a suicide substrate.

1.3.3. Drug-drug interactions and first-pass metabolism.

The levels of CYP3A4 activity may influence drug availability in the organism. If a drug is metabolised and inactivated by hepatic or enteric CYP3A4, an equivalent oral dose will remain in patients with a low metabolic ability for a longer time than in efficient metabolisers, potentially altering the clinical effect observed. Some authors report interactions between CYP3A4 inducers, when co-administered with other drugs. For instance the rate of cyclosporin A metabolism was altered by the presence of several classes of pharmaceutical compounds (e.g. rifampicin or cortisol), in primary cultures of human hepatocytes (Pichard et al., 1990). Such a situation may arise in vivo, for instance in organ transplant patients where the immunosuppressor drug cyclosporin A may be used to avoid an immune reaction against the foreign organ, in conjunction with rifampicin which may be administered to treat an infection. This may result in a decrease of cyclosporin A concentration in the plasma below the therapeutic level (because of the increased drug clearance) and may eventually lead to negative outcomes such as graft rejection. Some CYP3A4 protein inducers are also known to increase the level of intestinal P-glycoprotein which act as drug efflux pumps (Synold et al., 2001). Together with other drug transporters (Hall et al., 1999) these epithelial proteins contribute to the elimination of drugs and therefore their induction further increases the clearance of the substrate/inducers. This phenomenon, known as first pass metabolism (Hall et al., 1999), is of critical importance in drug therapy.

1.3.4. CYP3A4 gene structure.

The CYP3A4 gene as well as about 1kb of the 5’ flanking region was first described by Hashimoto and co-workers (Hashimoto et al., 1993). The gene is approximately 27kb long, the coding region is composed of 13 exons and the exon-intron organisation of CYP3A4 is very similar to that of CYP3A7 (expressed in human foetus) and CYP3A2
There is a 91% similarity between the first 1kb of regulatory region 5' to the transcription start site of \textit{CYP3A4} and \textit{CYP3A7} (Hashimoto et al., 1993). More recent studies report the cloning of 13 kb of \textit{CYP3A4} promoter (Goodwin et al., 1999) and 8.8 kb of \textit{CYP3A7} promoter (Bertilsson et al., 2001). Both of these latter groups describe functional steroid and xenobiotic receptor (SX) and constitutive androstane receptor (CAR) binding sites within a sequence named the xenobiotic responsive enhancer module (XREM, see sections 1.5.3.2 and 1.5.3.3) located approximately 8 kb upstream the transcription start. The promoter regions for both \textit{CYP3A4} and \textit{CYP3A7} share 90 % sequence identity up to -8.8kb (Bertilsson et al., 2001). Early \textit{in silico} analysis of the first kb of promoter suggested the presence of several putative binding sites for regulatory elements such as p53, HNF5, HNF4, ER (oestrogen receptor), GR (glucocorticoid receptor), AP3 (Activator protein 3), octamer 1 (Hashimoto et al., 1993).

1.4. Gene regulation mechanisms in eukaryotic organisms.

Eukaryotic gene expression is a complex process involving the transcription of RNA from the DNA template by the RNA polymerase II and the translation of the information on the RNA into the amino acid sequence of proteins. Transcription involves several steps from chromatin structure alterations which allow transcription initiation to the end of translation, all of which are likely to be subject to regulatory mechanisms (Latchman, 1998). Figure 1.3 provides an overview of the process.
Figure 1.3. Stages in eukaryotic gene expression where regulation can potentially take place (taken from Latchman, 1998).
1.4.1. Transcription activation.

Transcription activation represents one of the initial steps of gene expression and therefore its regulation can be critical for the subsequent steps of gene activation. This process is characterised by the recruitment of specialised proteins to the regulatory regions of genes (Tjian and Maniatis, 1994). This involves the assembly of the TATA-box binding protein with associated proteins (TFIID) together with another transcription factor, TFIIA. This is followed by the binding of more associated factors (TFIIB, F, H and E) and finally by the RNA polymerase II which becomes activated upon phosphorylation of its C terminal domain (Buratowski, 1994). The formation of the complex is favoured by the binding of transcription factors to regulatory sequences situated in various regions surrounding the coding region or even contained within the introns (Figure 1.4). These factors are thought to stabilise the transcription machinery (Latchman, 1998). The binding of these transcription apparatus proteins is strongly influenced by the chromatin structure surrounding the regulatory sites (Bonifer, 1999; Latchman, 1998). Chromatin contains the DNA molecules wrapped around histone proteins forming nucleosomes (see figure 1.4). This chromatin environment may not allow transcription when it is in a compacted form where the DNA is tightly associated with the nucleosomal proteins. However, prior to transcription and as part of a range of cellular processes, chromatin may undergo structural changes allowing access of the transcription machinery to the DNA regulatory regions and coding regions. This is carried out by ATP driven chromatin remodelling complexes, which bind to nucleosomes and modify the interactions between the histones and the DNA (reviewed in Sudarsanam and Winston, 2000).
1.4.2. Nuclear receptors.

Nuclear hormone receptors form a superfamily of transcriptional regulators involved in a variety of developmental and physiological processes (Mangelsdorf et al., 1995). These proteins are evolutionary conserved and a schematic representation of the common structural features is described in figure 1.5. Briefly, these receptors contain a DNA binding domain, a ligand binding domain, together with various other regulatory regions which allow for protein-protein interactions or co-activator recruitment (Aranda and Pascual, 2001). With the exception of the Constitutive Androstene Receptor (CAR, see section 1.5.3.3), which is constitutively active and requires no exogenous ligands to activate target genes (Baes et al., 1994), these receptors are ligand activated transcription factors which transcriptionally modulate target gene expression upon activation. They may act as monomers, homodimers or heterodimers with the retinoid X receptor and bind specific DNA response elements or interact with other regulatory pathways for instance through protein-protein interactions (Aranda and Pascual, 2001). Nuclear receptors interact and form complexes with an array of co-activator and co-repressor proteins (McKenna and O'Malley, 2002). The function of these co-regulators vary as they can sometimes be phosphorylated by MAP kinases, act as bridging factors between other
regulatory proteins or possess histone modifying activities (McKenna and O'Malley, 2002). For instance in an immunoprecipitation experiment using a tagged Peroxisome Proliferator Activated Receptor alpha (PPARα), a total of 25 peptides present in nuclear extracts from rat liver were found to interact with the liganded receptor (Surapureddi et al., 2002). Such proteins often contain a LXXLL motif (where L represents a lysine and X represents any amino acid), which allows for interaction with the AF2 domain in the C-terminal ligand binding domain of the nuclear receptor protein (see figure 1.5). These interactions bring further complexity to the gene regulation process (McKenna and O'Malley, 2002). In addition to their role in mediating signals from xenobiotic exposure, nuclear receptors may mediate the effect of several hormones (e.g. thyroid hormone, steroids) or cellular metabolites (e.g. some lipophilic molecules) (Aranda and Pascual, 2001) providing them with endogenous roles. Some receptors, known as ‘orphan nuclear receptors’, have no known endogenous ligands (Aranda and Pascual, 2001). As for the members of the cytochrome P450 superfamily, the increasing number of nuclear receptors identified prompted researchers to adopt a nomenclature system based on protein sequence homology (Nuclear Receptor Nomenclature Committee, 1999). The diagram in figure 1.6 shows an example of phylogenetic tree representing the sequence homology between 65 nuclear receptors from different families. For clarity the usual protein names will be used in this report instead of the nomenclature name.
Figure 1.5. Schematic representation of a nuclear receptor (taken from Aranda and Pascual, 2001). The DNA binding domain (region C) is involved in the recognition of specific DNA sequences (response elements). The region E and F contain the ligand binding domain as well as a dimerisation surface. Those two domains are connected by a linker domain (D). The ligand independent AF-1 transactivation domain is located in the NH$_2$ terminal region (region A/B) and the ligand dependent AF-2 transactivation domain is contained within the COOH-terminal of the receptor.
Figure 1.6. Phylogenetic tree connecting 65 known NR genes (taken from Nuclear Receptor Nomenclature Committee, 1999). The sequences clustered here are from vertebrate, arthropod and nematode species. The boxed numbers represent the bootstrap values for each subfamily.
1.4.3 CYP gene regulation.

1.4.3.1. Potential levels of CYP gene regulation.
Cytochrome P450 enzyme activities may vary in order to adapt to the organism physiological needs and/or environmental stimuli. A number of CYP genes are expressed constitutively but their basal level of expression may be altered following exogenous or endogenous signals (e.g. presence of a xenobiotic inducer or hormone). This therefore allows a rapid response to the change in chemical levels within the body. The inducibility of CYP genes may involve regulatory mechanisms at several levels of the gene expression process as described above (figure 1.4):
- at the transcriptional level (gene activation or repression),
- at the post-transcriptional level (mRNA splicing, transport, stabilisation, translocation),
- at the protein level (stabilisation of the protein without increase in mRNA).

1.4.3.2. An example of CYP gene regulation: induction of the CYP1A1 gene.
Although some regulation may take place at the post-transcriptional level, most of the CYP1A1 gene induction by aromatic hydrocarbons takes place at the transcriptional level (Denison and Whitlock, 1995). The diagram shown in figure 1.7 summarizes the induction process. The aryl hydrocarbon receptor (Ah receptor) is a cytosolic protein associated with heat shock protein 90 in the inactive state. It is activated upon binding of the inducer (e.g. 2,3,7,8 tetrachlorodibenzo-p-dioxin or TCDD) and following heterodimerisation with the Ah receptor nuclear translocator (Arnt protein) the receptor is translocated into the nucleus, where it binds the Xenobiotic Response Element (XRE) and activates gene transcription (Denison and Whitlock, 1995; Fujii-Kuriyama et al., 1992). Post translational modifications of the protein, such as phosphorylation may also alter the transcriptional effect of AhR (Mimura and Fujii-Kuriyama, 2003). Studies of the CYP1A1 gene promoter revealed that cis-acting DNA elements other than the XRE might be active. For instance, it appears that the induction and expression of the CYP1A1 gene is modulated by several other transcription factors such as the zinc finger protein Sp1.
(Fujii-Kuriyama et al., 1992). More recent work suggests that the local chromatin environment is modified by the liganded AhR, which allows other proteins such as Sp1 to bind the promoter (Mimura and Fujii-Kuriyama, 2003). Among the various AhR gene targets, including detoxification enzymes such as other CYP1 family members or the phase 2 enzyme UDP glucoronosyl transferase, a protein called Ah Receptor Repressor (AhRR) is also involved in the regulation process. It associates with the heterodimerisation partner Arnt, limiting the interactions with AhR, and also binds to the XRE acting as a repressor (Mimura and Fujii-Kuriyama, 2003) thereby creating a negative feedback loop.

**Figure 1.7. Mechanisms of transcriptional activation by AhR and negative feedback regulation of AhR by AhRR.** See text for a detailed discussion. (taken from Mimura and Fujii-Kuriyama, 2003)
1.5. Regulation of the CYP3A4 gene.

1.5.1. CYP3A4 is transcriptionally activated.

CYP3A genes are known to be induced by a variety of compounds including the synthetic glucocorticoid dexamethasone, the antibiotic rifampicin and pregnane compounds such as pregnenolone 16α carbonitrile (Gonzalez et al., 1993). Using nuclear run-on assays it has been reported that CYP3A1 is transcriptionally induced in immature rat livers (Telhada et al., 1992). Together with a substantial body of evidence (Honkakoski and Negishi, 2000), this suggests that the CYP3A4 gene induction is transcriptionally regulated. Although the induction mechanism has not been completely elucidated, some of the major regulatory proteins have been identified (see below).

1.5.2. CYP3A4 regulatory regions.

Following the cloning of CYP3A4, the gene structure within the genomic context has been analysed and descriptive studies of the promoter were reported (see above section 1.3.4) (Hashimoto et al., 1993). The CYP3A4 proximal promoter region was established to be an important regulatory region by reporter gene assay and sequence analysis revealed the presence of putative binding sites for several transcription factors such as the octamer binding protein (Oct 1), activator protein 3 (AP3), progesterone receptor / glucocorticoid receptor (PR/GR), hepatocyte nuclear factor 4 and 5 (HNF4 and 5), oestrogen receptor (ER), p53 and CCAAT-binding protein (CP-1) (Hashimoto et al., 1993). A distal enhancer (XREM) located from -7836 to -7607 in the 5' flanking region was later discovered which contains several functional binding sites for members of the nuclear receptor superfamily (Goodwin et al., 1999). The diagram presented in figure 1.8 represents the putative binding sites in the proximal promoter and the enhancer of the CYP3A4 gene (Gibson et al., 2002). It should be noted that the HNF5 factor was later found not to exist (Schrem et al., 2002) and that some of the sequences claimed to be similar to common response elements are not similar to consensus sequences (personal observation), such as those found in the Transfac database (see section 2.9.8.).
**Figure 1.8.** The *CYP3A4* promoter structure and putative transcription factor binding sites, as predicted by bioinformatic analysis (taken from Gibson *et al.*, 2002)

<table>
<thead>
<tr>
<th>Regulatory motif</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP-3</td>
<td>consensus sequence recognized by the gene-activating factor, activating protein 3</td>
</tr>
<tr>
<td>BTE</td>
<td>basal transcription element involved in maintenance of the basal level of transcription in many genes, possible through the BTE-binding protein (BTEB) in the liver</td>
</tr>
<tr>
<td>CAAT box</td>
<td>consensus sequence present in many genes and binds the liver-specific protein C/EBP (CCAAT enhancer-binding protein) and the ubiquitous transcription factor CP1 (CCAAT-protein-1), has a major role in determining the efficiency of the promoter</td>
</tr>
<tr>
<td>CACCC box</td>
<td>located upstream of the CAAT box and is thought to be involved in transcriptional activation; known to bind the transcription factor Sp1</td>
</tr>
<tr>
<td>DRn</td>
<td>a direct repeat with <em>n</em> bp spacing (AGTTCA-N<em>n</em>-AGTTCA) that binds PXR:RXR heterodimers and activates CYP3A-responsive genes</td>
</tr>
<tr>
<td>ERE/PRE/GRE</td>
<td>receptor-response elements with similar consensus sequences; binding of the ER, PR and GR receptors involved in both positive and negative regulation of gene transcription in response to oestrogens, progestogens and glucocorticoids</td>
</tr>
<tr>
<td>ER6</td>
<td>two everted copies of the AG(G/T)TCA motif, separated by six nucleotides; binds the PXR:RXR receptor complex and activates xenobiotic-dependent gene transcription in CYP3As; also locus of CAR binding</td>
</tr>
<tr>
<td>HNF-4 and HNF-5</td>
<td>regulatory motifs (hepatocyte-specific, nuclear transcription factors) that bind liver-specific transcription factors HNF-4 and HNF-5; also involved in the regulation of <em>CYP2B</em> genes (HNF-4) and transferrin gene (HNF-5)</td>
</tr>
<tr>
<td>OCT</td>
<td>conserved octamer motif present in many genes (including CYP1A1), binding members of the octamer factors family, resulting in the up/down regulation of responsive genes</td>
</tr>
<tr>
<td>TATA box</td>
<td>found in all eukaryotic genes and essential for transcription; recognized by the TATA binding protein (TBP)</td>
</tr>
<tr>
<td>P53</td>
<td>consensus sequence that binds the tumour suppressor protein, p53</td>
</tr>
</tbody>
</table>
1.5.3. Regulation of the CYP3A4 transcriptional activation by nuclear receptors

It is becoming increasingly evident that nuclear receptors are involved in the regulation of CYP genes. Additionally CYP enzyme products can in turn become nuclear receptor ligands and activate other genes (Honkakoski and Negishi, 2000). As will be shown below the regulation mechanisms can be direct or indirect and also allow for the coordinate regulation of several genes, producing a whole body of response to xenobiotic exposure.

1.5.3.1. The glucocorticoid receptor (GR, NR3C1).

In vivo and in vitro work on rat P450 revealed that the isoform initially identified as inducible by the antiglucocorticoid PCN (Lu et al., 1972) was also stimulated by glucocorticoid exposure (Schuetz and Guzelian, 1984) especially by dexamethasone. This P450 enzyme was in fact the product of a gene belonging to the CYP3A subfamily (Nelson et al., 1993; Nelson et al., 1996) and several of the subfamily members were found to be inducible by glucocorticoids in various cell models (Schuetz et al., 1993). Early attempts to provide a molecular mechanism for the induction process aimed to demonstrate the involvement of the glucocorticoid receptor (GR). Firstly two studies on the rat CYP3A1 and CYP3A23 promoters (Quattrochi et al., 1995 and Huss et al., 1996) failed to demonstrate a direct binding of the GR to the promoters. It should be noted that although these two genes have not been shown to be the same, their 5' regulatory regions are identical (Quattrochi and Guzelian, 2001) and therefore CYP3A1 and CYP3A23 are likely to be the same gene. Secondly it was attempted to involve the GR by demonstrating its activation by the non-steroidal CYP3A inducer rifampicin (Calleja et al., 1998). However, subsequent attempts to corroborate this finding were unsuccessful (Herr et al., 2000; Jaffuel et al., 1999). Examination of GR knock-out mice revealed that the presence of the receptor is not essential for cyp3a11 induction by glucocorticoids (Schuetz et al., 2000). In this study the authors do not report whether the expression of other receptors is affected by the loss of GR, therefore compensation mechanisms between different regulatory pathways may take place and the role of GR cannot be ruled out by this experiment. Although co-transfection experiments with a GR expression
plasmid support the hypothesis of GR involvement (El Sankary et al., 2000; Ogg et al., 1999), induction kinetic studies show that CYP3A activation does not resemble that of the tyrosine aminotransferase gene, a classical GR target (Schuetz and Guzelian, 1984) (also discussed in section 5.4). Together with the fact that the induction is not suppressed but enhanced by the addition of the anti-glucocorticoid RU486 (Williams et al., 1997), this suggests the existence of an alternative regulatory pathway.

1.5.3.2. The steroid and xenobiotic receptor (SXR, NR1I2)

The pregnane X receptor (PXR), a nuclear orphan receptor, was first discovered in the mouse in 1998 (Kliewer et al., 1998). Three groups subsequently reported the discovery of this protein in human (Lehmann et al., 1998: hPXR; Blumberg et al., 1998: Steroid and Xenobiotic Receptor, SXR; Bertilsson et al., 1998: Pregnan Activated Receptor, PAR). For the purpose of clarity the term SXR shall be retained for the rest of this thesis. It has been shown that SXR can bind a wide variety of known CYP3A4 inducers including rifampicin and dexamethasone (Lehmann et al., 1998). Northern blot analysis revealed that SXR is highly expressed in the liver and the small intestine (Lehmann et al., 1998), where most of CYP3A4 is expressed (Guengerich, 1999). Reporter gene assays using various inducers show that mPXR (the mouse SXR) and SXR are not activated to the same degree by compounds, providing a possible explanation for the inter-species variability in CYP3A induction (Kliewer et al., 1998; Lehmann et al., 1998). Similarly, the activation profile of SXR, GRα and ERα vary in the presence of various steroid and xenobiotic ligands (Blumberg et al., 1998).

Following activation by a ligand, SXR binds as an heterodimer with RXRα to a response element in the CYP3A4 promoter (Bertilsson et al., 1998; Blumberg et al., 1998; Kliewer et al., 1998). This element, called everted repeat 6 (ER6: TGAACCTaaggAGGTCA), is situated in the 5' region about 150bp upstream of the transcription start site. Additionally the receptor also binds to a direct repeat sequence (DR3: TGAACCTcaTGAACT) and can activate the rat CYP3A23 gene (Lehmann et al., 1998). Various endogenous and synthetic steroids, such as corticosterone and diethylstilbestrol, can bind to this receptor (Bertilsson et al., 1998; Blumberg et al., 1998; Kliewer et al.,
1998; Lehmann et al., 1998). However, the physiological concentrations of these steroids may be too low for activation (Kliwer et al., 1998). Therefore questions are raised about whether SXR is a general steroid sensor or if it has a specific endogenous ligand (Blumberg et al., 1998). However, recent evidence point towards a role of SXR as a sensor for the levels of bile acid and bile acid metabolites. Indeed lithocholic acid and 3-keto lithocholic acid were shown to bind the receptor and to induce cyp3a11 genes in mice (Staudinger et al., 2001; Xie et al., 2001).

The contribution of the SXR in gene activation by RIF, clotrimazole and RU486 has also been reported for CYP3A7 (Pascussi et al., 1999), which has a regulatory region very similar to the CYP3A4 promoter (91% similarity in the first 1 kb). Indeed a functional ER6 was found in the promoter region of the liver specific CYP3A7 (Pascussi et al., 1999). The effect of SXR on other genes has recently been demonstrated. For instance the MDR1 gene, which codes for the P-glycoprotein efflux transporter, was shown to be induced by various SXR ligands including the anticancer drug Paclitaxel (Taxol) (Synold et al., 2001). A study using cDNA and oligonucleotide microarrays with RNA from primary human hepatocytes demonstrated that several CYP genes involved in drug metabolism were induced by rifampicin, with CYP3A4 being the most extensively up-regulated (Rae et al., 2001). Whether SXR has a significant effect on genes not directly involved in drug metabolism still remains to be established.

It is also thought that SXR can be involved in xenobiotic mediated CYP3A4 repression (Takeshita et al., 2002). The anti-fungal drug ketoconazole was found to inhibit SXR-mediated transcription by corticosterone of a CYP3A4 promoter reporter gene construct. The authors suggest the mechanism may be due to a decreased SXR interaction with the coactivator SRC-1 and the repressor SMRT (Takeshita et al., 2002).

The structure of the human SXR ligand binding domain (LBD) has been determined by crystallography (Watkins et al., 2001). It appears that the ligand binding domain is larger in SXR than in most other nuclear receptors studied so far (e.g. Estrogen Receptor, ER) and that it is mainly lined with hydrophobic amino acids (20 out of 28 amino acids) with a few polar residues (4 out of 28 amino acids). Using site directed mutagenesis it was shown that the species-specific activation profile of the SXR could be affected by changing only four amino acid residues (Watkins et al., 2001). The mouse SXR can be
activated by the synthetic steroid 5-pregnen-3β-ol-20-one-16α-carbonitrile (PCN) but not by the human SXR ligand SR12813, whereas the human SXR cannot be activated by PCN. Following the mutation of four amino acids in the LBD of the mouse SXR into four polar amino acids known to interact with the high affinity ligand SR12813 in the human SXR, the mouse SXR was activated by SR12813 by not by PCN in reporter gene assays. This data suggests that species specificity in SXR activation profile may be caused by only a few residues. Studies of the SXR LBD in the presence of ligands (SR12813 and hyperforin, the active ingredient of Saint John’s Wort) revealed the high degree of flexibility of the protein structure in order to accommodate structurally diverse compounds (Watkins et al., 2003; Watkins et al., 2001). It was also shown that the ligand SR12813 could interact with the SXR LBD in three different positions (Watkins et al., 2001). Taken together these results shed light onto the mechanisms by which CYP3A genes can be activated by a wide range of chemicals though SXR and how the receptor can act as a broad chemical sensor.

1.5.3.3. CAR (NR1I3).

CAR is predominantly expressed in the liver and was found to constitutively transactivate (i.e. without xenobiotic ligand binding) retinoic acid response elements, controlling the expression of retinoic acid receptor β2 and alcohol dehydrogenase 3 (Baes et al., 1994). Like other members of the nuclear receptor superfamily, this Constitutively Active Receptor binds to direct repeat hexamers with variable spacers. It was initially shown that CAR could activate CYP2B6 through a phenobarbital response element (PBRE) located in the promoter, in response to phenobarbital and other xenobiotics (Sueyoshi et al., 1999). However, this response, as well as the constitutive activation, is inhibited by the addition of 3α androstenol, a steroid sex hormone metabolite, hence the alternative name for this receptor is Constitutive Androstane Receptor (Sueyoshi et al., 1999).

The role of CAR in the induction of the Cyp2b10 gene was demonstrated in vivo using knock out mice, where loss of CAR expression resulted in the abolition of Cyp2b10 induction by phenobarbital (PB) and 1,4-bis[2-(3,5-dichloropyridyloxy)] benzene (TCPOBOP) (Wei et al., 2000). Pharmacological experiments also demonstrated that
wild type animals were more resistant to zoxazolamine-induced paralysis (a Cyp2b10 substrate) than the CAR-deficient animals, following pre-induction of the metabolising enzymes with PB and TCPOBOP (Wei et al., 2000). The latter findings show that in the absence of CAR the xenobiotic metabolising enzymes could not be activated and therefore failed to protect the organism from the effect of the muscle relaxant zoxazolamine.

1.5.3.4. Retinoic X Receptor α (RXRα, NR2B1).

RXRα is an heterodimerisation partner for several nuclear receptors such as the vitamin D receptor and the thyroid hormone receptor (Kliewer et al., 1992), as well as CAR and SXR as shown above, but can also bind DNA response elements as a homodimer, although poorly. Experiments in genetically modified mice where the expression of this receptor was suppressed (Wan et al., 2000) resulted in the disruption of many hepatic gene functions (e.g. CYP3A1 and CYP2B10 gene expression, cholesterol homeostasis). Cellular amounts of RXRα may strongly influence the binding with their heterodimerisation partners to the regulatory sequences of the target genes. The role of RXRα in the xenobiotic response of the organism is crucial since it is an heterodimerisation partner for PPARα (CYP4A regulation), CAR and SXR (CYP2B and CYP3A regulation) (Waxman, 1999).

1.5.3.5. Hepatocyte nuclear factor 4 (HNF4, NR2A sub-family).

Several experiments report a role for this tissue specific factor in the regulation of the rat CYP3A23 gene (Huss and Kasper, 1998; Huss and Kasper, 2000). A functional binding site was identified using DNase I footprinting and transfection assays. An HNF4-like protein was identified which bound to the CYP2C2 gene promoter and HNF4 trans-activated constructs containing the binding site (Chen et al., 1994). A very recent report identified a functional binding site for HNF4α in the distal enhancer element (XREM) of the CYP3A4 gene, both in cell line and animal models (Tirona et al., 2003). This receptor is thought to enhance the CAR- and SXR- mediated regulation of the gene both at the
basal and drug induced level (Tirona et al., 2003). Therefore members of the HNF family seem to play a role in the liver specific expression of CYP genes. Additionally it was shown that hepatocyte nuclear factors plays a direct role in promoting SXR expression, since the nuclear receptor was not detected in HNF4α knock out mice, together with various other factors specific to mature liver metabolism (Li et al., 2000).

1.5.3.6. Nuclear receptor cross talk.

Using trans-activation assays and electromotility shift assay (EMSA), Xie and co-workers demonstrated that the human CAR is also involved in the regulation of the CYP3A4 gene, where it is constitutively active and further activated by its ligands (Xie et al., 2000a). This latter study indicates that SXR is also involved in the up-regulation of CYP2B genes via an element known as phenobarbital response element (PBRE), which contains a DR4 motif for nuclear receptor binding (direct repeat with 4 spacers). A similar observation was reported in a different experimental system (Smirlis et al., 2001). Using primary hepatocytes and in vivo liver transfection (using particle mediated delivery) it was shown that SXR can mediate the induction of the mouse Cyp2b and Cyp3a genes by synthetic steroids and that CAR could mediate the phenobarbital induction of both Cyp genes (Smirlis et al., 2001). This phenomenon was also observed in transgenic mice expressing a β-galactosidase reporter gene construct, where the induction of a CYP3A4 promoter construct by CAR was further confirmed in vivo (Goodwin et al., 2002). The fact that CAR and SXR bind the same response element demonstrates the existence of a cross talk between nuclear receptors involved in the xenobiotic response and suggests the presence of cellular regulatory networks. It is thought that these receptors act as part of a ‘metabolic safety net’, which has evolved to defend the organism from chemical insult (Xie et al., 2000a).

A recent study demonstrates that this cross talk results in the induction of several drug metabolism genes by both receptors (Maglich et al., 2002). Using ligands specific for either CAR or SXR, together with knock-out animals for each receptor, an overlapping set of genes coding for phase I and II enzymes as well as for drug transporters, was found to be co-regulated by the receptors. The results also indicate that some genes are only
induced by only one of the receptors, suggesting that CAR and SXR also have specific target activation patterns (Maglich et al., 2002).

1.6. Polymorphic expression of CYP3A4 and contribution to inter-individual variability in drug metabolism.

It has been estimated that two human genomes chosen at random within the population will differ at approximately $2.5 \times 10^5$ sites (Alberts et al., 2002). This means that a Single Nucleotide Polymorphism (SNP) may be found once every 1300 nucleotides in the genome (Alberts et al., 2002), with potential effects on the phenotype of the individual. There is extensive evidence that all cytochrome P450 genes involved in drug metabolism are expressed polymorphically (reviewed in Cholerton et al., 1992; Daly et al., 1998; Ingelman-Sundberg, 2002). Genetic polymorphisms may affect the coding region (information) or the regulatory region (expression of the information) of a gene. In the first instance, the protein structure and function may be affected and in the latter case the amount of gene expression at the basal and drug induced state may vary between individuals (e.g. Lamba et al., 2002). As discussed above for the CYP2D6 gene (section 1.1.3), population variations in the expression and sequence of CYP genes will result in the variability of the metabolic rates for the enzyme substrates (e.g. Gibson and Skett, 1994). This can lead to negative effects such as adverse drug reaction (ADR) or lack of therapeutic drug effect (Ingelman-Sundberg, 2002) if the plasma level in P450 enzyme substrate is too high or too low.

Various studies report a high inter-individual variability in CYP3A4 phenotype (enzyme level and activity) (e.g. reviewed in Lamba et al., 2002). Because CYP3A4 is transcriptionally regulated, variations in the sequence of the regulatory region of the gene could possibly be an explanation for this inter-individual variability. The fact that, unlike other CYP enzyme activities such as CYP2D6-mediated drug clearance (see section 1.1.3), CYP3A-mediated drug clearance seems to be unimodal rather than bimodal in the population suggests that several genetic polymorphisms or several environmental factors may influence the inter-individual variability (Lamba et al., 2002). A meta-analysis study
suggests that the genetic component of CYP3A4 variability ranges from 66 to 98%, with a median value of 83%, which provides an indication of the high importance of genetics in the variability in CYP3A4 expression (Ozdemir et al., 2000).

Rebbeck and co-workers were the first group to report a mutation in the 5' region of CYP3A4 (CYP3A4V, now called CYP3A4*1B), which consists of a substitution (A to G) -290 bp upstream of the transcription start (figure 1.9) (Rebbeck et al., 1998), situated in what was identified as the nifedipine response element, NFSE (Hashimoto et al., 1993).

Figure 1.9. The CYP3A4V mutation (CYP3A4*1B) located in the NFSE motif.
CYP3A4 wild type: 5'-AGGGCAAGGAG-3'
CYP3A4V: 5'-AGGGCAGGAG-3'

The presence of the CYP3A4V genetic variant was found to be associated with prostate cancer tumour grade in North American Caucasian patients (Rebbeck et al., 1998). The homozygous polymorphism was present in 9.6% of the population studied. The A to G substitution was not detected in any of the 128 DNA samples from Japanese volunteers (Ando et al., 1999). Another group reported the association between this polymorphism and the therapeutic outcome of patients undergoing chemotherapy with the anti-cancer drug epipodophyllotoxin (Felix et al., 1998). Individuals carrying CYP3A4V were less likely to suffer from de novo leukaemia caused by the genotoxic metabolites of the latter chemotherapeutic agent (Felix et al., 1998). In both prostate cancer and treatment induced leukaemia studies, the authors explained their finding by an altered expression of CYP3A4 which in turn would affect the metabolism of endogenous compounds (Rebbeck et al., 1998) or xenobiotics (Felix et al., 1998).

Although the presence of the polymorphism has since been confirmed in multiple studies, no group has yet confirmed any association between the CYP3A4V genotype and CYP3A4 enzyme activity or levels.

To assess CYP3A4V functionality, the genotype of 22 males and 24 females was established and CYP3A4 activity was measured as testosterone 6β hydroxylation (Westlind et al., 1999). This study revealed a high inter-individual variability and no correlation between genotype and activity was found. Gel shift assays revealed the low
significance of the polymorphism in terms of difference in cytosolic protein binding to an oligomer probe representing the wild type and the mutant sequence of the NFSE. Several other studies reported no significant correlation between this polymorphism and the metabolism of the CYP3A substrates alprazolam and trazodone (von Moltke et al., 2000), nifedipine (Ball et al., 1999) and cyclosporin (Rivory et al., 2000).

Two further mutations in the coding region of the CYP3A4 gene, in exon 7 and 12, were reported at the time of this study (Sata et al., 2000). The mutation in exon 7, termed CYP3A4*2A, was present in 2.7% of the Caucasian volunteers and was found to have a functional effect as it caused a decrease in nifedipine oxidation kinetics.

The three mutations described above were the only genetic variants in the coding and regulatory regions identified for CYP3A4 at the time this work was carried out. A more recent review reports a total of five mutations in the regulatory region of CYP3A4 (Lamba et al., 2002). A database of CYP gene mutations is frequently updated and is available on the Internet (http://www.imm.ki.se/CYPalleles/).
1.7. Project aims.

Previous work in our laboratory had focused on elucidating the mechanisms of induction of the human *CYP3A4* gene by hormonal and xenobiotic stimuli using reporter gene assays. The importance of a 1kb region within the proximal promoter was established for the drug response (Ogg *et al.*, 1999). Using co-transfection experiments with expression plasmids coding for GR and SXR, the role of both receptors in *CYP3A4* expression was experimentally confirmed (El Sankary, 2000; El Sankary *et al.*, 2001; El Sankary *et al.*, 2000). Additionally, a T>C transition -190bp from the transcription start was identified and this mutation significantly reduced the trans-activation of the promoter in response to dexamethasone and hydrocortisone (GR ligands) but did not affect the response to rifampicin (SXR ligand) (El Sankary, 2000).

This work raised several questions which were experimentally examined here.

- Is the ER6 the only *cis-*acting region of the *CYP3A4* promoter and by extension are SXR/CAR the only regulators of *CYP3A4* expression?

To identify further components of the regulatory network involved in the basal expression and drug-induced expression of *CYP3A4*, the transcription factor binding to the regulatory regions was examined by DNase I footprinting and computer analysis.

- Can the inter-individual variability in *CYP3A4* expression be explained by the presence of genetic variations within the proximal promoter, which could provide a molecular basis for this phenomenon?

By screening eleven phenotyped liver samples for which CYP3A activity has previously been assessed using an enzyme assay, the presence of further mutations was investigated within the 5' flanking region of the *CYP3A4* gene (about 1kb), using PCR and sequencing. The binding of transcription factors to newly identified genetic variants (Hamzeiy *et al.*, 2002) was assessed in order to investigate the possible mechanisms responsible for the inter-individual variability.
Are the putative transcription factor binding sites identified above (by DNase I and bioinformatics) functional and relevant to the CYP3A4 basal and drug-induced expression? The putative binding sites were mutated and the functional transcriptional effect assessed by reporter gene assay. The effect of the functional –190bp mutation on nuclear protein binding was also investigated.
CHAPTER 2 – MATERIALS AND METHODS

2.1. Materials.

2.1.1. General chemicals and plasticware.
Cell culture chemicals were purchased from Invitrogen. All other chemicals and reagents used in this work were of the highest purity available and were purchased from Sigma (St Louis MO) unless otherwise stated. The suppliers of all additional materials are listed below in table 2.1. Some specific molecular biology kits and instruments are described in the relevant areas of this chapter.
All work involving radiochemicals was carried out in accordance with the guidelines set out in the ‘Local Rules for Protection Against Ionising Radiations’ published by the University of Surrey Safety Office. Whenever possible all glassware and plasticware were autoclaved and solutions were autoclaved or filter sterilised. All nucleic acid work was carried out using DEPC treated water (0.1%).

2.1.2. Plasmids.
Maps of plasmids used in this work showing key restriction sites and functional features are presented in Figures 2.1 to 2.6. The plasmids contain the ampicillin resistance gene (β-lactamase) in order to enable selection for colonies containing the constructs during culture. The human SXR expression plasmid (hPXR-pSG5) was a gift from Dr Steven Kliewer (GlaxoSmithKline, Research Triangle Park, NC, USA). The reporter gene plasmids (pSEAPpro2 and pSEAP2basic vectors) were purchased from Clontech. The pSEAP pro2 plasmid (with an insert from the −301/+7 region of the CYP3A4 promoter) was engineered by Dr Wafaa El Sankary, University of Surrey. The pSEAP2-Basic plasmid DNA with a 300 bp insert from the CYP3A4 XREM region (−7972/−7673) as well as the same plasmid with an added 1141 bp (−1201/−61) of the CYP3A4 promoter
were kind gifts from Dr Hossein Hamzeiy, University of Surrey. The pSEAP2-Basic plasmid DNA with a 300 bp insert from the CYP3A4 XREM region (−7972/−7673) and the −301/+7 region of CYP3A4 promoter is hereafter termed pWT.
**Table 2.1. Suppliers of items used in this study.**

<table>
<thead>
<tr>
<th>Item or chemical</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial agar, phosphate buffer saline (PBS) tablets, Tryptone, Yeast extract</td>
<td>Oxoid</td>
</tr>
<tr>
<td>Agarose</td>
<td>Boehringer Mannheim</td>
</tr>
<tr>
<td>$\gamma^{32}$P ATP triethylammonium salt (370 MBq/ml, approx. 110 TBq/mmol)</td>
<td>Amersham</td>
</tr>
<tr>
<td>DNA oligomers for PCR and EMSA</td>
<td>MWG Biotech</td>
</tr>
<tr>
<td>DH5α or Top10F$^+$ E. coli competent bacterial cells</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Sequencing kits</td>
<td>Promega and USB</td>
</tr>
<tr>
<td>Restriction and other modifying enzymes and DNA size markers</td>
<td>Promega</td>
</tr>
<tr>
<td>Transformer site directed mutagenesis kit, mutS E. coli cells and</td>
<td></td>
</tr>
<tr>
<td>High Fidelity polymerase PCR kit, plasmids (pSEAP pro2 and pSEAP2-basic)</td>
<td>Clontech</td>
</tr>
<tr>
<td>PSG5 plasmid</td>
<td>Stratagene</td>
</tr>
<tr>
<td>0.22 µm filters</td>
<td>Nalgene</td>
</tr>
<tr>
<td>96 well Optiplates</td>
<td>Canberra Packard</td>
</tr>
<tr>
<td>Microbiology plasticware and cryovials</td>
<td>Greiner</td>
</tr>
<tr>
<td>Kits for DNA gel extraction, plasmid extraction (mini- and maxiprep)</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Disposable pipette tips</td>
<td>Alpha Laboratories</td>
</tr>
<tr>
<td>Microcentrifuge tubes</td>
<td>Scotlab</td>
</tr>
</tbody>
</table>
Figure 2.1. Map of the human SXR expression plasmid (hPXR-pSG5).

Figure 2.2. Map of the pSEAP2-promoter plasmid, an alkaline phosphatase reporter gene plasmid.

SV40: SV40 promoter, SEAP: secretory endothelial alkaline phosphatase, pUC ori: origin of replication, Amp r: lactamase conferring resistance to ampicillin. Important restriction enzyme cutting sites are also shown with their position (bp).
Figure 2.3. Map of the pSEAP2-promoter plasmid with the -301/+7 insert from the CYP3A4 proximal promoter.

SV40: SV40 promoter, SEAP: secretory endothelial alkaline phosphatase, pUC ori: origin of replication, Amp r: lactamase conferring resistance to ampicillin. Important restriction enzyme cutting sites are also shown with their position (bp).
Figure 2.4. Map of the pSEAP2-basic reporter gene plasmid with the distal *CYP3A4* gene enhancer XREM.

Important restriction enzyme cutting sites are shown with their position (bp).
Figure 2.5. Map of the pSEAP2-basic plasmid with the XREM insert and the -301/+7 insert from the CYP3A4 promoter (pWT).

Important restriction enzyme cutting sites are shown with their position. (bp)
Figure 2.6. Map of the pSEAP2-basic reporter gene plasmid with the distal enhancer XREM insert and the -1201/-61 insert from the CYP3A4 gene promoter.

Important restriction enzyme cutting sites are shown with their position (bp).
2.2. General microbiology techniques.

2.2.1. Transformation of bacteria.
20 to 50ng of plasmid or DNA from a ligation reaction were transformed into competent DH5α or Top10F' E. coli bacterial cells. A tube containing the frozen cells, stored at -80°C, was placed on ice for 5 minutes to thaw the suspension, DNA was added and the suspension was stirred gently using the pipette tip. The mixture was incubated on ice for 30 minutes and heat shocked at 37°C (DH5α) or 42°C (Top10F') for 30 s. LB broth was added and the tube was shaken at 225 rpm, 37°C for 1 hour. The bacteria were then spread on a LB agar plate containing ampicillin and incubated at 37°C or the volume of the suspension was made up to 5 ml with LB broth-ampicillin and shaking was resumed (in both cases the antibiotic concentration was 100 µg/ml). The cells were grown for 12-18 hours.

2.2.2. Preparation of glycerol stocks.
150µl of sterile glycerol were added to 850µl overnight cultures in LB media containing 100µg/ml ampicillin. The tube was briefly vortexed to ensure complete mixing and the stocks were frozen at -80°C.

2.2.3. Growth of bacterial cultures.
Transformed bacteria (from a glycerol stock or a transformation reaction) were spread on a LB agar plate containing ampicillin and grown at 37°C for 12-18 hours. A single colony was picked using a toothpick and inoculated into 5ml LB with 100 µg/ml ampicillin. This mini-culture was placed in a rotary incubator (37°C, 225rpm) for 12-18 hours. For a maxi-culture, 250ml LB containing 100 µg/ml ampicillin were inoculated with 250µl of the overnight culture then incubated (37°C, 225rpm) for 12-18 hours.
2.3. General molecular biology techniques.

2.3.1. Agarose gel electrophoresis.
DNA samples (~500 ng) were mixed to 5x loading buffer (50 % glycerol, 49 % nuclease free water and 1 % w/v Orange G dye) and run in 1 % w/v agarose gel made in TAE buffer at about 5 V/cm. Gels were pre-stained with ethidium bromide (0.5 μg/ml). DNA was then visualised under UV light and photographed (Polaroid films or Image Master imaging system from Pharmacia Biotech).

2.3.2. DNA extraction from agarose gel.
For salt sensitive applications such as sequencing or DNase I footprinting, DNA fragments (following amplification by PCR) were gel purified, to remove dNTPs and/or non-specific amplicons. The entire PCR reaction was run on a 1 % low melting point agarose, pre-stained with ethidium bromide, the band of interest was excised using a sterile scalpel and the DNA fragment was extracted using a gel extraction kit (Qiagen) following the manufacturer’s instructions. Briefly, 3 volumes of buffer GQ (solubilisation buffer) were added to one volume of the excised band and incubated until dissolution of the gel fragment was complete (typically 10 minutes at 65°C). One volume of isopropanol was added to increase the yield except for fragments of size raging from 500 to 4000 bp where this step has no effect. The mixture was applied to a QIAquick spin column and centrifuged for 1 minute. In order to remove traces of agarose, 500 μl of buffer QG were added and the column and centrifuged. The bound DNA was washed using 750 μl of buffer PE. In order to achieve a higher DNA purity (e.g. for sequencing) the buffer was incubated for 5 minutes at room temperature. Finally 30 to 50 μl water were added on the column, incubated at room temperature for 5 minutes and centrifuged.
2.3.3. DNA quantification.

The concentration of DNA preparations (e.g. purified PCR products, plasmids) was assessed spectrophotometrically with a Genequant II (Pharmacia Biotech). An aliquot was diluted in water and the absorbance at 260nm was measured against appropriate blank (water or TE buffer).

2.3.4. Restriction digest.

A typical restriction digest reaction consisted of 0.2 to 2 µg DNA (depending upon downstream application), 10 U of enzyme and the appropriate restriction enzyme buffer. For sensitive applications (probe labelling, site directed mutagenesis) BSA was added at a final concentration of 0.01 mg/ml. The reaction was incubated at 37°C for 2 hours.

2.3.5. Ethanol precipitation.

A typical ethanol precipitation consisted of the addition of 2 volumes of absolute ethanol to the sample which was then left for 5 to 30 minutes either at room temperature or in the cold (on ice, at -20°C or -80°C). The variety of conditions represents a compromise between the need to achieve high DNA recovery and the presence of salts which could co-precipitate at low temperatures. The DNA was pelleted by centrifugation, washed with 70 or 80% ethanol, air dried and resuspended in water. In some cases, DNA was also precipitated by adding 0.7 volumes of isopropanol, incubating the tube at room temperature and proceeding as for a classical ethanol precipitation.

2.4. Cloning.

2.4.1. PCR cloning.

When the ends of a DNA insert do not contain the suitable restriction sites for the fragment to be ligated at the desired location in a plasmid, those sites can be generated
using PCR. In this technique the PCR primers contain the recognition sequence of the relevant enzymes in their 5' ends. The primers were +7B (EcoRI) and -301A (Hind III) and they are described in table 2.5. PCR fragments were generated under normal amplification conditions (see section 2.10.3 and table 2.6) and prepared for the ligation by digesting them using the relevant enzymes (EcoRI and HindIII) in the appropriate buffers followed by gel purification.

2.4.2. Ligation.
Recipient vector was linearised using restriction enzyme(s) compatible with the insert DNA and gel purified. A ligation reaction contained 3U T4 DNA ligase, ligase buffer (30 mM Tris-HCl pH 7.8, 10 mM MgCl2, 10 mM DTT and 1 mM ATP), plasmid vector (40 ng) and insert DNA. The following equation was used to calculate the amount of insert to add to the ligation reaction:

\[
\frac{\text{vector}(\text{ng}) \times \text{size}(\text{kb})\text{insert}}{\text{size}(\text{kb})\text{vector}} \times \frac{\text{insert}}{\text{vector}}(\text{mol}) = \text{ng of insert}
\]

In order to ensure a high proportion of plasmids with inserts, the molar ratio of insert/vector used was set at around 20-25. The ligation reaction was then left overnight at 14°C. An alternative ligation protocol was also used, where the temperature cycled overnight between the optimal DNA ligase temperature (30°C, for 30 s) and a lower temperature (10 °C, for 30 s) more suitable to promote ligation of DNA ends (modified from Lund et al., 1996). The ligation reactions were then used to transform E. coli as described in section 2.2.1 and the bacteria plated on LB agar plates.

2.4.3. Selection.
In order to check individual colonies for the presence of plasmids with an insert, a quick and simple screening technique was used. Several colonies were picked with a toothpick and used to inoculate small volumes (200μl) of LB medium with ampicillin in a 96-well
plate. The bacteria were grown for 5 hours in a rotary incubator (37°C, 225rpm). Next 40µl of culture was mixed with an equal volume of phenol-chloroform and centrifuged (10000 xg for 2 minutes). The supernatant, consisting of the aqueous phase with all bacterial nucleic acids, was mixed with loading dye and run on a 1% agarose gel. The plasmid size was compared to that of a control transformation reaction containing the empty vector and any culture with a larger plasmid (i.e. migrating more slowly) was further studied by miniculture, restriction digest and sequencing.

2.4.4. Sequencing.
Fidelity of all constructs generated in this study was confirmed by sequencing using an ABI 373 automated sequencer (SBLS core facility, University of Surrey). Alternatively they were sequenced manually using a Thermo Sequenase™ cycle sequencing kit (USB) or finol (Promega), following the manufacturer’s instructions with minor modifications (see section 2.9.4 for protocol description).

2.5. Plasmid preparation.
Plasmid DNA was extracted from bacterial cultures (prepared as described above in section 2.2.3.) using Qiagen miniprep or Endofree maxiprep kits following the manufacturer’s instructions.

2.5.1. Minipreps.
Samples (1.5 to 5 ml) of overnight mini-cultures were spun in a microcentrifuge tube to collect the bacteria. The supernatant was removed and the cells resuspended in 250 µl buffer P1. Following addition of 250 µl buffer P2 the tube was mixed by inversion 5 times and the solution became clear and viscous. 350 µl buffer P3 were added and the tube was spun for 10 minutes at 10000 xg. The supernatant containing the plasmid DNA was applied to a QIAprep column and centrifuged for 1 minute. To increase yields the supernatant was sometimes re-applied to the column and spun a second time. The flow-
through was discarded and 750 µl of buffer PE were applied to was the column. In order to achieve higher plasmid purity (e.g. for sequencing) this buffer was incubated on the column for at least 5 minutes prior to column centrifugation. The flow-through was discarded and the tube was further centrifuged for 1 minute to remove residual traces of wash buffer. The column was then placed in a sterile microcentrifuge tube, water (30 to 100 µl) was applied and column was spun for 1 minute at 10000 xg. In order to increase the yields the water was sometimes incubated at room temperature for 5 minutes before centrifugation.

2.5.2. Endofree maxiprep.

A larger (250 ml) volume of overnight culture was centrifuged (6000 xg for 15 minutes at 4°C) to collect the bacterial cells, which were then resuspended in 10 ml buffer P1. Buffer P2 (lysis buffer) was added (10 ml) and the tubes were mixed by inversion 5 times. Following a 5 minute incubation at room temperature, 10 ml of chilled neutralisation buffer P3 were added and the tube was mixed by inversion. The lysate was poured into a QIAfilter cartridge, incubated for 10 minutes at room temperature and filtered into a collection tube where 2.5 ml of buffer ER were added. The tube was incubated for at least 30 minutes on ice in order to remove bacterial endotoxins which may impair the DNA transfection process. The lysate was transferred to a QIAGEN-tip equilibrated by the addition of 10 ml buffer QBT and the column was allowed to empty by gravity flow. The plasmid DNA bound to the column was washed twice with 2 x 30 ml buffer QC and eluted with 15 ml buffer QN. 10.5 ml isopropanol were added and the precipitated DNA was collected by centrifugation (6000 xg for 1 hour). The pellet was washed with 70% endotoxin-free ethanol, air dried and resuspended in 500 to 750 µl water.
2.6. General cell culture techniques.

2.6.1. Culture conditions.
Human hepatoma cells HepG2 (ECCAC No. 85011430) and HuH7 (a gift from Dr Steve Hood, GlaxoSmithKline UK) were grown in 75 cm² flasks with vented lids (Nunclon surface from Nalge Nunc International) in 5 % CO₂ in air at 37°C. HepG2 cells were cultured in 15ml Minimum Essential Medium (MEM), supplemented with 10 % (v/v) foetal calf serum, 1 % (v/v) non essential amino acids, 2 mM L-glutamine and 1 % (v/v) gentamycin. HuH7 cells were cultured in 15ml Dulbecco's modified Eagle medium (DMEM) supplemented with 10 % (v/v) foetal calf serum, 1 % (v/v) non-essential amino acids, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. The cells were subcultured every 3-5 days (~90% confluence) depending on the growth rate. The passage number was initially 78 for the HuH7 cells and they were subjected to a maximum of 10 further passages.

2.6.2. Subculture of cells.
When the cells reached more than 80% confluence, the medium was removed by aspiration, the cell layer washed with PBS and 1 ml trypsin-EDTA was added. The cell culture flask was gently rocked to ensure coverage of the cells by the protease and was either incubated at 37°C (HepG2 cells) or at room temperature (HuH7 cells) for 2-3 minutes. The cells were detached by tapping the sides of the flask and then 9 ml of medium added to dilute out enzyme activity. Usually 0.5 to 1 ml of this suspension was added to fresh media in a final volume of 15 ml in a new flask.

2.6.3. Storage and recovery of cells in liquid nitrogen.
HepG2 or HuH7 cells from near confluent flasks were trypsinised and spun at 1300 xg for 5 minutes. The resulting pellet was washed with PBS, the cells were resuspended in 91 % FBS and 9 % DMSO (concentration ≈ 4 ×10⁶ cells/ml) and aliquoted into cryovials.
The suspensions were frozen at -80°C for 1 hour and stored in liquid nitrogen. Upon recovery the tubes were thawed at 37°C for 5 minutes, the cells pelleted by centrifugation (5 minutes at 1300 xg) the supernatant removed, medium (MEM or DMEM) added and the cells were transferred to cell culture flasks.

2.7. Nuclear protein extract preparation.
Nuclear proteins were extracted from HepG2 cells grown as described in section 2.6.1.

2.7.1. HepG2 nuclear protein extraction.
The method used was derived from Dignam and co-workers (Dignam et al., 1983). Briefly, HepG2 cell suspensions were obtained by trypsinisation of approximately 90% confluent cells at passage 5 to 9. Cells were pelleted by centrifugation (1300 xg for 5 minutes) and washed in PBS. After another similar centrifugation, the cells were resuspended in 5 packed cell volumes (PCV, the volume of the cell pellet) of PBS and spun (1300 xg for 5 minutes). After removal of the supernatant, a suspension was made in 2 PCV of cold buffer A (10 mM Hepes-KOH, pH 7.9, 1.5 mM MgCl₂, 10 mM KCl and 0.5 mM DTT) and the cells were left to swell on ice. After 10 minutes, the cells were disrupted using 10 strokes of a Dounce homogeniser. The homogenate was then centrifugated at 2000 xg for 15 minutes. The difference between the volume of the supernatant and the initial volume was called the packed nuclear volume (PNV). The pellet was resuspended in 0.5 volumes of ice cold buffer C (25 % glycerol, 20 mM Hepes-KOH, pH 7.9, 1.5 mM MgCl₂, 0.2 mM EDTA, 20 mM NaCl, 0.5 mM DTT, 0.5 mM PMSF) and 0.5 PNV of buffer C high salt (buffer C containing 1.2 M NaCl) was added dropwise and with swirling. The suspension was homogenised using 10 strokes of Dounce homogeniser and then spun at 16000 xg for 30 minutes. The supernatant (nuclear proteins) was aliquoted and stored at -80°C.
2.7.2. Protein electrophoresis and Coomassie staining.

In order to assess the integrity of the extracted nuclear proteins, 10 µl of each sample were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie blue. The samples were mixed to an equal volume of loading dye and denatured at 95 ºC for 5 minutes. They were then loaded on a 6% SDS-polyacrylamide gel and run in an SDS containing buffer for 40 minutes at 200 V. After electrophoresis the gel was stained for 2-4 hours in Coomassie blue and destained overnight (10 % methanol and 10 % acetic acid).

2.7.3. Protein quantification.

Protein concentration was measured using a modification of the Lowry assay (Stoscheck, 1990). Duplicate aliquots (10 µl) of the extracted nuclear proteins were diluted in 390 µl water and mixed with 400 µl of freshly prepared 2x Lowry reagent (1 ml 2 % CuSO_4, 1 ml 1 % Na,K tartrate, 13 ml 10 % Na_2CO_3, 5 ml 1 % SDS and 5 ml 1 M NaOH). The solution was mixed by vortexing and incubated at room temperature for 10 minutes. After incubation, 200 µl of 0.2 N Folin-Ciocalteu’s phenol reagent was added slowly, while the sample was vigorously vortexed. The solution was left at room temperature for at least 30 minutes and the absorbance at 750 nm was determined using a Kontron spectrophotometer (Uvikon). Using the same method, a standard curve was generated (0, 25, 50, 75, 100 µg BSA in duplicate) and was used to calculate the unknown protein concentration by interpolation. The variability between duplicates was generally lower than 5%. All extracted nuclear protein concentrations were found to be within the same order of magnitude (2.5 to 4.0 µg/µl).

2.8. Electro mobility shift assay (EMSA).

Two protocols were used to investigate DNA:protein interactions. The initial EMSA study was designed to carry out a general investigation of protein binding to a large region of promoter -301/+7 (described in 2.8.1). The second EMSA study aim was to
assess the effect of mutations in short DNA oligomers on the binding affinity of HepG2 nuclear proteins (described in 2.8.2).

2.8.1. Electro mobility shift assay with the –301/+7 CYP3A4 proximal promoter fragment.

One hundred pmol of -301A primer (see table 2.5) was labelled at the dephosphorylated 5’ end with 2 μl 32P ATP (7.4x10^5 Bq) to generate -301*, using 10 U of T4 polynucleotide kinase (Promega). The reaction was incubated at 37°C for 60 minutes and to denature the enzyme, 80 μl TE buffer were added and the temperature was raised to 68°C for 10 minutes. The labelled primers were then used in a PCR reaction (described in section 2.10.3) with non-labelled +7B primers to generate a single strand labelled -301*/+7 probe. The product of the PCR reaction was then purified by ethanol precipitation or through a Qiagen spin column.
Attempts were made to generate labelled +7B primers. However, the incorporation efficiency was significantly lower than for the –301B primer. Hence this procedure was discontinued.

2.8.1.2. Binding reaction.
The binding reaction was carried out at room temperature (22 °C) and consisted of 10 μl of 2x binding buffer (40 mM Tris-HCl, pH 7.9, 100 mM NaCl, 20 % glycerol and 0.2 mM DTT), 1 μg of poly-dI:dC, 1-10 μl of protein extracts (representing 4 to 40 μg nuclear proteins) and the volume made up to 20 μl with nuclease-free water. After a 10 minute incubation, 2 μl of –301*/+7 probe were added and the reaction was further incubated for 30 minutes.
2.8.1.3. Gel electrophoresis.

2 μl loading buffer (0.5x TBE, 10 % glycerol, 0.25 % bromophenol blue and 0.25 % xylene cyanol FF) were added to the binding reaction and 5 μl of the mixture were loaded on a 4-6 % polyacrylamide gel, run for 2-4 hours at 200-600 V. To circumvent problems due to the presence of unwanted smears in the lanes, various electrophoresis conditions (gel concentration, running time and loading buffer with no dye) were initially examined. A competition experiment was carried out, using adding 1x, 10x and 100x excess unlabelled -301/+7 probe, in order to test if the binding was specific and reversible. The DNA was added in a concentrated form and the final reaction volume remained at 20 μl.

2.8.2. Gel shift assay to determine the effect of DNA mutations on nuclear protein binding.

2.8.2.1. Probe preparation.

Single stranded DNA oligomers (from MWG-Biotech) were annealed by heat denaturation of 100 pmol of both the sense and antisense strand in 1x annealing buffer (10mM Tris-HCl pH 8.0, 10 mM MgCl₂, 50 NaCl, 10μl final volume) for 2 minutes at 95°C followed by a decrease of temperature at the rate of 0.1°C per second down to 4°C. The concentration of the resulting double stranded oligomer was adjusted to 1pmol/μl with DEPC-water and 3.5pmol were 5' labelled in 1x kinase buffer (70 mM Tris-HCl pH 7.6, 10 mM MgCl₂ and 5 mM DTT) with 5 to 10U of T4 polynucleotide kinase and 1μl [γ³²P]-adenosine triphosphate (~10 TBq/mmol, Amersham Pharmacia Biotech, UK) for 30 minutes at 37°C. The resulting probe was diluted to 35 fmol/μl using TE buffer and stored at -20°C. It could be used for several weeks following labelling.

The oligomers used are described in table 2.2. Their length was kept to a minimum in order to avoid non-specific binding or binding of adjacent transcription factors.
Table 2.2. Oligomer sequences for the EMSA studies (the exact complementary nucleotide sequence was used for the opposite strand). The letters in underlined bold font represent the bases which differ between the wild type and the mutant oligomers.

<table>
<thead>
<tr>
<th>Oligomer name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP3A4*1B wt</td>
<td>5'-GAGACAAAGGGCAAGAGAGAGCCGAT-3'</td>
</tr>
<tr>
<td>CYP3A4*1B (-392 A→G transition)</td>
<td>5'-GAGACAAAGGGCAGAGAGAGCCGAT-3'</td>
</tr>
<tr>
<td>CYP3A4*1E wt</td>
<td>5'-TTAATAGATTATTTTGCCAATG-3'</td>
</tr>
<tr>
<td>CYP3A4*1E (-369 T→A transversion)</td>
<td>5'-TTAATAGATTATATGCCAATG-3'</td>
</tr>
<tr>
<td>CYP3A4*1F wt</td>
<td>5'-GTGTACAGCACCTGGTAGGGAC-3'</td>
</tr>
<tr>
<td>CYP3A4*1F (-747 C→G transition)</td>
<td>5'-GTGTACAGCAGCTGGTAGGGAC-3'</td>
</tr>
<tr>
<td>Wild-type (pWT)</td>
<td>5'-TTGATTGAGTTGTTTATGATACCT-3'</td>
</tr>
<tr>
<td>Mutant (pMUT)</td>
<td>5'-TTGATTGAGTTGCTTATGATACCT-3'</td>
</tr>
</tbody>
</table>

2.8.2.2. Binding reaction.

5-20µg nuclear proteins from DMSO or rifampicin treated HepG2 cells and 1-2µg poly[d(C)] were incubated in the presence or absence of various concentrations of unlabelled competitor probe (1 to 30 fold excess of either the wild type or mutant sequence) in binding buffer A (4% glycerol, 1mM MgCl₂, 0.5mM EDTA, 0.5mM DTT, 50mM NaCl, 10mM Tris-HCl (pH 7.5)) or B (5% glycerol, 1mM MgCl₂, 0.1mM EGTA, 40mM KCl, 25mM Hepes-NaOH (pH 7.9), 4% Ficoll 400) in a final reaction volume of 10µl. After 15 minutes on ice 3.5fmol of ³²P end-labelled double stranded DNA probe (representing the wild type sequence) were added to each tube and the incubation was carried out for a further 25 minutes.
2.8.2.3. Electrophoresis.

The binding reactions were resolved on a 4% non-denaturing polyacrylamide gel in 0.5x TBE at 150V for 25 minutes. The gels were exposed to Kodak X-Omat film overnight and quantified using video-based computerized densitometry on an MCID image analysis system (Imaging Research, Ontario, Canada). In some instances, additional experiments were carried out to demonstrate the interaction observed was indeed a DNA protein interaction. The proteins were treated with 30 micrograms proteinase K or heat denatured (90°C for 1 minute) before addition of the probe.

2.9. In vitro footprinting assays.

2.9.1. Overview of the method.

An overview of the three procedures used is given is described in Figure 2.7. Three methods were used, while attempting to carry out the in vitro footprinting of the proximal part of CYP3A4 promoter (-301/+7) but only the final protocol (Figure 2.7, protocol 3) was used for the promoter analysis. Briefly this technique consists in incubating a single end labelled DNA probe with nuclear proteins, DNase I is then added and randomly produces single strand cuts except in the regions protected from the enzyme by proteins bound to the DNA. After purification, the DNA is run on a denaturing polyacrylamide gel and the band pattern compared to that of a control lane of DNA digested in the absence of nuclear extracts. The protected areas can be observed because of the absence or faintness of the band pattern (footprint) indicating the presence of DNA binding proteins at a specific binding site.

After several unfruitful attempts using protocols described by Garabedian and co-workers (Garabedian et al., 1993) to carry out copper-phenanthroline (Figure 2.7, protocol 1) and DNase I (Figure 2.7, protocol 2) footprinting with a PCR labelled probe, a Promega Core Footprinting system was purchased. The experiments were carried out following the manufacturer’s instruction. Later all solutions were made in the laboratory, following the protocol (Figure 2.7, protocol 3).
Figure 2.7. Outline of the various in vitro footprinting procedures used.

1: In-gel footprinting  
   Label PCR primer  
   PCR with one labelled and one unlabelled primer.

2: Initial DNase I protocol  
   Label -301/+7 probe  
   Restriction digest of one end  
   Single end labelled probe  
   Incubation with nuclear proteins  
   Non-denaturing electrophoresis (EMSA)  
   Cut out the retarded bands  
   Chemical DNA cutting  
   DNase I digestion  
   Extraction of the DNA fragments  
   Denaturing gel electrophoresis  
   Autoradiography  

3: Final protocol
2.9.2. Probe preparation.

Four DNA probes (301/+7, -388/-83, -240/+69 and -388/+69) were amplified by PCR (as described in section 2.10.3) and gel purified. The primer sequences contained restriction sites towards their 5’ ends and are described in table 2.3 (the added restriction sites are shaded). The PCR conditions were the same as described in table 2.6 for -301/+7, except the annealing temperature were 56°C for -388/-83, 62°C for -240/+69 and 61°C for -388/+69.

Table 2.3. PCR primers used to generate the DNA probes for DNase I footprinting.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>-301A (Acc65 I)</td>
<td>5'-GGGGTACCCAGACAAGGGCAAGAGAGAGG-3'</td>
</tr>
<tr>
<td>+7B (Bgl II)</td>
<td>5'-GAAGATCTTGCACAGCAGTGATTCAGTGAGG-3'</td>
</tr>
<tr>
<td>-388A (EcoRI)</td>
<td>5'-CCGGAATTCCACAGGCCACACTCCAGGCAT-3'</td>
</tr>
<tr>
<td>-83B (BamHI)</td>
<td>5'-CGCGGATCCCTCTGCCTGACAGTTGGGAAGAG-3'</td>
</tr>
<tr>
<td>-240A (EcoRI)</td>
<td>5'-CCGGAATTCTTGATAAGAACCCAGAACC-3'</td>
</tr>
<tr>
<td>+69B (BamHI)</td>
<td>5'-CGCGGATCCCTGCCTTTCAGCTCTGTGT-3'</td>
</tr>
</tbody>
</table>

The amplicons were labelled using 1 μl 32P (3.7x10^5 Bq) and 10 U T4 polynucleotide kinase for 30 minutes and an extra 10 U PNK were added for another 30 minutes. One end of the probe was then removed by digestion for one hour, using 40 U of the appropriate restriction enzyme (Bgl II or BamHI for the sense strand and Acc65I or EcoRI for the antisense strand). The single end labelled probe was extracted using an equal volume of phenol/chloroform, then another volume of chloroform, ethanol precipitated and resuspended in TE buffer.
2.9.3. *In vitro* footprinting.

An incubation reaction was set up with 25 μl 2x binding buffer (50 mM tris-HCl pH 8.0, 12.5 mM MgCl₂, 1 mM EDTA, 20 % v/v glycerol, 1 mM DTT and 25 or 100 mM KCl depending on whether the buffer had a low or high salt content), 5 μl DNA probe (approx. 7 kBq), nuclear proteins (5 to 80 μg) and water up to 50 μl. The mixture was left on ice for 10 minutes. Then 50 μl of Ca²⁺/Mg²⁺ buffer was added (5 mM CaCl₂ and 10 mM MgCl₂), incubated at room temperature for 1 minute and the DNA digested with 0.15 U DNase I for 1 minute. The reaction was stopped by the addition of 90 μl of warm (37°C) stop solution (200 mM NaCl, 30 mM EDTA, 1 % w/v SDS and 9 μg tRNA). DNA fragments were then purified by phenol-chloroform extraction and ethanol precipitated. The resulting pellet was washed in 70 % ethanol and resuspended in loading dye (95 % formamide, 20 mM EDTA, 0.05 % xylene cyanol and 0.05 % bromophenol blue) before denaturing gel electrophoresis (described in section 2.9.6). An alternative loading dye (1:2 0.1 M NaOH:formamide v/v, 0.1 % xylene cyanol and 0.1 % bromophenol blue) was also used and gave equivalent results.

Positive control reactions were carried out using a 342 bp DNA probe, representing the SV40 (Simian virus 40) promoter region (with a *Hind* III end), which was incubated with 2 μl of recombinant AP2 protein (Promega). This DNA probe contains a consensus binding sequence for AP2 (CCCCAGGC).

2.9.4. Sequencing reactions.

In order to be able to identify the base sequence of the protected region during the gel analysis, sequencing reactions were run alongside the DNA samples. These sequencing reactions were carried out using a Thermo Sequenase™ cycle sequencing kit (USB), following the manufacturer’s instructions with minor modification. Briefly, a primer (-301A or +7B depending on which strand the footprinting analysis was applied to) was labelled using 1 μl ³²P (3.7x10⁵ Bq) in the presence of 10 U T4 polynucleotide kinase for 30 minutes and four sequencing reactions were set up for each DNA template. Each tube contained a different chain terminating dideoxynucleoside triphosphate (ddA, ddC, ddT or ddG) in the presence of deoxynucleoside triphosphate (dNTP). Thermo Sequenase
DNA Polymerase in reaction buffer was added and the tubes were placed on a thermal cycler (denatured at 95°C for 2 minutes, then 95°C for 30 seconds (denaturation), 42°C for 30 seconds (annealing), 70°C for 1 minute (extension) for 30 to 40 cycles). Loading buffer (10mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol) was then added and the reactions kept at -20°C for up to 2 weeks.

2.9.5. Cerenkov counting.

Preliminary experiments revealed the importance of an equal radioactivity loading between each well, to ensure the most accurate detection of the protected areas, which are deduced by the comparison of band intensities between different lanes of the gel. The method used here allows an indirect measurement of high-energy β-emitters such as ³²P. The Cerenkov effect is the generation of photons when a charged particle (e.g. electron) passes through a dielectric medium (e.g. air) (Parker, 1974) and was measured by placing a tube containing the radioactive sample dissolved in loading dye, in a scintillation vial. The activity was then semi-quantitatively assessed using a Wallac 1410 scintillation counter. Preliminary experiments revealed that the response (luminescence) was linear and very well and significantly correlated ($r^2=0.97$, $p<0.0001$) to the amount of radioactivity measured, within the range of activities used for a DNase I footprinting experiment (data not shown).

2.9.6. Denaturing gel electrophoresis.

The samples and the sequencing reactions were loaded on at the same time, following heat denaturation (95°C for 2 minutes) and 2 minutes in ice. Electrophoresis (45W constant power, 1600-2000V) was carried out in 6% polyacrylamide gels containing 7 M urea in 0.5xTBE and pre-run for about one hour prior to loading.

2.9.7. Autoradiography.

After electrophoresis the apparatus was dismantled, the gel was transferred to a Whatmann filter paper and wrapped in cling film. It was then opposed to Kodak film for
1-2 days (with intensifying screen) or 3-7 days (without intensifying screen) at -80°C to stop diffusion.

2.9.8. Binding site analysis.
Computer analyses were performed to identify potential transcription factors binding sites corresponding to DNase I protected areas of the -301/+7 region of the CYP3A4 promoter. MatInspector 2.2 (available at transfac.gbf.de/cgi-bin/matSearch/) and AliBaba2 (www.uni-magdeburg.de/~grabe/alibaba2) were used to interrogate the Transfac 4 database and searches were made using various levels of similarity to the consensus sequences to gain the most useful information.

2.10. Mutation detection in 11 human liver samples.

2.10.1. Liver samples used in this study.
The eleven human liver samples used were a gift from Dr D. Moore (Hoffman La Roche, Nutley, NJ, USA) and were obtained from Keystone tissue bank (PA, USA). They had previously been phenotyped for CYP3A and among the donors (5 males and 6 females), there was a 14.1-fold difference in midazolam metabolism (Table 2.4).
Table 2.4. Human liver samples used in this study.

<table>
<thead>
<tr>
<th>Donor code</th>
<th>Age</th>
<th>Ethnicity</th>
<th>Gender</th>
<th>CYP3A4 activity *</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL 02</td>
<td>24</td>
<td>Caucasian</td>
<td>Male</td>
<td>350</td>
</tr>
<tr>
<td>HL 03</td>
<td>21</td>
<td>Caucasian</td>
<td>Male</td>
<td>286</td>
</tr>
<tr>
<td>HL 13</td>
<td>11</td>
<td>African-American</td>
<td>Female</td>
<td>626</td>
</tr>
<tr>
<td>HL 14</td>
<td>12</td>
<td>African-American</td>
<td>Male</td>
<td>442</td>
</tr>
<tr>
<td>HL 15</td>
<td>28</td>
<td>Caucasian</td>
<td>Female</td>
<td>759</td>
</tr>
<tr>
<td>HL 16</td>
<td>45</td>
<td>African-American</td>
<td>Female</td>
<td>1873</td>
</tr>
<tr>
<td>HL 17</td>
<td>27</td>
<td>Caucasian</td>
<td>Female</td>
<td>286</td>
</tr>
<tr>
<td>HL 18</td>
<td>48</td>
<td>Caucasian</td>
<td>Female</td>
<td>183</td>
</tr>
<tr>
<td>HL 19</td>
<td>36</td>
<td>Caucasian</td>
<td>Male</td>
<td>518</td>
</tr>
<tr>
<td>HL 20</td>
<td>40</td>
<td>Caucasian</td>
<td>Male</td>
<td>133</td>
</tr>
<tr>
<td>HL 21</td>
<td>61</td>
<td>Caucasian</td>
<td>Female</td>
<td>746</td>
</tr>
</tbody>
</table>

*: The CYP3A4 activity was measured using a midazolam hydroxylation assay (CYP3A specific) and the results are expressed in pmol metabolite/min/mg.

2.10.2. Genomic DNA extraction from human liver.

Some DNA samples had previously been extracted by Dr Wafaa El-Sankary (University of Surrey). Gel electrophoresis and PCR were used to assess the integrity of the genomic DNA. Missing DNA samples, or samples of unsuitable quality were re-extracted from the human liver samples, using a QIAamp DNA mini kit (Quiagen), following the manufacturer’s instructions.

2.10.3. Polymerase Chain Reaction (PCR).

The primers for PCR were obtained from MWG Biotech (Milton Keynes, UK) and are described in Table 2.5. For CYP3A4 promoter region analysis they were designed using Vector NTI 4.0 (Informax, USA) whereas for the exon 7 analysis, sequences were obtained from the literature (Sata et al., 2000).
Table 2.5. PCR primers used in the mutation detection study. Subsequent inspection of the sequences revealed that primer +7B actually goes up to +9 in its 5’ region and that -301A reaches −300 bp only. The highlighted sequences represent the added restriction sites and the +7B (EcoRI) and -301A (Hind III) were used for subsequent cloning of the PCR products (see section 2.4.1).

<table>
<thead>
<tr>
<th>Primer*</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>+7B</td>
<td>5’-CACAGCAGTGATCCAGTGAGG-3’</td>
</tr>
<tr>
<td>-301A</td>
<td>5’-AGACAAAGGGCAAGAGAGAGG-3’</td>
</tr>
<tr>
<td>+7B (EcoRI)</td>
<td>5’-CCGGAATTC CACAGCAGTGATCCAGTGAGG-3’</td>
</tr>
<tr>
<td>-301A (Hind III)</td>
<td>5’-CCCAAGCTTAGACAAGGGCAAGAGAGAGG-3’</td>
</tr>
<tr>
<td>-609A</td>
<td>5’-TAGACTATGCCCTTGAGGAGC-3’</td>
</tr>
<tr>
<td>-556B</td>
<td>5’-TGTTTCCACGCTGTGTTCCC-3’</td>
</tr>
<tr>
<td>-1086A</td>
<td>5’-ATCATGCTGGCTGAGGTGG-3’</td>
</tr>
<tr>
<td>Exon7A</td>
<td>5’-CCTGTTGCATGCATAGAGG-3’</td>
</tr>
<tr>
<td>Exon7B</td>
<td>5’-GATGATGGTCACACATATC-3’</td>
</tr>
</tbody>
</table>

* A: sense primer, B: antisense primer.

The Clontech Advantage® HF and HF-2 kits, which contain a high fidelity polymerase, were used. Each reaction mixture contained 40 to 200 ng genomic DNA, 100 pmol of each primer, 5 µl of 10x dNTP mix, 5 µl of 10x reaction buffer, 1 µl of HF polymerase and the volume was adjusted to 50 µl using nuclease free water. The PCR reactions were performed in a PTC-200 thermal cycler (MJ Research). The cycling conditions are described in Table 2.6.
Table 2.6. Cycling conditions for the PCR reactions.

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Denaturation</th>
<th>Annealing</th>
<th>Extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>-301/+7</td>
<td>94 °C for 10 s</td>
<td>60 °C for 30 s</td>
<td>72 °C for 30 s</td>
</tr>
<tr>
<td>-609/+7</td>
<td>94 °C for 10 s</td>
<td>60 °C for 30 s</td>
<td>72 °C for 40 s</td>
</tr>
<tr>
<td>-1086/-556</td>
<td>94 °C for 10 s</td>
<td>65 °C for 30 s</td>
<td>72 °C for 45 s</td>
</tr>
<tr>
<td>Exon7A/B</td>
<td>95 °C for 30 s</td>
<td>55 °C for 30 s</td>
<td>72 °C for 60 s</td>
</tr>
</tbody>
</table>

Prior to cycling the tubes were incubated at 94 or 95 °C for 1 minute to fully denature the DNA template. Following 30-35 cycles, a final extension step was carried out (72°C for 5 minutes).

2.10.4. Sequencing.

2.10.4.1. CYP3A4 promoter region.

Automated DNA sequencing was performed by Rachel Anderson (GlaxoSmithKline) on an ABI prism 377 (Applied Biosystems). 200 to 400 ng of template DNA (purified PCR product) were mixed with 3pmol of the appropriate forward or reverse primers (described in Table 2.5.). An additional internal primer was used for the reverse sequencing of the -609/-234 region in the -609/+7 fragment (-234B: 5'-CTTATCAGAAACTCAAGTG-3').

2.10.4.2. Exon 7.

Automated DNA sequencing was performed by Ms D. Robillard (Endocrinology research group, University of Surrey), using a Thermo Sequenase™ Cy™ 5 dye terminator kit (Amersham Pharmacia Biotech) and read with a Long Read Tower™ system V 3.1, by Visible Genetics.

Some of the coding region analysis was carried out by Ms Donna McGowan (MSc Toxicology, 1999-2000) as part of her research thesis.
2.10.5. Sequence analysis.

The sequences were then aligned against one of the published CYP3A4 sequence (Goodwin et al., 1999), available on the Internet at NCBI under accession number AF185589 (http://www.ncbi.nlm.nih.gov/entrez/). The alignment analysis was carried out using ClustalW available on the Internet server of the Pasteur Institute, Paris (http://bioweb.pasteur.fr/seqanal/interfaces/clustalw-simple.html).

2.11. Site Directed Mutagenesis.

Two methods were used in order to modify the −301/+7 fragment of CYP3A4 promoter cloned into pSEAP plasmids using the unique site elimination method (Deng and Nickoloff, 1992). Briefly, the plasmid is denatured to separate both DNA strands and two oligomers are then annealed: one mutagenic primer containing the desired SDM modification and one selection primer mutating a unique restriction site within the template plasmid. The oligomers are used as primers for the DNA strand synthesis reaction containing a DNA polymerase and ligase. The plasmids which sequence has been altered are then amplified using DNA repair deficient bacteria and selected using the fact that non modified plasmids can be linearised by restriction digest at the unique restriction site and have a very low transformation efficiency compared to uncut supercoiled mutated plasmids with a mutated unique restriction site. Initially a Clontech Transformer site directed mutagenesis kit was used and the procedure was mainly carried out using manufacturer’s instruction (section 2.11.1) using pSEAP pro2 with a −301/+7 insert as a template plasmid. Then, due to the lack of success in generating certain mutations with this method, a new protocol was developed (section 2.11.2) using pSEAP2-basic with XREM and −301/+7 insert as a template. The mutated promoter fragments generated in the pSEAP pro2 vector were then cloned into pSEAP2-basic with an XREM insert (procedure described in section 2.4). During cloning of the −301/+7 fragment it was noticed that an insert contained a T→C substitution located at position −193 disrupting a putative HNF3 binding site. This was most likely to be a PCR artifact. This fragment was called WT* and it was decided to investigate the effect of the SDM
generated modifications in this vector in comparison to mutants generated in the true wild type (WT) plasmid (i.e. possible interactions between the mutations). A diagram presented in chapter 5 (figure 5.6) describes all the constructs generated in the SDM experiments, they are also listed in the two bottom rows of table 2.9.

The mutagenic and selection primers are described in Table 2.7 and Table 2.8. Before use 1 μg or 10 μg of oligomer was phosphorylated using 10 U of T4 polynucleotide kinase, 1 mM ATP in kinase buffer (70 mM Tris-HCl pH 7.6, 10 mM MgCl₂ and 5 mM DTT) in a final volume of 20 μl at 37°C for 60 minutes. The enzyme was heat denatured at 65°C for 10 minutes.

Table 2.7. SDM mutagenic primers. The letters in underlined bold font represent the bases which are altered by the mutagenesis.

<table>
<thead>
<tr>
<th>Putative site altered</th>
<th>Sequence</th>
<th>Alteration (WT→mut)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER6</td>
<td>CATAGAATATGTTCTCAAAGGAGAACGTGAGTGG</td>
<td>AA→TT and GT→AA</td>
</tr>
<tr>
<td>Sp1</td>
<td>GGAGAAGCCTCTTTAACTGCGAGCGAG</td>
<td>CC→TT</td>
</tr>
<tr>
<td>HNF3</td>
<td>GATTGATTGAGTTGTGGATGATACCTCATA</td>
<td>TT→GG</td>
</tr>
<tr>
<td>C/EBP</td>
<td>GTGTGTGTGATTTCGAGGCCACTTCCAAGG</td>
<td>TTT→GAG</td>
</tr>
</tbody>
</table>

Table 2.8. SDM selection primers.

<table>
<thead>
<tr>
<th>Name (vector mutated)</th>
<th>Restriction site modified</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection1 (pSEAP2-pro and pSEAP2-basic)</td>
<td>Eco47 III</td>
<td>GTGCCGGCACTAGTCTTTCCGCTTC</td>
</tr>
<tr>
<td>Selection2 (pSEAP2-pro)</td>
<td>Xba I</td>
<td>CCCGGTTACTCAGTGTCGGGGC</td>
</tr>
<tr>
<td>Selection3 (pSEAP2-basic)</td>
<td>Sal I</td>
<td>AAGGATCCGTATACCGATGCC</td>
</tr>
</tbody>
</table>
2.11.1. Initial SDM protocol.

Gel purified WT* or WT plasmid (100 ng) was added to annealing buffer (20 mM Tris-HCl, pH 7.5, 10 mM MgCl₂ and 50 mM NaCl) and 100 ng of both selection and mutagenic primer. The volume was adjusted to 20 µl with water. The annealing reaction was then heated for 3 minutes at 100°C and immediately transferred to an ice water bath for 5 minutes. After the tube was spun down, T4 DNA ligase (2-4 units) and T4 DNA polymerase (2-4 units) were added to synthesis buffer (proprietary composition, contains a nucleotide mixture in Tris-HCl, pH 7.5) in a final volume of 30 µl. The synthesis reaction was incubated at 37°C for 2 hours. The DNA was then digested using 5-10 units of the appropriate restriction enzyme (see table 2.8) with 0.01 µg/ml BSA in a final volume of 50 µl at 37°C for 2 hours and DNA repair deficient mutS E. coli cells were transformed with 10 µl of the reaction (equivalent to 20 ng plasmid DNA from the initial reaction). After overnight culture in LB ampicillin the plasmid DNA was recovered by a miniprep procedure and 100 ng were digested again with the same restriction enzyme. A 25 ng fraction of the digest was used to transform Top10F' bacteria as described in section 2.2.3. The transformation reaction was plated on LB agar with ampicillin and the mutant selection was carried out by growing individual colonies in LB ampicillin, extracting the DNA (miniprep) and digesting the plasmid with the restriction enzyme which unique recognition site is modified by the selection primer. Positive (uncut) constructs were sequenced and it was found that the fraction of false positive (with a modified restriction site but with no mutation in the CYP3A4 promoter fragment) varied greatly between experiments. In order to increase the mutation efficiency of certain primer combinations, the amount of mutagenic primer was increased from 100 ng to 1 µg in order to achieve a 10-fold excess compared to the amount of selection primer (Hutchinson and Allen, 1997). This procedure only worked for the generation of the WT* mutants in the initial stages of the project.
2.11.2. Second SDM protocol.

A second method was developed in order to generate mutations in the WT fragment and this proved more efficient at generating mutations. It is a combination of the Clontech Transformer site directed mutagenesis kit and Promega Altered sites II system and is described below. The plasmid template was initially denatured before the annealing. This was done by incubating 2.5 µg DNA with 400 mM NaOH (final volume 20 µl) for 10 minutes at room temperature. The reaction was neutralised by the addition of 4 µl 3 M NaOAc (pH 4.8). The plasmid was ethanol precipitated, washed in 70 % ethanol and resuspended in 20 µl water. A sample was run on a 1 % agarose gel to assess the DNA concentration.

The annealing reaction consisted of Clontech annealing buffer (20 mM Tris-HCl, pH 7.5, 10 mM MgCl₂ and 50 mM NaCl), 75 ng selection primer, 1 µg mutagenic primer and 100 ng denatured DNA. The reaction was heated at 75°C for 5 minutes and cooled down to room temperature at the rate of 1°C/s using a thermal cycler. The synthesis reaction was then carried out using Promega ligase buffer (30 mM Tris-HCl pH 7.8, 10 mM MgCl₂, 10 mM DTT and 1 mM ATP), 10 U T4 DNA polymerase, 3 U T4 DNA ligase, 0.5 mM dNTP in a final volume of 30 µl for 2 hours at 37°C. The enzymes were denatured at 70°C for 5 minutes then digested using the appropriate restriction enzyme. A fraction of the digest representing 20 ng of DNA was used to transform mutS E. coli cells and the remainder of the protocol was carried out as described above (section 2.11.1). It was found that the second protocol generated a higher proportion of mutant plasmids than the first one used (about 10 times more efficient). It should be noted that in spite of the differences in procedure or selection primers, the resultant plasmids were identical apart from a few bp located in the unique restriction site used for mutant selection and these were specifically located on the opposite end of the plasmid to the CYP3A4 promoter fragment of interest. These mutations should therefore not interfere with DNA:protein interactions within the promoter region or affect the coding region for the reporter gene product.
2.11.3. Generation of competent *E.coli* BMH 71-18 *mutS* bacterial cells.

The cell stock was streaked on a LB agar plate containing 50 µg/ml tetracycline and were grown overnight at 37°C. A single colony was picked and grown overnight in LB broth (50 µg/ml tetracycline). 1ml of this miniculture were then used to inoculate 100 ml of LB broth in a sterile flask and the cells were grown at 37°C until the OD at 600 nm of the suspension reached 0.5±0.03 (usually 3 hours post inoculation). The flask was chilled on ice for 20 minutes and cells were then collected by centrifugation (1200 x g for 5 minutes at 4°C) and resuspended in ice cold 10 ml TSS (Transformation and Storage Solution: 85% LB medium, 10% w/v PEG molecular weight 8000, 5% DMSO and 50mM MgCl₂ (pH6.5)). The competent *mutS E. coli* were then used for a transformation reaction (described in section 2.2.1) or aliquoted in sterile tubes and stored at -80°C for later use up to 2 months after preparation.

2.12. Transient transfection protocol for the reporter gene assay.

The protocol was taken from Dr Hossein Hamzeiy (Hamzeiy, 2002) with minor modifications.

2.12.1. Plasmids.

Following preliminary transfection experiments all -301/+7 promoter fragments generated by SDM in pSEAP2-pro were cloned into pSEAP2-basic containing the XREM (distal enhancer) insert. Briefly, in this system it was found that inducibility of the reporter gene construct was only seen in the presence of the XREM when cloned next to the promoter of interest. During preliminary experiments a comparison of the induction profile was carried out between the -1201/-61 region of CYP3A4 promoter versus the -301/+7 region in plasmids containing the XREM element (-1201/-61 is described in
(Hamzeiy et al., 2002)). The plasmids used in this study are listed in table 2.9 and a general map of the vector is described at the beginning of this chapter (figure 2.5).

Table 2.9. Plasmids used in the transient transfection experiments.

<table>
<thead>
<tr>
<th>Reference plasmid</th>
<th>Test plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>-1201/-61</td>
</tr>
<tr>
<td>WT*</td>
<td>pER6*, pSp1*, pHNF*, pCEBP*</td>
</tr>
<tr>
<td>WT</td>
<td>pSp1, pHNF, pCEBP</td>
</tr>
</tbody>
</table>

e.g. pER6* represents the pSEAP2-basic vector with XREM insert, containing the -301/+7 insert of promoter (WT* version) modified with the ER6 mutation described in table 2.7. The corresponding reference plasmids contain the wild type -301/+7 insert (WT) or the -301/+7 insert with the -193 T>C substitution (WT*).


Near confluent (~90%) HuH7 cells were harvested by trypsinisation as previously described. The suspension was spun for 5 minutes at 1000rpm, the supernatant removed by aspiration and the pellet washed in 5ml PBS. The cells were resuspended in 5ml DMEM and counted using a haemocytometer. The suspension was then diluted using the appropriate amount of DMEM to a concentration of 10⁵ cells/ml. The cells were seeded in a 96 well plate (120µl/well, equivalent to 12000 cells/well), avoiding the use of the outside wells because of the presence of an edge effect caused by evaporation detected in preliminary experiments. The outside wells were filled with PBS (120µl/well). The plates were incubated in a humidified container for 48 hours at 37°C.
2.12.3. Transfection.

Previous transfection protocols in our laboratory relied on the calcium phosphate precipitation method to achieve transfer of the plasmid DNA into cells. (e.g. Ogg et al., 1999). In this work Fugene 6 Transfection Reagent (Roche) offered a more efficient alternative and was used following manufacturer's recommendations. The components of the incubation mixture were added to a sterile microcentrifuge tube and incubated for at least 30 minutes at room temperature. The order of addition was critical (based on the manufacturer's instructions) and was as follows: serum free DMEM (for a final volume of 100μl), 10 μl of Fugene6 (care was taken to minimise contact between plastic surfaces and undiluted reagent), 2 μg of plasmid DNA, 1.5 μg of hPXR expression plasmid. After incubation 5 ml of serum free DMEM were added. The amounts of reagent and DNA were adjusted proportionally to this final volume of medium according to the needs of the experiments.

The DMEM from the HuH7 seeded plates was removed by inverting and shaking the plates in order to avoid disruption of the cell layer by the pipette tips. The transfection mixture was then added to the cells (120μl/well) and the plates incubated at 37°C for 24 hours. It is worth noting that during initial experiments it was found that a shorter (6 hours) incubation period could also be used as this was sufficient to produce good levels of transfected cells as determined by the SEAP levels on day 1.

2.12.4. Dosing of the cells with rifampicin (day 1 and day 3).

During preliminary experiments it was found that the different mutated plasmids, as well as plasmid from different maxipreps, varied in terms of transfection efficiency. In order to assess this phenomenon, a 10μl sample was taken from each well to measure the levels of alkaline phosphatase on day 1 (prior to drug treatment) and this value was used to normalise the reporter gene levels observed in day 3. It was assumed the measurement of SEAP production in the medium by the transfected constructs reflects the amount of
plasmids in the cells of a given well. It also takes into account any endogenous activation (e.g. endogenous SXR ligands). This was preferred to the use of co-transfection with another plasmid (e.g. β galactosidase) because upon xenobiotic treatment the transcription levels of the latter may be altered and the induction of the two plasmids may not be relative. Additionally the transfection efficiency of the two plasmids may differ.

The remaining medium was removed by inverting and shaking the plates and replaced with fresh medium containing various concentrations of xenobiotic dissolved in DMSO (usually 0.5 to 10μM rifampicin). The solvent concentration was kept constant between treatments to 0.1%. The plates were incubated at 37°C for 48 hours and 10μl samples were taken to measure the levels of alkaline phosphatase on day 3 in order to assess induction of the constructs caused by rifampicin treatment.

2.12.5. Chemiluminescent detection of SEAP protein activity.

The levels of alkaline phosphatase secreted in the cell culture medium were assessed using a chemiluminescent alkaline phosphatase assay (Aurora kit, ICN Pharmaceuticals) with minor modifications to the manufacturer’s protocol. Eighteen microliters of 1x dilution buffer were added to 10μl of medium in each well of a 96-well optiplate (polystyrene microplates). The mixture was incubated at 65°C for 30 minutes in order to remove any endogenous alkaline phosphatase activity and then cooled on ice to room temperature. 24μl of assay buffer were added and the plates were incubated for 5 minutes at room temperature. Then 24μl reaction buffer (containing the substrate) were added and the plates were incubated at room temperature for 20 minutes. The plated were read on a Lumicount plate reader (Packard) with an optimal gain and photomultiplier settings automatically determined by the machine after screening the plate for the maximum and minimum luminescence values.

The day 1 and day 3 values were corrected for gain difference by multiplying the day 3 value by the ratio of the gain on day 3 / gain on day 1. The values were then corrected for transfection efficiency (value on day 3 / value on day 1) and averaged (n=6).

The induction was modeled using the following hyperbola equation (GraphPad Prism\textsuperscript{TM}, version 2.00):

\[
Y = \frac{I_{\text{max}}X}{(EC_{50}+X)}
\]

where \(Y\) is the induction, \(X\) the xenobiotic concentration, \(I_{\text{max}}\) the maximal induction and \(EC_{50}\) the concentration of ligand required to reach half-maximal induction. The 95% confidence intervals for the variables was calculated using the software. Following non-linear regression analysis, \(r^2\) values ranged from 0.40 to 0.85, meaning that by using this model, 40 to 85% of the induction was explained by the increasing amounts of rifampicin. In order to further investigate the effect of the mutations, individual data points were compared between constructs transfected in cells exposed to similar rifampicin concentrations. The data normality was first confirmed by a Kolmogorov-Smirnov test and the comparisons were made using an unpaired t-test with Welch correction.
CHAPTER 3 - CHARACTERISATION OF DNA:PROTEIN INTERACTIONS IN THE CYP3A4 PROXIMAL PROMOTER.

3.1. Introduction.

In order to elucidate the molecular mechanisms of CYP3A4 gene regulation it is necessary to identify which proteins may bind to the regulatory regions to further determine the factors driving expression (e.g. hormone specific and tissue specific signals). Several reports have demonstrated that the first 300 bp of CYP3A4 proximal promoter may act as a basal transcription unit conferring responsivity to several classical CYP3A4 inducers in reporter gene assays (Barwick et al., 1996; El Sankary et al., 2002). Within this region an everted repeat (ER6) was identified as the binding site for the SXR, an important regulator of CYP3A4 induction (Lehmann et al., 1998). To date this is the only cis-acting element that has been experimentally defined within the proximal promoter.

In addition to the proximal promoter a second region with an important role in CYP3A4 expression has previously been characterised. The distal enhancer (XREM) was shown to be located from in the 5' flanking region of the transcription start (-7836 to -7607) and nuclear receptor binding sites for SXR as well as possible binding sites for other transcription factors were identified with DNase I footprinting and reporter gene assays (Goodwin et al., 1999).

To further examine DNA:protein interactions within the proximal promoter it was decided to carry out a mapping study to determine other cis-acting elements apart from the ER6 within this important regulatory region. As the hepatoma cell line HepG2 had successfully been used in reporter gene assays in our laboratory (El Sankary et al., 2002; El Sankary et al., 2001; El Sankary et al., 2000; Ogg et al., 1999) it was hypothesised the transcription factors necessary for CYP3A4 regulation are expressed within this cell line and nuclear extracts were prepared. DNA:protein interaction were first confirmed using
EMSA and HepG2 nuclear. Following this the nature of these interactions was examined using DNase I footprinting and bioinformatics analysis.

3.2. Results.

3.2.1. Electrophoretic motility shift assay for the proximal promoter (-307/+7 region).

The photograph in figure 3.1 (panel A) shows a typical EMSA gel where the -301/+7 DNA probe was incubated with varying amounts of HepG2 nuclear extract. Three retarded bands were observed with higher molecular weight complexes being observed with the highest protein concentrations. It should be noted that for in the lanes with the highest protein amounts some of the probe was left in the wells where the samples were loaded and did not fully enter the gel. A competition experiment carried out using unlabelled -301/+7 as a competitor shows that the binding is reversible following the addition of 100x excess cold probe (see panel B in figure 3.1).

3.2.2. Footprinting analysis.

In order to identify the protein binding pattern in each DNA-protein complex seen in the EMSA experiments, in vitro copper-phenanthroline footprinting was attempted (see figure 2.7, protocol 1). This technique consists in the chemical digestion of DNA from individual bands cut out from a polyacrylamide gel. However, as several attempts to carry out this assay were unsuccessful this procedure was discontinued and DNase I footprinting was carried out as an alternative. The results presented below were generated using a method adapted from a commercial kit (figure 2.7, protocol 3).
3.2.2.1. Positive control using recombinant AP2 and SV40 enhancer element.

A protected area was observed when the single end labelled SV40 probe was incubated with a recombinant AP2 protein and digested with DNase I (data not shown). This proved the protocol used allowed for the identification of protein binding sites in the test DNA sequence.
Figure 3.1. Autoradiograms of EMSA experiments. A: a ³²P radiolabelled –301/+7 DNA probe was incubated with increasing amounts of HepG2 nuclear protein. B: the same labelled probe was incubated with increasing amounts of HepG2 nuclear protein and the same reaction containing 40 µg proteins was spiked with increasing amounts of unlabelled probe to compete the binding. This experiment was carried out four times and was reproducible.
3.2.2.2. Optimisation of *in vitro* DNase I footprinting of the -301/+7 fragment.

An experiment was carried out in order to optimise DNase I cutting with a range from 0.05 to 0.60 U of enzyme (data not shown) and the final DNase I activity chosen was 0.15U. In several subsequent experiments binding buffers containing high and low salt concentration were used (25 and 100 mM KCl), as well as various amounts of nuclear proteins in order to investigate the effect of these two variables on the protected area pattern on both DNA strands. It also appeared that varying the electrophoresis time allowed for a better detection of protected areas located towards the extremities of the probe (i.e. longer electrophoresis for the larger fragments and short electrophoresis for the small fragments). The initial methods included the use of poly-dI:dC, which prevents non-specific binding of proteins to DNA. However, the binding buffer in the Promega Core Footprinting kit purchased did not contain any poly-dI:dC. In an experiment using the optimised assay conditions, 50μg of protein and 3μg poly-dI:dC were included in a footprinting reaction and the presence of this compound did not appear to affect the protection pattern. Hence poly-dI:dC was not included in further experiments.

3.2.2.3. Footprinting of 3 overlapping fragments and composite model of nuclear protein binding to the CYP3A4 -301/+7 proximal promoter fragment.

A representative DNase I footprinting experiment using the -301/+7 probe is shown in figure 3.2. Comparison of the digestion pattern in the presence or absence or proteins allowed for the identification of several DNA:protein interaction sites within the fragment. Surprisingly all footprinting experiments failed to demonstrate any protein concentration-dependent effect on the binding pattern in contrast to what was suggested by the presence of several complexes (states of binding) observed in the EMSA experiment (see section 3.2.1). Also footprinting experiments using nuclear extracts from rifampicin or dexamethasone treated HepG2 cells failed to reveal any qualitative change in the binding pattern (not shown). Protected areas near to the ends of the probe could not be assessed because the bands were too faint or the pattern was not found to be reproducible. Because of the lack of resolution on the edges of the DNA fragment it was decided to amplify 2 overlapping probes from genomic DNA and to carry out the DNase
I footprinting protocol. Fragments representing the -388/-83 and -240/+69 regions of the CYP3A4 promoter were processed as described for the -301/+7 fragment. Unexpectedly the protein binding pattern from those 2 new probes did not exactly match that of the first probe. It was then decided to generate a model representing the protein binding site pattern in agreement with at least 2 of the DNA probes used. Following footprinting experiments on both strands, a composite diagram of protected areas was constructed (figure 3.3). Putative transcription factor binding sites were assigned using the Alibaba2 search programme to interrogate the Transfac 4.0 database, with assignment based upon liver specificity of transcription factors or known involvement in CYP3A gene regulation. It can be noted that some binding regions are far wider than the putative binding sites suggesting the presence of multiple overlapping regulatory sites. It is interesting to note that the site containing the previously identified T>C mutation (-190 bp from the transcription start, see chapter 5) was protected, indicating the binding of one or several types of HepG2 nuclear proteins to this region. According to this model the potential transcription factor binding sites were the ER6 (where CAR, SXR may bind) as well as response elements for Sp1, Oct1, HNF3, AP2 and C/EBPα.
Figure 3.2. Typical autoradiogram of an *in vitro* DNase I footprinting experiment on the $-301/+7$ region of the *CYP3A4* gene. The $^{32}$P end-labelled probe was incubated with increasing amounts of nuclear proteins from HepG2 cells. This experiment was repeated four times and was reproducible.
Figure 3.3. Composite model of the protein binding to the -301/+7 region of the CYP3A4 gene using *in vitro* DNase I footprinting on the -301/+7, -388/-83 and -240/+69 promoter fragments.

The transcription putative binding site (underlined) are described in the text and were assigned using the Alibaba2 programme to interrogate the Transfac 4.0 database.
3.3. Discussion.

3.3.1. Electrophoretic motility shift assay for the proximal promoter (-307/+7 region).

This experiment was carried out to confirm that under the experimental conditions examined HepG2 nuclear proteins could interact in vitro with the CYP3A4 promoter fragment. The results indicate that nuclear proteins do bind to this regulatory region and that this binding seems to be influenced both quantitatively and qualitatively by the protein concentration. Indeed three retarded bands in total were observed, suggesting the presence of at least three different DNA-protein complexes. The molecular weight of each type of complex is inversely proportional to the distance of migration and heavier DNA:protein complexes were observed for increasing amounts of protein in the binding reaction. The nuclear proteins bind with high affinity because some of the complexes started disappearing in the presence of a 10 fold molar excess of unlabelled oligomer and were completely competed out with 100 fold excess unlabelled probe. This therefore demonstrates the specificity of the observed DNA:protein interactions.

3.3.2. In vitro DNase I footprinting of the -301/+7 fragment: method optimisation and technical considerations.

Buffer salt concentration (K+ at 25 or 100 mM) did not seem to have an effect on the protection pattern (data not shown). This is surprising since the ionic strength of the buffer should modify the protein binding kinetics and the affinity (Garabedian et al., 1993). There is a certain lack of correlation between the EMSA and footprinting results since no concentration effect was seen on the protein binding in the latter assay. Addition of increasing amount of proteins had no qualitative or quantitative effect on the apparent
binding. The reason(s) for this discrepancy are unknown. It could be because the differing assay conditions may favour alternate interactions.

The protection pattern showed little binding at both extremities of the -301/+7 fragment. Such an effect may be as a result of several reasons. Firstly DNA:protein interactions may not be stable near the termini of the fragments, preventing full DNase I protection. Secondly technical considerations such as compression at the top of the gel and diffusion (resulting in weaker signal) near the bottom of the gel may make identification of protected regions more difficult. Due to this it was decided to examine three overlapping fragments and combine the protection patterns seen in the centre of each fragment to obtain an overall pattern.

3.3.3. General footprinting pattern analysis of the 3 overlapping fragments and composite model of nuclear protein binding to the CYP3A4 proximal promoter.

The lack of concordance in the protein binding between the three fragments tested (-388/-83, -301/+7, -240/+69) was unexpected since the regions of overlap between the probes contain the same DNA sequences and therefore should be recognised by the same factors. Besides the fact that this phenomenon could be due to a technical error, this result may possibly reflects protein-protein interactions occurring within the CYP3A4 promoter. The binding of a given factor may require, or be enhanced by, the presence of other DNA binding proteins. For example, members of the C/EBP family can interact with other transcription factors such as NF-κB via their leucine zipper motif (LeClair et al., 1992) and hence influence their binding to DNA. When comparing the different probes used in this study it may be that some of the binding sites occupied in one fragment may be unoccupied in the others because some DNA sequence necessary for the binding of an interacting protein is absent in that particular fragment. To circumvent such potential problems DNA:protein interactions were only assigned if they were present in two of the three fragments tested (i.e. they showed a high confidence of being real).
The footprinting pattern shows large footprints covering several putative transcription factor sites (figure 3.3). This suggests the presence of composite binding sites due to association of several factors. The high degree of binding to this region is not surprising since its importance in trans-activation assays has been shown to be critical (Hashimoto et al., 1993) and it also confers responsiveness to many CYP3A4 inducers in primary hepatocytes and hepatoma cell lines (e.g. Barwick et al., 1996; El Sankary et al., 2002). The fact that several of the putative binding sites are recognised by factors from the HNF and C/EBP families is consistent with the liver expression of these factors (Xanthopoulos et al., 1991). In fact a link has been suggested between the low cytochrome P450 content and the low abundance of liver enriched transcription factor in de-differentiating hepatocytes and cell lines (Rodriguez-Antona et al., 2002).

In concordance with the published literature for the XREM region (Goodwin et al., 1999), the observed binding pattern to the 300 bp 5' of the transcription start was not changed by the use of nuclear extracts from xenobiotic treated HepG2 cells. Footprinting patterns of the rat CYP3A1 promoter with extracts from rat liver from control and dexamethasone-treated animals were also found to be identical (Quattrochi et al., 1995), suggesting this may be a general phenomenon of CYP3A proximal promoters and not a species specific effect. Because the drug concentration and cell treatment were similar to conditions where a CYP3A4 promoter construct can be trans-activated in a reporter gene assay (El Sankary et al., 2000; Ogg et al., 1999), it can be hypothesised that all transcription factors necessary for promoter activation are present within the nuclear extracts. This could suggest either that this technique is not suitable for measuring quantitative changes in receptor association with a promoter or that the receptor association is not altered by xenobiotic exposure and instead other regulatory proteins interact with the co-associated protein to increase their ability to activate transcription. The former hypothesis is confirmed by work described in chapter 5 where a single base pair mutation which had an effect on glucocorticoid mediated activation of a -301/+7 construct did not cause any noticeable change on the nuclear protein binding pattern in footprinting experiments. Instead the change in affinity could be measured by EMSA in a competition assay with unlabelled competitor probes (El Sankary et al., 2002). The latter
hypothesis may also be valid in this situation since nuclear receptors are known to recruit co-activators upon ligand binding and to interact with repressor proteins when inactive (e.g. Jenster, 1998). Indeed following xenobiotic activation SXR has been shown to recruit the SRC1 co-activator protein in a concentration dependent fashion (Kliewer et al., 1998). Additionally SXR has been shown to interact with the transcription repressor protein SMRT (Silencing Mediator for Retinoid and Thyroid receptors) (Takeshita et al., 2002). The fact that SXR (and potentially other transcription factors) associates with repressors but may still bind to promoters in the basal state and thereby inhibit the transcription machinery is also suggested by experiments in SXR knockout mice. Homozygous knockout mice expressed cyp3a11 at a higher level (~4 fold higher) than the wild type animals (Staudinger et al., 2001) suggesting that removal of SXR may actually increase transcription levels in some circumstances. However, in another study on PXR knock-out mice (Xie et al., 2000b) ablation of the receptor was not reported by the authors as being the cause for an up regulation of the cyp3a11 gene. However, closer inspection of the Northern blot data from this paper suggests up-regulation may be indeed occurring (figure 1c in Xie et al., 2000b). Alternatively the discrepancy between the papers could be explained by differences in the choice of sequence used to target and disrupt the SXR gene, although this is unlikely.

3.3.4. Transcription factors potentially binding to the CYP3A4 promoter region.

Based on an assessment of the current literature and the data presented above the following factors have been selected as potential candidate transcription factors involved in CYP3A4 regulation.

3.3.4.1. Specificity protein 1 (Sp1).

This transcription factor is part of the Sp-protein family which mainly binds to GC/GT boxes in the promoters of several genes (Suske, 1999). Sp1 is involved in the regulation of cell cycle regulated or hormonally activated genes and has been shown to be crucial in
the developmental process (Suske, 1999). Indeed disruption of Sp1 binding to the promoter of genes involved in the development by thalidomide is a candidate mechanism for this compound toxicity (Stephens et al., 2000). This transcription factor has also been found to interact with Pit-1 with both having the ability to bind to their overlapping binding sites (Schaufele, 1990). Sp1 recognises sequences in the proximal promoter and the enhancer (PBRE) of the CYP2B1 gene (Muangmoonchai et al., 2001). Experiments in rat hepatocytes have demonstrated that this factor enhances the basal activity of the CYP2B1 proximal promoter and also increases both the basal and phenobarbital-induced activation in the PBRE regulatory region (Muangmoonchai et al., 2001).

Sp1 and Sp3 could both bind a NF-κB-like response element in the promoter of the CYP3A7 gene (GGCAAGTCCC) (Saito et al., 2001). However, this would not occur in the same region of the CYP3A4 promoter (GTCAAGTCCC) due to a 1 bp change in sequence. This element is located at approximately 2300bp away from the transcription start of CYP3A7 but does not exclude that closer putative binding sites may be functional. Indeed, using gel shift assays with unlabelled competitors and specific antibodies, Sp1 was found to bind to the BTE region of the CYP3A5 gene and to activate reporter gene constructs (Iwano et al., 2001). Also in the same study the authors claim proteins from the Sp family also binds to the same BTE element of the CYP3A4 and CYP3A7 genes (unpublished data in Iwano et al., 2001). This contradicts findings from this study since the model in figure 3.3 does not indicate any binding to this site.

3.3.4.2. Hepatocyte Nuclear Factor 3 (HNF3) and members of the HNF family.

HNF3 is a transcription factor involved in the tissue specific expression of numerous genes (e.g. Darlington, 1999). Therefore HNF3 and other members of the HNF family are likely to play a role in the expression of the CYP3A4 gene.

HNF4 is another member of the HNF family and it has been shown to be involved in the regulation of the rat CYP3A23 gene (Huss and Kasper, 2000) and members of the CYP2C gene sub-family (Chen et al., 1994). Using adenoviral vectors coding for antisense HNF4 RNA resulted in a loss of receptor protein and a decrease in the expression of CYP3A4.
and various other CYP genes (Jover et al., 2001). If HNF4 is not directly involved in CYP3A4 regulation, data from HNF4 knock out mice indicate that this transcription factor is involved in maintaining mouse PXR expression (Li et al., 2000), which in turn would have an effect on CYP3A target genes.

It is interesting to note that some reports have proposed a role for HNF4 as a possible mediator of the effects of nitric oxide (NO) on cytochrome P450 gene expression. This molecule is involved in signalling during inflammation or infection. Reporter gene assays demonstrated that CYP2D6 gene down regulation following NO treatment was caused by the loss of protein binding at an HNF4 site (Hara and Adachi, 2002). It was suggested that this could take place through S-nitrosylation of cysteine residues at the level of the DNA binding domain thereby reducing the DNA binding activity of the receptor (Vossen and Erard, 2002).

3.3.4.3. CCAAT/enhancer binding protein alpha (C/EBPα).

As with the members of the HNF family, C/EBPα is part of a gene family involved in the regulation of liver specific gene expression (Darlington, 1999). It is mainly expressed in the liver but is also expressed in adipose tissue (Lekstrom-Himes and Xanthopoulos, 1998). C/EBPα is generally thought to be involved in the control of genes involved in energy homeostasis. Mice lacking C/EBPα were found to die within 8 hours of birth due to their inability to store glucose through down-regulation of a number of key genes including the neoglucogenic enzyme phosphoenolpyruvate carboxkinase (PEPCK) (Wang et al., 1995). C/EBPα has been reported to increase the activity of the -169/+11 fragment of the CYP3A4 promoter together with another liver specific element (D element binding protein) (Ourlin et al., 1997). However, because this effect was cell line dependent it cannot be excluded that the activation observed was due to increased production of another CYP3A4 regulating transcription factor whose expression was dependent on C/EBPα. It should be noted that in the same study a construct containing the CYP3A7 regulatory region was not activated by the D element binding protein. This could constitute a molecular basis for the developmental pattern of CYP3A gene expression.
A very recent report suggests the presence of two functional binding sites for C/EBPα in the -350/-311 and -628/-608 regions of the rat CYP3A1 gene promoter (Rodrigues et al., 2003). Although the rat and human CYP3A regulatory sequences are different it is not unreasonable to assume that they may share a certain degree of resemblance in terms of the transcription factors affecting the proximal promoter activity, thus strengthening the case for the involvement of C/EBPα in the regulation of CYP3A genes.

3.3.4.4. Factors binding to the everted repeat 6 (ER6).

As expected from previous experimental evidence (e.g. reviewed in Honkakoski and Negishi, 2000) the ER6 site of the CYP3A4 promoter was occupied by nuclear proteins from HepG2 cells. However, this is the first time this binding has been observed within a sequence larger than the ER6 element itself, taking into account potential interactions within the promoter that might affect this binding.

Two of the nuclear receptors occupying this site are most likely to be CAR and SXR, since both have been shown to bind to the ER6, as heterodimers with RXR, using transfection assays and EMSA (Xie et al., 2000a) and were also shown to bind several CYP3A4 inducers (Moore et al., 2000). This cross talk between the two nuclear receptors is well documented and it is now known that SXR can activate what were thought to be CAR target genes and vice-versa CAR can activate what were known to be SXR target genes (Smirlis et al., 2001).

Recent research has shown that the ER6 sequence is also a functional binding site for the vitamin D receptor (VDR). The earliest report demonstrates a dose dependant induction of CYP3A4 mRNA by the natural VDR ligand 1α,25 dihydroxyvitamin D₃ (1,25-(OH₂)D₃) in the intestinal Caco2 cell line(Schmiedlin-Ren et al., 1997). The vitamin D analogue 19-nor-1α, 25 dihydroxy vitamin D2 was later shown to cause an induction in this cell line although this was to a lesser extent (Thummel et al., 2001). CYP3A catalytic activity and protein content, as well as CYP3A4 mRNA were induced by 1,25 dihydroxy-vitamin D3 in another colon cell line (LS180), a pancreas derived cell line (HPAC) and human hepatocytes. However, the liver (HepG2) and stomach (Hs746T) cell lines did not support induction (Schmiedlin-Ren et al., 2001). Experimental evidence suggests a role...
for the VDR since this receptor binds to the ER6 sequence in vitro and in a reporter gene assay exposure to 1,25-(OH$_2$)-D$_3$ was shown to activate an ER6 construct in the presence of a VDR expression plasmid but not in the presence of a PXR expression plasmid (Thummel et al., 2001). Using antisense RNA technology the role of VDR was further established as the induction was suppressed in Caco2 cells failing to express the receptor (Hara et al., 2002).

VDR is also involved in the regulation of other CYP genes such as CYP2C9 and CYP2B6 through response elements that were previously shown to be transactivated by CAR and SXR (Drocourt et al., 2002). In addition to inducing CYP3A4, CYP2C9 and CYP2B6 mRNA in primary human hepatocytes, 1,25-(OH$_2$)-D$_3$ transactivated the xenobiotic response element of all three genes in a reporter gene assay. Gel shift assays provided a molecular basis for this observation since recombinant VDR was shown to specifically interact DNA probes representing the three response elements. The VDR-dependent induction of CYP3A4 was found to be influenced by the phosphorylation state of the receptor in that transcription can be modulated via protein kinase C and tyrosine kinase (Hara et al., 2002).

Experiments with SXR knock out mice demonstrate the in vivo relevance of the above observations (Makishima et al., 2002). Besides the VDR was found to bind lithocholic acid and its metabolites and upon receptor activation the cyp3a11 gene was induced in the liver and intestine. This is thought to be a protective mechanism to detoxify the organism from the effects of potential hepatotoxins and enterocarcinogen bile acid metabolites.

In conclusion HepG2 cells may not provide a suitable cellular environment for the observation of this phenomenon (Schmiedlin-Ren et al., 2001) and VDR may not be present in the protein complexes binding to the ER6 sequence in the -301/+7 fragment but VDR appears to have an emerging role in CYP3A4 regulation.
3.4. Conclusion.

In this chapter the binding of nuclear protein extracts on the CYP3A4 proximal promoter was investigated for the first time. A model of protein binding sites in the −301/+7 region was produced in order to speculate on the nature of the transcription factors regulating the transcriptional activation of this gene. As expected the model suggests the presence of nuclear receptors binding to the ER6 region as well as a role for several types of liver specific transcription factors. The observations from the DNase I footprinting experiment are corroborated by a wealth of information describing the involvement of the putative factors described above as regulators of CYP gene activation and basal activity.
CHAPTER 4 - NATURAL POLYMORPHISMS WITHIN THE CYP3A4 GENE.

4.1. Introduction.

Insights into the mechanisms of gene regulation can be gained from the discovery of genetic variants in the population. Mutations which result in a measurable effect on the phenotype can be informative within the context of mechanistic studies since they highlight a region of the genome involved in a qualitative or quantitative manner with the phenotypic variable measured. In addition the relationship between altered genotype and phenotype may provide a clue on the in vivo relevance of the mechanisms studied.

Owing to their importance in drug metabolism, inter-individual differences in CYP3A expression have been well documented (e.g. reviewed in Lamba et al., 2002). For instance in a recent study the hepatic CYP3A content was found to vary by a factor of approximately 40 fold (Lin et al., 2002). Because CYP3A genes are transcriptionally regulated (Telhada et al., 1992), efforts have focused on the analysis of the 5' flanking region of the gene. The first identified mutation reported was an A>G transversion located 290bp upstream of the transcription start site (Rebbeck et al., 1998) and was thought to be related to the incidence of prostate cancer as the mutation was found to be statistically associated with men suffering from tumours of a higher clinical grade. This mutation was also shown to be statistically associated of a higher incidence of treatment related leukemia (Felix et al., 1998). However, such associations are only speculative as no mechanistic information was provided to link CYP3A4-mediated effects with these disease states. Indeed CYP3A activity or protein levels were not assessed in either study.

At the time the work in this study was carried out, no further variants in the promoter region had been reported but two new mutations in the coding region had been identified (Sata et al., 2000). CYP3A4*2 is the most significant allele reported in this study in terms of effect and incidence. It caused a serine to proline change at position 222 of the amino
acid chain in exon 7 and exhibited an altered nifedipine metabolism but no statistically significant change in testosterone metabolism was observed (Sata et al., 2000).

The work presented in this chapter firstly aimed to identify novel mutations in the CYP3A4 proximal promoter (-1086/+7) and exon 7 of the CYP3A4 gene, within a panel of 11 liver samples known to exhibit variability in CYP3A content and catalytic activity. Secondly, oligomers representing a previously identified (Rebbeck et al., 1998) and two novel (Hamzeiy et al., 2002) genetic variants in the CYP3A4 regulatory regions were used to assess the effect of these mutations on the DNA binding of nuclear proteins from HepG2 cells.

4.2. Results.

4.2.1. Mutation detection.

4.2.1.1. Sequencing of the CYP3A4 promoter region (-1086 to +7bp).

Alignment of the three CYP3A4 5' flanking region sequences available online at the time of this study (see section 2.10.5) revealed some variations in the first kb between the sequences reported by Hashimoto (Hashimoto et al., 1993), Goodwin (Goodwin et al., 1999) and Amirimani (respectively under accession number D11131, AF185589, AF181105) and they are described in Table 4.1. These errors were either due to sequencing errors or mutations in the donor DNA, since they were not in the 11 eleven liver samples used in this study (unless all donors in this study were carriers of the same mutation).
Table 4.1. Variation in 3 reported promoter sequences and comparison with the 11 liver samples from the current study.

<table>
<thead>
<tr>
<th>Consensus from 2 sequences</th>
<th>Nature and position of the sequence variation</th>
<th>Sequence observed in the 11 liver samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>•A at –502bp in Hashimoto and Amirimani</td>
<td>G at –502bp in Goodwin</td>
<td>A at –502bp</td>
</tr>
<tr>
<td>•ACAT at –715/-718bp in Goodwin and Amirimani</td>
<td>CATA at –715/-718bp in ACAT at –715/-718bp</td>
<td>Hashimoto</td>
</tr>
<tr>
<td>•G at –722bp in Goodwin and Amirimani</td>
<td>Deletion at position –722bp</td>
<td>G at –722bp</td>
</tr>
<tr>
<td></td>
<td>in Hashimoto</td>
<td></td>
</tr>
</tbody>
</table>

Because of the high sequence homology (91%) reported by Hashimoto *et al.* between the first kb of the promoters of CYP3A4 and CYP3A7 it was difficult to obtain primers highly specific to CYP3A4. There was therefore a possibility that unwanted CYP3A7 promoter sequences could be co-amplified during PCR. Analysis following the sequencing of all fragments was carried out and confirmed that only the CYP3A4 promoter was amplified on all occasions. Hence, analysis for the presence of mutations occurred in the CYP3A4 promoter only.

Complete sequencing of all 11 individuals over the –1086/+7 region revealed no reproducible mutations although occasional sequencing artefacts (i.e. one strand only) were observed. Therefore it can be concluded that the –1086/+7 region of the CYP3A4 promoter sequenced from the 11 liver samples contained no novel mutations, or any of the previously identified polymorphisms (Rebbeck *et al.*, 1998; Sata *et al.*, 2000).

4.2.1.2. Sequencing of the CYP3A4 exon 7.

Sequencing of the DNA from the eleven liver samples revealed wild type sequence at serine 222, with no cases of the published polymorphism CYP3A4*2 (Sata *et al.*, 2000).
4.2.2. Analysis of the binding affinity of nuclear proteins to natural mutations in the CYP3A4 promoter.

Whereas the screening of 11 phenotyped liver samples revealed no novel mutations in the CYP3A4 proximal promoter, screening of a large population by Dr Hossein Hamzeiy allowed the discovery of novel polymorphisms (Hamzeiy, 2002). A population of 101 subjects (41 females and 60 males of mainly Caucasian or Iranian origin) was screened for variations to the wild type sequence in both the proximal promoter and the XREM region of CYP3A4 using a non-radioactive SSCP technique (Hamzeiy et al., 2002). This work revealed the presence of a known variant (CYP3A4*1B) within the population but also allowed for the identification of 3 novel mutations (table 4.2 and figure 4.1).

To study the functional effect of these polymorphisms, DNA from these mutant promoters was cloned into a pSEAP plasmid and their response to classical CYP3A4 inducers studied in a reporter gene assay (Hamzeiy et al., submitted). The drug concentration choice was based on preliminary experimental evidence (Hossein Hamzeiy, personal communication), where the compounds were found to fully induce the constructs without causing cytotoxicity. Typical experiments are shown in figure 4.2 and the data demonstrate a clear effect on the transactivation of the reporter gene constructs. Although all mutations affected the induction of the constructs in a statistically significant manner, for at least one of the four xenobiotic inducers used, CYP3A4*1B and CYP3A4*1F appeared to have the strongest effects. Overall, it should be noted that, although the statistical analysis demonstrates a difference in induction to a matched wild type promoter sequence, the effect of the mutations doesn’t exceed a 2.15 times increase in induction. Therefore, using the experimental system described in figure 4.2, the genetic variants appear to have a limited effect on the activation of the CYP3A4 promoter reporter gene construct.

The effects of three of these mutations on protein binding were examined by EMSA in order to investigate possible mechanistic links between the altered transcription rates and the binding of transcription factors to the mutation sites (described in figure 4.1). Due to time limitation the protein binding to the CYP3A4*15B allele was could not be investigated.
Table 4.2. Novel alleles identified following screening of a population of 101 subjects (data taken from Hamzeiy *et al.*, 2002)

<table>
<thead>
<tr>
<th>Allele name</th>
<th>Position from ATG start codon</th>
<th>Modification</th>
<th>Observed incidence in 101 subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP3A4*1B</td>
<td>-392</td>
<td>A→G</td>
<td>9</td>
</tr>
<tr>
<td>CYP3A4*1E</td>
<td>-369</td>
<td>T→A</td>
<td>1</td>
</tr>
<tr>
<td>CYP3A4*1F</td>
<td>-747</td>
<td>C→G</td>
<td>17</td>
</tr>
</tbody>
</table>
| CYP3A4*15B | • Insertion between -845 and -844
• -392
• Position 485 on cDNA | ATGGAGTAG A→G G→A | 1                                 |

In this section the work on mutation detection and functional analysis by transfection assay was carried out by Dr Hossein Hamzeiy.
Figure 4.1. Novel CYP3A4 alleles within the context of the putative response elements (adapted from Hamzeiy et al., 2002)

<table>
<thead>
<tr>
<th>Wild type</th>
<th>CYP3A4*1B</th>
</tr>
</thead>
<tbody>
<tr>
<td>-392</td>
<td>-392</td>
</tr>
<tr>
<td>ATAGAGACAAGGCAAGGAGAGGCGATT</td>
<td>ATAGAGACAAGGCAAGGAGAGGCGATT</td>
</tr>
<tr>
<td>NFSE</td>
<td>NFSE</td>
</tr>
<tr>
<td>=========</td>
<td>=========</td>
</tr>
<tr>
<td>Spt-1</td>
<td>Spt-1</td>
</tr>
<tr>
<td>=========</td>
<td>=========</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wild type</th>
<th>CYP3A4*1E</th>
</tr>
</thead>
<tbody>
<tr>
<td>-369</td>
<td>-369</td>
</tr>
<tr>
<td>ATAGAGACAAGGCAGAAGGCGATT</td>
<td>ATAGAGACAAGGCAGAAGGCGATT</td>
</tr>
<tr>
<td>CAAT box</td>
<td>CAAT box</td>
</tr>
<tr>
<td>=========</td>
<td>=========</td>
</tr>
<tr>
<td>NF-1</td>
<td>NF-1</td>
</tr>
<tr>
<td>=========</td>
<td>=========</td>
</tr>
<tr>
<td>Oct-1</td>
<td>Oct-1</td>
</tr>
<tr>
<td>=========</td>
<td>=========</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wild type</th>
<th>CYP3A4*1F</th>
</tr>
</thead>
<tbody>
<tr>
<td>-747</td>
<td>-747</td>
</tr>
<tr>
<td>CTGTGTACAGCACGGCTGGTAGGG</td>
<td>CTGTGTACAGCACGGCTGGTAGGG</td>
</tr>
<tr>
<td>Spt-1</td>
<td>Spt-1</td>
</tr>
<tr>
<td>=========</td>
<td>=========</td>
</tr>
</tbody>
</table>
Figure 4.2. Reporter gene assay experiments in HepG2 cells demonstrating the effect of the novel mutants on the transcriptional activation by xenobiotics (data taken from Hamzeiy, 2002). HepG2 cells were transfected with the wild type (CYP3A4*1A) or mutation-containing pSEAP2 reporter plasmid DNA (with XREM insert) in the presence of pSG5-hPXRΔATG. After DNA transfection, cells were exposed to CYP3A4 inducers at 10 μM (phenobarbital, 1mM) and incubated for a further 60 hours prior to chemiluminescent SEAP assay. (DEX: dexamethasone, PB phenobarbitone, MIF: mifepristone, RIF: rifampicin). The results are derived from 4 experiments in which each transfection and exposure was performed six times (n=24). Relative inductions of mutation-containing plasmid constructs compared to the wild type are presented in the table. Differences were analysed by two-sided t-test (*, P<0.05; **, P<0.01; ***, P<0.001 comparing mutant to wild type). The p values relate to the wild type promoter.

* P<0.05, ** P<0.01, *** P<0.001. (comparing mutant to wild type allele)
**Figure 4.2.** (continued)

<table>
<thead>
<tr>
<th>Drug Allele</th>
<th>DEX</th>
<th>PB</th>
<th>MIF</th>
<th>RIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3A4*1B</td>
<td>1.47 ***</td>
<td>1.20 *</td>
<td>2.15 ***</td>
<td>1.59 ***</td>
</tr>
<tr>
<td>3A4*1E</td>
<td>1.22 *</td>
<td>0.95</td>
<td>1.12</td>
<td>1.10</td>
</tr>
<tr>
<td>3A4*1F</td>
<td>1.62 ***</td>
<td>1.23 *</td>
<td>1.89 ***</td>
<td>1.57 ***</td>
</tr>
<tr>
<td>3A4*15B</td>
<td>1.36 **</td>
<td>1.04</td>
<td>1.32 *</td>
<td>1.32 *</td>
</tr>
</tbody>
</table>

* P<0.05, ** P<0.01, *** P<0.001.
All DNA oligomers, both wild type and variant tested, bound proteins from the HepG2 nuclear extract and this was shown to be specific as the retarded complex disappeared with a competitor concentration of approximately 30 fold excess to that of the labelled probe (figure 4.3). A single non-specific complex was seen, with CYP3A*1F, which was only competed out by the presence of 1200 fold excess competitor probe (data not shown). All binding reactions were optimised in order to observe an effect of the mutations and the differences in protocol probably reflect relative difference in affinity. For instance the competition assay using the probe representing the CYP3A4*1B allele was only observed in binding reactions containing higher levels of nuclear proteins. The three mutations tested seemed to affect nuclear protein binding, with CYP3A4*1E causing an increase in affinity for 15 fold excess unlabelled competitor whereas CYP3A4*1B and CYP3A4*1F were responsible for a decrease in affinity for competitor concentrations of 15 and 5 to 30 fold respectively. Overall, the CYP3A4*1F allele had the most pronounced effect on protein binding.

In order to confirm the findings for the probe representing CYP3A4*1B an additional experiment was carried out where protein binding was abolished by digestion with proteinase K at 37°C for 20 minutes and heat denaturation of the protein extract at 95°C for 20 seconds together with the appropriate controls (data not shown).
Figure 4.3. Effect of the CYP3A4*1B, *1E and *1F alleles on HepG2 nuclear protein binding. Error bars indicate standard error of the mean (±SEM, n=3). † = non-specific binding.
4.3. Discussion.
To date only 5 variant alleles with mutations in the 5' flanking region of CYP3A4 have been reported (reviewed in Lamba et al., 2002) and these are all located between -747 and -62 bp from the transcription start. In a study of 300 samples, no mutations were found in the ER6 element located in the proximal promoter as well as within the SXR binding site located in the distal enhancer XREM (Zhang et al., 2001). This is consistent with another study of 101 subjects where no mutations were found in both the proximal ER6 and a 300 bp region of the XREM (Hamzeiy, 2002). This raises the possibility that the ER6 or the distal SXR binding sites are so essential for CYP3A4 gene regulation by nuclear receptors and for general xenobiotic metabolism that any alteration at the level of these regions may impair survival of individual carriers and therefore such variants may not propagate in the population.

4.3.1 Mutation detection in the promoter of CYP3A4.
The study of the -1086/+7 region of the CYP3A4 promoter, for 11 phenotyped liver samples exhibiting a 14.1 fold variability in CYP3A activity revealed no inter-individual genetic differences as determined by sequencing analysis.
There is a possibility that all or part of the phenotype variability observed may be caused by several factors such as the degradation of the samples or the lack of specificity of the method used to determine the phenotype (the hydroxylation of the probe substrate midazolam is also a CYP3A5 marker). Due to technical difficulties relative to antibody quality and probe substrate the identity assessment of the exact CYP3A can prove unreliable at the level of the protein content and the metabolic activity. According to a recent study it appears that CYP3A5 product content represented more than half of the hepatic CYP3A content in a third of a Caucasian population sample and in more than half of an African American population sample. This indicates that the levels of this enzyme may have been previously underestimated (Kuehl et al., 2001). Some reports describe a strong inter-individual variability in CYP3A level and activity but suggest this could be due to the polymorphic expression of CYP3A5 (e.g. Kuehl et al., 2001 and reviewed in Lamba et al., 2002). However, Western blotting analysis revealed the levels of CYP3A4 immuno-reactive proteins varied within the samples although this did not correlate with
the observed enzyme activity (Nick Plant, personal communication). This highlights the difficulties in finding a reliable phenotypic marker to carry out phenotype-genotype correlation studies.

Rebbeck et al. (1998) previously reported a -290 A>G (later known as allele CYP3A4*1B) mutation and the population frequency for Caucasian subjects was 9.6% (homozygous) and 12.8% (heterozygous) on a sample of 94 patients. Another study also reveals a rather low frequency of the mutation in Caucasian patients (GG: 1%, AG: 6%, n=117) and a higher frequency in healthy African-Americans (GG: 53%, AG: 28%, n=116) (Paris et al., 1999). This inter-ethnic difference is in agreement with the findings from another study (Sata et al., 2000).

In the present study 8/11 samples originated from Caucasian subjects and 3/11 from Afro-American subjects therefore the results observed are not inconsistent with the published literature. In a study using 39 liver samples exhibiting 31-fold variability in testosterone metabolism (CYP3A4 marker), no further mutations were found within the -490/+10 region of the promoter (Westlind et al., 1999).

Although the first 300bp are known to be very important in the regulation of the CYP3A4 gene (Hashimoto et al., 1993), this work demonstrates that other factors must explain the polymorphic expression of CYP3A4 since no mutations were found in spite of a significant variability in the CYP3A phenotype (activity and protein level by Western blot).

4.3.2. Mutation detection in the coding region: exon 7

The study of 11 liver samples available reveals an absence of the previously reported mutation (Sata et al., 2000) in exon 7. This suggests that the inter-individual differences in CYP3A catalytic activity observed were probably not due to mutations in the coding region causing the production of enzyme with a lower activity. The result is not surprising since the population frequency reported was very low, 2.7% (Sata et al., 2000) and may therefore probably not be very significant for the population as a whole.

Attempts to amplify exon 12, which was also reported to contain a mutation (Sata et al., 2000), failed because it could not be amplified using PCR, following the published
primer sequence and reaction conditions. This is probably because one of the PCR primer Sata and his group used was significantly shorter than the other in the pair (only 17bp), indeed a longer primer allows satisfactory amplification (Hossein Hamzey, personal communication). This mutation was only found in one subject (Sata et al., 2000)

4.3.3. Alternative explanations to inter-individual differences in CYP gene expression.

4.3.3.1. The regulators.

This chapter has so far focused on how mutations in the regulatory regions of the CYP3A4 gene may provide a rational basis to explain inter-individual differences in phenotype. However, it cannot be excluded that that the genetic component of CYP3A4 variability may be due to quantitative or qualitative variations of the regulatory proteins driving its expression rather than variations in the regulatory regions of this gene. For instance six mutations in the gene coding for SXR have been identified each with a frequency of less than 3% apart from a P27S change affecting 14.9% of patients from African origin (Hustert et al., 2001). Using functional assays, the latter authors reported that the frequent P27S had no effect on gene activation; however, some of the variant proteins did alter the response (increased or decreased) of reporter gene constructs containing ER6 or DR3 elements known to bind SXR suggesting that some of these variants may contribute to the variation in CYP3A expression. In a larger study, more mutants were identified including variants in the 5' flanking region of the SXR gene disrupting putative transcription factor binding sites, such as for instance a C/EBP binding site (Zhang et al., 2001). Taken together these results demonstrate that SXR ligand and DNA binding activity, as well as expression levels, may be subject to inter-individual variability, which in turn would be expected to affect target gene expression levels. However, a direct association between altered SXR protein and target gene expression would be surprising since this would imply that CYP3A4 expression (or for that purpose the expression of any SXR target gene) is solely dependent on SXR activation. As seen in chapter 5 and in the literature, CYP3A4 expression is not solely
driven by SXR but by other factors acting in the promoter and at the level of the ER6 (Smirlis et al., 2001; Xie et al., 2000a) and any of these other factors may also be variable.

4.3.3.2. Role of epigenetics.

Another possible explanation for CYP3A variability may come from the emerging area of epigenetics (also discussed in chapter 6). Genomic sequencing of individuals’ regulatory regions does not allow the identification of variation in DNA methylation or chromatin structure. Chromatin structure can be responsive to signals following a range of post-translational modifications to the histone proteins which are known as the ‘histone code’ (Jenuwein and Allis, 2001) and in some cases deregulation in the dynamics of histone modifications can be related to carcinogenesis (Schneider et al., 2002).

DNA base modifications such as methylation of cytosine (\(^{\text{5}}\)CpG) can also alter the levels of gene expression since they can affect transcription factor binding. Indeed the binding of the Sp1 protein to a DNA response element comprising methylated bases has been found to be reduced (Clark et al., 1997). Of particular relevance is the work of Takahashi and co-workers on CYPIAI expression in rabbit (Takahashi et al., 1998). They reported that methylation of a sequence binding the AhR (aryl hydrocarbon receptor) transcription factor caused a decrease in affinity for this protein, which explains the cell type specific expression of the cytochrome P450 gene since the methylation pattern was only found in certain cell types and was responsible for silencing of the genes.

To date there is no evidence that these mechanisms are directly responsible for inter-individual CYP3A4 variation. It cannot be excluded, however, that within the in vivo situation epigenetic factors may affect CYP3A4 regulatory regions or expression of CYP3A4 regulatory proteins.

4.3.4. Binding of nuclear proteins to mutant CYP3A4 alleles.

Because the regions altered in the variant alleles were beyond the area (-301/+7) covered in the study presented in chapter 3 no data on binding site localisation was available. However the fact that the oligomers tested bound HepG2 nuclear proteins and that this
binding could be removed via competition with a DNA oligomer (figure 4.3) or digestion with proteinase K does suggest that the sequences are representative regions where DNA protein interactions take place.

The CYP3A4*1B allele is located within a region known as the NFSE (Nifedipine specific element) originally studied because it was thought to be responsible for the CYP3A7/CYP3A4 transition during development, since the NFSE is only present in the promoter of CYP3A4 but not of CYP3A7 (Hashimoto et al., 1993). Incubation of a DNA probe representing the NFSE with human adult and foetal liver extracts revealed a DNA:protein interaction which was considered to be non-specific. Another study focusing on the functional effects of the mutated allele showed this region could bind several proteins or protein complexes present in human liver nuclear protein extracts, but none of these interactions was of high affinity since a high amount of protein and low stringency buffer had to be used (Westlind et al., 1999). In this study, high amounts of protein were also required to observe a difference in competitor affinity for the wild type and mutant DNA oligomers but because as little as 10-fold excess unlabelled oligomer were necessary to remove binding the interaction is thought to be specific. Gel shift assays previously carried out on this region (Hashimoto et al., 1993; Westlind et al., 1999) used an oligomer centred on the NFSE region, however, as shown in figure 4.4 the oligomer design in this study differed. Analysis of the region surrounding the NFSE suggested the presence of 2 putative Sp1 sites (as determined by the Transfac database, figure 4.4). The oligomer used in this study was designed to incorporate these sequences and it may be that the protein binding was enhanced by this modification since both putative sites would be covered and surrounded by a flanking base.
This study: \[ 5'-\text{GAGACAGGGCAAGAGAGGCCGAT}-3' \]
Previous studies: \[ 5'-\text{AGAGACAGGGCAAGAGAGAGG} \]

It is interesting that while discussing their results Hashimoto and co-workers (Hashimoto et al., 1993) refer to unpublished data showing the BTE sequence could compete out protein binding to the NFSE. The BTE is a GC-rich box known to bind Sp1 proteins in the promoter of the CYP1A1 gene and may act as a repressor element (Kaczynski et al., 2002). This corroborates the hypothesis that the NFSE may bind Sp1.

The variant allele which showed the greatest change in binding during the EMSA experiments was \( \text{CYP3A4*1F} \). Disruption of a putative Sp1 site in the variant allele (Figure 4.1) results in a decrease in affinity for nuclear protein (Figure 4.3) whereas the transcription rate is increased (figure 4.2). This would suggest a repressor role for the transcription factor binding to this site. As described above Sp1 has previously been shown to act as a transcriptional repressor for CYP genes (\( \text{CYP1A1} \)) through binding to the BTE. This would provide a molecular basis for the observations described above and would explain the changes seen at the level of the nuclear protein binding and at the transcriptional level for the \( \text{CYP3A4*1B} \) and \( \text{CYP3A4*1F} \) alleles.

Gel shift assay using either recombinant Sp1 protein or antibody raised against this factor would confirm this hypothesis.
4.4. Conclusion.

Sequencing of the *CYP3A4* proximal promoter in a panel of 11 liver samples showing variability in CYP3A protein levels and catalytic activity revealed no deviation from the wild type sequence. However, genetic variations could be detected in another study using a larger sample cohort. Two of these novel variants and a previously reported allele, which are all known to alter the response to xenobiotics in reporter gene assays, altered protein binding of HepG2 nuclear extracts to synthetic oligomers representing the mutated alleles. Analysis of these mutations in the regulatory region of *CYP3A4* has given some insight into the mechanisms that may drive the expression of this gene. However, *CYP3A4* inter-individual variability is not only caused by variations in the regulatory regions of the gene but also probably caused by variations in the quantity and properties (i.e. in terms of DNA binding, ligand binding) of the regulatory proteins. Therefore in order to discover predictors of inter-individual variability it is necessary to identify the factors driving hepatic *CYP3A4* expression and this will be described in the next chapter.
5.1. Introduction.

Following the mapping of DNA:protein interactions with the CYP3A4 proximal promoter by DNase I footprinting (chapter 3) a logical step is to examine the role of the protein binding regions identified in the regulation of the CYP3A4 gene, both during basal and chemically induced gene transcription. As discussed earlier (chapter 3) there is currently no empirical evidence for the involvement of factors other than the proteins binding at the ER6 in CYP3A4 regulation. However, there does exist some circumstantial evidence that factors such as tissue specific transcription factors may be involved in the regulation of the hepatic expression of this gene (Ourlin et al., 1997).

The aim of the work presented in this chapter was to mutate putative protein binding sites and to examine the functional effect under basal and drug-induced conditions using a reporter gene assay. From this important information on other factors involved in CYP3A4 expression can be determined. Four artificial mutations known to interrupt putative DNA:protein sites of interaction were tested.

Firstly during routine cloning of the −301/+7 promoter fragment a T>C substitution was identified at −190 bp by Dr Wafaa El Sankary in our laboratory. This mutation was found to decrease dexamethasone by not rifampicin mediated induction of a −301/+7 construct of the CYP3A4 promoter. This was further investigated in the current study by the use of DNase I footprinting (see section 3.2.2.3) to characterize the protein binding at this site and gel shift assays to assess the effect of the mutation on protein binding.

Secondly following the DNase I footprinting on the −301/+7 region three sites were selected on the basis of the protein binding pattern and computer analysis of the DNA fragment. Using site directed mutagenesis, the DNA sequence in these regions was altered and the functional effect of the mutations was tested using a reporter gene assay.
5.2. Results: Glucocorticoid-mediated induction of \textit{CYP3A4} is decreased by disruption of a protein:DNA interaction distinct from the ER6.

In this section the initial cloning work and the functional analysis by transfection assay were carried out by Dr Wafaa El Sankary (as described in El Sankary \textit{et al.}, 2002). Following amplification of the \text{-301/+7 CYP3A4} promoter fragment by PCR and cloning of the product, a single base pair mutation (T>C transition, -190bp) was detected which was probably caused by polymerase infidelity. This mutation disrupted a putative HNF/C/EBPα binding site, as assessed by DNase I footprinting and bioinformatic analysis (chapter 3). Using a reporter gene assay in HepG2 cells it was found that although the rifampicin induction remained unchanged compared to the wild type promoter (pWT), the induction by dexamethasone and hydrocortisone was markedly reduced (figure 5.1) in the mutant promoter (pMUT). In order to investigate the possible effects of this mutation on the transcription factors binding to this site, DNase I footprinting was carried out using a wild type and a mutant probes representing the \text{-301/+7} region of the promoter incubated in the presence of HepG2 nuclear proteins. The resulting protein binding pattern (figure 5.2) demonstrated that the mutation did not result in any qualitative change. The relative affinities of the wild type and mutant DNA oligomers representing the site of DNA protein interaction was next examined using EMSA in the presence of unlabelled competitors. The results shown in figure 5.3 demonstrate that both fragments bind a protein present in HepG2 cell nuclear extracts and that its binding affinity to the wild type sequence is higher than to the mutant sequence. This would suggest that DNA:protein interactions are reduced in the mutant although the exact nature of this interaction has not yet been identified.
Figure 5.1. A single base pair mutation alters responsiveness of the -301/+7 region of the CYP3A4 promoter to synthetic and endogenous steroids but not to rifampicin (this experiment is described in El Sankary et al., 2002).

The wild type (pWT) and mutant (pMUT) plasmids were transfected in HepG2 cells and instead of measuring the ‘induction’ as described in the method section (chapter 2), the degree of reporter gene production was assessed using a similar measurement, the Specific Chemical Effect (SCE, Plant et al., 2000). The SCE allows a thorough examination of the reporter gene assay data by taking into account experimental variables such as drug-induced proliferative effects or plasmid transfection efficiency.
Figure 5.2. The -190 T>C substitution does not alter the binding pattern of HepG2 nuclear proteins to -301/+7 fragment of the *CYP3A4* promoter.

<table>
<thead>
<tr>
<th>Wild Type</th>
<th>Mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>G A T C</td>
<td>+ + + - G A T C</td>
</tr>
</tbody>
</table>

DNA probes representing the -301/+7 region of the *CYP3A4* promoter were prepared by labelling at one 5' end using $^{32}$P following amplification by PCR from the wild type (pWT) and mutant (pMUT) plasmids. The probes were incubated in the presence (+) or absence (-) of nuclear proteins from HepG2 cells, digested with DNase I and the DNA was run on a denaturing polyacrylamide gel. The mutation site (arrow) disrupts a DNA protein interaction (black line) encompassing a putative HNF/C/EBP binding site.
Figure 5.3. Binding of HepG2 nuclear proteins to the WT oligomer is reduced by the T>C –190 transition.

The oligomers represent the site of DNA:protein interaction in the wild type (WT) and mutant (MUT) plasmids. Error bars indicate standard error of the mean (± S.E.M.) for three repeat experiments.
5.3. Results: Mutagenesis of putative regulatory protein binding sites.

5.3.1. Reporter gene assay optimization and technical considerations.

Preliminary experiments using a construct containing the -1078/-57 fragment of the CYP3A4 promoter (p3A4-CMV-cSPAP as described in El Sankary et al., 2000) and HepG2 cells failed because of a very low transfection efficiency. On the other hand it was found that HuH7 cells could be transiently transfected in a reproducible manner using Fugene 6 (Roche).

The pSEAP pro2 vector with -301/+7 insert was used in the initial work but although the levels of alkaline phosphatase were higher than in the empty pSEAP pro2 vector the construct was not responsive to rifampicin. However, when a similar construct containing the distal enhancer XREM (pSEAP2-basic with XREM insert and -301/+7 insert) was transfected the levels of reporter gene could be detected and could be induced in a concentration-dependent manner. It was also noted that in order to be able to observe high levels of basal expression and a sufficient response to rifampicin, co-transfection with a SXR expression plasmid was necessary (data not shown). Finally it was found the construct could be activated by rifampicin concentration as low as 0.5 μM (but not 0.1 μM, as shown for instance on figure 5.4) and that the dose–response curve generally reaches a plateau for concentrations between 1 and 5 μM (e.g. see figure 5.4).

5.3.2. Comparison of the -301/+7 versus -1201/-61 CYP3A4 promoter fragments.

During preliminary work, the induction of the -301/+7 promoter region was compared to that of a longer promoter fragment (-1201/-61). Both fragments were cloned at the same position in a pSEAP2 promoter plasmid containing the -7972/-7673 XREM fragment. It was found that the levels of reporter gene transcription were significantly higher for the short fragment (figure 5.4), for all doses of rifampicin and the DMSO treatment. However, the level of induction (defined as the ratio between the rifampicin induction and the DMSO solvent control values) were similar. This is explained by the fact that the
-301/+7 fragment exhibited a higher basal level of SEAP expression than the -1201/-61 fragment.

Figure 5.4. Comparison of the rifampicin induction profile for the pWT plasmid (with a -301/+7 insert) and a longer CYP3A4 promoter fragment (-1201/-61).

Both constructs contained the distal enhancer XREM. Induction for each rifampicin concentration was calculated as described previously (section 2.12.6) and is presented as the mean ± standard error of the mean (n=6). The statistical analysis is described in chapter 2 (*: p<0.05, **: p<0.01). This experiment was carried out in triplicate and was reproducible.
5.3.3. Comparison WT versus WT*.

The WT and WT* constructs from which all the SDM plasmids are derived were tested using transfection assays in order to assess the possible effect of the T>C substitution at -193bp present in the pWT* plasmid. The results are presented in figure 5.5. Both promoter had very similar activation profiles and from this experiment it can be concluded that the mutation has very little or no detectable effect on rifampicin mediated reporter gene activation.

5.3.4. Site Directed Mutagenesis (SDM) procedure.

In spite of several attempts to mutate the ER6 motif in both the wild type fragment (WT) and the wild type fragment containing the T→C substitution located at position -193 (WT*), SDM experiments only resulted in one of the half site modified with a AA→TT transversion within the WT* promoter. The use of a longer mutagenic primer, which should have improved the annealing to the plasmid template, proved unsuccessful. Attempts to generate ER6 mutations in the WT fragment repeatedly failed and the procedure was discontinued because of lack of time. The diagram in figure 5.6 describes the plasmids generated during the SDM experiments.
**Figure 5.5.** Comparison of the rifampicin induction profile for pWT and pWT* plasmids.

Summary of the $I_{\text{max}}$ and EC$_{50}$ values generated by the regression analysis presented with the 95% confidence interval for the pWT and pWT* plasmids.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>$I_{\text{max}}$</th>
<th>EC$_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pWT*</td>
<td>WT: 19.29 (13.86 to 24.73)</td>
<td>WT: 0.12 (0.16 to 0.40)</td>
</tr>
<tr>
<td></td>
<td>WT*: 19.59 (14.73 to 24.45)</td>
<td>WT*: 0.18 (0.10 to 0.46)</td>
</tr>
</tbody>
</table>

Induction for each rifampicin concentration was calculated as described previously (section 2.12.6) and is presented as the mean ± standard error of the mean (n=6). The statistical analysis is described in chapter 2. This experiment was carried out in duplicate and was reproducible.
Figure 5.6. Summary of the plasmids generated during the SDM experiments.

WT*  WT

Putative Sp1 (-97bp) CTTCCAAC → CTTTTAAC

Putative HNF3 (-188bp) TGTTTATG → TGTGGATG

Putative C/EBPa (-127bp) TCTTTGC → TCGAGGC

ER6 half site (-166bp) ATGAACTC → ATGTTCTC

The circles represent the mutations and an approximation of their relative location on the -301/+7 fragment. As described in the method chapter pSpl, pHNF, pCEBP were created from the wild type (pWT) plasmid and pSpl*, pHNF*, pCEBP* and pER6 were created from pWT* which contains a T>C substitution at -193bp.
5.3.5. Effect of the mutations at the basal level in HuH7 cells.

All SDM generated mutations were tested against their respective original template (WT or WT*). It should be noted that the plateau values of the WT and WT* plasmid induction may vary between experiments. This is essentially because of inter-assay variations in the detection of the alkaline phosphatase (SEAP). Indeed, although sample levels of reporter gene levels can be reproducibly determined the accuracy is highly sensitive to several exogenous factors (e.g. incubation time, temperature) and all experiments provide comparative rather than absolute readings of SEAP levels.

Initially the possible effect of the SDM fragments on basal reporter gene expression was examined. The SEAP gene production pattern was consistent for fragments generated from both WT (figure 5.7 panel A) and WT* (figure 5.7 panel B) once again suggesting the T>C mutation at position -193 was silent. The ER6, HNF and C/EBP mutants all exhibited a statistically significant reduced basal activity compared to control, while the Sp1 SDM plasmids (pSp1 and pSp1*) showed no statistical change in basal expression levels.
Figure 5.7. Effect of the mutations in pWT (A) and pWT* (B) on the basal levels of reporter gene expression upon transfection in HuH7 cells.

A: pWT vector.

B: pWT* vector

HuH7 cells were transfected with the mutant plasmids and the levels of reporter gene expression were measured 48 hours later. Differences were tested by ANOVA (*: p<0.05) n=6 representative experiment. The experiment was carried out in duplicate and was reproducible.
5.3.6. Effect of the mutations in rifampicin treated HuH7 cells.

The experiments with transfected pER6*, pHNF*, pHNF, pCEBP* and pCEBP were carried out 3 to 4 times and provided highly reproducible results. The graphs in figures 5.8 to 5.11 represent the results for one representative experiment.

The pER6* plasmid (figure 5.8) exhibited a lower $I_{\text{max}}$ in response to rifampicin (95% CI). The HNF mutants (figure 5.9) had a different response to rifampicin mediated gene activation, since the pHNF plasmid was induced to the same level as the pWT plasmid but the mutation in pHNF* caused a significant $I_{\text{max}}$ decrease. Both pCEBP and pCEBP* (figure 5.10) were affected in a similar way by the mutation as the $I_{\text{max}}$ in response to rifampicin were significantly decreased (based on the 95% CI) to about a third of the value for the control plasmids. For the five SDM constructs described above the EC$_{50}$ 95% confidence intervals overlapped demonstrating that there was no significant effect of the mutations on this variable (at p<0.05). The comparison of the SEAP levels between the wild type and mutant sequences for all rifampicin doses allowed identical conclusions apart from pHNF* where the significant difference was found at the 0.5, 5 and 10 μM doses.

Unlike all the other constructs used, the pSp1 and pSp1* plasmid transfections appeared to be generally less reliable since there was a high inter- and intra-experiment variability. The regression analysis of each individual experiment is presented in figure 5.11 and demonstrates the inconsistency of the results. Overall, there was no statistically significant effect on the $I_{\text{max}}$ value of the mutants using both the 95% CI of the values generated by the regression analysis and the t-test. Therefore it appears that similarly to what was found for the experiments on the expression at the basal level, the Sp1 mutation has no effect on the activation of the $-301/+7$ region of the CYP3A4 promoter.
Figure 5.8. Rifampicin response of the ER6 mutant construct upon transfection in HuH7 cells.

A

![Graph showing induction profile for pWT* and pER6* plasmids](image)

B

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>$I_{\text{max}}$</th>
<th>$EC_{50}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pER6*</td>
<td>WT: 45.52 (39.61 to 51.44)</td>
<td>WT: 0.18 (0.00 to 0.35)</td>
</tr>
<tr>
<td></td>
<td>MUT: 15.63 (12.79 to 18.46)</td>
<td>MUT: 0.23 (-0.04 to 0.50)</td>
</tr>
</tbody>
</table>

A: Induction profile for the pWT* and pER6* plasmids presented as induction (± SEM, n=6) and the statistical analysis is described in chapter 2 (**: p<0.001).

B: Summary of the $I_{\text{max}}$ and $EC_{50}$ values generated by the regression analysis presented with the 95% confidence interval for the mutated plasmid (MUT : pER6*) against the value for the respective wild type control (WT : pWT*).
Figure 5.9. Rifampicin response of the HNF mutant constructs upon transfection in HuH7 cells.

A

B

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>$I_{\text{max}}$</th>
<th>EC$_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pHNF*</td>
<td>WT: 34.69 (29.81 to 39.57)</td>
<td>WT: 0.55 (0.21 to 0.88)</td>
</tr>
<tr>
<td></td>
<td>MUT: 21.93 (17.76 to 26.11)</td>
<td>MUT: 0.36 (0.01 to 0.72)</td>
</tr>
<tr>
<td>pHNF</td>
<td>WT: 68.33 (53.83 to 82.84)</td>
<td>WT: 0.12 (-0.13 to 0.38)</td>
</tr>
<tr>
<td></td>
<td>MUT: 76.98 (65.33 to 88.62)</td>
<td>MUT: 0.35 (0.08 to 0.63)</td>
</tr>
</tbody>
</table>

A: Induction profile for the pWT* and pHNF* plasmids and the pWT and pHNF plasmids presented as induction (± SEM, n=6) and the statistical analysis is described in chapter 2 (*: p<0.05).

B: Summary of the $I_{\text{max}}$ and EC$_{50}$ values generated by the regression analysis presented with the 95% confidence interval for the mutated plasmid (MUT: pHNF* or pHNF) against the value for the respective wild type control (WT: pWT* or pWT).
Figure 5.10. Rifampicin response of the C/EBP mutant constructs upon transfection in HuH7 cells.

A: Induction profile for the pWT* and pC/EBP* plasmids and the pWT and pC/EBP plasmids presented as induction (± SEM, n=6) and the statistical analysis is described in chapter 2 (*: p<0.05, **: p<0.001).

B: Summary of the Imax and EC50 values generated by the regression analysis presented with the 95% confidence interval for the mutated plasmid (MUT: pC/EBP* or pC/EBP) against the value for the respective wild type control (WT: pWT* or pWT).

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Imax</th>
<th>EC50 (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pC/EBP*</td>
<td>WT: 75.94 (63.91 to 87.97)</td>
<td>WT: 0.47 (0.13 to 0.81)</td>
</tr>
<tr>
<td></td>
<td>MUT: 25.84 (21.07 to 30.61)</td>
<td>MUT: 0.41 (0.05 to 0.78)</td>
</tr>
<tr>
<td>pC/EBP</td>
<td>WT: 29.79 (26.03 to 35.55)</td>
<td>WT: 0.40 (0.15 to .64)</td>
</tr>
<tr>
<td></td>
<td>MUT: 11.30 (10.04 to 12.55)</td>
<td>MUT: 0.27 (0.09 to 0.45)</td>
</tr>
</tbody>
</table>
Figure 5.11. Summary of the transfection experiments for the Sp1* (A) and pSp1 (B) plasmids.

A: pSp1* plasmid.

![Graph showing Imax and 95% CI](image1)
![Graph showing EC50 and 95% CI](image2)

The EC₅₀ (μM) and Imax values, and their 95% confidence interval, were calculated using a non-linear regression model described in section 2.12.6. The experiments were carried out in triplicate and were reproducible.
The EC_{50} (µM) and Imax values, and their 95% confidence interval, were calculated using a non-linear regression model described in section 2.12.6. The experiments were carried out four times and were reproducible.
5.4. Discussion: Glucocorticoid-mediated induction of \textit{CYP3A4} is decreased by disruption of a protein:DNA interaction distinct from the ER6.

The data from the reporter gene assay (figure 5.1) and gel shift assay (figure 5.3) demonstrate that a single base pair mutation can alter the response of the \(-301/+7\) promoter fragment to hydrocortisone and dexamethasone (thought to act as GR ligands) but not to rifampicin (thought to be mainly a SXR ligand) and that this is probably due to the mutation decreasing the binding of a protein present in HepG2 nuclear extracts. DNase I footprinting experiments (chapter 3) suggests the factor involved may bind to a putative HNF3 or C/EBP\(\alpha\) response element. This work therefore suggests that glucocorticoids may act at least partly through a SXR independent pathway and that the ER6 may not be the only region of the promoter involved in the xenobiotic response of the \textit{CYP3A4} promoter.

Early reports demonstrated that an alternative mechanism to the classical GR-mediated gene regulation (e.g. as for the tyrosine aminotransferase gene) may be operative for dexamethasone induction of CYP3A genes (Schuetz and Guzelian, 1984). Those results suggested the presence of an alternative receptor. Glucocorticoid mediated induction of CYP3A genes remained unexplained for several years until the discovery of SXR (Bertilsson et al., 1998; Blumberg et al., 1998; Kliewer et al., 1998) which has the ability to bind classical CYP3A inducers including glucocorticoids. However, in co-transfection experiments, GR addition was still found to have a positive effect on drug response in spite of the absence of any consensus GR binding sites in the proximal promoter region of \textit{CYP3A4} (Ogg et al., 1999). A more probable scenario to the direct binding of GR was then suggested by experimental data on the regulation of the level of the xenobiotic nuclear receptors by glucocorticoid. It was found that glucocorticoid treatment of human hepatocytes transcriptionally increases the levels of SXR (and CAR) gene expression at the mRNA and protein levels as well as the levels of the heterodimerization partner RXR\(\alpha\) (Pascussi et al., 2000a; Pascussi et al., 2000b). The same group went on to demonstrate the relevance of this observation by showing that \textit{CYP3A4} induction in human hepatocytes is enhanced by glucocorticoids at low concentrations by increasing
the levels of nuclear receptors (SXR, CAR, RXRα) and at higher concentrations by activating SXR (Pascussi et al., 2001). Therefore this work demonstrates a further level of complexity than proposed by earlier models by showing that xenobiotic may regulate the levels of nuclear receptors prior to activating them through binding. This suggests an alternative pathway and it provides some insight into possible mechanisms of CYP3A4 gene regulation by glucocorticoids.

The fact that the mutation affects putative binding sites for HNF3 and/or C/EBPα is not inconsistent with earlier findings that these factors may be involved in interactions with GR. If no direct binding of the GR to the CYP3A4 proximal promoter takes place the experimental results reported here could be explained by protein-protein interactions as GR has been shown to interact with members of the C/EBP family to enhance transcription without directly binding to the DNA and instead acting via the Activation Function 2 (AF2) of both receptors (Boruk et al., 1998). Alternatively the effect of the mutation could be due to a modification of the levels of HNF3 or C/EBPα by glucocorticoids in the same manner as RXR, SXR and CAR are upregulated by glucocorticoids in human hepatocytes (Pascussi et al., 2000a; Pascussi et al., 2000b; Rodrigues et al., 2003). Members of the HNF3 family (α and β) have been shown to be up-regulated by dexamethasone treatment in rat (Imae et al., 2000).

The nuclear protein(s) involved in the phenomenon described above have not yet been identified and this would be a logical next step. Initial attempts to produce recombinant C/EBPα and HNF3 or carry out supershift assays were unsuccessful and could not be completed due to time constraints.

5.5 Discussion: Mutagenesis of putative regulatory protein binding sites.

5.5.1 Reporter gene assay optimization and technical considerations.
Numerous attempts to carry out functional analysis of the CYP3A4 promoter using transfection assay as previously described by others in our laboratory (e.g. El Sankary et al., 2000; Ogg et al., 1999) were unsuccessful in this study. The method employed here
differs from previous methods used in our laboratory in a number of ways. First the cell line used (HuH7) was different to the previous hepatoma cell line used (HepG2). This was because the cells were found to be easier to maintain in culture (personal observation) and were readily transfected using Fugene 6 (instead of the previously used calcium phosphate method) at a sufficient level to allow for reproducible measurements of reporter gene levels. Also the reporter gene construct was different to the one previously developed and used in our laboratory (Ogg et al., 1999). The main advantage of the new construct engineered by Dr Hossein Hamzeiy (Hamzeiy, 2002) is that it included the native distal enhancer element XREM (Goodwin et al., 1999) suggesting that within the experimental set-up reported here this enhancer region may be critical for the induction process. Indeed in preliminary experiments, constructs containing the -301/+7 promoter region in the presence and absence of the XREM were found to produce detectable levels of reporter gene expression but plasmids with an XREM insert provided a stronger response to CYP3A4 inducers than those without (data not shown). It is worth mentioning that there appears to be a discrepancy in terms of the rifampicin concentrations necessary to achieve full induction in various experimental systems. In the present study a significant level of reporter gene activation can be achieved at a concentration as low as 0.5 µM and the increase peaks for rifampicin concentrations between 5 to 10 µM, which is not inconsistent with findings in HepG2 cells using a similar construct (Goodwin et al., 1999). However, using a construct with a different enhancer to the XREM and the -301/+7 fragment of CYP3A4 the concentration needed to reach maximum induction appeared to be higher than 100 µM (El Sankary et al., 2002) or higher than 20 µM with a similar construct containing the -1201/-61 region of promoter (Hamzeiy, 2002). The causes for this variability are unknown and are probably due to variations in the constructs (enhancing effect of the XREM), cell line sub-clone populations (which may slightly vary in the transcription factor content) or individual experimenter variation. The comparison of the pWT and pWT* induction profiles presented in figure 5.5 demonstrates that the mutation (T>C, -193bp) has no effect on the rifampicin activation of the -301/+7 fragment of CYP3A4. It is possible, however, that the response to other inducers may be altered as it was the case for the glucocorticoid induction of the pMUT plasmids (see above section 5.4).
Finally it should be noted that during preliminary experiments (not shown) the induction of the construct was only achieved in the presence of co-transfected SXR expression plasmid. This is consistent with the fact that the levels of SXR and several other nuclear receptors are present in HuH7 cells at lower levels than in human liver (Anna Phillips, University of Surrey, personal communication).

5.5.2. Comparison of the $-301/+7$ fragment versus $-1201/-61$.

The data from this experiment show a difference in the reporter gene expression between the 2 constructs suggesting the presence of factors within HuH7 cells acting as $CYP3A4$ repressors between $-1201/-301$ and/or acting as activators between $-61/+7$. The ratio of reporter gene production for rifampicin dosed cells against the respective DMSO solvent control are similar between the two constructs, suggests that in the $-1201/-301$ and/or the $-61/+7$ fragment additional factors interact to alter the basal level of $CYP3A$ expression, but that these do not interfere with the rifampicin mediated $CYP3A$ expression. It could be that because HuH7 cells exhibit a foetal phenotype with $CYP3A7$ expression but no $CYP3A4$ expression (Saito et al., 2001) the regulatory regions of the $CYP3A4$ gene may interact with endogenous transcription factors repressing the promoter activity. The DNase I footprinting data presented in chapter 3 suggest that besides the TATA box there may be at least 2 sites where proteins (or protein complexes) may interact with the $-61/+7$ region of the $CYP3A4$ promoter. Unfortunately the nature of the possible proteins binding at those sites could not be determined experimentally. There are, however, several possible candidate proteins including the nuclear receptor COUP-TF which is known to act as a transcriptional repressor and is also known to be present at high level in foetal tissues and cell lines (Qiu et al., 1994).

5.5.3. Effect of the mutations on the basal expression of reporter gene constructs in HuH7 cells.

In this experiment the level of reporter gene expression was measured by transfecting the wild type and mutated plasmids in HuH7 cells without adding any exogenous SXR ligands. However in this context it is probable that the co-transfected SXR may have
been activated by endogenous cellular metabolites or steroids potentially present in the cell culture media. Indeed because of a large ligand-binding domain (Watkins et al., 2001) the promiscuous PXR ligand binding domain can interact with a very wide variety of structurally unrelated compounds (Moore et al., 2000). Therefore basal state will be defined here as the absence xenobiotic inducers and in the presence of endogenous activators. The results presented in figure 5.7 suggest that factors binding to the ER6 are involved in the basal expression of the CYP3A4 gene, within the experimental system used in this study. A similar effect has been previously reported, together with data demonstrating the importance of the other nuclear receptor binding sites located within the XREM region (Goodwin et al., 1999). This is in contrast with results from experiments using SXR knock out mice which failed to demonstrate a down regulation of cyp3a11 gene in the untreated animals (Staudinger et al., 2001; Xie et al., 2000b). This is probably because in knock out mice the product of the missing SXR gene can be replaced by CAR binding to the intact cyp3a11 regulatory regions (Smirlis et al., 2001; Xie et al., 2000a) or because SXR in the in vivo basal state may act as a transcriptional repressor due to its interaction with co-repressors (discussed in section 3.3.3). This result highlights the very important role of the ER6 as a cis-acting element. The fact that the HNF and C/EBP SDM constructs exhibited a lower reporter gene activity in non-drug treated cells suggests that the proteins binding to these sites are involved in the basal expression of the CYP3A4 gene. This is in agreement with the fact that proteins from both the HNF and C/EBP families are mainly expressed in the liver (Xanthopoulos et al., 1991) and are involved in the tissue specific expression of several hepatic genes (Diehl, 1998; Hayashi et al., 1999; Lekstrom-Himes and Xanthopoulos, 1998). Hence it would not be surprising if these factors played a role in the tissue specific basal expression of CYP3A4 in the liver. The pSp1 and pSp1* plasmids, however, did not display any change in the level of SEAP production suggesting a lack of effect of these mutations on the basal promoter activity. This is in contradiction with the fact the Sp1 is involved in the basal regulation of many genes (Suske, 1999) and that it was predicted to play a similar role herein. However, as will be discussed below, the responses of the Sp1 SDM constructs are variable suggesting a complex role.
5.5.4. Effect of the mutations in rifampicin treated HuH7 cells.

The results generated by the reporter gene assay study following the mutagenesis of regions from the -301/+7 fragment of CYP3A4 proximal promoter indicate that some of the mutated regions are located within important cis-acting elements with regards to the rifampicin mediated induction of CYP3A4 gene expression.

As expected the pER6* plasmid which contains two base pair mutations (AA>TT, -166bp) within the ER6 region exhibited a lower response to rifampicin treatment than its wild type (pWT*) counterpart. This experiment acted as a positive control to assert that mutations to important regulatory elements could affect the transcriptional activation of the constructs generated within the experimental set-up presented here. Data from the literature suggest that if only one of the ER6 half sites is mutated there is still possibility for in vitro synthesised SXR:RXRα protein heterodimers to recognise the DNA sequence (Lehmann et al., 1998). Because it could not be determined whether transcription factor binding to this ER6 site was completely abolished, this experiment could not prove that the activation of a CYP3A4 reporter gene construct could be achieved without binding of proteins (SXR-RXR heterodimers in this case) to this region.

The responses of the pMUT, pWT*, pHNF and pHNF* plasmids to rifampicin raises interesting questions about the mechanisms of CYP3A4 gene regulation. The sequence variations of these plasmids are summarized in figure 5.12. Previous work demonstrated that the pMUT construct responds to rifampicin in the same way as a plasmid containing the wild type promoter but that the response to GR ligands is altered (figure 5.1). The similar responses to rifampicin observed in the pHNF and pWT* constructs compared to that of the pWT plasmid were therefore consistent with the published data (El Sankary et al., 2002). However, the pHNF* plasmid which contains 3 base pair mutations responded in a different manner as rifampicin induction was decreased.
Figure 5.12. Representation of the locations where bases are modified in the pMUT (El Sankary et al., 2002) (A), pHNF (B), WT* (C) and pHNF* (D) plasmids.
This could be because the cumulative effect of both mutations affects the protein-binding site in a more pronounced manner (e.g., the putative HNF3 binding site) or because the combination of 2 mutations (pWT* and pHNF) affects two protein-binding sites which disruption results in a loss of SXR mediated response. Testing the pHNF* plasmid with other xenobiotics (e.g., glucocorticoids) will shed light on this phenomenon. This work is currently undertaken in our laboratory.

The inconsistency of the responses seen for the pSp1 and pSp1* plasmids as well as the high intra-experiment variability are unlikely to be due to problems with the plasmid preparations since they were re-extracted from bacterial cultures. Additionally, problems with the plasmid transfection efficiency would have been noticed from unusual levels of SEAP production in the cell media prior to xenobiotic dosing (measured at ‘day 1’) and this did not appear to be the case.

The results for the Sp1 mutant plasmids were not significant. It was expected the activation profile would be disrupted since Sp1 is known to be involved as a repressor in the regulation of other CYP genes (discussed in section 4.3.4, Kaczynski et al., 2002; Muangmoonchai et al., 2001). The negative result could be due to the fact that the nuclear protein binding to this region is not significantly involved in the regulation of the CYP3A4 gene or because the mutation was not sufficient to disrupt the DNA:protein interaction. In order to test this hypothesis, further disruption of the putative protein binding site by SDM would be necessary.

The most significant result was the clear reduction of the I_{max} observed in plasmids where a putative C/EBPα binding site had been mutated (pCEBP and pCEBP*). As discussed in chapter 3, previous reports demonstrate that the -169/+11 region of the CYP3A4 promoter as well as the -350/-311 and -628/-608 of the CYP3A1 promoter can be up regulated by co-transfection of a C/EBPα expression plasmid (Ourlin et al., 1997; Rodrigues et al., 2003). In a functional study of the CYP3A5 proximal promoter it was reported that a member of the Nuclear Factor-Y (NF-Y) family bound to a CAAT box (Iwano et al., 2001). Therefore care should be taken in interpreting the computer analysis of the -301/+7 with the Transfac database. Indeed although data from the literature strongly suggests that the CAAT box binding factor is C/EBPα this has not been
demonstrated in this report and further work is required in order to determine the exact nature of the factor binding at this location.

In this attempt to identify factors involved in the regulation of the CYP3A4 gene the exact nature of the mechanisms involved in the altered activation of the constructs caused by the mutations remains to be tested. Indeed it is possible that altered binding of the factors described above may cause the effect seen without affecting the binding to the ER6 but it could also be that the changes seen are due to changes in the binding at the ER6 site caused by modifications of the protein-protein interactions within the promoter fragment. The former hypothesis is consistent with the interactions described in the mechanisms of CYP3A5 gene regulation (Iwano et al., 2001) where members of the Sp and NF-Y families act synergistically to activate reporter gene constructs. The latter hypothesis would fit with the model of interactions described in a study of the PECK gene promoter (Stafford et al., 2001) where binding of the GR is enhanced by the presence of 'accessory' transcription factors (HNF3 and COUP-TF) binding in the flanking region of the GRE.

Due to lack of time it was not possible to further identify the factors binding to the mutated sites and attempts to generate in vitro transcribed proteins were unsuccessful. In spite of this uncertainty it should be noted that any effect of the mutations on the reporter gene construct basal activity and activation strongly suggests that in this proximal promoter context and in this HuH7 cellular environment the affected region is involved in the transcriptional activity.

Also due to time limitation a study of the effect of the mutations on HepG2 and HuH7 nuclear protein binding could not be optimised. The data suggested binding to the response elements thought to be involved, however, quantification of differential affinity of the proteins could not be made. Therefore there is theoretically a possibility that the alterations generated by mutagenesis may have enhanced nuclear protein binding or may have created novel protein binding sites.
5.6. Conclusion.

The work described in this section confirms some of the hypotheses generated by the DNase I footprinting mapping experiment described in chapter 3. Using site directed mutagenesis and reporter gene assays the role of $cis$-acting elements other than the ER6, including a putative HNF3 and C/EBP$\alpha$, has been demonstrated for the first time. The exact nature of the mechanisms and the identity of the transcription factors involved remain to be investigated. These findings demonstrate that the molecular mechanisms driving the expression of the $CYP3A4$ gene are more complex than initially thought, as they are likely to results from the integration of xenobiotic-, hormone- and tissue-specific signals.
CHAPTER 6 – DISCUSSION

6.1. Introduction.

Following years of investigation into the molecular mechanisms of the CYP3A4 gene regulation, the cloning of the SXR provided answers to several enigmas such as the wide range of compounds capable of inducing this important drug-metabolizing enzyme or the fact that several of these inducers had apparent contradictory therapeutic effects (i.e. glucocorticoids and antiglucocorticoids) (Bertilsson et al., 1998; Blumberg et al., 1998; Kliewer et al., 1998). The discovery of this receptor also raised questions and for instance there was a need to investigate the role of other transcription factors in CYP3A4 basal expression and activation. Indeed the interaction with other regulators is likely to play a role in modulating this response, since SXR is a good xenobiotic sensor (Jones et al., 2000) but its very promiscuity causes problems with the fine tuning of the target gene response. The present study was aimed at investigating the transcriptional regulation of the CYP3A4 gene in the light of these recent advances, focusing on the proximal promoter region. Although, due to technical problems and time restrictions, the identity of the regulatory factors contributing to the induction process is not fully understood, a DNase I footprinting mapping study coupled with functional assays of transcription strongly suggest that factors others than SXR acting at the ER6 can influence the basal level of expression as well as the degree of induction of an heterologous reporter gene construct by rifampicin. The cis-acting elements involved are putative binding sites for members of the HNF and C/EBP families. Additionally a study of the effect of naturally occurring functional mutations on the binding of nuclear proteins to DNA oligomers suggested that further regions are likely to be involved in the hepatic regulation of the CYP3A4 gene and that the most likely transcription factor candidate is Sp1.

This chapter describes how these findings may have potential implications on the current understanding of the hepatic CYP3A4 gene regulation by pathophysiological and environmental signals other than xenobiotics and on the understanding of inter-individual
variability in CYP3A4 expression and drug-activation. This section will be concluded by a discussion on how recent discoveries in the field of eukaryotic gene regulation could be applied to the study of this important gene.

6.2. Liver specific transcription factors and CYP3A4 gene expression.

Although this work does not provide unequivocal evidence for the identity of specific factors, it demonstrates that regions in the CYP3A4 proximal promoter can bind transcription factor(s) present in hepatic cell lines and that alterations in their binding sites can impact on the basal and rifampicin induced expression of this gene. Together with recent evidence published in the literature (e.g. Huss and Kasper, 2000; Jover et al., 2001; Ourlin et al., 1997; Rodrigues et al., 2003) this work suggests that members of the C/EBP and HNF family possibly contribute to CYP3A4 gene expression.

A study of the rat CYP3A23 gene promoter indicates that a functional HNF4 binding site is located between bases –85 to –110 (from the transcription start site) (Huss and Kasper, 1998). Although this site is not necessary for the expression and induction of CYP3A23, the presence of the HNF4 liver specific transcription factor is required in order to achieve full induction of the gene (Huss and Kasper, 1998; Huss and Kasper, 2000). In a human hepatocyte model, adenovirus mediated transfection of the HNF4 gene demonstrated the role of this factor in the maintenance of CYP3A4 gene expression (Jover et al., 2001).

Two studies point towards a role of C/EBPα in the regulation of the CYP3A4 and CYP3A1 genes. Using co-transfection experiments, the expression of reporter genes was increased by C/EBPα expression plasmids (Ourlin et al., 1997; Rodrigues et al., 2003). Additionally, this transcription factor was found to interact in vitro with CYP3A1 regulatory sequences (Rodrigues et al., 2003).

The transfection assay data reported in chapter 5 unequivocally demonstrate that a transcription factor interacting with a binding site for C/EBPα in the CYP3A4 proximal promoter impacts on both the basal and rifampicin induced expression of the gene. Additionally an HNF3 response element is also involved in at least the basal expression of an heterologous –301/+7 construct of the CYP3A4 proximal promoter.
This section highlights that besides their involvement in sustaining the tissue-specific expression of the \textit{CYP3A4} gene, liver specific transcription factors are probably involved in mediating gene expression in response to pathological states and might be important variables to explain the polymorphic liver specific expression of the \textit{CYP3A4} gene.

\textbf{6.2.1. Liver specific transcription factors and inflammation.}

Several reports describe a change in P450 levels or activity following infection or inflammatory reaction (reviewed in Morgan, 1997). For instance, \textit{CYP3A4} expression has been reported to be decreased by interleukin 6 (IL6) exposure in human hepatocytes (Muntane-Relat \textit{et al.}, 1995). This phenomenon could be due, at least partly, to an altered expression of CAR and PXR, since IL6 has been reported to cause a decrease in the level of the nuclear receptors in human hepatocytes (Pascussi \textit{et al.}, 2000c). This observation was also made in mice injected with bacterial lipopolysaccharides, responsible for inducing acute inflammation, where \textit{Cyp2b10} and \textit{Cyp3a} mRNA levels were rapidly decreased in parallel with the mRNA levels of CAR, PXR and RXR\textalpha{} (Beigneux \textit{et al.}, 2002). This would suggest a mechanistic link between pathological states and inflammation-induced reduction of P450's. There is, however, conflicting evidence supporting which liver specific transcription factors may transduce the effect of the signaling molecules produced during inflammation or infection and result in decreased \textit{CYP} gene expression. As discussed earlier in section 3.3.4.2 there is a possibility that the effects of nitric oxide (NO) may be mediated by HNF4. Also, a member of the C/EBP family (C/EBP\textbeta{}) is known to be involved in the response to IL6 (Lekstrom-Himes and Xanthopoulos, 1998). Therefore it cannot be excluded that members of the C/EBP family can directly alter \textit{CYP3A4} expression by acting at the promoter level. Indeed it has been reported that IL 6 causes an increase in the levels of a low molecular form of C/EBP\textbeta{} (C/EBP\textbeta{}-LIP), which reduces the activating effect of C/EBP\textalpha{} and has a repressive effect on \textit{CYP3A4} expression levels in human hepatocytes (Jover \textit{et al.}, 2002).

It could be that both mechanisms are involved in the response to inflammation. However, in order to reconcile the two models it may be hypothesized that liver specific
transcription factors mediate the effect of inflammation mediators and then impact the level of the liver expressed CAR, SXR and RXRa, which in turn have an effect on the levels of the target gene CYP3A4. This hypothesis remains to be tested.

In conclusion this work provides a possible mechanistic link between certain pathologies and their effect on CYP3A4 gene expression. Knowledge of the involvement of liver specific transcription factors in CYP3A4 regulation may provide valuable mechanistic clues to explain these observations and help design treatments and compounds with a view to avoid drug-drug interactions and adjust doses of therapeutic drugs as the metabolism of CYP3A4 substrates and/or inducers may be altered in patients suffering from infection or inflammatory reactions.

6.2.2. Liver specific transcription factors and inter-individual variability in drug metabolism.

The mutation screening of 11 human liver samples with a 14.1-fold variability in CYP3A activity (phenotype) revealed that the first kilobase of proximal promoter contained no mutations (as discussed in chapter 4). Although the study of a larger population sample (101 subjects) revealed that 28 individuals contained mutations in the CYP3A4 regulatory region (Hamzeiy et al., 2002), it is possible that the inter-individual variability in drug metabolism displayed in the population may also be due to variability in the levels of regulatory proteins or epigenetic features such as DNA methylation (see section 4.3.3 and below). In addition to qualitative and quantitative variations in the factors binding at the ER6 (SXR and CAR), there is evidence that the expression of liver specific transcription factors may vary because of genetic or environmental factors.

6.2.2.1. Mutations in liver specific transcription factors.

Some hepatic transcription factors are known to contain mutations in their exon, intron and promoter regions. For instance a total of 65 mutations have been reported for the HNF1α coding region, as well as 6 promoter variants (Ellard, 2000). Some of these mutations are thought to be associated with the occurrence of certain forms of diabetes,
the maturity onset diabetes of the young (MODY). Another gene which mutations are linked to MODY is that coding for HNF4α (Schrem et al., 2002), where various mutations have been discovered that were found to disrupt various domains of the nuclear receptor. In addition to members of the HNF family, there is a report of mutations in the gene coding for C/EBPα (Gombart et al., 2002). However, because these 12 mutations were detected in tumour tissue, it is not clear whether these may be present in healthy tissue or within a healthy population.

Although no work so far has revealed an association between mutations in liver specific transcription factors and CYP gene expression, it can be hypothesized that genetic variability in these hepatic regulators may result in variability in the expression level of their target genes (namely CYP3A4).

6.2.2.2. Variability in the parameters which regulate liver specific transcription factor expression.

Besides genetic alterations which may affect expression and protein function between individuals, endocrine and environmental factors can possibly alter the expression levels of liver specific transcription factors. For instance in a study of human liver samples taken by biopsy, the mRNA and protein levels of C/EBPα and β were shown to vary between patients (which suggests polymorphic expression or the influence of environmental factors) and their levels were found to be greatly affected by plating in primary cultures suggesting an influence of stress and cell proliferation signals in regulating the levels of expression (Ferrini et al., 2001). In a study of HNF3 family member mRNA expression in rat liver (Imae et al., 2000) it was reported that a low protein diet increased HNF3γ, dexamethasone treatment increased HNF3α and β and that diabetic animals exhibited an increased HNF3β and γ expression which was reversed upon insulin injection. An enhancer element in the HNF4α regulatory region allows transcriptional regulation of the gene by glucocorticoids (Bailly et al., 2001). C/EBPα has also been shown to be affected by synthetic glucocorticoids, since it is up regulated following dexamethasone treatment in rat hepatocytes (Rodrigues et al., 2003).
together these results suggest that the hormonal status may alter the levels of liver transcription factors thereby possibly affecting the CYP3A4-mediated drug metabolism. A recent study suggests that xenobiotics may also alter hepatic transcription factor content. Treatment of rat primary hepatocytes with the polychlorinated biphenyl (PCB) mixture Aroclor 1254 resulted in an increase in mRNA for C/EBPα and members of the HNF1 and 4 families (Borlak et al., 2002). These changes had an effect on the secretion of albumin, which was raised concomitantly and which is a known target for these transcriptional regulators. Additionally, because the transcriptional regulation of some of these factors is influenced by synthetic steroids, it is possible that inter-individual variations in steroid metabolism and in the levels of plasma steroids (which in turn may be linked to the environment and ageing) may impact upon transcription factor levels and the expression of downstream target genes. There is substantial evidence that HNF4α is involved in the regulation of several other liver specific transcription factors (Hayashi et al., 1999; Schrems et al., 2002), as well in the regulation of other nuclear receptors such as SXR (Li et al., 2000). Therefore modifications in the regulatory networks formed by these factors could possibly alter the CYP3A4 phenotype.

In conclusion, since CYP3A4 expression is likely to be driven partly by liver specific transcription factors (C/EBPα and HNF family members), variability in the functionality or expression of these regulators (through genetic or environmental factors) is likely to result in variability in the expression of the CYP3A4 gene.
6.3. Role of epigenetics on gene regulation.

In cell nuclei the DNA is associated with histone proteins and forms chromatin fibers. These fibers are formed by nucleosome units which consist of 146bp of DNA wrapped around an octamer formed by 2 of each core histone proteins (H2A, H2B, H3, H4) (Zhang and Reinberg, 2001). This chromatin template undergoes structural modifications associated with various processes such as cell division, DNA replication and repair as well as gene regulation (Zhang and Reinberg, 2001). The core histone proteins consist of a globular C-terminal around which the DNA is wrapped and a N-terminal tail protruding outward from the nucleosome unit. The histone tails may undergo covalent post-transcriptional modifications (acetylation, methylation, ubiquitination, phosphorylation) by specific enzymes, which are part of multi-protein complexes recruited to various regions of the genome following intra- and extra-cellular signals (Jenuwein and Allis, 2001). It has been suggested that these modifications form an histone code (Jenuwein and Allis, 2001; Turner, 2002) and depending on the nature and the position of the group added, these histone tails establish which area of the chromatin are transcriptionally active (euchromatin) or transcriptionally repressed (heterochromatin). For instance acetylation of the lysine 9 residue and methylation of the arginine 3 residue on the tail of histone 4 are associated with gene activation, whereas methylation of lysine 9 on the H3 tail is associated with transcriptional repression (Berger, 2002).

Histone modifying enzymes are part of multiprotein complexes which are specifically recruited to promoters prior to transcription initiation. For instance the histone methyltransferase PRMT1 has been shown to physically interact with the transcription factor Ying Yang 1 (YY1), as part of a complex which binds to a YY1 activated promoter, and methyle the arginine 4 residue on the histone H4 tail (Rezai-Zadeh et al., 2003). This modification resulted in an increased transcription of the target gene. As mentioned in the introduction chapter (section 1.4.2), nuclear receptors are capable of interactions with other proteins (coregulators), bringing an additional level of complexity to the regulation of target genes (McKenna and O'Malley, 2002). SXR is known to recruit SRC-1 (Lehmann et al., 1998) and SMRT (Takeshita et al., 2002), which act respectively as coactivator and corepressor. Some of these coregulator proteins (e.g. SRC-1) possess
the ability to acetylate histones (Jenster, 1998) thereby modifying the chromatin structure to make it more accessible for the transcription machinery. Attempts to stimulate the genomic transcription of CYP3A4 in HuH7 and HepG2 cells (which express CYP3A4 at very low levels), using over-expression of SXR in order to overcome the low expression of the regulatory protein, have been unsuccessful (Anna Philips and Nick Plant, personal communication). This observation lends support to the hypothesis that chromatin structure may be an important factor in CYP gene regulation. There is clearly a scope for further work on the histone modifying enzymes involved in CYP3A gene regulation in vivo. For instance since post-translational modifications of histone tails may have a role in cellular memory processes and in the long-term activation or repression of genomic regions (Jenuwein and Allis, 2001; Turner, 2002), it is possible that such mechanisms may explain the variations in drug metabolism enzymes observed during development (Lacroix et al., 1997). The study of the histone modifications and nucleosomal structure at the CYP3A locus may provide insight into the ontogeny of CYP3A gene expression (CYP3A4 Vs CYP3A7) as well as possible phenotype alterations in CYP genes occurring during disease and xenobiotic exposure.

6.4. Conclusion.

This work demonstrates that the ER6 is not the only cis-acting element involved in the transcriptional activation of the CYP3A4 gene. Mutations affecting putative protein binding sites which are response elements for a variety of factors can alter the binding of nuclear proteins and/or the in vitro response of the proximal promoter in reporter gene assays. This finding demonstrates that the molecular mechanisms of the CYP3A4 gene expression are far more elaborate than originally anticipated, since there are several layers of complexity in the regulation process. The diagram in figure 6.1 presents a model that summarises the factors involved in the regulation of this important hepatic drug metabolism enzyme, in the light of the results presented in this thesis and in the recent literature. It appears that questions relating to the transcriptional regulation of the CYP3A4 gene in the human liver are in fact addressing wider issues such as the regulatory networks formed by xenobiotic and tissue specific transcription factors.
Recent advances in genetic and molecular biology techniques (e.g. transgenic technology and DNA microarrays) will undoubtedly contribute towards a more complete understanding of this system and while allowing scientists to uncover fascinating biological mechanisms, they will eventually allow the design of safer drugs and therapy designs.

Figure 6.1. Summary of the mechanisms involved in the hepatic regulation of the \textit{CYP3A4} gene. The xenobiotic and liver specific transcriptions factors regulate the expression of the \textit{CYP3A4} gene. Both the \textit{CYP3A4} genes and its regulators are subject to the effect of epigenetic mechanisms. The various gene regulation events and protagonists are influenced by a variety of exogenous factors.

\textbf{Transcription Regulators:}
- liver specific (HNF and C/EBP families) and xenobiotic transcription factors (CAR and SXR)

\textbf{Epigenetics:}
- chromatin structure and DNA methylation

\textbf{Environment:}
- Diet, Drugs, Hormones, Pathology

Form a network of inter-regulated genes

\textit{CYP3A4} gene
6.5. Future directions.

Following this study of the transcriptional mechanisms of the CYP3A4 gene regulation, it would be interesting to identify the transcription factors binding to the *cis*-acting elements identified here. This could be done by a number of complementary approaches including reporter gene assay experiments with co-transfection of expression plasmids for factors thought to be interacting with the regulatory region as well as DNA:protein interaction studies using recombinant proteins and/or antibodies specific to the protein bound to the labeled DNA oligomer.

Once the involvement of those transcription factors has been established monitoring their levels (as well as that of any other factor known to be involved in the regulation of CYP genes or xenobiotic nuclear receptors) under various experimental treatments as well as in liver tissue samples from several individuals would provide insight into the mechanisms of drug-drug interaction and inter-individual variability in drug metabolism. This could be done alongside the measurement of several other target genes for instance as part of a gene array experiment (RNA levels) or a proteomics experiment (measurement of global changes in protein levels).

Because there is evidence from this work and the literature that SXR may bind unliganded to the CYP3A4 promoter and repress transcription and also interacts with a variety of coactivators, it would be interesting to investigate the *in vivo* associations formed between the nuclear receptor and other nuclear proteins. This could be done by analyzing the proteins present in complexes found in liver extracts which would have been immunoprecipitated.

Finally, as discussed above, the study of the ‘histone code’ in place at the nucleosomal structure of the CYP3A locus may provide insight into the ontogenic, basal and xenobiotic induced regulation of the CYP3A4 gene.
7. BIBLIOGRAPHY.


Beigneux, A. P., Moser, A. H., Shigenaga, J. K., Grunfeld, C., and Feingold, K. R. (2002). Reduction in cytochrome P-450 enzyme expression is associated with repression of CAR (constitutive androstane receptor) and PXR (pregnane X receptor) in mouse liver
during the acute phase response. Biochemical and Biophysical Research Communications 293, 145-149.


regulate overlapping but distinct sets of genes involved in xenobiotic detoxification. Molecular Pharmacology 62, 638-646.


hepatocytes: Consequences on cytochrome P450 gene regulation. Molecular Pharmacology 58, 1441-1450.


genetic polymorphism in the 5'-upstream regulatory region. Biochemical and Biophysical Research Communications 259, 201-205.


