NANO-SCALE LIQUID CHROMATOGRAPHY ELECTROSPRAY TANDEM MASS SPECTROMETRIC IDENTIFICATION OF MULTIPLE CYTOCHROME P450 ISOFORMS IN TISSUES

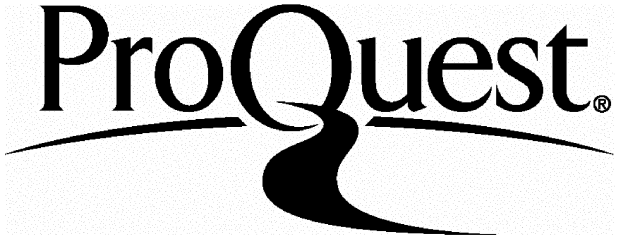
By

Sonia Nisar

A Thesis Submitted for the degree of Doctor of Philosophy of the University of London

Department of Pharmaceutical and Biological Chemistry
The School of Pharmacy
University of London
29-39 Brunswick Square
London
WC1N 1AX

February 2005



Abstract

The cytochromes P450 (CYPs) are membrane bound proteins that collectively carry out a wide range of biological oxidations important in the metabolism of steroids, ecosanoids and xenobiotics including anticancer agents. A method of CYP separation and analysis has been developed using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and nano-scale liquid chromatography-electrospray-tandem mass spectrometry (LC-ESI-MS/MS) for the direct identification of multiple CYPs found in rat liver, human tissues and tumours. Endoplasmic reticulum from selected tissues was prepared as microsomes using ultracentrifugation and subjected to SDS-PAGE. The proteins present in the 48-62 kDa gel band region were excised, cut into small pieces and digested with trypsin. The resultant peptides were solvent extracted, separated and analysed by LC-ESI-MS/MS. CYP identifications were made by searching the MS/MS data using Sequest software against a protein database (obtained from National Centre for Biotechnology Information). Twenty-four CYP isoforms from the sub-families 1A, 2A, 2B, 2C, 2D, 2E, 3A, 4A, 4F, CYP17 and CYP19 were positively identified in rat liver, fourteen isoforms of which were shown to be gender biased consistent with previous studies on sexual dimorphism or gender predominance. Fifteen CYP isoforms namely, 1A2, 2A6, 2B6, 2C8, 2C9, 2C17, 2C18, 2C19, 2D6, 2E1, 3A4, 3A5, 4A11, 4F1 and 4F3 were positively identified to be present in normal human liver. CYPs were also identified from paired samples of human liver: 1A2, 2A6, 2B6, 2C8, 2C9, 2C17, 2C19, 2D6, 2E1, 3A4, 4A11, 4F2 and 4F8 and colon metastasis in the liver: 1A2, 2A6, 2C8, 2C9, 2D6, 2E1, 3A4, 4A11, 4F2 and 4F8. In normal colon and colon primary tumour only the CYP3A subfamily was identified. These results confirm the use of an MS based proteomic approach to the identification of multiple CYPs that obviates the need to predetermine which CYPs are present. They also suggest that the liver tissue environment is significantly increasing the expression profile of CYPs in colon tumours that metastasise to the liver.

Contents

Chapter 2: Materials and methods

Chapter 3: Results

Chapter 4: Discussion

Appendices

List of Figures

List of Tables

List of Abbreviations

Full name	Abbreviations
Absolute quantification	AQUA
Acetonitrile	ACN
Adenosine triphosphate	ATP
Ammonium persulphate	APS
Anthraquinone-di-N-oxide	AQ4N
Aryl hydrocarbon receptor	AHR
Aryl hydrocarbon receptor nuclear translocator	ARNT
Atmospheric pressure ionization	API
Bovine serum albumin	BSA
Collision induced dissociation energy	CID
Complementary deoxyribonucleic acid	cDNA
Constitutively active receptor	CAR
Cytochrome P450	CYP
Deoxyribonucleic acid	DNA
Dithiothreitol	DTT
Eagles minimum essential medium	EMEM
Electron transfer chain	ETC
Electrospray ionisation	ESI
Endoplasmic reticulum	ER
Flavin adenine dinucleotide	FAD
Flavin mononucleotide	FMN

Foetal calf serum	FCS
Fourier transform ion cyclotron	FT-ICR
General protein/mass analysis for windows	GPMAW
Gram	g
Heptafluorobutyric acid	HFBA
High performance liquid chromatography	HPLC
Isotope-coded affinity tags	ICAT
Immobilised polyacrylamide gel	IPG
Internal diameter	i.d.
liquid chromatography	LC
Isoelectric focusing	IEF
Mass spectrometry	MS
Mass to charge ratio	*m/z*
Matrix assisted laser desorption ionisation	MALDI
Messenger ribonucleic acid	mRNA
Methionly-arginyl-phenylalanyl-alanine acetate	MRFA
Mobile phase	MP
Nanoelectrospray ionisation	NSI
Nano-liquid chromatography	Nano-LC
Nicotinamide adenine dinucleotide phosphate reduced form	NADPH
Non ionic detergent	NP-40
Normalized display mode	NL
One/two dimensional	1 or 2 D

Peptide mass fingerprints	PMF
Phenylmethanesulphonyl fluoride	PMSF
Polyacrylamide gel	PAGE
Polycyclic aromatic hydrocarbon	PAH
Polyvinylidene fluoride	PVDF
Post source decay	PSD
Potassium chloride	KCl
Pregnane X receptor	PXR
Radio frequency	RF
Selected ion monitoring	SIM
Single nucleotide polymorphism	SNPs
Single reaction monitoring	SRM
Sodium dodecyl sulfate	SDS
2,3,7,8-Tetrachlorodibenzo-p-dioxin	TCDD
N,N,N',N'-Tetramethylethylenediamine	TEMED
2,3,7,8-Tetrachlorodibenzo-*p*-dioxin	TCDD
Time-of-flight	TOF
Trifluoroacetic acid	TFA
Tris(hydroxymethyl)methylamine	Tris
Tris-saline	TS
UK Human Tissue Bank	UKHTB
Ultra-violet	UV

Amino acids

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic acid
Cys, C	Cysteine
Glue, E	Glutamic acid
Gln, Q	Glutamine
Gly, G	Glycine
His, H	Histidine
Ile, I	Isoleucine
Leu, L	Leucine
Lys, K	Lysine
Met, M	Methionine
Phe, F	Phenylalanine
Pro, P	Proline
Ser, S	Serine
Thr, T	Threonine
Trp, W	Tryptophan
Tyr, Y	Tyrosine
Val, V	Valine

Acknowledgments

I wish to record my sincere thanks to my supervisors Professor Laurence H. Patterson, Dr. Kevin J. Welham and Dr.William Griffiths for providing me with the opportunity to carry out research in mass spectrometry and for their guidance and support throughout my PhD. Sincere thanks to Dr. Andrew F. Wilderspin for his help running gels.

Thanks for the financial support provided from EPSRC and CASE as well as the continual support, help, advice and encouragement from all my friends and colleagues. In particular I would especially wish to thank Mr. Emmanual Samual for his advice on LC system and Dr. Catherine S. Lane for her help using LCQ.

I would like to extend my thanks to the following:

- Dr Steven A. Everett (Grey Cancer Institute) for providing me with an opportunity to work within the GCI laboratories and harvesting the MCF-7 cell lines.
- UK Human Tissue Bank for providing me with human liver samples.
- Professor MC Bibby (University of Bradford) for providing me with human tumour xenografts.
- Dr. Barry Fuller and Professor Brian Davidson (Royal free hospital) for providing me with colorectal metastases.

Dedicated to my mother and father

Chapter 1

Introduction

1.1. Cytochromes P450

The first report on the existence of a cytochrome P450 (CYP) enzyme or a "microsomal carbon monoxide-binding pigment", as it was called at that time, was published in 1958 by Klingenberg *et al.* This enzyme gave a unique 450-nm absorption peak, and when its haemoprotein nature was recognized, it was given the name cytochrome P450 (Omura, 1999).

CYP enzymes are expressed ubiquitously in different life forms: they have been found in animals, plants, fungi, and bacteria (Nelson *et al.,* 1996). They seem to be indispensable for eukaryotic species, but not for prokaryotes, since some bacteria lack CYP enzymes (Nelson, 1999). Eukaryotics need CYPs for the biosynthesis of sterols, which are constituents of plasma membrane (Omura, 1999). Eukaryotic CYP enzymes are membrane-bound, mostly localized to the endoplasmic reticulum (ER), but some CYPs are also present in mitochondrial inner membranes.

There are over 280 different families of CYPs, and currently more than 1925 sequenced and named isoforms (drnelson.utmem.edu/Cytochrome P450.html). In the rat, a species traditionally important in drug development studies, there are about 50 CYP genes as identified from the UNIGENE database (drnelson.utmem.edu/ UNIGENE.RAT.html). Humans are estimated to have approximately 100 different CYP genes, of which 59 are thought to be expressed based on transcript identification (Ingelman-Sundberg, 2004). The notable diversity of CYP enzymes has given rise to a systematic classification of individual forms into families and subfamilies. The protein sequences within a given gene family are at least 40% identical (e.g. CYP2A6 and CYP2B6), and the sequences within a given subfamily are > 55% identical (e.g. CYP2A6 and CYP2A7) (Nelson *et al.,* 1996).

There are 18 different CYP families currently known in humans. The enzymes in the families 1-3 are mostly active in the metabolism of xenobiotics, whereas for example, CYP5, CYP11, CYP17, CYP19, CYP21 and CYP27 are involved in cholesterol and steroid hormones metabolism (Table 1.1).

Table 1.1: Human CYP families and their main functions

CYP family	***Main functions***
CYP1	*Xenobiotic metabolism*
CYP2	*Xenobiotic metabolism, arachidonic acid metabolism*
CYP3	*Xenobiotic and steroid metabolism*
CYP4	*Fatty acid hydroxylation*
CYP5	*Thromboxane synthesis*
CYP7	*Cholesterol 7α-hydroxylation*
CYP8	*Prostacyclin synthesis*
CYP11	*Cholesterol side-chain cleavage, steroid 11β – hydroxylation, aldosterone synthesis*
CYP17	*Steroid 17α-hydroxylation*
CYP19	*Androgen aromatization*
CYP21	*Steroid 21-hydroxylation*
CYP24	*Steroid 24-hydroxylation*
CYP26	*Retinoic acid hydroxylation*
CYP27	*Steroid 27-hydroxylation*
CYP39	*Oxysterol 7α - hydroxylation*
CYP46	*Cholesterol 24-hydroxylation*
CYP51	*Sterol biosynthesis*

Data adapted from Gonzalez, 1992; Nelson, 2003

Three genes, CYP1A1, CYP1A2 and CYP1B1, are the members of the CYP1 family. All three genes share the main features of regulation and are all transcriptionally controlled by the AHR-ARNT (aryl hydrocarbon receptor-aryl hydrocarbon receptor nuclear translocator) pathway (Schmidt & Bradfield, 1996). They are induced by polycyclic aromatic hydrocarbons (PAH), 2,3,7,8-Tetrachlorodibenzodioxin (TCDD) (Schmidt & Bradfield, 1996), and smoking (Willey *et al.*, 1997; Zevin & Benowitz, 1999). However, there is variation in the extent and cell specificity of their expression

and induction. Importantly, they all are active in the metabolism of PAHs into reactive intermediates that can bind to DNA and, if the damage goes unrepaired, may produce mutations involved in neoplasmic transformation (Shimada *et al.,* 1996). Thus, they have been implicated in the formation of chemically caused cancers (Nebert *et al.,* 1996). CYP1A1 is a major extrahepatic CYP enzyme (Raunio *et al.,* 1995a). Its level of expression in human liver is very low (Edwards *et al.,* 1998). Because of the significance of CYP1A1 in the activation of procarcinogens, there have been active efforts to link the polymorphisms of the CYP1A1 gene with the individual susceptibility to chemically induced cancers, especially lung cancer (Raunio *et al.,* 1995b).

The expression of CYP1A2 appears liver-specific, since no CYP1A2 protein has been detected in any other tissue (Raunio *et al.,* 1995a). Similar to CYP1A1, CYP1B1 is also an extrahepatic CYP form expressed in almost every tissue, including kidney, prostate, mammary gland, and ovary (Sutter *et al.,* 1994; Shimada *et al.,* 1996a; Tang *et al.,* 1999). In general, CYP1B1 basal expression is higher compared to CYP1A1 (Shimada *et al.,* 1996a; Eltom *et al.,* 1998). The expression of CYP1B1 in human liver is highly debated (Chang *et al.*, 2003). It has been suggested to be overexpressed in tumours (Murray *et al.,* 1997). The CYP1B1 gene has several alleles, and interestingly, the functionally impaired alleles have been shown to be linked with human primary congenital glaucoma (Stoilov *et al.,* 1997). This finding demonstrates that even the CYPs classified as "xenobiotic-metabolizing" enzymes may have important functions in modulating growth and differentiation. CYP1B1 allelic variants that affect the rate of conversion of estradiol into 4-hydroxyoestradiol have been described (Shimada *et al.,* 1999; Li *et al.,* 2000).

The human CYP2 family is a heterogeneous group of enzymes. It contains the subfamilies CYP2A, CYP2B, CYP2C, CYP2D, CYP2E, CYP2F, and CYP2J (Nelson *et al.,* 1996). CYP2B6, CYP2D6, CYP2E1, CYP2F1, and CYP2J2 are the only functional members in their respective subfamilies, whereas the CYP2A subfamily contains two and CYP2C four functional enzymes. The human CYP2A subfamily contains three genes i.e. CYP2A6, CYP2A7, and CYP2A13. The CYP2A6 protein has been detected in liver (Yun *et al.,* 1991). There is a significant interest in CYP2A6, due to its major role in the

metabolism of nicotine *in vitro* (Nakajima *et al.,* 1996). The genetic polymorphisms of the CYP2A6 gene have been associated with interindividual differences in smoking behavior (Pianezza *et al.,* 1998). It is induced *in vivo* by phenobarbital and other antiepileptic drugs (Sotaniemi *et al.,* 1995). In human hepatocytes, CYP2A6 is induced by phenobarbital and rifampicin (Dalet-Beluche *et al.,* 1992, Rodríguez-Antona *et al.,* 2000). Relatively high levels of CYP2A13 mRNA have been detected in human lung and adult and fetal nasal mucosa (Koskela *et al.,* 1999). CYP2B6 is a minor CYP form in human liver. CYP2B is inducible by barbiturates in rodents. This induction is mediated by nuclear receptor CAR (constitutively active receptor) (Sueyoshi *et al.,* 1999).
The human CYP2C subfamily contains four highly homologous genes: 2C8, 2C9, 2C18 and 2C19. CYP2C9 is the main CYP2C in human liver, followed by CYP2C8 and CYP2C19 (Edwards *et al.,* 1998). Polymorphisms in CYP2C19 gene cause a higher incidence of poor drug metabolizer phenotypes in Asians (23%) than Caucasians (3-5%).
The CYP2D subfamily consists of one gene, CYP2D6. CYP2D6 is perhaps the best studied P450 with a drug metabolism polymorphism. This enzyme is responsible for more than 70 different drug oxidations. Since there may be no other way to clear these drugs from the body, poor metabolizers may be at severe risk for adverse drug reactions. It has profound effects on the metabolism of several commonly used pharmaceuticals, including tricyclic antidepressants, haloperidol, metoprolol, propranolol, codeine, and dextromethorphan (Pelkonen *et al.,* 1998). CYP2E1 has been studied extensively due to its role in the metabolism of ethanol and also as an activator of chemical carcinogens (Lieber, 1997). This is another P450 enzyme that may be related to smoking induced cancer.

The human CYP3 family contains only one subfamily (Nelson *et al.,* 1996). CYP3A includes four genes, CYP3A4, CYP3A5, CYP3A7, and CYP3A43. The CYP3A4 enzyme is the most abundant CYP in human liver. It has been estimated that about 50% of drugs metabolised by CYPs are metabolised by CYP3A4 (Bertz & Granneman, 1997). The transcription of CYP3A monooxygenases are stimulated by Pregnane X receptor (PXR). CYP3A5 is expressed polymorphically in human liver (Wrighton *et al.,* 1989), in lung (Kivistö *et al.,* 1996), colon (Gervot *et al.,* 1996), kidney (Schuetz *et al.,* 1992, Haehner *et al.,* 1996), oesophagus (Lechevrel *et al.,* 1999), and anterior pituitary gland

(Murray *et al.,* 1995), demonstrating CYP3A5 to be a more extrahepatic CYP3A form. CYP3A7 is mainly expressed in human foetal liver, where it is the major CYP form (Kitada & Kamataki, 1994). Low levels of CYP3A7 mRNA have been detected in adult liver (Hakkola *et al.,* 1994; Schuetz *et al.,* 1994). CYP3A7 has similar catalytic properties compared with other CYP3A enzymes, including testosterone 6β - hydroxylation (Kitada *et al.,* 1985).

The CYP4 family encodes several cytochrome P450 enzymes that are capable of hydroxylating the omega-carbon and, to a lesser extent, the (omega-1) position of saturated and unsaturated fatty acids, as well as enzymes active in the omega-hydroxylation of various prostaglandins (Simpson, 1997). CYP4B1 is the only CYP4 family member with activity towards xenobiotics. It was isolated from a lung cDNA library, and the mRNA was found to be expressed in human lung, but not in liver (Nhamburo *et al.,* 1989). mRNA expression has also been demonstrated in human colon (McKinnon *et al.,* 1994) and placenta (Yokotani *et al.,* 1990; Hakkola *et al.,* 1996). Heterologously expressed CYP4B1 catalyzes 6β -hydroxylation of testosterone, a typical CYP3A reaction.

The Human Genome Project has revealed new CYP genes not previously discovered. The Cytochrome P450 homepage provided by Dr. David R. Nelson (drnelson.utmem.edu/CytochromeP450.html) lists four new genes (CYP2R1, CYP2S1, CYP2U1 and CYP3A43) in the families 1-3.

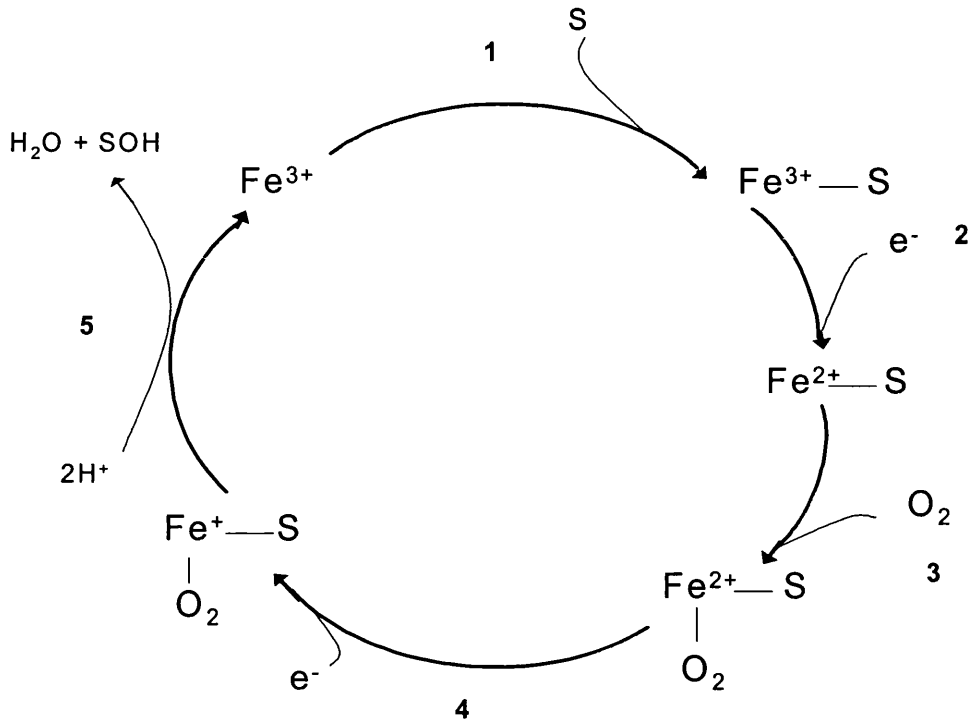
1.2. Catalytic cycle and structure of CYP450

CYPs catalyze many types of oxidation reactions, but the one that is most important is hydroxylation. These enzymes are called mixed function oxidases or monooxygenases, because they incorporate one atom of molecular oxygen into the substrate and one atom into water. They differ from dioxygenases that incorporate both atoms of molecular oxygen into the substrate.

Overall Oxidation Reaction:

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

The sequence of events leading to CYP catalysed oxidation reactions is shown in Fig 1.1



Fe, haem iron of CYP; S, substrate; O oxygen

Figure 1.1: Reaction cycle of CYP450

Modified from: Lewis (1992)

Fig 1.1 is a schematic representation of the generally accepted sequence of reactions involved in the oxidation of CYP substrates. The initial step of CYP catalysis is the rapid binding of the substrate to the CYP in its ferric form. The binding of the substrate to CYP results in a conformational change in the protein and an elevation of the redox potential of the haem iron. This renders the CYP susceptible to a one e^- reduction to a ferrous CYP substrate complex. In the next step, molecular oxygen interacts with the reduced CYP substrate complex giving rise to an oxy-CYP substrate complex.

Subsequently the oxy-CYP substrate complex receives a second electron and this complex degrades into the reaction products, which are water and a hydroxylated substrate (Ortiz Montellano, 1995).

CYPs add a hydroxyl group in the phase 1 step of drug metabolism. The hydroxyl then serves as the site for further modification in phase II drug metabolism.

For CYP to function, they also need a source of electrons. The electrons are donated by another protein that binds briefly to the CYP and transfer an electron from a prosthetic group. This transfer of an electron between proteins is called an electron transfer chain (ETC). In the endoplasmic reticulum, the source of electrons is Nicotinamide adenine dinucleotide phosphate (NADPH)-cytochrome P450 reductase (Omura, 1999). NADPH CYP reductase is membrane-bound by an N-terminus, which crosses the ER membrane once. The bulk of this protein is on the cytosolic side of the ER membrane (see Fig 1.2.). This protein has two domains each contains one flavin. Two electrons are acquired from NADPH to migrate from Flavin adenine dinucleotide (FAD) to Flavin mononucleotide (FMN), then to the CYP haem iron.
In mitochondria, the ETC is slightly different where electrons are transferred from NADPH by redoxin reductase to redoxin and then to CYP (Gonzalez, 1990).

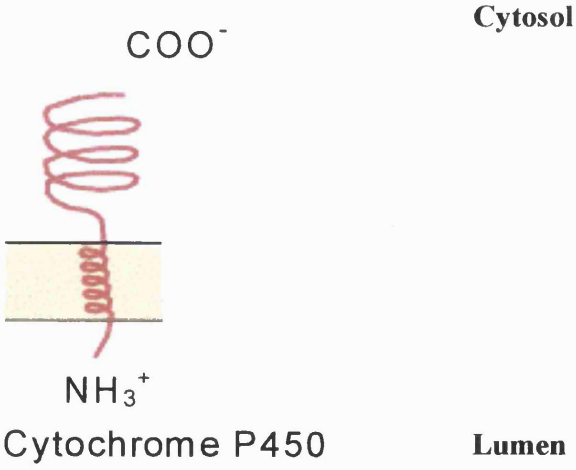


Fig 1.2: Model depicting integral membrane topologies for the mammalian CYPs of the ER.

Source: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=mcb.figgrp.4776

CYPs are membrane bound, hydrophobic in their character and therefore difficult to crystallize. These haemoproteins are monomeric (Mr 50+/- 7 kDa) containing one haem per molecule of enzyme (Ortiz Montellano, 1995). The haem-iron is coordinated to the thiolate of a cysteine residue, which endows the enzyme with its characteristic oxidative properties. Despite their sometimes minimal sequence similarity, all CYPs have a similar structural fold with a highly conserved core (Graham & Peterson, 1999). The first mammalian CYP crystallised was rabbit CYP2C5. CYP2C5 was crystallised after the removal of the N-terminal anchor peptide and replacement of an internal hydrophobic sequence with a more water soluble sequence from a related enzyme. The X-ray structure has been solved (Williams *et al.*, 2000). CYP2C5 has provided useful blueprints for homology modeling and comparative sequence alignment of the human CYPs. Despite, the report of the crystal structure of a mammalian CYP2C5, insights into how the human CYPs are able to recognize structurally diverse ligands remain elusive. Using the method of crystallization of CYP2C5 the crystal structures of CYP2C9 (Williams *et al.* 2003) and 3A4 (Williams *et al.*, 2004) were resolved (see Fig 1.3 and Fig 1.4). The structure of CYP3A4 revealed an active site, with little conformational change associated with the binding of substrates. The heme of CYP3A4 has greater accessibility to the active site than does that of CYP2C9, which could then allow two substrate molecules to have access to the reactive oxygen, consistent with data indicating that CYP3A4 is able to bind and metabolize multiple substrate molecules simultaneously (Korzekwa, 1998).

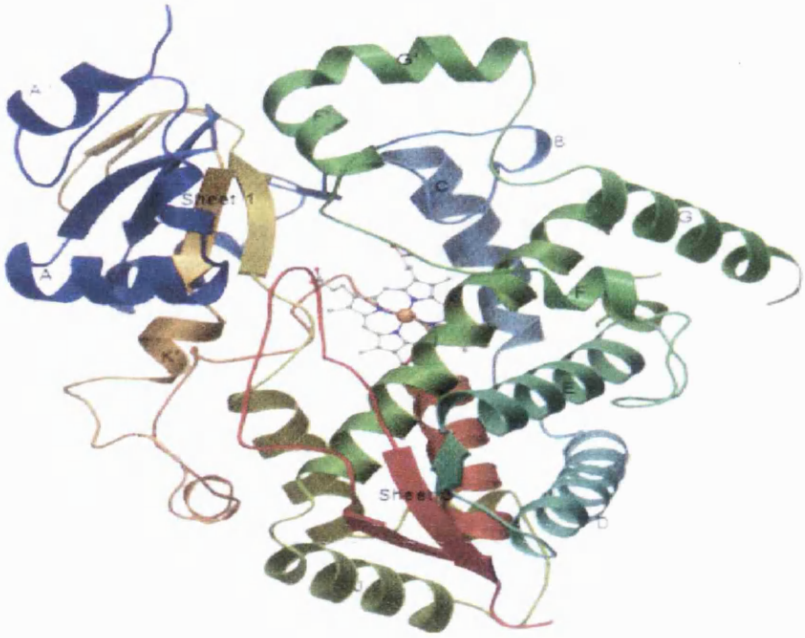


Fig 1.3: Crystallised structure of a human CYP3A4. Overall fold of CYP3A4 coloured from blue at the N-terminus to green, to yellow, to red at the C-terminus.

Source: Williams et al. (2004)

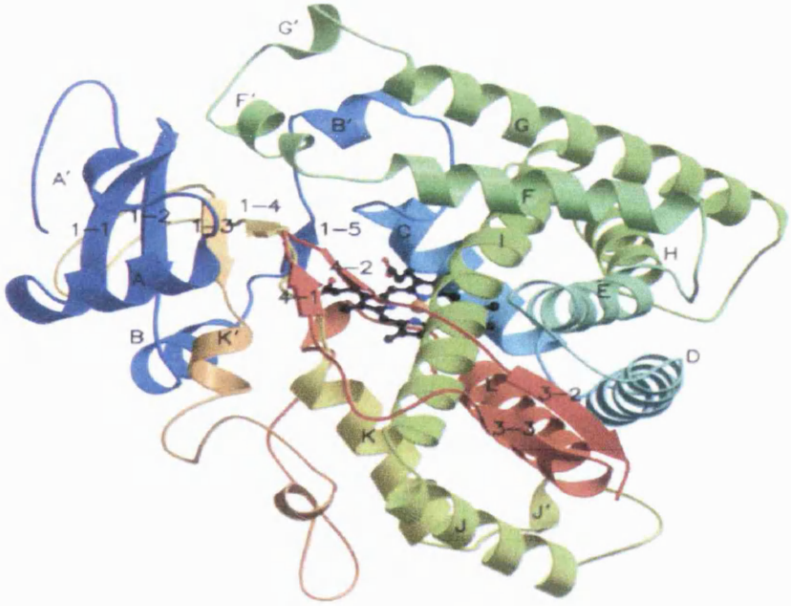


Fig 1.4: Crystallised structure of a human CYP2C9. Overall fold of CYP2C9 coloured from the blue at the N-terminus to red at the C-terminus. The haem group is depicted as a ball-and-stick model in the centre of the molecule, flanked by helices I and L.

Source: Williams et al. (2003)

1.3. Expression of CYPs in tissues

Most studies have focused on hepatic forms of CYPs firstly due to their importance to drug metabolism and secondly, due to their role as a port of entry for all ingested substances. CYPs are also expressed in extrahepatic tissues including the intestine, kidney, lung, brain, adrenal gland, skin and placenta (McKinnon and McManus, 1996; Ding and Kaminsky, 2003). On xenobiotic exposure, organisms can increase their metabolic activities to eliminate the chemicals effectively from the body, usually by activating the transcription of the CYP genes (Gonzalez, 1989). Whereas the liver is the main organ to metabolise xenochemicals, the metabolising enzymes in the exposed tissues such as lung, skin, and nasal mucosa play important roles in reducing the toxicity of the chemicals before they enter the body circulation.

Expression of some CYP isoforms is a risk factor in certain cancers since these enzymes can convert procarcinogens to carcinogens. Therefore, several studies were carried out to show the significant differences in the expression of CYPs between tumour and peritumour tissue. Examples include a differential expression of CYP1A, 3A4 and 2E1 between normal and tumour tissue in the stomach (Murray *et al.,* 1998). CYP 1A1 is expressed in oesophageal tumours (Nakajima *et al.,* 1996) and CYP1B1 mRNA is expressed in both normal breast tissue and breast tumours (Huang *et al.*, 1996). A study performed using reverse transcriptase polymerase chain reaction and immunoblotting has indicated that the CYP1B1 protein is expressed in breast tumour (McKay *et al.,* 1995). A study performed on endocrine regulation of CYP isoforms in the rat breast (Hellmold *et al.,* 1995) has showed that the CYPs content of breast tissue is 1000 fold less than in the liver. The catalytic properties of the polymorphic human CYP 1B1 variant has been studied by Shimada (1999), suggested that polymorphism in the human CYP1B1 gene may cause some alterations in catalytic function towards procarcinogens and steroid hormones. From this it has been suggested that CYP1B1 may contribute to susceptibilities of individuals towards mammary and lung cancers in humans. In prostate cancers, the levels of CYP1A and CYP3A expression are significantly increased (Murray *et al.*, 1995). CYP3A is also consistently expressed in kidney tumours (Murray *et al.*, 1999). In contrast to the above studies evidencing an increased CYP expression in

tumours, some investigations have demonstrated that there is undetectable expression of CYP in tumours especially CYP3A and CYP1A compared to normal tissues (Massaad *et al.*, 1992).

Cancers of the alimentary tract, in particular stomach and colorectum are among the most common cancers world wide, with localized, surgical resection remaining the only curative approach. Colon cancer is the most common malignant tumour of the gastrointestinal tract and the commonest cause of death from cancer. Colorectal cancer is a disease that results from the growth and reproduction of abnormal cells beginning to form in the lining of the colon or rectum. Cancer cells can also break away and spread to other parts of the body (such as liver and lung) where new tumors form. The spread of colon cancer to distant organs is called metastasis of colon cancer (see Fig 1.5).

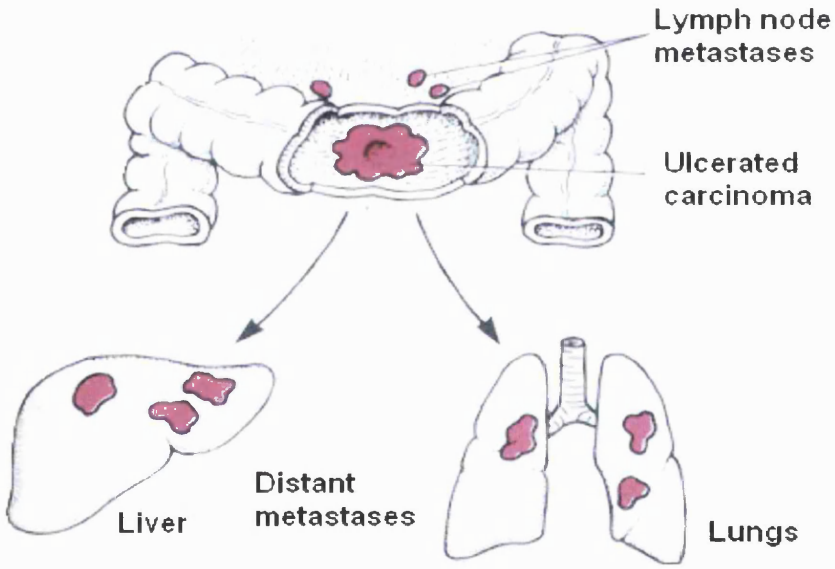


Fig 1.5: Medical illustration of a colon cancer and its spread to liver, lungs and lymph nodes

Modified from: Rubin and Farber (1988)

Less than 10% of colorectal tumours can be explained by inheritable susceptibilities, whose genetic basis at the DNA level is already well understood. It has been estimated that environmental factors contribute to the causation of sporadic colorectal cancer

(Lichtenstein *et al.,* 2000). Colonic carcinomas are often intrinsically resistant to anti-cancer drugs. Evidence has showed that the expression of enzymes participating in biotransformation may play a part in tumour drug-resistance. Moreover, the same group of xenobiotic metabolising enzymes are probably involved in the development of colonic carcinoma (McKay *et al.,* 1993). About 50% of patients with colorectal cancer are cured from their disease. However, for patients with metastatic disease at diagnosis or at relapse, the 5-year survival is less than 5% (Doherty & Michael, 2003).

The variable expression of CYPs between tumour and normal tissue can provide a basis for relative sensitivity to anticancer drugs, thereby localising drug actions to tumours (Patterson *et al.,* 1999). A potential anticancer therapeutic approach resides in the differential expression of CYPs between normal and tumour tissue. Thus, elevated levels of CYPs in tumours relative to levels in normal tissues could convey enhanced selectivity and sensitivity to a prodrug, which is dependent on CYP-mediated bioactivation.

1.4. Development of an antitumour Prodrug

The expression of drug metabolising CYPs in tumours makes these enzymes an attractive target for cancer therapy.

Physiological studies of tumours have showed that solid tumours are poorly vascularised resulting in regions of acute and chronic hypoxia (*Patterson et al.,* 2000). This hypoxic condition may contribute to the diminished oxidative activity of CYP enzymes (Shan *et al.*, 1992). The idea of hypoxia in tumours lead to the development of antitumour drugs, an example of which is AQ4N (Anthraquinone-di-N-oxide), an antitumour prodrug designed to be excluded from a cellular nuclei until metabolised in hypoxic cells to give AQ4 (Patterson *et al.,* 2000). AQ4N is a stable, O_2-insensitive cytotoxic metabolite (Smith *et al.*, 1997). As described earlier, CYP function by accepting two electrons from NADPH via CYP reductase, which readily binds to molecular oxygen. One oxygen atom is then inserted into a drug substrate while the other oxygen atom is reduced to H_2O.

While carrying out experiments using human liver microsomes, Raleigh *et al.* (1997) showed that AQ4N is reduced by a membrane bound CYP3A family in a process that is inhibited by air. Thus, AQ4N could be used to determine the expression of CYP in hypoxic regions of solid tumours.

1.5. Polymorphism

Major drug metabolism differences in human are caused by polymorphisms in CYPs. A polymorphism is a difference in DNA sequence found at 1% or higher in a population. These differences in DNA sequence can lead to differences in drug metabolism, so they are important features of CYP genes in humans. In Caucasians, the polymorphism for the poor metaboliser phenotype is only seen in 3% of the population. However, it is seen in 20% of the Asian population. Because of this difference, it is important to be aware of a person's ethnic origin when drugs are given that are metabolized differently by different populations. CYPs have been shown to be polymorphic as a result of single nucleotide polymorphism (SNPs), gene deletions, and gene duplications (Table 1.2).

Table 1.2: Examples of human CYP polymorphisms

CYP	*Variant*	*Mechanism*	*Consequence*
CYP2A6	CYP2A6*2	$Leu^{160} \rightarrow His^{160}$	Inactive enzyme
	CYP2A6*4	Gene deletion	No enzyme
	CYP2A6*5	$Gly^{479} \rightarrow Leu^{479}$	Defective enzyme
CYP2B6	CYP2B6*5	$Arg^{487} \rightarrow Cys^{487}$	Reduced enzyme activity
CYP2C9	CYP2C9*2	$Arg^{144} \rightarrow Cys^{144}$	Reduced affinity for CYP reductase
	CYP2C9*3	$Ile^{359} \rightarrow Leu^{359}$	Reduced V_{max} / K_m ratio
	CYP2C9*4	$Ile^{359} \rightarrow Thr^{359}$	Reduced V_{max} / K_m ratio
	CYP2C9*5	$Asp^{380} \rightarrow Glu^{380}$	Reduced V_{max} / K_m ratio
CYP2C19	CYP2C19*2	Splicing defect	Inactive enzyme
	CYP2C19*3	Premature stop codon	Inactive enzyme
	CYP2C19*4	GTG initiation codon	Inactive enzyme
	CYP2C19*5	$Arg^{433} \rightarrow Trp^{433}$	Inactive enzyme
	CYP2C19*6	$Arg^{132} \rightarrow Gln^{132}$	Inactive enzyme
	CYP2C19*7	Splicing defect	Inactive enzyme
	CYP2C19*8	$Trp^{120} \rightarrow Arg^{120}$	Reduced enzyme activity
CYP2D6	CYP2D6*2Xn	Gene duplication	Increased enzyme activity
	CYP2D6*4	Splicing defect	Inactive enzyme
	CYP2D6*5	Gene deletion	No enzyme
	CYP2D6*10	$Pro^{34} \rightarrow Ser^{34}$	Unstable enzyme
		$Ser^{486} \rightarrow Thr^{488}$	
	CYP2D6*17	$Thr^{107} \rightarrow Ile^{107}$	Reduced V_{max} / K_m ratio
		$Arg^{298} \rightarrow Cys^{298}$	
		$Ser^{486} \rightarrow Thr^{486}$	
CYP3A4	CYP3A4*2	$Ser^{222} \rightarrow Pro^{222}$	Reduced V_{max} / K_m ratio

Source: Rodrigues and Rushmore (2002)

It has been recognized that numerous factors contribute to intersubject variability in drug response. These include gender, age, diet, exposure to environmental factors and polymorphism (De Wildt, 1999).

Age: the activity of CYP enzymes decreases with advancing age in humans. Metabolism of antipyrine (metabolised by at least 10 CYP isoenzymes, CYP1A2, 2A6, 3A4, 2B6, 2C8, 2C9, 2C18, 2C19, 2D6 and 2E1), lidocaine, diazepam and theophylline decreases in the elderly (Sotanieui *et al.,* 1997). In vivo activities of CYP1A2, 3A4, 2C9 and 2D6 reported to be low at birth, but maximally increased at the young adult stage and decreased in old age (Tanaka, 1983).

Gender: gender based differences in metabolic activity of hepatic CYPs have been identified in humans. Women exhibit higher baseline 3A4 activity than men and therefore a greater extent of interactions on average. It has been shown that the clearance of diazepam and prednisolone is more in women, but clearance of propranolol is more in men (Tanaka, 1983).

Nutrition: starvation and obesity are known to induce CYP enzymes in rodents, but in humans these conditions inhibit CYP enzymes. Obesity has been reported to increase metabolism of enflurane and sevoflurane in humans (O'Shea *et al.,* 1994).

Environmental factors: Cigarette smoking is known to induce CYP enzymes. Smoking has been found to increase clearance of phenacetin and theophylline (Sarkar and Jakson, 1994).

The metabolizing capabilities of CYPs, their expression in cancerous tissues, polymorphism, and alteration in drug metabolism lead investigation into the presence of CYPs in tissues using different techniques.

1.6. Why use proteomics to study CYPs?

Proteomics is the systematic analysis of proteins for their identity, quantity and function (Pennington & Dunn, 2001). In contrast to a cell's static genome, the proteome is both complex and dynamic.

In the search for tumour drug targets, there has been a concerted effort to define gene expression levels at the transcript level (Ross, 2000; Scherf, 2000). However, it is evident that mRNA expression data alone are insufficient to determine functional outcomes of a cell. For example, mRNA data provide little information about protein activation state, or posttranslational modification (Simpson & Dorow, 2001).

Traditional methods for the detection of CYP proteins have relied on techniques such as immunoblotting and activity assays, or on the detection of CYP mRNA (Anderson and Seilhamer, 1997). These techniques have significant limitations. Western blots, whilst being very sensitive, rely on the availability of isoform-specific antibodies and it is necessary to pre-select which CYPs are to be analysed for, and to identify each isoform in turn. Activity assays that are geared to interrogate the activity of a CYP isoform invariably require multiple analysis techniques, and different assays must be developed for different target substrates, and even then they may not be totally isoform-specific. Measurements at the mRNA level are fraught with uncertainty since the presence and abundance of a particular type of mRNA does not necessarily indicate a similar presence and abundance of the corresponding protein; there are numerous reports highlighting the disparity between mRNA transcript and protein expression levels (Cai and Guengerich, 2001; Chen *et al.*, 2003).
Mass spectrometry is a direct method for the analysis of expressed proteins and offers uniquely the ability to detect low levels of multiple proteins in a single run.

The proteome of a differentiated cell is estimated to consist of thousands to tens of thousands of different types of proteins including many with covalent post-translational modifications, i.e. phosphorylation, which may be transient. It is estimated that intracellular proteins are expressed over a dynamic range of at least six orders of

magnitude and as many as twelve orders of magnitude in circulating plasma. However, proteomic analysis of peptides derived from proteolysis of proteins rather than proteins themselves is important because it is difficult to obtain sequence information from proteins at the level of sensitivity available for peptides (Forbes *et al.*, 2001).

The most commonly used proteomic approach is to separate proteins isolated from a biological mixture by 1 or 2 dimensional gel electrophoresis, cut the gel region of interest, digest the proteins contained within the gel, and analyse the extracted peptides by HPLC and mass spectrometry (see Fig 1.6.).

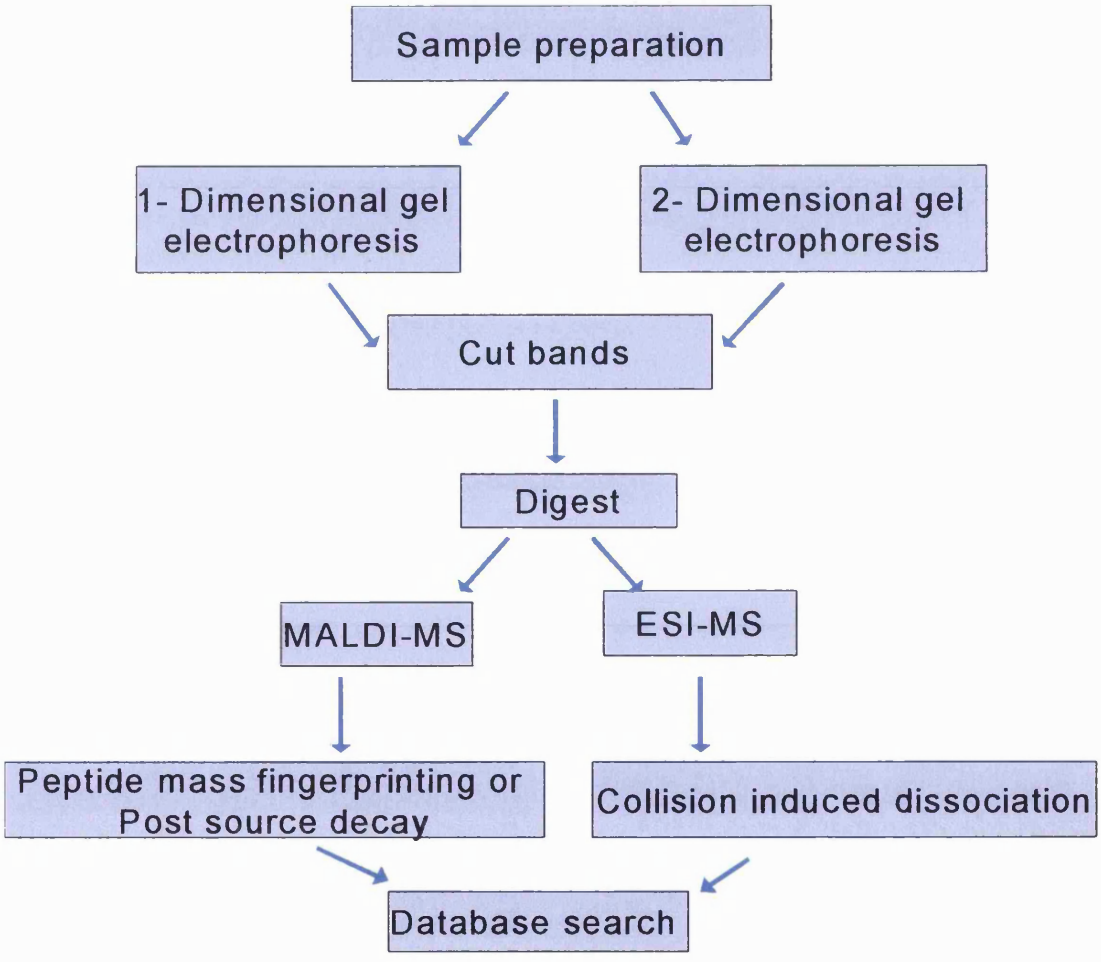


Fig 1.6: Schematic outline of the current protocol showing the recovery of peptides and proteins from gel separations for analysis by mass spectrometry

Discovery based proteomic tools

There are three major proteomic tools

1.6.1. Gel electrophoresis (one dimensional or two dimensional)

1.6.2. High performance liquid chromatography (HPLC)

1.6.3. Mass spectrometry (MS)

Matrix assisted laser desorption ionisation-time of flight (MALDI-TOF)

Electrospray ionisation-MS (ESI-MS)

Tandem MS (MS/MS)

1.6.1. Gel electrophoresis (one dimensional or two dimensional)

Electrophoresis is a simple and a rapid tool to separate proteins based on the fact that charged molecules will migrate through a matrix upon application of an electric field. Generally, the sample is run in a support matrix such as agarose or polyacrylamide gel. Agarose is mainly used to separate larger macromolecules such as nucleic acids whereas polyacrylamide gel is widely used to separate proteins. The compounds used to form polyacrylamide are monomeric acrylamide and N, N'-methylene-bisacrylamide. Chemical polymerisation is initiated by TEMED (tetramethylethylenediamine) and ammonium persulfate (APS). When APS is dissolved in water, it forms persulfate free radicals, which in turn, activate the acrylamide monomer. TEMED is added to serve as a catalyst to accelerate the polymerisation reaction due to its ability to carry electrons.

Hames (1998) suggested that factor which affect the rate of polymerisation, include type and concentration of initiators, purity of reagent, pH, temperature, oxygen and concentration of monomers.

Sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) or one dimensional electrophoresis, originally described by Laemmli (1970), is the system in which proteins are fractionated strictly by their size. Prior to SDS-PAGE, sample is heated at 100°C, in

the presence of excess SDS, this denatures the protein mixture, and a thiol reagent is employed to reduce disulfide bonds. Under these conditions, all reduced polypeptides bind the same amount of SDS on a weight basis (1.4g SDS/g polypeptide) independent of the amino acid composition and sequence of the protein. The SDS:protein complexes are negatively charged have a similar charge density and thus can be separated based on their size only.

In an SDS-PAGE discontinuous buffer system, the samples are loaded directly onto the large pore gel, called a stacking gel, polymerised on top of the resolving gel. Stacking gels have different pH (pH 6.8), compared to both running buffer (pH 8.3) and resolving gel (pH 8.8) and contain no glycine ions. Separation of the stacked proteins is accomplished as the proteins enter the resolving gel because of the decreased mobility of proteins and increased mobility of the trailing glycine ions. The former is achieved by the increased gel concentration so that molecular sieving is enhanced, and the latter is achieved by the increase in the pH from 6.8 in the stacking gel to 8.8 in the resolving gel, since the mobility of glycine ions is pH dependent.

SDS-PAGE is used for the estimation of the protein size, protein purity, protein quantitation, monitoring protein integrity, comparison of the polypeptide composition of different samples, analysis of the number and size of polypeptide subunits.

Although SDS-PAGE is a commonly used gel electrophoresis system for analysing proteins, it cannot be used to analyse intact protein complexes and proteins, whose biological activity need to be retained for subsequent functional testing. In these situations, it is essential to use a non-denaturing system. A further limitation of SDS-PAGE is that different proteins with the same size are unlikely to be resolved by this technique. Any single band may be composed of multiple components of similar molecular mass. A non-denaturing gel system separates native proteins; separation is based not only on protein size but also on protein charge and shape.

The coupling of isoelectric Focusing (IEF) (first dimension) with SDS-PAGE (second dimension) results in a 2-dimension (2DE) method that separates proteins according to

two independent parameters that is charge and size respectively. The steps involved in 2DE are as follows: perform 1-D IEF; exchange the buffer in the focusing gel for the SDS buffer; place the 1-D gel in direct contact with the second-dimension SDS gel; perform SDS electrophoresis; detect the separated protein spots.

Electrophoretically resolved proteins can be detected using organic dyes (Coomassie Blue), silver stains or fluorescent stains. In addition to these staining methods for the detection of proteins, physical methods such as autoradiographic methods can be used for protein visualisation. By Coomassie blue staining around 500 spots can be visualised at the picomole level with silver staining sensitivity can be increased by one or two orders of magnitude, allowing the visualisation of up to 10000 spots (Hames, 1998). New fluorescent methods may give the attomole sensitivity.

Although 2-DE provides the highest resolution, there is growing evidence that hydrophobic proteins, proteins that are very acidic or very basic, as well as proteins that are present in very low abundance may not be amenable to 2-DE (Harry *et al.* 2000). Separation of low copy number proteins in amounts sufficient for post separation analysis continues to present a challenge for 2D techniques, and preconcentration is required in this situation. Santoni *et al.* (1999) proposed that fractionation of proteins by Triton X-114 combined with solubilisation with CHAPS resulted in the inability to detect hydrophobic proteins on 2-DE gels. Rabilloud *et al.,* (1997) suggested that IEF leads to severe quantitative losses of hydrophobic proteins but a denaturing mixture containing urea, thiourea and detergent (both non-ionic and zwitterionic) has been shown to improve protein solubility. However, it has also been suggested that thiourea as other sulphur containing compounds, strongly inhibits acrylamide polymerisation and is therefore not compatible with the standard set up for carrier ampholyte-IEF in tube gels. It has been estimated that only about 1% of integral membrane proteins are actually resolved on current 2DE (even when thiourea is used in the lysis buffer) (Garrels *et al.*, 1997). Therefore, SDS-PAGE had been regularly used to separate very hydrophobic proteins especially integral membrane proteins (Kashino *et al.,* 2001).
Once the proteins have been separated by SDS-PAGE the proteins in the gel can be transferred onto a membrane made of nitrocellulose or polyvinylidene fluoride (PVDF),

by applying a current, a process called electroblotting. Western blot can be used to detect a certain protein in a sample by using an antibody specific to that protein.

Alternatively, after separation the proteins can be subjected to in-gel digestion and extraction. For digestion cyanogen bromide, trypsin or other endopeptidases can be used, however, trypsin is most frequently used (Hames, 1998). The amide bond specificity of trypsin is Arg-X, Lys-X. Peptides generated by proteolysis can be separated using HPLC techniques as described below.

1.6.2. High Performance liquid chromatography (HPLC)

The digestion of an unfractionated protein mixture greatly increases the number of components to be analysed and condenses the resultant peptides into a relatively narrow mass range, thereby placing great demand on the performance of the HPLC column used for peptide separation. For peptide mixtures, derived from abundant proteins separated by SDS-PAGE, reversed phase (RP) HPLC approaches have been used to monitor changes in protein expression.

RP-HPLC is a form of partition chromatography between the polar mobile phase and non-polar stationary phase of the column. Typical stationary phases are nonpolar hydrocarbons, waxy liquids or bonded hydrocarbons (such as C_{18}, C_8, C_4, etc.) and the solvents are polar aqueous-organic mixtures such as methanol-water or acetonitrile-water. The C_{18}, C_8, and phenyl-bonded phases are most often used in the RP mode. There are no significant limitations on the type of RP column used with LC-MS interfaces or mass analyzers (see Table below).

Table 1.3. Types of RP column

C-18	Non-Polar
Phenyl	
C-8	
C-4	
CN	More Polar

Separation mechanism

Separations are based on relative polarity between the stationary and the mobile phase. For an initial separation on a narrow bore, column flow rates of 200nl/min (75μm i.d x 15mm) on RP column using a formic acid and acetonitrile (ACN) gradient system is recommended. A typical RP gradient run consists of loading a sample on a column equilibrated with polar solvent (for example 0.1% (v/v) formic acid in water). The mobile phase polarity is then changed by introducing a less-polar solvent (for example 60% ACN containing 0.1% (v/v) aqueous formic acid). Since the stationary phase is relatively non-polar, very polar species are not retained on the column and quickly elute, while nonpolar species are retained on the column until the mobile phase becomes more non-polar. Alteration in mobile phase components and gradient slope can alter peptide retention and selectively (Mahoney & Hermodson, 1980).

Studies have been carried out to optimise HPLC chromatographic peptide separation with alternative mobile and stationary phases (Young and Wheat, 1990). A study by Young and Wheat suggested that peptides separated using RP-HPLC using increasing concentration of ACN in the presence of formic acid could be improved by adjustment of gradient slope, temperature and ion-pair reagents (Medzihradszky *et al.,* 2001). Peptide mapping at high temperature leads to reduced retention for all peptides.

The disadvantage of RP-HPLC is that very hydrophilic peptides will not stick to the column and thus will elute in the column flow-through. Conversely, very hydrophobic peptides might not elute until the end of a gradient or not at all. It is therefore possible that some of the peptides in a sample will go undetected.

There are two goals, which provide the primary impetuses behind column miniaturization.

A). Improved separation efficiencies
B).To interface HPLC with MS

The trend toward miniaturization has had significant impact in the field of HPLC. The decrease in column internal diameter (i.d) from 4.6 mm to 320 μm (packed capillary columns) reduces the optimal flow rate from 1 ml/min to ~ 50 μl/min. However, further reduction in column i.d to 100 μm or less (nanoscale) reduced the optimal flow rates to ~ 100 nl/min. The reduced column size leads to reduce relative loading capacity. On the other hand, the concentration at the detector is inversely proportional to the sample amount injected. Therefore, the reduced loading capacity is compensated for by the increased relative concentration (Tomer, 2000). This is an important consideration for ESI-MS.

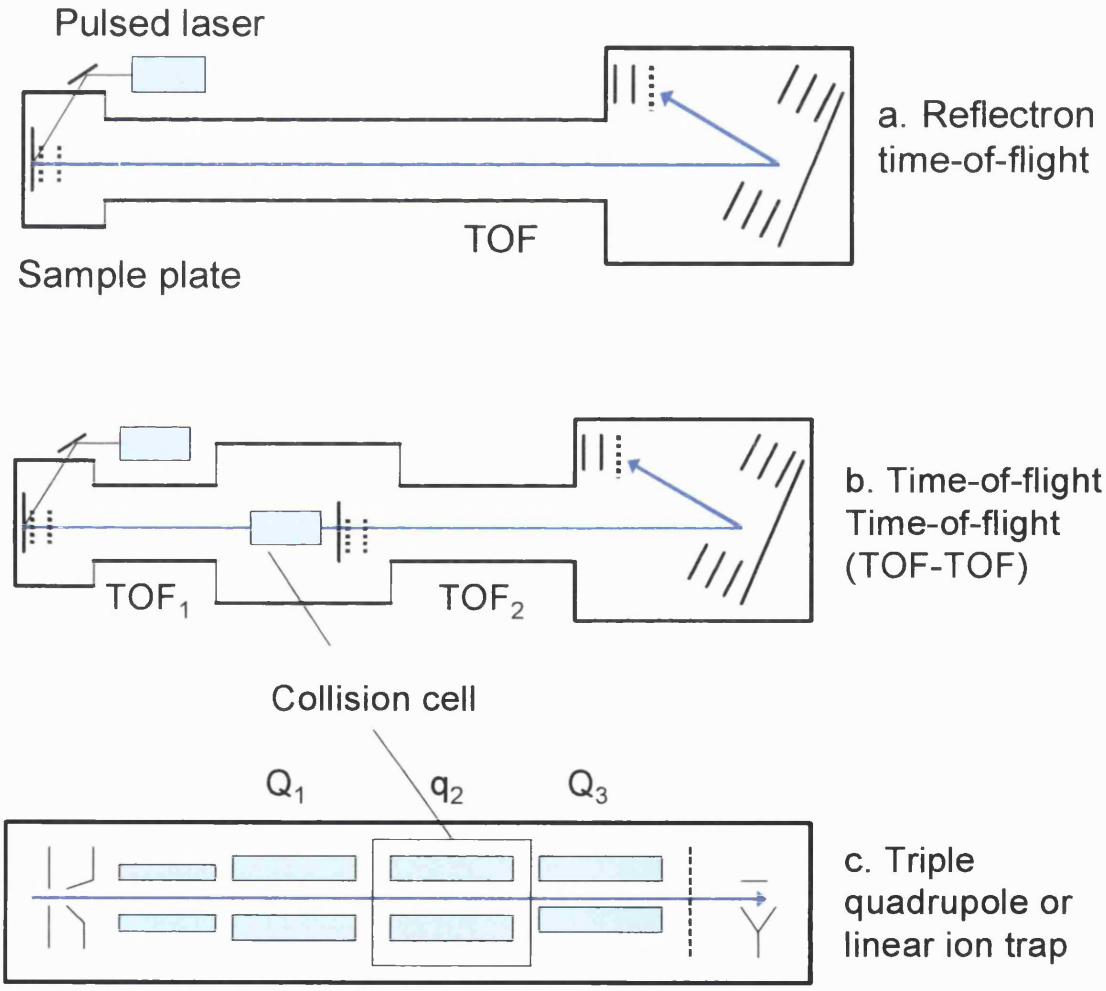
1.6.3. Mass Spectrometry

Mass spectrometry is now an indispensable tool for rapid protein and peptide structural analysis and the widespread use of MS is a reflection of its ability to solve structural problems not conclusively determined with other techniques. By definition, a mass spectrometer consists of an ion source, a mass analyser that measures the mass-to-charge ratio (*m/z*) of the ionised analytes, and a detector that registers the number of ion at *m/z* value.

The mass analyser is central to the technology. There are four basic types of mass analyser currently used in proteomics research. These are ion trap, TOF, quadrupole and Fourier transform ion cyclotron resonance (FT-ICR) analysers. They are different in design and performance. These can be stand alone or, in some cases, put together in tandem to take advantage of the strengths of each (see Fig 1.7). Fig 1.7a shows a reflectron TOF instrument, where ions are accelerated to high kinetic energy and are separated along a flight tube as a result of their different velocities. The TOF-TOF instrument incorporates a collision cell between two TOF sections, fragmentation occurs in the collision cell, and the fragments are separated in the second TOF section. Quadrupole mass spectrometers select ions by time-varying electric fields between four rods, which permit a stable trajectory only for ions of a particular desired *m/z*. The quadrupole TOF, combines the front part of a triple quadrupole instrument with a

reflectron TOF section. Ion traps trap ions, fragment them and scan out the fragments to generate a MS/MS spectrum. FT-ICR also traps the ions but in a strong magnetic field.

In ion-trap analysers, the ions are first captured or 'trapped' for a certain time interval and are then subjected to MS or MS/MS analysis. A disadvantage of ion traps is their relatively low mass accuracy, due in part to the limited number of ions that can be accumulated at the point-like centre of the trap before space-charging distorts their distribution and thus the accuracy of the mass measurement. The 'linear' or 'two-dimensional ion trap is an exciting recent development where ions are stored in a cylindrical volume.



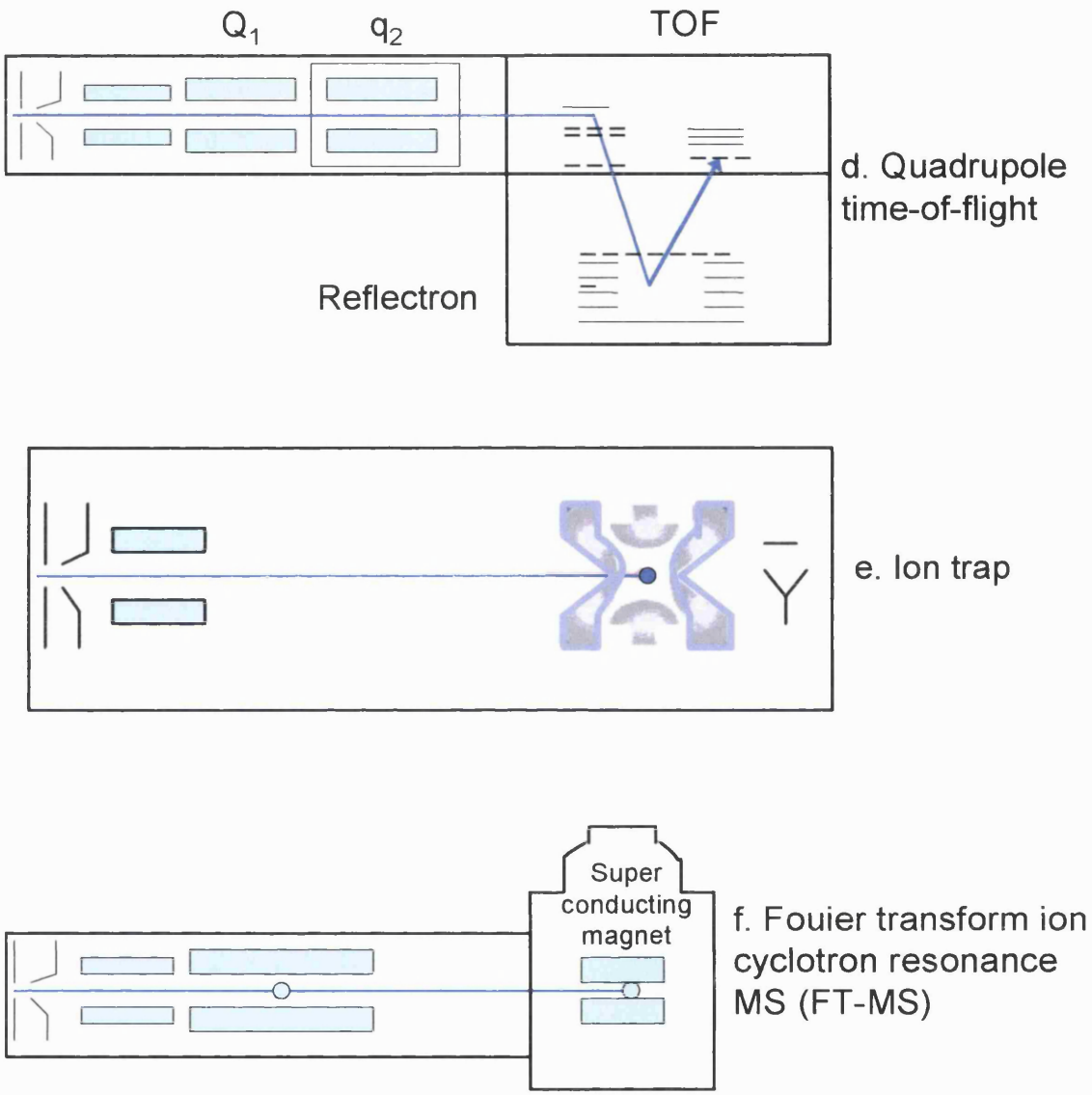


Fig 1.7: Different instrumental configurations

Modified from Aebersold and Mann (2003)

Electrospray ionisation (ESI) and matrix assisted laser desorption ionisation (MALDI) are the two commonly used methods to volatise and ionise proteins or peptides for mass spectrometric analysis (Fenn *et al.*, 1989; Karas and Hillenkamp, 1988). The difference is MALDI ionizes the sample out of a dry, crystalline matrix via laser pulses while ESI

ionizes the analytes out of a solution and is therefore readily coupled to liquid-based (for example, chromatographic and electrophoretic) separation tools.

1.6.3a. MALDI-TOF

MALDI was introduced in the late 1980s by the group of Hillenkamp (Karas *et al.*, 1987).

The principals involved in MALDI are the sample is dispersed in a large excess of matrix material, which will strongly absorb the energy from a UV or IR laser (UV at 337 nm, IR at 2.94 μm). The matrix contains a chromophore to absorb the laser light and since the matrix is in a large molar excess it will absorb essentially the entire laser radiation. The matrix isolates sample molecules in a chemical environment that enhances the probability of ionisation without fragmentation.

The MALDI source has traditionally been coupled to TOF mass analysers because of its pulsed nature. Short pulses of laser light focused onto the sample spot cause the sample and matrix to volatilize and produce matrix neutrals, positive, negative ions and sample neutrals (Fig 1.8). Ions formed in the MALDI process are accelerated by a high voltage supply (for example, +20-30KV) and then allowed to drift down a flight tube where they are separate according to *m/z*.
Ions travel with velocity v = d/t; (v: velocity, d: tube distance, t: time). All ions will leave the source at the same time with the same kinetic energy due to their having been accelerated through the same potential difference. The relationship between kinetic energy, mass and velocity is given by this equation,

$E = ½ mv^2$
E = kinetic energy, m = mass of ions, v = velocity of ion

The time-of-flight of the ions produced will only be dependent on the mass and the charge of the produced ions. The larger the *m/z* of an ion, the slower its velocity and thus

the longer it takes to traverse the field-free drift zone. The above equation indicates that the mass and velocity are inversely proportional.

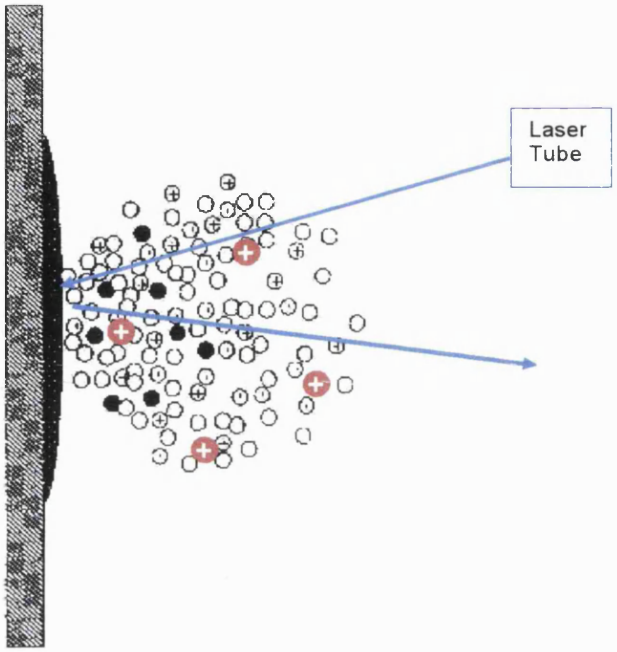


Fig 1.8: Showing a laser flash producing matrix neutrals, positive, negative ions and sample neutrals

Procedure

A matrix concentration of 1mg/ml is ideal and usually from one to ten picomoles of sample is required for analysis. This is spotted onto the sample position on the metal strip and then 0.5 µl of matrix (usually 50mM) is applied to the sample position as well. Evaporation of the solvent in air results in the formation of analyte:matrix co-crystals. The plate containing the analyte:matrix is inserted into the vaccum chamber ($10^{-5} - 10^{-8}$ torr) of the MS.

Many different matrices can be used for MALDI. Some of the most common matrices are given in Fig 1.9.

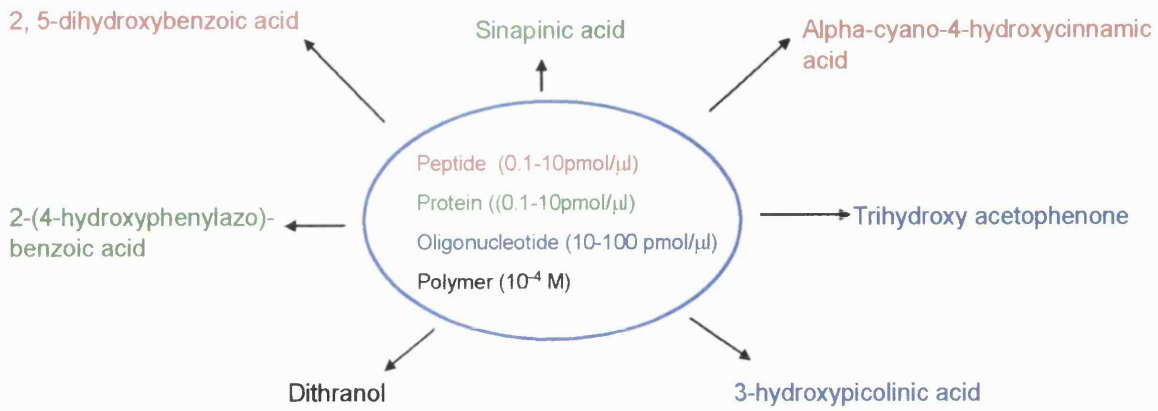


Fig 1.9: Matrices used for MALDI-TOF

Source: Applied Biosystems

It is best to remove buffer salts and detergents (e.g. by dialysis) prior to analysis of samples and to dissolve the sample in a suitable solvent (e.g. 0.1% TFA/water), which will not interfere with the spectrum. If there is too much salt in a sample, it would effectively suppress the sample signal, giving no sample spectrum. It may then be possible to dilute the sample to the point where the contaminants will have little effect on the spectrum. Aebersold and Goodlett (2001) suggested that samples isolated from the biological sources analysed by MALDI are compatible with biological buffers such as phosphate and tris and low concentration of urea, non-ionic detergent and some alkali metal salts.

In early studies molecules were ionised by MALDI and *m/z* ratio measured in a linear TOF mass analyser (Fig 1.10). MALDI-TOF is still used to identify proteins by what is known as peptide mapping, also referred as peptide-mass mapping or peptide-mass fingerprinting. In this method, a protein is identified by matching a list of experimental peptide masses with the calculated list of all peptide masses of each entry in the database (for example a comprehensive database). Because mass mapping requires an essentially purified target protein, the technique is commonly used in conjunction with prior protein

fractionation using 2D electrophoresis. The addition of sequencing capability to the MALDI, makes protein identification more specific than those sequence obtained by simple peptide mass mapping. It should also extend the use of MALDI-MS to the analysis of more complex samples, thereby uncoupling MALDI-MS from 2D electrophoresis. However, if MALDI-MS/MS is to be used with peptide chromatography, the effluent of a liquid chromatography run must be deposited on a sample plate and mixed with the MALDI matrix.

In ordered to obtain protein identification from tryptic peptides, firstly, a peptide mass fingerprint (PMF) is generated by MALDI, which then provides a percentage of sequence coverage to identify the protein. Secondly, a peptide can be selected for PSD to produce N and C-terminal ions. The fragmentation of peptides by post source decay (PSD) can be carried out using reflectron TOF instruments.

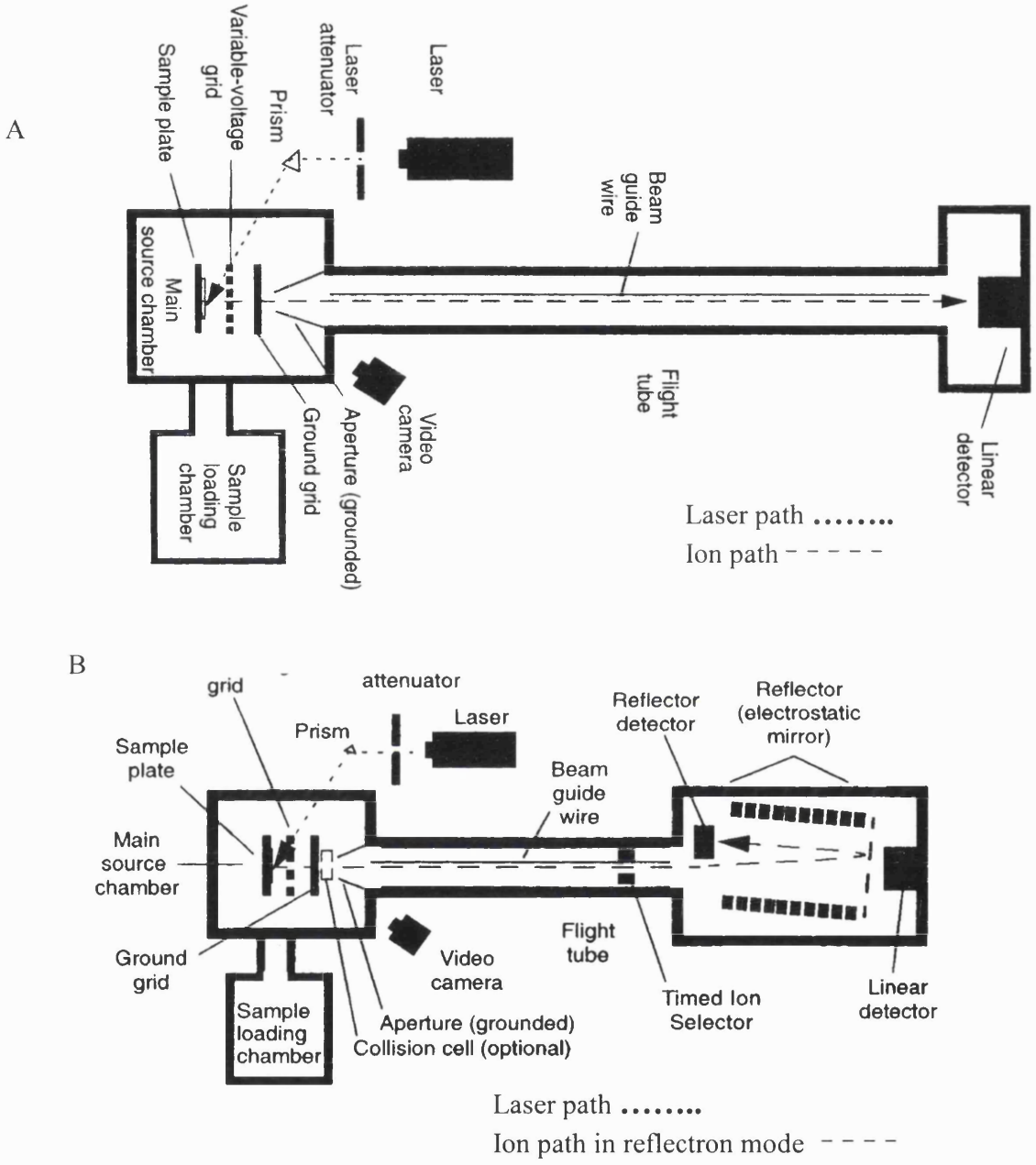


Fig 1.10: A. MALDI-MS with a linear detector B. MALDI-MS with linear and reflectron source

Source: Applied Biosystems

1.6.3b. Electrospray ionisation tandem mass spectrometry (ESI-MS/MS)

This is one of the "so-called" atmospheric pressure ionisation (API) techniques and came to prominence in the early 1990s. While MALDI is used for ionizing peptides; ESI is a method of ionization/vaporization for the widest range of polar biomolecules.
The sample is dissolved in a solvent, such as a one-to-one mixture of water and either acetonitrile or methanol, and pumped through a fine needle. The tip of the needle is at atmospheric pressure and is floated at a high potential (for example, 3-4 kV) relative to a counter electrode. The high potential difference causes the analyte flow to disperse as a fine spray of highly charged droplets which is directed across a small inlet orifice of a plate held at a lower potential (~+100-1000 V) at the front end of the mass spectrometer. The orifice is the interface between the ion sources, and the mass analyser. A simplified schematic diagram of an ES interface is shown in Fig 1.12.

As the droplets traverse from the tip of the needle to the counter electrode, they are desolvated by a counter current of gas such as argon or nitrogen. The process of generating ions from electrospray involves multiple steps of fissioning and evaporation. This fissioning process has been documented photographically, and it is well known that when the parent droplet reaches the Rayleigh limit and fission, multiple, small, highly charge droplets are formed (Fig. 1.11).

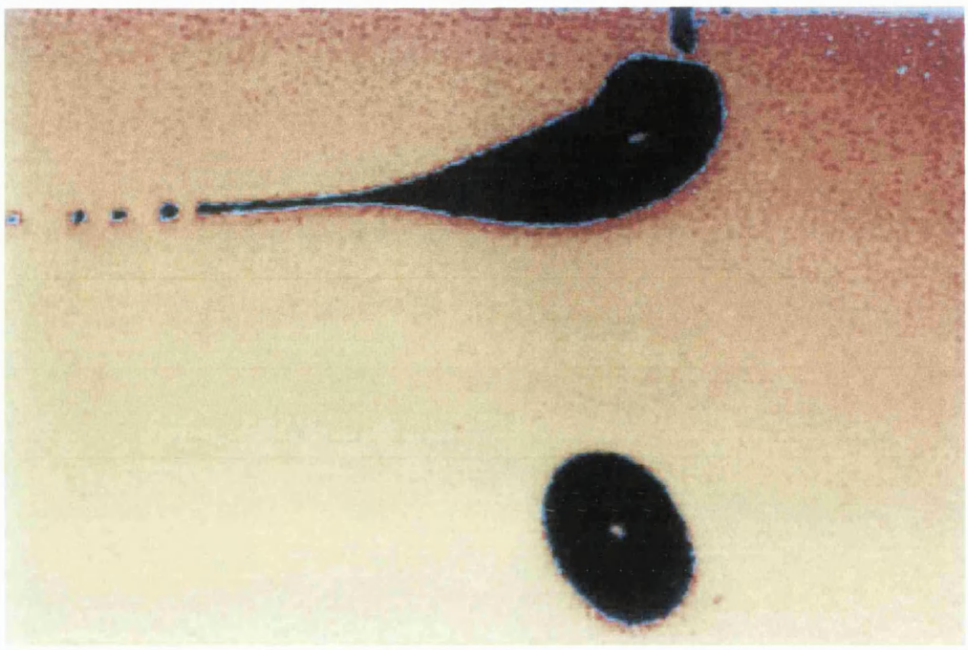


Fig 1.11: Photograph of a charged droplet in the act of fissioning. The droplet distorts, creating a miniature Taylor cone, and a string of offspring droplets are ejected.
Source: (Gomes & Tang, 1994)

The solvent must contain the means to add or remove protons. During the process of desolvation the increasing charge density in the shrinking droplet results in a coulombic explosion releasing ions from the liquid phase of the droplet into the gas phase (Carr & Annann, 1997). It is characteristic of ESI to produce peptide ions bearing multiple charges. The number of charges is a function of the number of ionisable groups in a peptide. If tryptic peptides are analysed by ESI-MS in +ve ion mode, most of the peptides will carry at least 2-charges, one at the N-terminal NH_2 group and one on the side chain of the basic amino acid positioned by trypsin at the C-terminal of the peptide. The mechanism of ES is represented in Fig below.

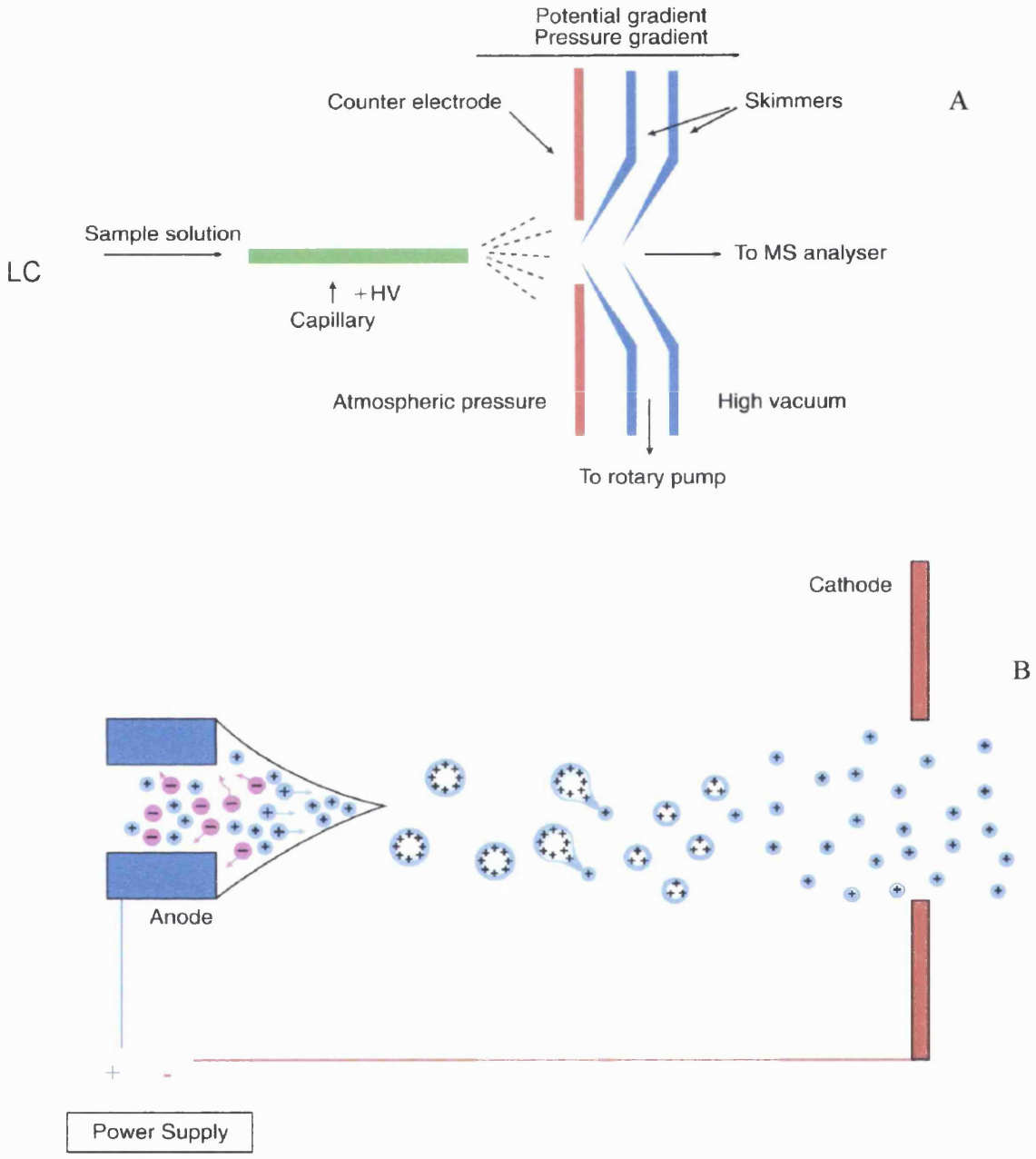


Fig 1.12. A, Features of an ES interface; B, mechanism of ES

Source: Griffiths et al. (2001)

Once the ions are formed, in the ion-trap instrument they are subjected to oscillating electrical fields applied by a radio frequency voltage applied to the ring electrodes. The radio frequency (RF) field traps the ions near the center of the ring electrodes

(Willoughby et al., 1998). The ion trap behaves like a high pass filter. As the RF voltage on the ring electrode is ramped, low masses become unstable and exit the trap to the detector.

ES is a concentration - rather than a mass-dependent process, and improved sensitivity is obtained for high-concentration low-volume samples (Fenn *et al.,* 1989).
The coupling of HPLC to electrospray allows the mass spectrometric analysis of complex mixtures without lengthy purification, extraction and derivatisation steps, using LC to clean up the sample before its separated components are subjected to MS.

The most common modes of acquiring LC-MS data are:
(1) Full scan acquisition resulting in the typical total ion current plot (TIC)
(2) MS/MS product ion scan
(3) Selected Ion Monitoring (SIM)
(4) Selected Reaction Monitoring (SRM) or multiple reaction monitoring (MRM). MRM and SRM are essentially the same experiment (see Fig 1.13).

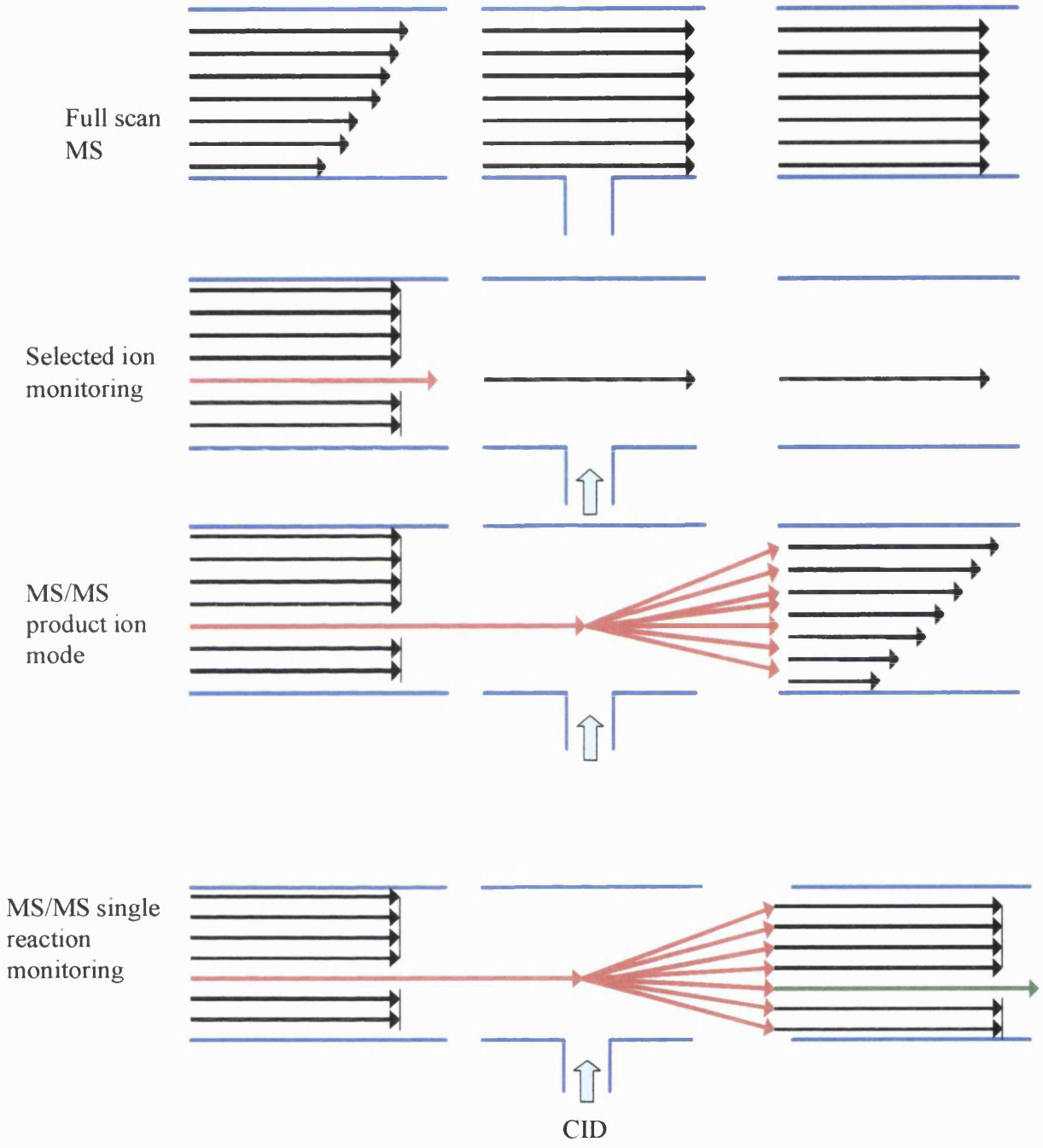


Fig 1.13: Modes of acquiring LC-MS data

Typically, the mass spectrometer is set to scan a specific mass range. This mass scan can be wide as in the full scan analysis or can be very narrow as in selected ion monitoring. A single mass scan can take anywhere from 10 ms to 5 s depending on the type of scan. During an LC-MS analysis, many scans can be acquired, and are represented by adding up mass scans and plotting the total ion current as an intensity point against time.

In SIM, the mass spectrometer is set to scan over a very small mass range, typically one mass unit. Only compounds with the selected mass are detected and plotted. The SIM experiment is more sensitive than the full scan experiment because the mass spectrometer can dwell for a longer time over a smaller mass range. SRM delivers a unique fragment ion that can be monitored and quantified in the midst of a very complicated matrix. The SRM experiment is accomplished by specifying the parent mass of the compound for MS/MS fragmentation and then specifically monitoring for a single fragment ion. One could think of this operation as the SIM of a fragment ion. The specific experiment is known as a "transition" and can be written parent mass → fragment mass.

For proteomic analysis, as the peptide elute into the mass spectrometer, data-dependent acquisition can be performed. The mass spectrometer is set to acquire a scan over the *m/z* 400-2000. Ions detected which are above a preset ion-current threshold, are then automatically selected and a tandem mass spectrometry (MS/MS) experiment performed.

For tandem MS such as in triple quadrupole, or ion traps, fragment ion spectra are generated by a process called collision-induced dissociation (CID) in which the ion to be analysed is isolated and fragmented in a collision cell and the fragment ion spectrum is recorded. The ion trap is a tandem MS in which ions can be accumulated and stored prior to analysis. In the ion trap the central part of this instrument is both a mass analyser and collision cell. C terminal ions (x, y, z) and N-terminal ions (a, b, c) are generated as a result of peptide backbone fragmentation (Fig 1.14). Ammonium ions arise by the loss of carbon monoxide from the b-ions with the formation of a double bond between amide nitrogen and the alpha carbon.

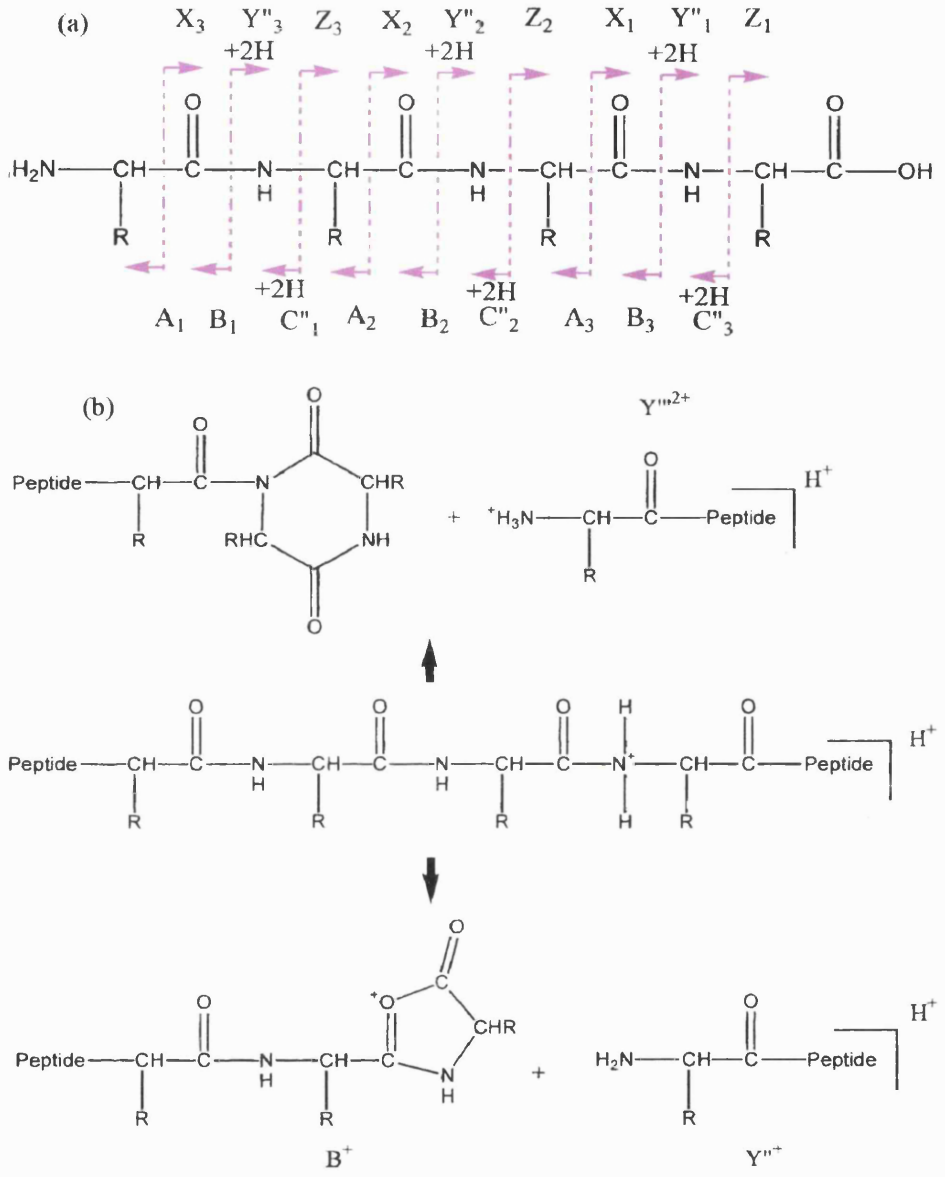


Fig 1.14: (A) Nomenclature for peptide fragmentation (Roepstorff and Fohlman 1984) and (B) schematic representation of the cleavage of peptide bonds via the mobile proton model

Source: Griffiths et al. (2001)

Protein identification using peptide CID spectra are clearer than those achieved by mass mapping because, in addition to the peptide mass, the peak pattern in the CID spectrum also provides information about peptide sequence. These spectra are searched against database using different search engines.

1.7. Peptide sequencing via database searching of tandem mass spectra

With the increasing amount of protein sequence information available in databases, protein identification using database search routines have become possible (Yates, 1993). The peptide mass fingerprinting or mass mapping procedure to identify a protein, is to subject it to trypsin digestion, and to record the mass spectrum of the resulting peptides. The masses of the proteolytic peptides are then searched against a protein or DNA database. The search algorithm then theoretically digests all appropriate proteins in the database with the specified enzyme and matches the theoretical and experimental peptide masses. The best matches are then displayed. Somewhat similar approaches are available for database searching based on MS/MS data. The most useful software algorithms allow for the identification of peptide sequences directly from an LC-MS/MS analysis. For example, the software algorithm Sequest matches a peptide sequence with a tandem mass spectrum using the following steps,

(1) Peptides with molecular masses matching that of the peptide ion sequenced in the tandem mass spectra are extracted from a protein database.

(2) Each peptide is given a preliminary score by examining the number of predicted fragment ions from the database peptide that match the acquired fragment ions in the tandem mass spectrum.

(3) The top best-matching peptides undergo an ion-matching algorithm that generates a cross-correlation score. The most commonly used search database is National Centre for Biotechnology Information, NCBInr a non-identical protein database maintained by NCBI for use with their search tools BLAST and Entrez.

A large number of internet sites have been established, which can be used to identify proteins using both PMF and MS/MS data. Few examples of these sites are given in Table 1.4.

Table 1.4: Internet sites used for proteomic analysis

Internet site	Useful utilities	Site address
Protein prospector	Search database with PMF and MS/MS data. Provide theoretical data for a known peptide sequence	http://prespector.ucsf.edu
PROWL	Search database with PMF and MS/MS data. Provide theoretical data and useful protocols	http://prowl.rockefeller.edu/PROWL/prowl.html
EXPASY	PMF database searches, theoretical tools and links to many other sites	http://us.expasy.org
PeptideSearch	Search PMF data or a sequence tag Search database with PMF and MS/MS data and	http://www.mann.embl-heidelberg.deGroupPages
Mascot	sequence tags Search database with PMF data	http://www.matrixscience.com
PepMAPPER	Search database with PMF data	http://wolf.bms.umist.ac.uk/mapper/
MOWSE	Gives theoretical fragments and data on a	http://srs.hgmp.mrc.ac.uk/cgi-bin/mowse
PredictProtein server	protein amino acid sequence Search sequence tags	http://www.ebi.ac.uk/~rost/predictprotein/
BLAST		http://www.ncbi.nlm.nih.gov/BLAST

Source: Newton et al. (2004)

In addition to qualitative protein characterization, measurement of relative and absolute protein expression between two different sample states is also important.

1.8. Quantitative proteome analysis

The goal of quantitative analysis in mass spectrometry is to correlate the intensity of the signals with the quantity of the compound present in the sample.

2DE-based proteome analysis can provide information about protein abundance at the gel level by comparing staining intensities. However, when peptide mixtures are analyzed directly by LC-MS/MS techniques, the original quantitative information is lost.

Recently, quantitative proteome analysis techniques based on direct mixture analysis have been introduced incorporating stable isotope labeling. Aebersold and co-workers (1999) have developed a method for labeling peptides with isotope-coded affinity tags (ICAT), which allow accurate quantitation of differences in the level of expression of proteins. The structure of the ICAT reagent is shown in Fig 1.15. The reagent consists of three elements: an affinity tag (biotin), which is used to isolate ICAT-labeled peptides; a linker, which can incorporate stable isotopes; and a reactive group with specificity toward thiol groups (i.e. to cysteines). The reagent exists in two forms: heavy (containing eight deuteriums) and light (containing no deuterium).

The method allows the analysis of a single protein (shown here as a protein expressed in one cell state at 1 copy/cell; and in another cell state at 3 copies/cell), but is equally applicable to total cell lysates. The proteins from cell state 1 and cell state 2 are harvested, denatured, reduced, and labeled at cysteines with the light or heavy ICAT reagents, respectively. The samples are then combined and digested with trypsin. ICAT-labeled peptides are isolated by biotin-affinity chromatography and analysed by HPLC-tandem mass spectrometer. The ratio of the ion intensities for an ICAT labeled pair quantifies the relative abundance of its parent protein in the original cell state. In addition, the tandem mass spectrum reveals the sequence of the peptide and unambiguously identifies the protein. This strategy results in the quantification and identification of all cysteine containing protein components in a mixture. It is, in theory, applicable to protein mixtures as complex as the entire proteome.

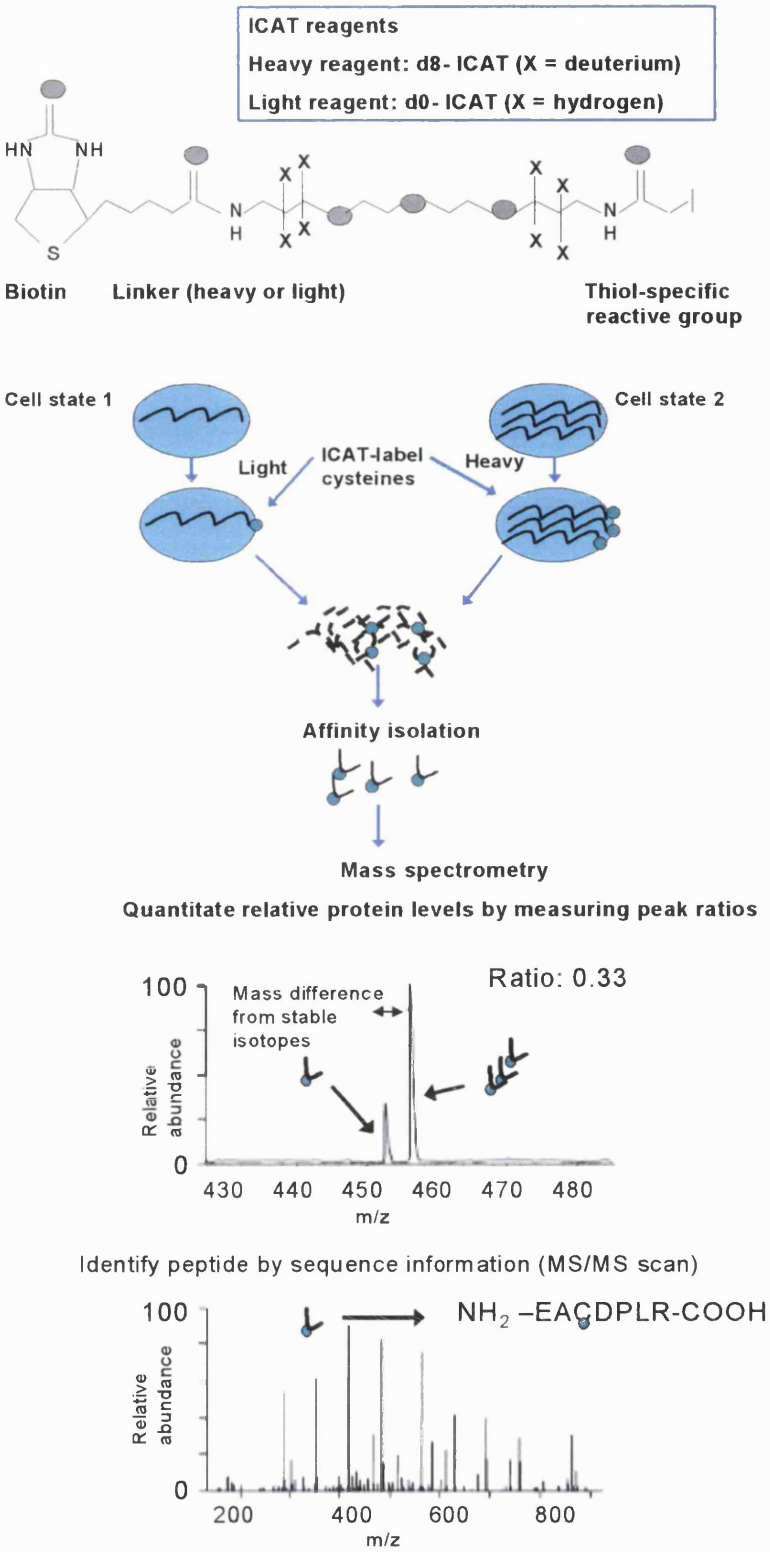


Fig 1.15: The ICAT strategy for quantifying differential protein expression

Modified from: Aebersold et al. (1999)

In contrast to the relative quantification strategies just described, Gerber *et al.* (2003) have introduced absolute protein quantification (AQUA) technique to quantify low abundance yeast proteins involved in gene splicing. In this method synthetic peptides with incorporated stable isotopes were used.

Although proteomic technologies have made rapid progress in the analysis of soluble proteins in recent years, membrane proteins which are estimated to be 20-30% of the human genome (Stevens & Arkin, 2000), has been less extensively analysed. CYPs are membrane bound proteins and to date there have been relatively few reports of the analysis of CYP proteins by mass spectrometry, the majority of MS reports have been on activation-based studies concentrating on the observation of substrates and metabolites involved in CYP reactions rather than the CYPs themselves. However, recent reports show that LC-ESI-MS/MS is applicable to the analysis of CYPs directly; for example CYP2C9 and 2E1 and their covalent adducting metabolites (Koenigs *et al.,* 1999; Regal *et al.,* 2000). MALDI-TOF mass spectrometry has been used to determine the molecular masses of closely related CYP2B1 and 2B2 (Lewis *et al.,* 1993). MALDI-TOF analysis of peptides generated by cyanogen bromide cleavage of CYP3A4 (He *et al.,* 1998; Lightning *et al.,* 2000) and trypsin proteolysis of CYP2E1 (Cai and Guengerich, 2001) have been described.

With regard to multiple protein identification, 2DE followed by MALDI-TOF and PMF has become a principal approaches for the proteomic profiling of various *in vitro* and *in vivo* biological systems (Henzel *et al.,* 1993). However, PMF is most reliable when samples which contain only a few proteins, are analysed. The presence of multiple proteins within a SDS-PAGE band may lead to spurious results. In the present work we have analysed tryptic peptides from liver microsome separated into SDS-PAGE bands, however, in our work we have performed LC-ESI-MS/MS analysis. Although 2DE followed by PMF has become a principal approach for the proteomic profiling of various *in vitro* and *in vivo* biological systems (Galeva and Alterman, 2002), a major drawback of the 2DE approach is its low performance in separation of membrane proteins, e.g., CYPs (Galeva *et al.,* 2003). Alternatively, SDS-PAGE has been used for separating proteins from complex tissue samples, but this adds the requirement of HPLC separation

and MS/MS. Further MS/MS provides amino acid sequence information, which in combination with PMF data provides a secure identification of proteins.

1.9. Aims

The problem of identifying membrane bound proteins is challenging and there are a variety of techniques available to identify these proteins. However, up to now there has been no systematic approach to the identification of multiple CYPs from complex biological systems using nano LC-MS/MS.

The aim of this study is to establish a MS based method to measure multiple CYPs in biological systems and to apply this method to the identification and the quantification of CYP protein expressed in normal tissues and tumour.

The specific objectives of this project are

1. To develop the nano LC-MS/MS method for the detection of CYPs.
2. To apply this technique to the analyses of multiple CYPs in tissues known to be rich in these proteins, specifically rat and human liver.
3. To investigate the gender-biased male and female rat liver CYP expression by nano LC-MS/MS.
4. To compare the CYP content of metastases grown in the liver with those in the surrounding liver tissue.
5. To investigate the presence of CYPs in normal colon tissue and primary colon tumours and then to compare with colorectal metastases for CYP expression.
6. To extend these studies to the measurement of CYPs in MCF-7 cell lines and human tumour xenografts.
7. To develop further methods to maximize the total sequence coverage for the identified proteins.
8. To develop LC-MS/MS quantitation methods and to determine if LC-MS/MS data of tryptic digest of proteins can be used for quantitation, when using heavy isotope labeled peptides.

Chapter 2

Materials and Methods

Materials

Coomassie brilliant blue G-250, potassium chloride (KCl), coomassie brilliant blue R-250, horse heart myoglobin, phenylmethanesulphonyl fluoride (PMSF), ethylenediaminetetraacetic acid (EDTA), 3,3'-diaminobenzidine (DAB) peroxidase substrate, glycine, sodium dodecyl sulphate (SDS), ammonium bicarbonate, ammonium persulphate (APS), tetramethylethylenediamine (TEMED), methionyl-arginyl-phenylalanyl-alanine acetate (MRFA), tris (hydroxymethyl) methylamine, non ionic detergent (NP-40) and 3,3'-diaminobenzidine (DAB peroxidase substrate) were purchased from Sigma. HPLC grade acetonitrile, methanol, water were from Fisher Scientifics and Ultra pure protogel was obtained from National Diagnostics, UK. Hoeffer Mini-gel apparatus for SDS-PAGE was from BioRad. Dithiothreitol (DTT) was from Melford Laboratories. Nitrocellulose membrane and antibodies were obtained from Amersham Life Science. Prestained protein markers were purchased from New England Biolabs. Complete protease inhibitor cocktail tablets were from Roche Diagnostics. 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD, dioxin), Eagles minimum essential medium with Eagles salt (EMEM), foetal calf serum (FCS) non-essential amino acid (NEAA), L-glutamine and sodium bicarbonate were purchased from Invitrogen, UK. Bovine serum albumin (BSA) was from Fluka. Phosphate buffered saline from Oxoid, UK. JA-20 rotor, J2-21 centrifuge, Beckmann L8-60M, Ti-65 rotor and 70.1-Ti rotor were purchased from Beckman. Ultra Turax T25 electronic homogeniser was from Jake & Kunkel. 6K10 centrifuge and 11162 rotors were obtained from Sigma. SpeedVac was purchased from Labconco Corporation, Kansas City, Missouri.

Wistar rats were obtained from Bantin and Kingman, Hull, UK. Human Caucasian breast adenocarcinoma cells, MCF-7, were obtained from European Collection of Cell Cultures, (Salisbury, UK). Human liver and metastasis were supplied by the UK Human Tissue Bank (UKHTB, Innovation Centre, Leicester LE1 5XY).
Human xenografts (excised from mice) were supplied by Professor MC Bibby (University of Bradford, UK). Heavy isotopes labeled peptides were purchased from Thermo Electron (Hemel Hempstead).

Methods

2.1. Cell culture

MCF-7 cell lines were grown and induced with TCDD at Grey Cancer Institute (Mount Vernon Hospital Northwood, Middlesex, UK). The method used is described below.

Passage of cell lines

Human caucasian breast adenocarcinoma cells, MCF-7, were obtained from European Collection of Cell Cultures, (Salisbury, UK). Cells were cultured in a media which was prepared by adding 50 ml of 10 x Eagles minimum essential medium with Earles salt (EMEM), supplemented with 50 ml of FCS (10%), 5 ml NEAA, 5 ml of 200mM L-glutamine, and 12 ml of sodium bicarbonate (7.5 %) in 500 ml sterile water. All cells were sub-cultured under sterile conditions in 75 cm^2 (T75) vent/close flasks (Helen Biosciences, UK).

The cultured cells were washed with twice with 10 ml of phosphate buffered saline (PBS), to remove any FCS which inhibits trypsin, and passaged by adding 3-5 ml of 1:10 trypsin/EDTA solution to the cultured cells. The cells were incubated at 37°C (for 5 min) in a humidified atmosphere of 5% CO_2 / 95% air. After the cells were detached from the flask, an appropriate volume (10 ml for a 75 cm^2 flask) of medium containing 10% FCS was added to the cells to stop further action of the trypsin. The cells were resuspended in 10 ml sterile PBS, pH 7.3 (Oxoid, UK) and pelleted at 1000 x rpm in a Denley^{TM} BR40 centrifuge at room temperature for 5 minutes to remove trypsin. The cell suspensions were split into 1:4 into 75 cm^2 flasks containing a 15 ml of medium. The cells were induced with 10 nM TCDD and incubated for 21 hours at 37°C in a humidified atmosphere of 5% CO_2 / 95% air.

2.2. Preparation of microsomes

2.2.1. Preparation of microsomes from MCF-7 cell lines

Following TCDD induction the cells were washed twice with 10 ml of ice cold PBS. Instead of using trypsin the cells were manually scraped into 10 ml 0.01M Tris-HCl containing 0.25 M sucrose and 15% glycerol pH 7.4. A protease inhibitor cocktail (complete inhibitor tablets), prepared by adding one tablet to 50 ml volume (manufacturers instructions) in 0.25 M sucrose / 0.05 M Tris HCl buffer pH 7.4 was used as a homogenisation buffer. The cells were then homogenised making sure there was minimal frothing as this indicates protein denaturation. The homogenate was then centrifuged at 9,000 rpm (Beckman J2-21 centrifuge, JA-20 rotor) for twenty minutes at 4°C. The resultant supernatant was then centrifuged at greater than 100,000 x g (44,000 rpm) (Beckman L8-60M ultracentrifuge, Ti-65 rotor) for 1h at 4°C. The pellet obtained after centrifugation was re-suspended in 0.1M Tris-HCl, containing 15% glycerol and 1mM EDTA pH 7.4.

2.2.2. Preparation of rat liver microsomes

Fresh liver (9-10 g) from sacrificed male and female outbred 10 week Wistar rats (Bantin and Kingman, Hull, UK) was removed. The excised liver was placed into a pre-weighed glass beaker (kept on ice). The liver was weighed and washed with cold isotonic saline (0.9% sodium chloride w/v, 4 °C) to remove blood. The whole rat liver was placed onto a clean glass plate and chopped finely and all connective tissue was excised. Liver was homogenised using an electronic homogeniser (Ultra Turax T25 from Jake & Kunkel) in 100 ml of 0.01 M Tris-HCl buffer, pH 7.4, containing 0.25 M sucrose, 15% glycerol and 0.67 mM PMSF. PMSF diminishes proteolytic action during microsome preparation. PMSF is not stable in aqueous solution and was added at the time of use from a 0.67 M ethanol solution, while stirring. (The ethanol solution of PMSF is stable at 4 °C for 2-3 months). Microsomes were prepared using differential centrifugation as follows: an initial centrifugation at 2,400 x g (Sigma 6K10 centrifuge, 11162 rotor) for 10 min at 4 °C was used to sediment the cell debris, nuclei, and unbroken cells. To enhance the yield

of microsomes, the sediment was rehomogenised in a homogenisation buffer (100 ml of 0.01 M Tris-HCl buffer, pH 7.4, containing 0.25 M sucrose, 15% glycerol and 0.67 mM PMSF) and the washings combined with the original supernantant and recentrifuged at 2,400 x g for 10 min at 4 °C. The supernatant obtained from this step was centrifuged at 12,000 x g (Beckman J2-21 centrifuge, JA-20 rotor) for 20 min at 4 °C to sediment mitochondrial fragments. Supernatant from this step were centrifuged at 180,000 x g using the Beckmann L8-60M centrifuge (Ti-65 rotor) for 1h at 4 °C. The resultant microsomal pellets were suspended in 0.1 M Tris-HCl, containing 15% glycerol and 1mM EDTA, pH 7.4, and then recentrifuged at 180,000 x g using the Beckmann L8-60M (Ti-65 rotor) centrifuge for 1 h to obtain the washed microsomal fraction. The final pellet was resuspended in 10 ml 0.1 M Tris-HCl, containing 15% glycerol and 1 mM EDTA, pH 7.4, aliquoted and stored at -80 °C.

2.2.3. Preparation of human liver microsomes

Human liver microsomes from 6 donors were obtained from redundant material procured from multi-organ donors. Consent from both the legal authorities (LREC) and from the families concerned was obtained.

Human microsomes were isolated and stored by United Kingdom Human Tissue Bank members. This procedure was performed on melting ice to ensure that the sample remained cool during preparation.

The liver tissue was scissor-minced and suspended in 3 volumes of ice cold KCl Phosphate-EDTA buffer (11.5 g KCl, 10 ml of 100 mM EDTA, 1 mM DTT, and 100 ml of 100 mM potassium phosphate buffer pH 7.5 made up to 1 litre with sterile water).

Sample were then homogenised using a tissue grinder. The resultant homogenate was subjected to high-speed centrifugation at 14,500 x g (20 minutes, 4°C) to remove cell debris. The supernatant was transferred into ultracentrifuge tubes and spun in an ultracentrifuge at 105,000 x g (1 hour, 4°C) to precipitate the microsomes. The pellet (microsomes) was resuspended in an ice cold KCl phosphate-EDTA and recentrifuged at

105,000 x g (40 minutes, 4°C). The final pellet was resuspended in 100 mM TrisHCl pH 7.4 and 250 mM sucrose. The microsomal solution was aliquoted in small volumes (1ml) and a small sample retained for protein content determination. All samples were stored at -80°C.

2.2.4. Preparation of microsomes from human colorectal metastasis and paired normal liver

The human colorectal metastasis and liver samples were received from consenting patients by the Royal free hospital and then provided to the School of Pharmacy on dry ice. Tissues for the study were accessed from the resection and discarded masses of tumour and surrounding liver, which were removed as part of the surgical treatment for hepatic metastases arising from colon cancers. The samples were stored at –80 °C until used.

Grinding of tissues under liquid nitrogen and subsequent homogenization

The tissue samples were weighed and immersed in liquid nitrogen. A percussion mortar and pestle, forceps and spatula were also placed in the liquid N_2 in polystyrene container and allowed to cool. The percussion mortar and pestle was removed from the liquid nitrogen and the sample was placed into the mortar and the steel pestle placed into the mortar. The entire mortar and pestle was sealed in a plastic sealable bag; and placed on the floor and the pestle was hit with a rubber mallet to grind the tissue to a powder.

The ground tissue was transferred into an ice cold homogeniser along with 1 mL of homogenisation buffer (0.25 M sucrose, 50 mM Tris-HCl, pH 7.2, 4°C and complete protease inhibitor cocktail tablets (one tablet for every 50 ml volume) (Roche Diagnostics) for every 0.1 g of ground tissue. The ground tissue was then homogenised using ~30 strokes of a polytetrafluoroethylene-head pestle. The homogenised tissue was transferred to an Oak Ridge centrifuge tube using a Gilson pipette. If samples were to be stored at this stage, then they were snap frozen in liquid nitrogen and then loaded into centrifuge tube placed inside a universal tube before storing at -80 °C.

Equipment, which was to be immediately reused, for example, the mortar and pestle, was rinsed with water. The waste was autoclaved upon completion of the procedure.

The centrifuge tube containing homogenized tissue was placed into a 50 mL Falcon centrifuge tube and centrifuged at 2400 x g (Sigma 6K10 centrifuge) for 10 min to sediments the cell debris, nuclei and unbroken cells. The supernatant poured off into Beckman centrifuge tube. The tube was filled right to the top with homogenisation buffer. The centrifuge tube containing cell debris and unbroken cells was autoclaved. The supernatant was centrifuged at 12,000 x g (i.e. 11,500 rpm) for 20 minutes at 4 °C, using the 70.1-Ti rotor in the Beckmann L8-60M centrifuge. This centrifugation step sediments the mitochondrial fraction and any broken mitochondrial fragments. The rotor was removed from the centrifuge and transferred to the biological safety cabinet. The supernatant was decanted into Beckman centrifuge tube and centrifuged at 180,000 x g (i.e. 45,000 rpm) for 1 hour at 4 °C, using the Beckmann L8-60M centrifuge. This pellets the microsomal fraction. The supernatants were discarded to the waste container in the safety cabinet while microsomal pellets were resuspended in 0.1 M Tris-HCl pH 7.4, containing 15% glycerol, 1 mM EDTA and one complete protease inhibitor cocktail tablet.

The microsomal suspension obtained previously was recentrifuged at 180,000 x g for 1 hour at 4 °C. The supernatant was decanted and the microsomal pellet was resuspended in 0.1 M Tris-HCl pH 7.4, containing 15% glycerol and complete protease inhibitor cocktail tablet. The microsomes were aliquoted and stored at -80 °C until required. Figure 2.1 illustrates the preparation of microsomal fractions.

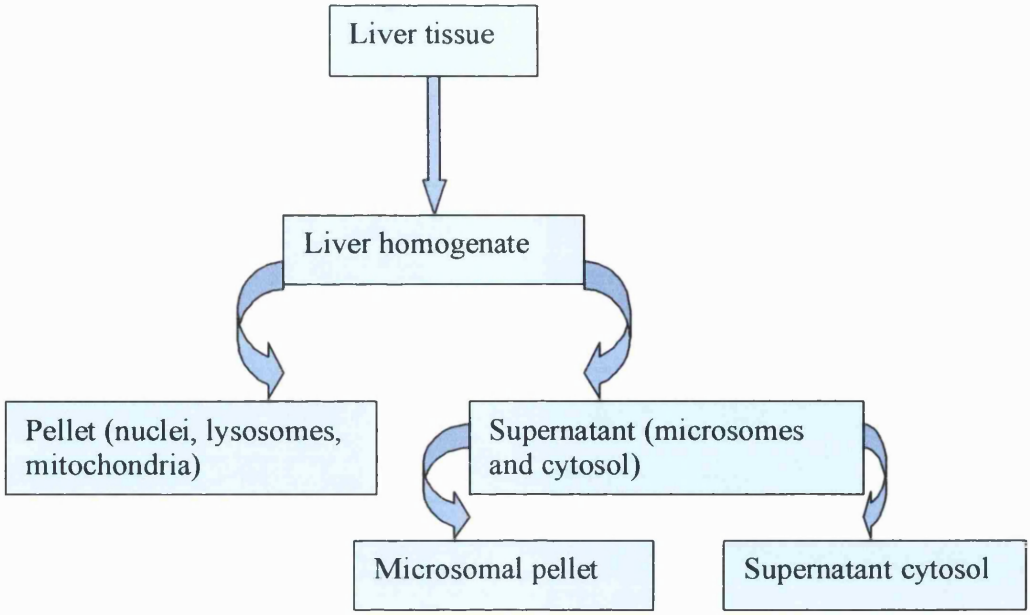


Fig 2.1: Diagrammatic representation of the subcellular fractions prepared by differential centrifugation of liver tissue: homogenate, S9 fractions, microsomes, and cytosol

2.2.5. Preparation of microsomes from Xenograft samples:

Human xenografts (excised from mice) were supplied by Professor MC Bibby (University of Bradford). The xenografts included IG (ovary), RT, BE (brain), COLO (colon) and MDA-MB (breast). Microsomes were isolated from xenografts using the method described in section 2.2.4.

2.3. Bradford Assay

The protein content of microsomes isolated from rat liver, human liver, colorectal metastasis, xenograft samples and MCF-7 cells was determined using the Bradford assay (Bradford, 1976). This method involves the binding of Coomassie brilliant blue G-250 to protein, which causes a shift in the absorption maximum of the dye from 465 to 595 nm, and it is the increase in absorption at 595 nm which is monitored. A stock solution was prepared by dissolving 30 mg of Coomassie brilliant blue G-250 in 100 ml of absolute

ethanol and 50 ml of concentrated phosphoric acid (10 M). The resulting solution was diluted to a final volume of 1litre with distilled water. This Bradford reagent was filtered before use to remove any undissolved dye.

The spectrophotometer was set to A_{595} and zeroed using a blank. A blank was made using 100 µl of water with 1 ml of Bradford reagent. Bovine serum albumin (BSA) was used to make up standards of varying concentrations. The concentrations used were 10 µg/ml, 20 µg/ml, 30 µg/ml, 40 µg/ml, 50 µg/ml, and 60 µg/ml. 100 µl of aliquoted BSA standard were mixed with 1 ml of Bradford reagent and the absorbance measured. A calibration curve was constructed using BSA protein. The gradient was calculated. Unknown protein samples were diluted with water until the protein concentration fell within the concentration range of the BSA standard curve. 1 ml of Bradford reagent was then added to the sample solution and mixed. The absorbance was measured at 595 nm and the protein concentration calculated using the gradient from the calibration curve.

(A595 of unknown sample / Gradient from calibration Curve) x dilution factor = protein concentration (mg/ml)

The above equation was used to calculate the protein concentration.
25-30 µg of microsomal proteins was separated on a SDS-PAGE.

2.4. Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

One dimensional SDS-PAGE was performed using standard methods on the Hoefer Mighty Small gel system (Amersham Biosciences, Buckinghamshire, UK). The method of Laemmli (1970) was used routinely for the analysis of proteins.

Reagents used for SDS-PAGE

A stock solution of Tris-glycine buffer (500 ml) was prepared by dissolving 7.55 g Tris-HCl, 47 g of glycine and 25 ml of 10% SDS were mixed in ~400 ml of distilled water, and then made up 500ml with water. Tank buffer was prepared by adding160 ml of stock Tris-glycine buffer in distilled H_2O to give an 800 ml final volume.

1M Dithiothreitol (DTT) stock solution was made up by mixing 154 mg of DTT in 1 ml of distilled water. This was stored in 25 µl aliquots at -20°C. Ammonium persulphate (10%) was prepared by weighing 0.1 g ammonium per sulphate in 1 ml water. This was stored in 400 µl aliquots. Solubilisation buffer was prepared by dissolving 2% SDS, 20% glycerol and 0.01 % bromophenol blue in 50 mM Tris-HCl, pH 6.8. The solubilisation buffer was filtered using a 0.22 µm filter to remove any undissolved dye. Ultra pure protogel (30% acrylamide (w/v), 0.8% (w/v) bis-acrylamide (37.5:1) was purchased ready mixed from National Diagnostics.

Table 2.1: Formulation for SDS-PAGE Resolving and Stacking Gels:

Reagents	*Resolving gel*		*Stacking gel*
Monomer Concentration	10%	12%	4.0%
Protogel	5 ml	6 ml	867 µl
Distilled H_2O	3.75 ml	2.75 ml	3.5 ml
1.M Tris-HCl, pH 8.8	6 ml	6 ml	-
1M Tris-HCl, pH 6.8	-	-	676 µl
10% SDS	120 µl	120 µl	52 µl
10% ammonium per sulphate	120 µl	120 µl	52 µl
TEMED	12 µl	12 µl	5.2 µl

A vertical gel apparatus was used to separate proteins. The electrophoresis separation was carried out at room temperature.

The glass plates to run the SDS-PAGE were washed in detergent, rinsed with tap water, distilled water, methanol and allowed to air dry. The resolving gel was prepared by mixing 1 M Tris buffer (pH 8.8), acrylamide/bisacrylamide (protogel), distilled water, 10% SDS, 10% ammonium per sulphate and TEMED as shown in table 2.3.1. The solution was immediately poured to a level ¾ of the height of the glass plates. A layer of water saturated n-butanol was immediately introduced above the gel mixture to ensure a flat interface between the resolving gel and the stacking gel after polymerisation and to prevent the trapping of O_2 molecules that in turn prevents the polymerisation process. On completion of the polymerisation process, the stacking gel consisting of Tris buffer (1 M, pH 6.8), protogel, distilled water, 10% SDS, ammonium per sulphate and TEMED was poured on top of the resolving gel. A comb was introduced into the upper gel before polymerisation to form the sample wells. The comb was carefully removed after the completion of polymerisation.

Whilst the gel was setting, the samples were prepared. Samples were boiled at 95 °C for 8 min in a solubilisation buffer (2% SDS, 20% glycerol, 0.01 % bromophenol blue and 50 mM Tris-HCl, pH 6.8) and 0.1 M dithiothreitol, microcentrifuged at 10,000 xg, cooled and applied to the sample wells. Tank buffer was added to the upper and lower reservoirs of the electrophoresis tank and air bubbles trapped under the gel were removed using a syringe.

Electrophoresis was carried out at 20 mA and voltage of 70 V until the bromophenol blue indicator dye had crossed the stacking gel. At that point, the current was increased to 30 mA and the voltage increased to 110 V and the gel left to run until the dye had reached the base of the resolving gel.

Gels were stained with a Coomassie brilliant blue R-250 buffer (0.2 g Coomassie brilliant blue R-250 (0.2%), 30 ml methanol (30%), 10 ml acetic acid (10%) in 100 ml distilled water) for 1 hour. Staining of the protein was achieved by immersing the gel in destaining solution which consists of 30% methanol and 30% methanol with 10% acetic acid.

2.5. Western blotting

The proteins from the gel were transferred onto a nitrocellulose membrane by electrophoresis. All components of the blotting apparatus were equilibrated in tris-glycine buffer containing 20% methanol, including the gel and the nitrocellulose membrane. A 'sandwich' was assembled in a support cassette which consisted of a piece of pre-soaked sponge overlaid with two pieces of pre-soaked 3MM paper followed by the SDS-PAGE gel then the nitrocellulose membrane. A further pre-soaked 3MM paper was placed over this and then finally another piece of sponge (Fig 2.2). The blot 'sandwich' was electro-blotted for 1 hour at 150 mA and 100 V.

After 1hour blotting, the membrane was blocked in 10 mg/ml BSA in tris-saline (TS) overnight at 4°C.

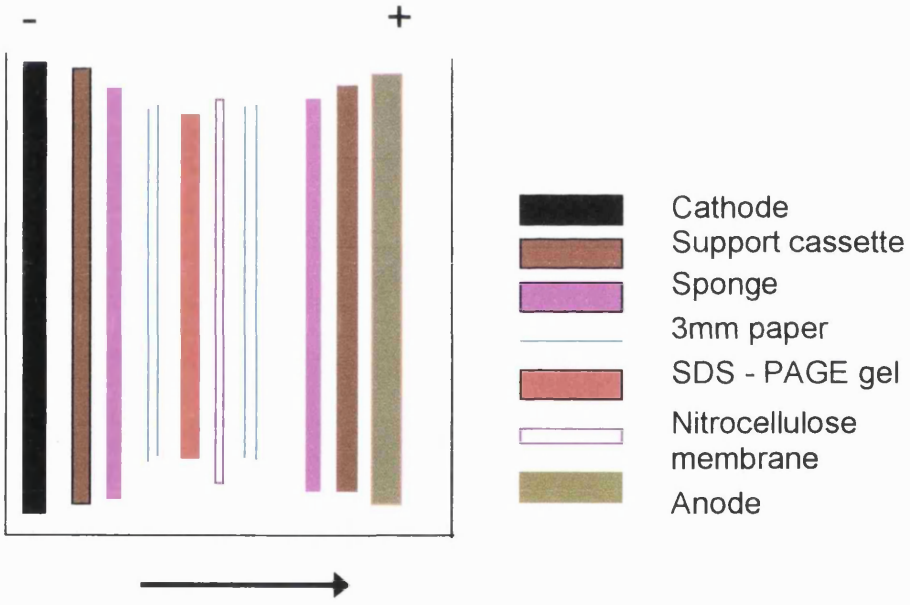


Fig 2.2: Apparatus used for the Western blot procedure.

Primary antibody solution was prepared as follows: 1/100 dilution of mouse anti CYP1B1 was made in 10 mg/ml BSA Tris saline with 0.05% NP 40. The blocking solution covering the membrane was then discarded and the membrane covered with the primary antibody. The blot was incubated with the primary antibody for 2 hours at room temperature. The primary antibody solution was discarded and the blot rinsed with TS

for 10 min. The TS was discarded and replaced with TS + 0.05% NP 40 and rinsed for 10 min. This was repeated. The blot was then washed in TS for 10 min. The secondary antibody solution was prepared by diluting a rabbit anti mouse horseradish peroxide conjugate by a factor of 1000 in 10 mg/ml BSA TS with 0.05% NP 40. The TS was discarded and replaced with TS + 0.05% NP 40 and rinsed for 10 min. This was repeated. The blot was then washed in TS for 10 min.

The blot was developed using DAB and urea H_2O_2 tablets.

2.6. Trypsin digestion and extracting peptides from polyacrylamide gel

The following protocols were used for trypsin digestion and extracting peptides from polyacrylamide gels.

2.6.1. Preparation of gel slices for tryptic digest:

The stained gel was washed with HPLC grade water and the band of interest was excised from the gel using a one-sided stainless steel razor blade. The protein band was cut as close to the stained band as possible to reduce the amount of "background" gel. A gel piece of approximately the same size was excised from a nonprotein-containing region of the gel as a control. The protein bands were transferred into 1.5 ml microfuge tubes. The gel slices containing the protein of interest were washed with 100 µl aliquots of water (with vortexing) repeatedly until the pH was neutral as estimated by spotting samples on a pH indicator paper. A rapid destain solution (100 µl, 50 mM NH_4HCO_3 in 40% ethanol, pH 7.8) was added, vortexed briefly, and then let stand at room temperature for 15 min. This step was repeated to allow the Commassie stain to be completely removed (as observed visually). The gel pieces were removed from the tube and placed onto the side of a new safety razor blade. HPLC grade H_2O (2-3 drops) was added to the gel slice to make it stick to the blade. The gel was chopped up with another razor into fine pieces. The chopped pieces were transferred into a microfuge tube with 25 µl of acetonitrile,

then vortexed and microcentrifuged. The tubes were incubated at room temperature for 15 min, and then the acetonitrile discarded. Fresh acetonitrile was added to the gel pieces. The acetonitrile was removed and the gel pieces were dried in a SpeedVac for 40 min.

2.6.2. Preparation of trypsin stock solution

Twenty microfuge tubes (0.5 ml) were washed several times with Millipore-grade water and labelled as 1 μg/5μl. Promega trypsin (sequencing grade modified trypsin provided in 20 μg aliquots) was prepared in 100 μl of 25 mM NH_4HCO_3. This 100 μl solution was divided into the labeled tubes in aliquots of 5 μl (1 μg each), immediately snap-frozen in liquid nitrogen and stored at –20°C.

2.6.3. Digestion of gel slices with trypsin

Dried gel particles were rehydrated with diluted trypsin. This was prepared by adding NH_4HCO_3 (25 mM) to an aliquot of 1 μg/5μl trypsin stock to give 75 ng/μl solution. Rehydration was performed while keeping tubes on ice (35 min) to avoid autolysis of the trypsin and to allow the enzyme to become fully absorbed into the gel. After incubation of samples containing trypsin on ice, 25 μl of NH_4HCO_3 (25 mM) solution was added to the trypsin-soaked gel pieces ensuring the gel was immersed. The samples were incubated overnight at 37°C.

2.6.4. Extracting digested peptides from trypsinised gel pieces

25-30 μl of a solution of TFA (5%) and acetonitrile (50%) in dH_2O was added to each digest, mixed well and microfuged for 2 min. The samples were sonicated for 6 min to extract the peptides. The tubes were microcentrifuged for a further 5 min to pellet the gel particles. The supernatant was transferred to a fresh tube. These steps were repeated, and the extracts pooled. The pooled extracts dried in a SpeedVac (approximately 3 hr).

2.6.5. Desalting the digest samples

The dried-down digests were rehydrated with a solution of 10 µl of water containing 0.2 µl of heptafluorobutyric acid, and sonicated for 5 min. A C18 Zip Tip (Millipore) was further wetted with 50% acetonitrile in water (aspirated a few times) using a 10 µl Gilson pipette. The Zip Tip was washed with 10 µl of methanol 3 times and equilibrated in freshly prepared 0.05% TFA in water. Sample was loaded onto the tip, pipette in from the top, and solvent expelled and collected. This process repeated for 10 times. Water-washed 0.5 ml microfuge tubes were already prepared for collecting Zip Tip eluate and washes. The sample loaded tip was washed with 10µl 0.05% TFA in water. This process was repeated 7 times. Finally, peptides eluted from the tip using 30 % of methanol, 0.5% formic acid into clean tube.

The procedure used for in-gel digestion is shown in the Fig below.

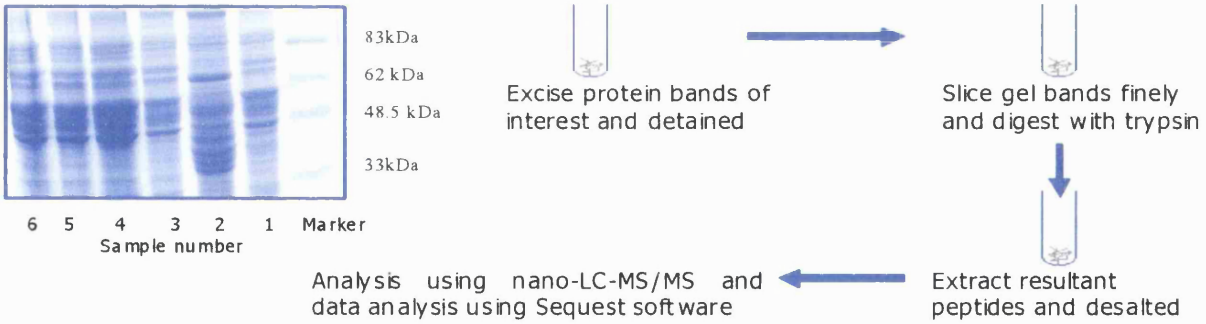


Fig 2.3: Procedure for in-gel digestion of proteins.

2.7. Nano-HPLC

Nano-scale LC was performed using an LC Packings Ultimate Capillary HPLC system with FAMOS autosampler (Dionex, Camberley, Surrey, UK). A separate Ultimate Micropump was employed as a loading pump. A 2.4 μL injection needle (Dionex) was used to inject 1 μL of sample into a 10 μL sample loop, and then onto a 1 mm x 300 μm PepMap™ C18 guard column (5 μm) (LC Packings) at 20 μL/min flow rate using 0.1 % TFA in water as loading solvent. The sample absorbed on the guard column was washed with 0.1% TFA (from the loading pump) for 3.5 min before the column was switched in-line and the sample eluted at a flow rate of 200 nL/min onto a 15 cm x 75 μm PepMap™ C18 column (3 μm) (LC Packings) previously equilibrated with 95% mobile phase A (5% acetonitrile containing 0.1% formic acid) and 5% mobile phase B (80% acetonitrile containing 0.1% formic acid). Five minutes after sample injection, the proportion of mobile phase B was increased linearly to 50% over 30 min, and then stepped to 95% and maintained at this level for 10 min (wash phase). The column was then re-equilibrated for 20 min with 95% mobile phase A, 5% mobile phase B (see table 2.2).

Table 2.2: Gradient programme used for the separation of tryptic digests.

Time (min)	*Flow rate (μL/min)*	*Mobile phase A Composition (%)*	*Mobile phase B Composition (%)*
0.00	0.20	95	5
5.00	0.20	95	5
30.00	0.20	50	50
30.02	0.20	5	95
40.00	0.20	5	95
40.02	0.20	95	5
60.00	0.20	95	5

The valve-switching diagram is shown in Fig 2.4 is a typical setup for 1-D nano LC.

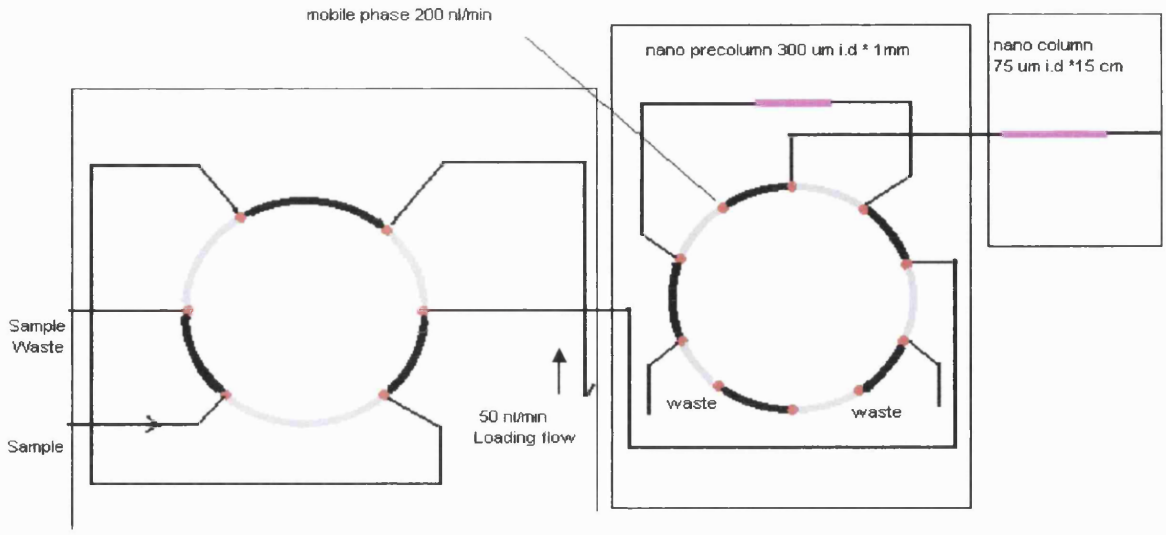


Fig 2.4: One-dimensional nano-LC column-switching flow diagram. The FAMOS is used for loading samples into the sample loop. The SWITCHOS is used for preconcentrating samples onto a precolumn. The Ultimate pump provides the solvent to elute the sample components from the precolumn onto the nano LC column.

The column effluent was continuously directed into an LCQ^{duo} mass spectrometer fitted with a nano-ESI source (ThermoElectron, Hemel Hempstead, UK) and spectra were recorded.

2.8. LCQ

The LCQ has an atmospheric pressure ionization source that was set at ESI (Electrospray Ionization) mode. The LCQ was operated with different scan functions such as Full Scan (from 50 to 2000 *m/z*), Selected Ion Monitoring (SIM), Selected Reaction Monitoring (SRM), MS/MS and Zoom Scan.

Static nanospray was performed automatically, this was done using Big 3 mode over 1 hour. In automatic mode the scan range used was *m/z* 400-2000. Data was collected in the full scan and data dependent MS/MS modes; 3 microscans (ms) were performed, with the maximum ion injection time of 200 ms. After a full scan, the data dependent

MS/MS was performed firstly on the most intense ion from that scan, then on the second most intense ion and lastly on the third most intense ion from the initial scan. While collecting MS and MS/MS data some contaminants peaks were regularly observed in the spectra. These contaminant peaks were added to a reject mass list: *m/z* 445.90, 415.95, 515.40, 610.99 and 536.95. Therefore, when the scan was performed these mass peaks were ignored.

The data system of the LCQ used a Finnigan LCQ data processing and instrument control software called Xcalibur®. Xcalibur® software is a flexible Windows based data system which integrates instrument setup, acquisition, data processing and the data files generated can be quickly reviewed with the Qualitative browser.

Mass spectrometer conditions were optimised using in-solution tryptic digests of purified recombinant CYP isoforms 1A2, 2E1 and 3A4, obtained from PanVera (Madison, Wisconsin USA).

Nano electrospray was performed in two modes: Static nanospray, and LC-ES-MS/MS

2.8.1. Static nanospray

The static nanospray source is fashioned after the work of Wilm and Mann, and allows analysis of low volume solutions over extended periods of time. The static nanospray was performed using 0.5 μm (tip ID) glass capillaries (New Objective, Inc., Woburn, MA, USA). Sample volumes of 5-10 μL were loaded in to these glass tips using Gel-Saver pipette tips (Fisher Scientifics). The LCQ was operated in the positive ion mode and the parameters used when data collected manually were: spray voltage, 1-1.2 kV; capillary temperature 180 °C; mass range 400-2000; sheath gas and auxiliary gas were kept zero. Static nanospray was used for two major purposes.

Firstly, to carry out diagnostic tests to ensure that LCQ was working optimally using a L-methionly-arginyl-phenylalanyl-alanine acetate (MRFA) tuning solution, 5 pmol/uL MRFA in 1% acetic acid methanol/water (50:50). Secondly, to optimise the tuning

conditions using the auto tune function. In order to optimise relevant tuning conditions CYP2E1 trypsin digest was used.

2.8.2. LC-ES-MS/MS

This mode was used while connected to nano LC columns. The emitter tips used for this application were 50 μm fused silica (15 μm tip diameter). In this mode the data was collected as full scan of *m/z* 400-2000 and then the data dependent MS/MS on the first, second, and third most intense ions from each mass scan.

ESI was performed under the following conditions: positive ionisation mode; spray voltage, 1.8 kV; capillary voltage, 28 V; capillary temperature, 180 °C; no sheath or auxiliary gas used. Data was collected in the full scan and data dependent MS/MS modes; 3 microscans (ms) were performed, with the maximum ion injection time of 200 ms. In the full scan mode, ions were collected in the *m/z* range 400 to 2000. The MS/MS collision energy was set to 35%.

2.9. Protein Identification

MS/MS spectra containing fragmentation information were searched using SEQUEST.

SEQUEST:

When LC-ESI-MS/MS was performed for the analysis of trypsin-digested proteins, peaks were selected automatically for MS/MS. MS/MS spectra were searched using Sequest Browser software (Eng *et al.,* 1994; Yates *et al.,* 1995), against a rat protein database containing 21576 entries including 47 CYPs or a human protein database containing 196,000 entries of which 55 were CYPs (obtained from National Centre for Biotechnology Information (NCBI), Bethesda, MD, USA, July 2003). By definition, Sequest correlates uninterpreted tandem mass spectra of peptides with amino acid sequences derived from protein and nucleotide databases. Sequest determine the amino

acid sequence and thus the protein(s) and organism(s) that correspond most closely to the mass spectra being analysed.

Sequest requires FASTA formatted databases. These databases are in ASCII text format. There is a single header/description line per sequence entry. This header line is denoted by the first character of the line being the greater than '>' sign. The end of header is denoted by a carriage return-line feed (see Fig 2.6).

Sequest dta files created and the following parameters were specified for searching,

Scan limits (set by Experiment): scans to process

Mass range (set by Experiment): precursor *m/z* range to search Bottom and Top molecular weight (400-4000 Da)

Minimum TIC ($5.0e^5$) intensity of the MS^2 TIC

Minimum ions (15) the number of ion present in the MS/MS spectrum

Intermediate scans (25) the difference between the first and the last scan numbers that can be grouped to creat a dta file, grouped scans (1) the mimimum number of MS/MS scans for the same precursor ion, which should be used to create a dta file. One dta file created for all the scans generated for one mass peak. The minimum group count was set at 1 (the number of scans of the same peak) and minimum ion count at 15 (how many peaks must appear in the MS/MS to accept it).

The point behind grouping scans is to decrease search time and increase the quality of the MS/MS data by co-adding scans. There is a slight possibility that two different peptides that elute very close in time with similar precursor *m/z* could be grouped together mixing their MS^2s.

In ordered to run a Sequest database search an enzyme has to be specified, for example for a search against the human database using trypsin as an enzyme. Other search parameters, which need to be specified before the search starts, are, peptide mass tolerance (1) (this is the error in peptide mass measurement), maximum internal cleavage

(2) and precursor mass tolerance (1.4) mass measurements (± value) error to allow grouping.

The above parameters are recommended by the manufacturers (Thermo finnigan).

The main aim of the Sequest approach is to find the peptide sequence in a database that best explains the fragment ions present in a spectrum. Candidate sequences are found in the database based on intact peptide masses, and the complete or partial spectra expected to result from the fragmentation of these candidate peptides are generated and compared to the experimental spectrum. The Sequest search results were assessed by examination of the XCorr (cross correlation) and delCn (delta-normalized correlation) scores. The XCorr function measures similarity between the mass to charge ratios (*m/z*) of the observed ions with those predicted for amino acid sequences within the database. The delCn is a measure of the difference between the first and second ranked amino acid sequences predicted to correspond to the experimental spectrum. Proteins that were matched by two or more peptides with Xcorr values ≥2.5 and a delCn score greater than 0.1 were considered conclusively identified, provided that the peptides were unique to that protein in the database (Ducret *et al.,* 1998). Such positive identifications were confirmed by manual interpretation of the spectra. Once an amino acid sequence was identified from an MS/MS spectrum its protein of origin was established.

Sequest generates the results in form of a summary. Sequest_Summary has two applications. The first application of Sequest_Summary is to give a concise overview of a batch of search results without the necessity of having to look at each individual Sequest output files. The second application is to perform protein identification by noting which proteins are most prevalent in a set of Sequest output results. Each top scoring peptide is displayed along with the input mass calculated from the spectrum and the actual mass of the matched peptide, correlation scores, preliminary score and rank, number of ions matched in the preliminary scoring, and the protein reference/accession number. A sample output of Sequest_Summary appears as follows.

SEQUEST SUMMARY

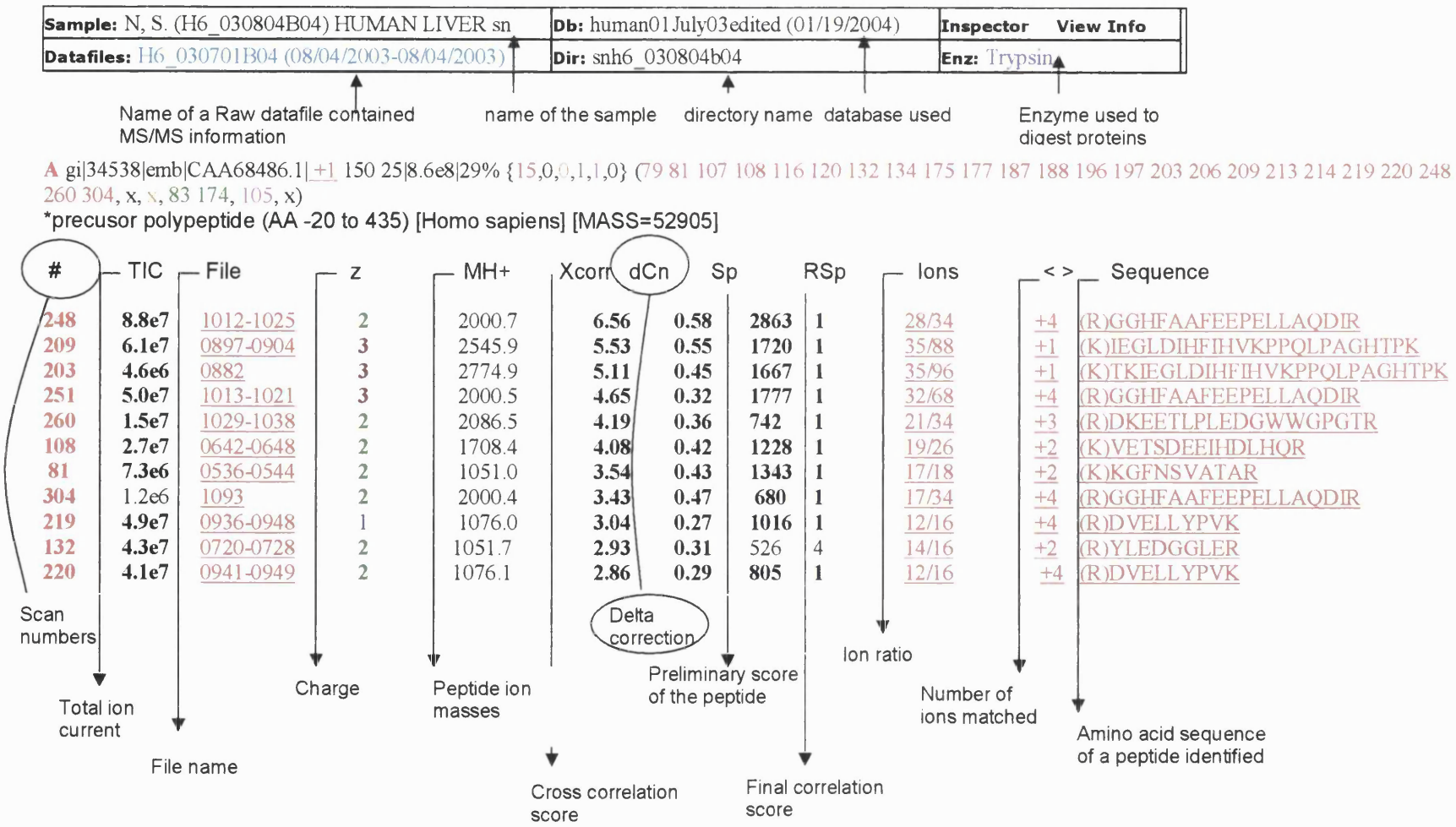


*Fig 2.5: A sample output of a Sequest_*Summary

In the above figure the MH+ column lists the peptide ion masses, z indicates the charge (singly, doubly or triply charged ion), RSp shows the final correlation score rank, Sp column indicates the preliminary score of the peptide and < > number of ions matched.

The individual protein sequence information with total % of protein identification is contained within the summary and an example of this summary is shown in Fig below.

>gi|**13699818**|ref|NP_000762.2| cytochrome P450, family 2, subfamily C, polypeptide 9; cytochrome P450, subfamily IIC (mephenytoin 4-hydroxylase), polypeptide 10; mephenytoin 4-hydroxylase; microsomal monooxygenase; xenobiotic monooxygenase; flavoprotein-linked monooxygenase; cytochrome P450, subfamily IIC (mephenytoin 4-hydroxylase), polypeptide 9 [Homo sapiens] gi|6686268|sp|P11712|CPC9_HUMAN Cytochrome P450 2C9 (CYPIIC9) (P450 PB-1) (P450 MP-4) (S-mephenytoin 4-hydroxylase) (P-450MP) gi|87271|pir||B38462 S-mephenytoin 4-hydroxylase (EC 1.14.14.-) cytochrome P450 2C9 - human gi|359735|prf||1313295A cytochrome P450 [MASS=55628]

MDSLVVLVLC LSCLLLLSLW RQSSGRGKLP PGPTPLPVIG NILQIGIKDI SKSLTNLSKV YGPVFTLYFG LKPIVVLHGY EAVKEALIDL GEEFSGRGIF PLAERANRGF GIVFSNGKKW KEIRRFSLMT LRNFGMGKRS IEDRVQEEAR CLVEELRKTK ASPCDPTFIL GCAPCNVICS IIFHKRFDYK DQQFLNLMEK LNENIKILSS PWIQICNNFS PIIDYFPGTH NKLLKNVAFM KSYILEKVKE HQESMDMNNP QDFIDCFLMK MEKEKHNQPS EFTIESLENT AVDLFGAGTE TTSTTLRYAL LLLLKHPEVT AKVQEEIERV IGRNRSPCMQ DRSHMPYTDA VVHEVQRYID LLPTSLPHAV TCDIKFRNYL IPKGTTILIS LTSVLHDNKE FPNPEMFDPH HFLDEGGNFK KSKYFMPFSA GKRICVGEAL AGMELFLFLT SILQNFNLKS LVDPKNLDTT PVVNGFASVP PFYQLCFIPV

Mass (average): 55627.9 **Identifier:** gi|13699818 **Database:** C:/Xcalibur/database//human01July03edited.fasta
Protein Coverage: 121/490 = **24.7%** by amino acid count, 13219.2/55627.9 = **23.8%** by mass

Fig 2.6: An example of an amino acid sequence generated for individual proteins in a FASTA format

The sequence coverage obtained for CYPs identified from rat, liver and tumours is the result of three LC runs each searched with Sequest. The total sequence coverage was calculated by summing the three Sequest summaries generated by three LC-runs obtained from the LCQ.

2.10. Optimisation of protein identification

Methods have been developed to improve sequence coverage for the CYP proteins. The main reason behind this aim was to improve protein characterization.

Multiple injections

A sample is repeatedly injected and analysed with data dependent MS/MS methods using a wide *m/z* range (400-2000) in five experiments. In between each injection a blank was run. In ordered to search five data files, five directories were made and dta files were created in the following manner,

1st directory	contained MS/MS information obtained from 1st injection only
2nd directory	1st and 2nd injection were grouped under this directory
3rd directory	1st, 2nd and 3rd injection were grouped
4th directory	1st, 2nd, 3rd and 4th injection were grouped
5th directory	1st, 2nd, 3rd, 4th and 5th injection were grouped

Narrowing scan range

A sample is repeatedly analysed by data dependent MS/MS methods using three narrow *m/z* ranges (e.g., 400-1000, 1000-1500 and 1500-2000) as survey scans from which to select ions for CID rather than a single full range (400-2000). The three data files were grouped under the same directory and Sequest searched.

Results from these three analyses were compared to those from triplicate LC-ES/MS/MS analysis over the same total *m/z* range, 400-2000 discussed in result.

2.11. Absolute Quantitation (AQUA) of proteins

In ordered to carry out quantitative studies, the AQUA strategy for the absolute quantitation of proteins (Gerber *et al.*, 2003) is now being used (work in progress).

In this method the in-gel digestion was carried out as described in 2.9 with the exception that the protein was proteolysed with ~1 μg trypsin (per band) in the presence of the AQUA internal standard peptide. The dried gel particles were rehydrated using AQUA internal standard peptide dissolved in the trypsin buffer and left for another 30 min to

allow the enzyme to become fully absorbed into the gel. After incubation overnight (~16 hours), gel pieces were centrifuged at 10,000 rpm for 5 min, and then the peptides were extracted twice with 5% formic acid / 50% acetonitrile. There were three slightly different procedures which were used to improve peptide recovery from the gel pieces.

The peptide was generated by solid-phase peptide synthesis such that one residue (for example leucine) was replaced with the same residue containing stable isotopes (for example six ^{13}C and one ^{15}N atoms). The result is a peptide that is chemically identical to its native counterpart formed by proteolysis, but is distinguishable by MS via a 7 Da mass shift. An abundant fragment ion for sets of native and internal standard peptides is chosen and then specifically monitored using a single reaction monitoring (SRM) method.

In ordered to validate the Gerber method horse heart myoglobin was used and peptide selected for analysis is LFTGHPETLEK. Two SRM transitions were monitored in *m/z* 636.4 → 716.4 (native peptide) and 639.9 → 723.4 (internal standard peptide).

An experiment was performed to show that the amounts of intact and the labelled peptide were accurately measured to run on SDS-PAGE. This was done using solution digests of intact myoglobins with AQUA peptide of the same concentrations were monitored using SRM.

The in-solution digestion of myoglobin (500 fmol) was carried out in a solution of 80 % acetonitrile and 25 mM ammonium bicarbonate buffer, pH ~ 8. Trypsin was added to the protein in a molar ratio of 1:50. The digestion reaction was carried out in an incubator at 37 °C (Russell *et al.,* 2001). The digest samples were then run on LC-ESI-MS/MS in a SRM mode.

Chapter 3

Results

Identification of Cytochrome P450 Isoforms in Tissues by Nano-scale Liquid Chromatography Electrospray Ionisation Tandem Mass Spectrometry

The aim of this study was the application of nano-scale LC-ESI-MS/MS to the analysis of tissue endoplasmic reticulum with emphasis on the identification of CYPs. Recognising the limitations of 2-D gel electrophoresis separation of membrane CYP proteins (Harry *et al.,* 2000), an alternative method of CYP separation was established. In this method SDS-PAGE was used for separating the endoplasmic reticulum proteins into discrete bands, followed by trypsin digestion of selected bands. Nano-scale LC separation of peptides followed by ESI-MS/MS was used to obtain amino acid sequence information. This sequence information was combined with the peptide mass data to identify proteins.

Validation of the method for identifying CYPs

An evaluation of the performance of the SDS-PAGE LC-ESI-MS/MS procedure for the detection and identification of CYPs was carried out. Known amounts of purified recombinant CYP isoforms 1A2, 2E1 and 3A4 were loaded onto the SDS-PAGE gel and proceeding through the analytical cycle allowed the identification of the proteins. Using this approach it was found that these three CYPs could be identified down to 1 pmol concentration on the gel. These recombinant CYPs were also used to optimise the mass spectrometer conditions including, CID voltage (35 V), spray voltage (1.9 kV), capillary temperature (180°C). Successful protein identification was defined as one in which two or more unique peptides were found with Sequest Xcorr values $\geq$ 2.5.

3.1. Rat Liver

In ordered to illustrate the use of this technique both male and female rats were investigated to determine whether this technique can distinguish gender specific multiple CYP protein expression.

Fig. 3.1 shows a SDS-PAGE gel of rat liver microsomes. Five bands of approximately equal size covering the molecular weight range 48-62 kDa were cut out and subjected to in-gel digestion with trypsin as described in method 2.6. The resultant peptides were extracted and analysed by LC-ESI-MS/MS.

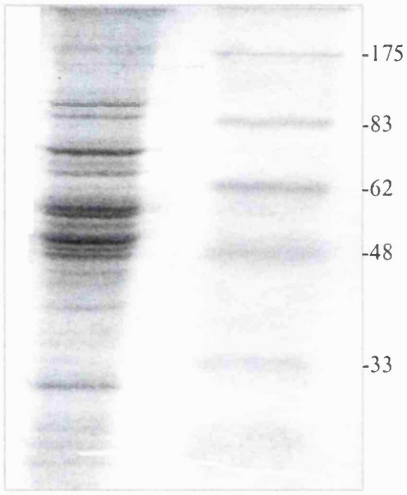


Figure 3.1: SDS-PAGE separation of rat liver microsomal proteins.

Fig 3.3A shows a base peak chromatogram of tryptic peptides from band 4 (52-56 kDa). The peptides as they eluted from the analytical column were directly infused into the mass spectrometer where the peptide ions were selected for fragmentation. To identify the proteins present the MS/MS spectral information were evaluated by the Sequest algorithm (Eng *et al.*, 1994). Sequest then identified the tryptic peptides by matching their MS/MS spectra against *in silico* generated theoretical spectra from the database. This is illustrated in Fig. 3.3C, which shows the MS/MS spectrum of the doubly charged peptide of *m/z* 815.2. Sequest determined the amino acid sequence of the peptide to be ***GTAVLTSLTSVLHDSK*** identified as a component of the CYP2C12 protein. The full amino acid sequence for 2C12 is given in Fig 3.2.

>gi|6978745|ref|NP_036841.1|| cytochrom P450 15-beta gene [gi|205932|gb|AAA41784.1| (M33550) cytochrome P450 15-beta (CYP2C12) [Rattus norvegicus] [gi|226780|prf||1605163A cytochrome P450 15beta [Rattus norvegicus] [MASS=55894]

MDPFVVLVLS LSFLLLLYLW RPSPGRGKLP PGPTPLPIFG NFLQIDMKDI RQSISNFSKT YGPVFTLYFG
SQPTVVLHGY EAVKEALIDY GEEFSGRGRM PVFEKATKGL GISFSRGNVW RATRHFTVNT LRSLGMGKRT
IEIKVQEEAE WLVMELKKTK GSPCDPKFII GCAPCNVICS IIFQNRFDYK DKDFLSLIEN VNEYIKIVST PAFQVFNAFP
ILLDYCPGNH KTHSKHFAAI KSYLLKKIKE HEESLDVSNP RDFIDYFLIQ RCQENGNQQM NYTQEHLAIL
VTNLFIGGTE TSSLTLRFAL LLLMKYPHIT DKVQEEIGQV IGRHRSPCML DRIHMPYTNA MIHEVQRYID
LAPNGLLHEV TCDTKFRDYF IPKGTAVLTS LTSVLHDSKEFPNPEMFDPG HFLDENGNFK KSDYFMPFSA
GKRKCVGEGL ASMELFLFLT TILQNFKLKS LSDPKDIDIN SIRSEFSSIPPTFQLCFIPV

Fig 3.2: Amino acid Sequence for CYP2C12

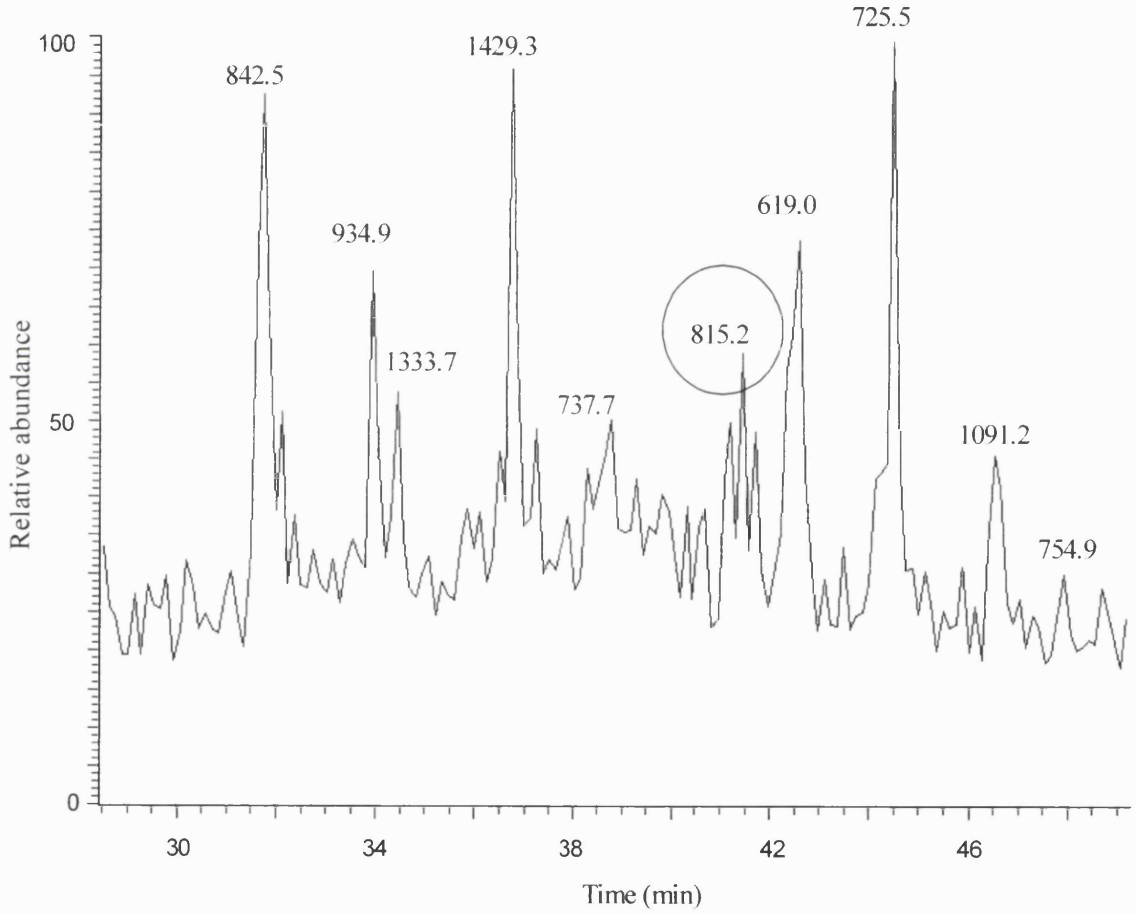
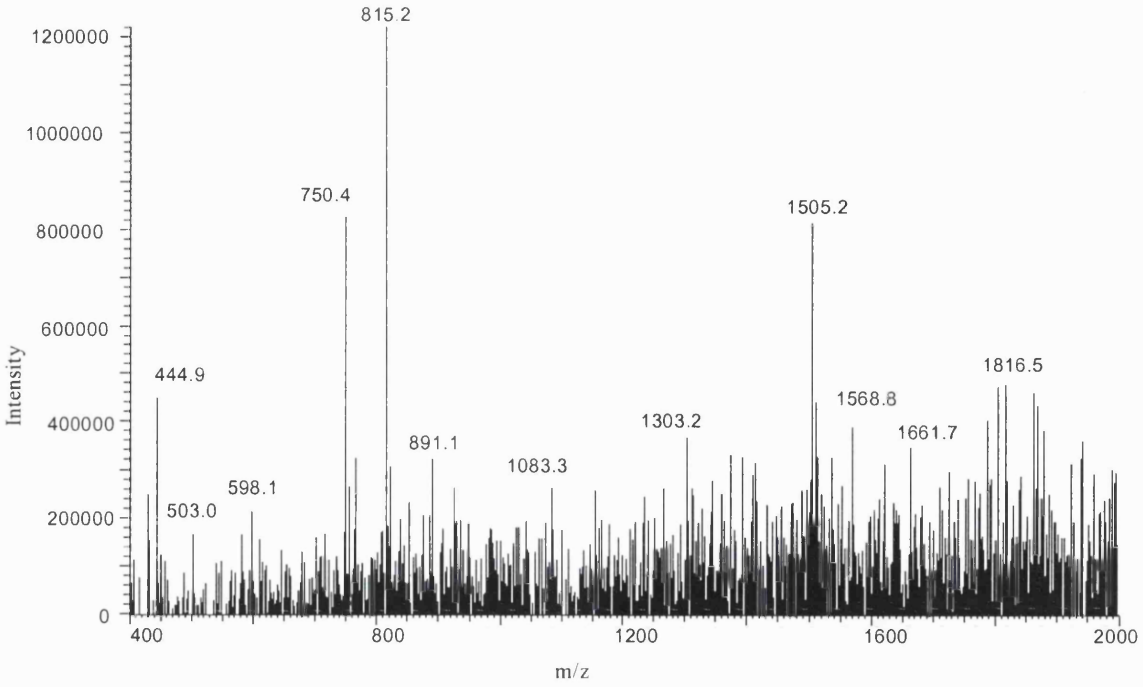
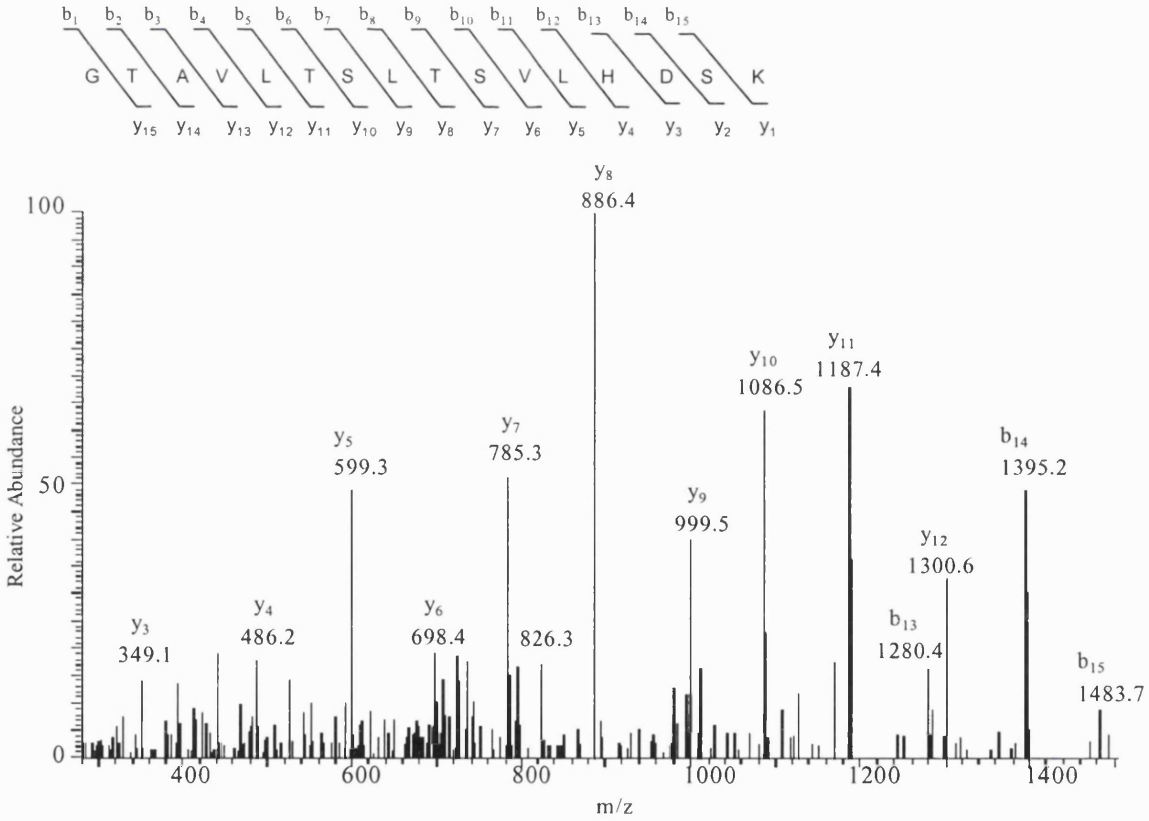


Fig 3.3 A. Base peak chromatogram of the chromatographic peak containing the tryptic peptide of m/z 815.2 is shown encircled.



B. MS spectrum of the chromatographic peak eluting at 41.2 min.



C. MS/MS spectrum for 815.2

*Fig 3.3: A. LC-ESI-MS base peak chromatogram of tryptic peptides from the 52- 56 kDa band in Fig 3.1, B. MS spectrum of the chromatographic peak eluting at 41.2 min, C. MS/MS spectrum of m/z 815.2 $[M + 2H]^{2+}$ identified as **GTAVLTSLTSVLHDSK**, to originate from CYP2C12. Shown in the inset of the MS/MS spectrum is the peptide's amino acid sequence.*

Similarly, the tryptic digest samples from male rat liver were analysed. The MS/MS spectrum of a doubly charged peptide of *m/z* 876.7 is shown in Fig 3.5D. Sequest determined the amino acid sequence of the peptide to be ***YIDLVPTNLPHLVTR***, from the CYP2C11 protein (Fig 3.4).

>gi|117228|sp|P08683|CPCB_RAT CYTOCHROME P450 2C11 (CYPIIC11) (P-450(M-1)) (P450H) (P450-UT-A) (UT-2) gi|92127|pir||A26685 cytochrome P450 2C11 - rat gi|203868|gb|AAA41062.1| (J02657) cytochrome P-450(M-1) [Rattus norvegicus] [MASS=57181]

MDPVLVLVLT LSSLLLLSLW RQSFGRGKLP PGPTPLPIIG NTLQIYMKDI GQSIKKFSKV YGPIFTLYLG
MKPFVVLHGYEAVKEALVDL GEEFSGRGSF PVSERVNKGL GVIFSNGMQW KEIRRFSIMT LRTFGMGKRT
IEDRIQEEAQ CLVEELRKSKGAPFDPTFIL GCAPCNVICS IIFQNRFDYK DPTFLNLMHR FNENFRLFSS PWLQ
VCNTFP AIIDYFPGSH NQVLKNFFYIKNYVLEKVKE HQESLDKDNP RDFIDCFLNK MEQEKHNPQS EFTLESL
VAT VTDMFGAGTE TTSTTLRYGL LLLLKHVDVTAKVQEEIERV IGRNRSPCMKDRSQMPYTDAVVHEIQRYID
LVPTNLPHLV TRDIKFRNYF IPKGTNVIVS LSSILHDDKEFPNPEKFDPG HFLDERGNFK KSDYFMPFSA
GKRICAGEAL ARTELFLFFT TILQNFNLKS LVDVKDIDTTPAISGFGHLPPFYEACFIPV QRADSLSSHL

Fig 3.4: Amino acid Sequence for CYP2C11

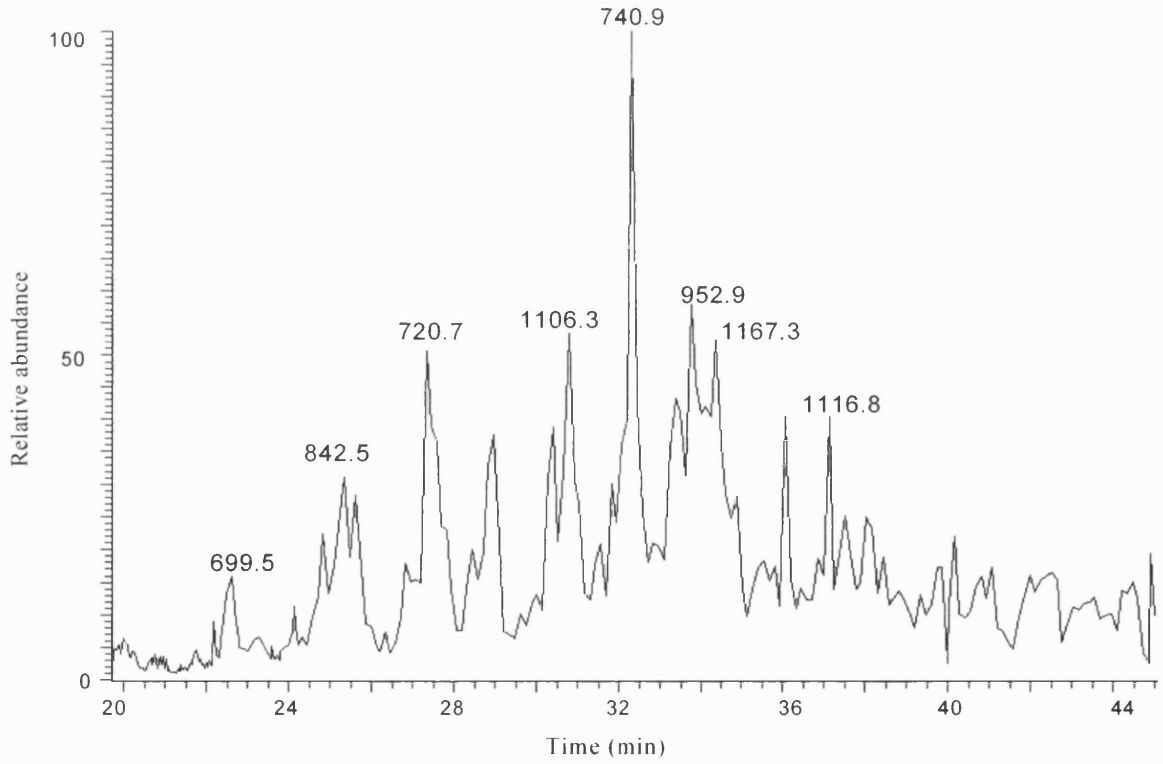


Fig 3.5 A. Base peak chromatogram of a tryptic digest obtained from male rat liver

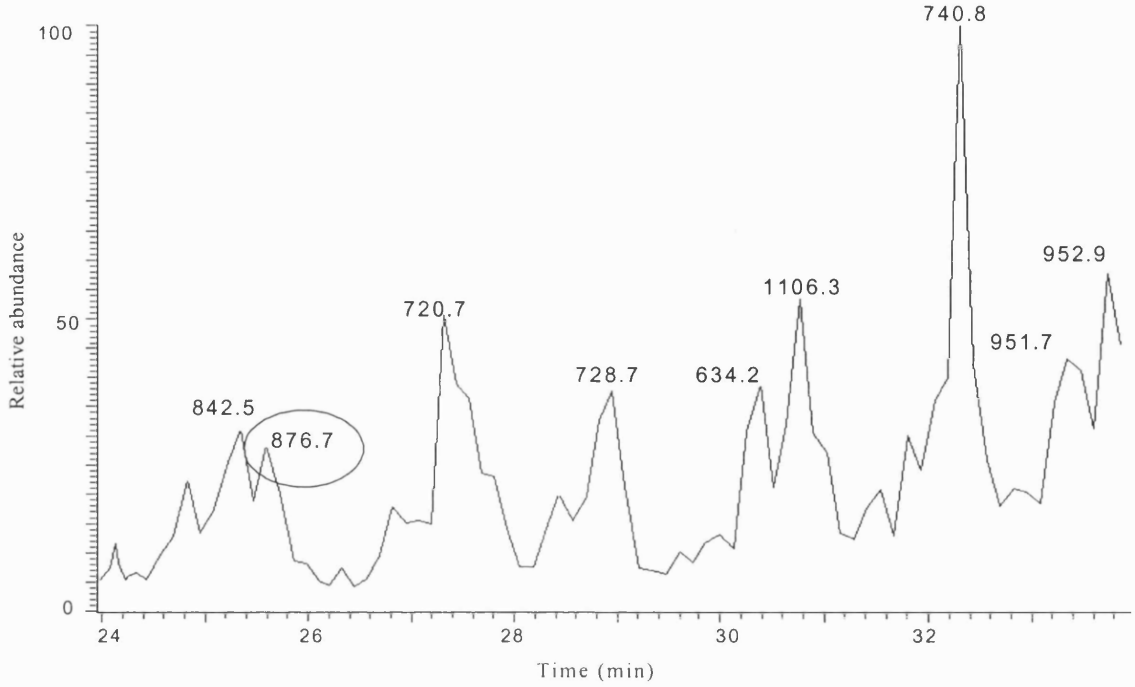


Fig 3.5 B. Expanded area showing base peak chromatogram where a peptide of m/z 876.7 (encircled) elutes at retention time of 25.7 min

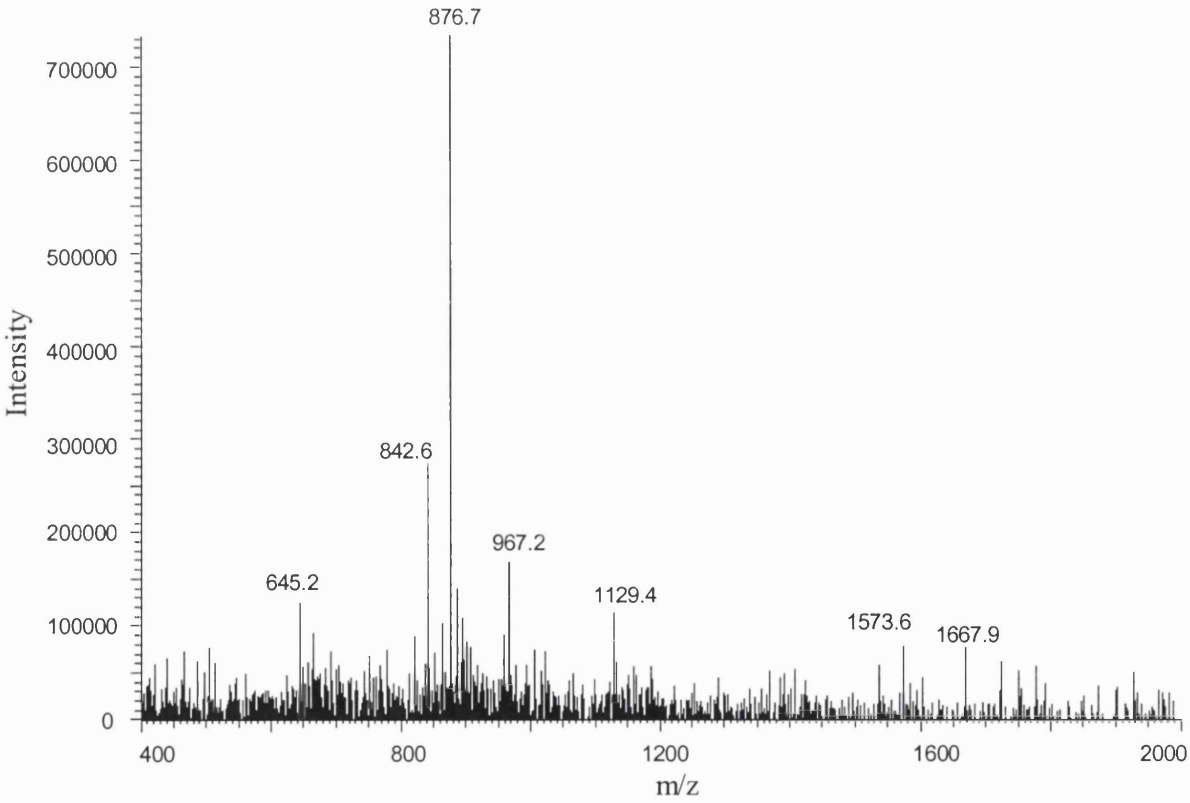
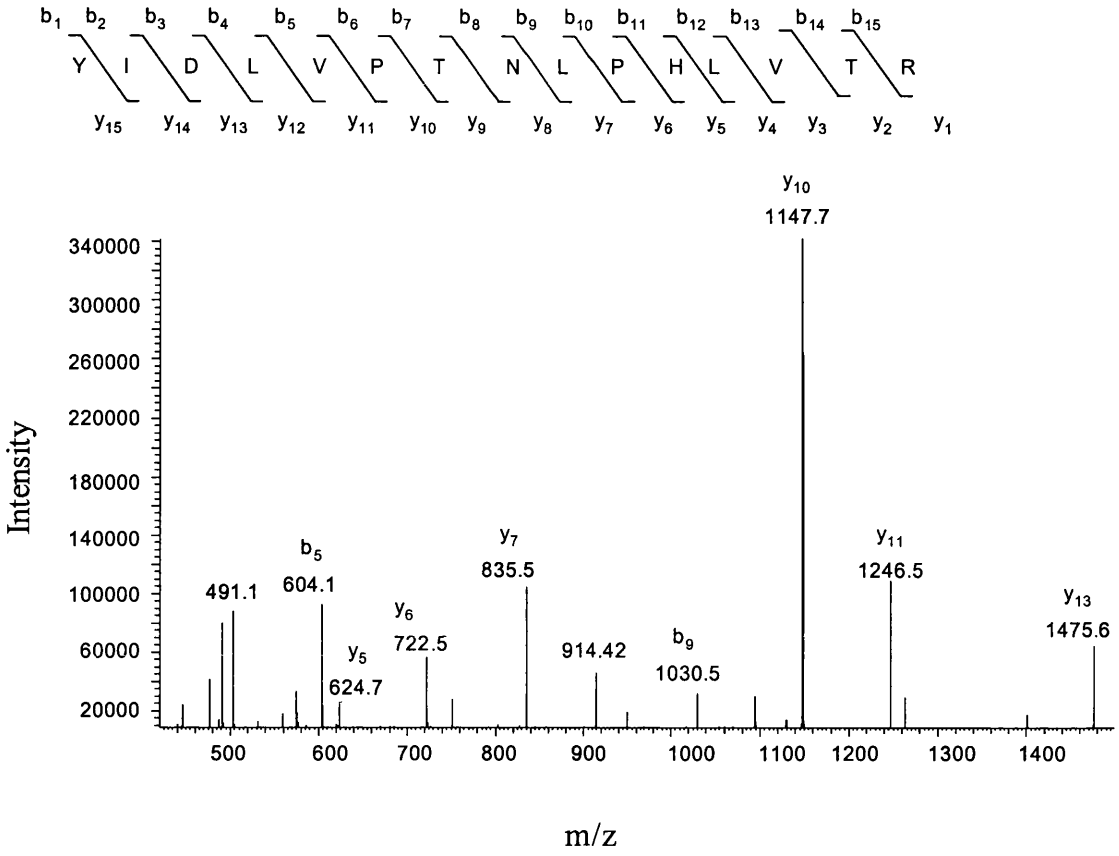


Fig 3.5 C. MS spectrum of the chromatographic peak eluting at 25.7 min



D. MS/MS spectrum of 876.7

Fig. 3.5: A. Base peak chromatogram, B. Expanded base peak chromatogram showing the elution of peptide m/z 876.7 at 25.7 min, C. MS spectrum of the chromatographic peak eluting at 25.7 min, D. MS/MS spectrum of the tryptic peptide ***YIDLVPTNLPHLVTR*** *$[M + 2H]^{2+}$ ion of m/z 876.7, identified to originate from a male specific CYP2C11 isoform.*

Using this approach a total of 24 CYP proteins, about half of the entire known rat CYPs, were identified in male and/or female Wistar rat liver as shown in Table 1.

For each of the CYP isoforms found, at least 2 unique peptides were identified with Xcorr scores of ≥2.5. It was possible to differentiate between closely related CYPs for example homologous CYP2D2 and 2D3 (77.8 %) (Fig 3.6). The differentiation was based on the peptides, which were unique to CYP 2D2 and 2D3. The tryptic peptides ^{92}ELLVTYGEDTADRPLLPIYNHLGYGNK^{118}, ^{170}EAEHPFNPSILLSK^{183}, ^{184}AVSNVIASLVYAR^{196}, ^{197}RFEYEDPFFNR^{207}, and ^{273}DMTDAFLAEMQK^{284} are unique to CYP2D2, whereas the tryptic peptides ^{273}DLTDAFLAEIEK^{284} and ^{287}GNPESSFNDANLR^{299} are unique to CYP2D3.

```
2D2 MGLLIGDDLWAVVIFTAIFLLLVDLVHRHKFWTAHYPPGPVPLPGLGNLLQVDFENMPYS 60
2D3 MELLAGTGLWPMAIFTVIFILLVDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLCNMPYS 60
     *  * **  ***   *  *     *  ***  **       * *        **

2D2 LYKLRSRYGDVFSLQIAWKPVVVINGLKAVRELLVTYGEDTADRPLLPIYNHLGYGNKSK 120
2D3 MYKLQNRYGDVFSLQMGWKPVVVINGLKAVQELLVTCGEDTADRPEMPIFQHIGYGHKAK 120
     *   **         **             *     *        **  ** *   * *

2D2 GVVLAPYGPEWREQRRFSVSTLRDFGVGKKSLEQWVTEEAGHLCDTFAKEAEHPFNPSIL 180
2D3 GVVLAPYGPEWREQRRFSVSTLRNFGVGKKSLEQWVTDEASHLCDALTAEAGRPLDPYTL 180
                           *            *  *    ****  ** ** **

2D2 LSKAVSNVIASLVYARRFEYEDPFFNRMLKTLKESFGEDTGFMAEVLNAIPILLQIPGLP 240
2D3 LNKAVCNVIASLIYARRFDYGDPDFIKVLKILKESMGEQTGLFPEVLNMFPVLLRIPGLA 240
     *   *      *     * *  * ***  *    *  *  ***    ** *  *    *

2D2 GKVFPKLNSFIALVDKMLIEHKKSWDPAQPPRDMTDAFLAEMQKAKGNPESSFNDENLRL 300
2D3 DKVFPGQKTFLTMVDNLVTEHKKTWDPDQPPRDLTDAFLAEIEKAKGNPESSFNDANLRL 300
    *    **** ***  ****    *   *     *       **            *

2D2 VVIDLFMAGMVTTSTTLSWALLLMILHPDVQRRVHEEIDEVIGQVRRPEMADQARMPFTN 360
2D3 VVNDLFGAGMVTTSITLTWALLLMILHPDVQCRVQQEIDEVIGQVRHPEMADQAHMPFTN 360
      *   *       *  *             *  **          *       *

2D2 AVIHEVQRFADIVPTNIPHMTSRDIKFQGFLIPKGTTLIPNLSSVLKDETVWEKPLRFHP 420
2D3 AVIHEVQRFADIVPMNLPHKTSRDIEVQGFLIPKGTTLIPNLSSVLKDETVWEKPLRFHP 420
                  * *  *     **

2D2 EHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARMELFLFFTCLLQRFSFSVLAGRPRPST 480
2D3 EHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARMELFLFFTCLLQRFSFSVPTGQPRPSD 480
                                                    ** *    *

2D2 HGVYALPVTPQPYQLCAVAR 500
2D3 YGVFAFLLSPSPYQLCAFKR 500
    *   * **** *      **
```

**Different amino acid sequence*
Highlighted sequences were identified for CYP2D2 and 2D3 using LC-MS/MS

Figure 3.6: Sequence alignment and sequence coverage map for CYP2D2 and 2D3.

This methodology revealed many differences in the detection of CYP protein expression based on rat gender (see Table 3.1).

Table 3.1

CYPs identified by MS/MS to be present in rat liver microsomes

Cytochrome P450 Identified	*SWISS-PROT Accession Numbers*	*Number of matched peptides (% sequence coverage by amino acid count)*	
		Male	*Female*
1A2	P04799	8 (19)	11 (28)
2A1	P11711	5 (16)	11 (30)
2A2	P15149	8 (15)	nd
2B3	P13107	nd	6 (17)
2C6	P05178	4 (15)	6 (18)
2C7	P05179	2 (11)	8 (19)
2C11	P08683	6 (18)	nd
2C12	P11510	nd	7 (25)
2C13	P20814	6 (18)	nd
2C22	P19225	4 (16)	nd
2C23	P24470	nd	6 (16)
2C24	P33273	nd	2 (8)
2D1	P10633	6 (17)	13 (35)
2D2	P10634	9 (24)	12 (34)
2D3	P12938	4 (12)	6 (18)
2D4	P13108	nd	5 (15)
2D5	P12939	7 (16)	nd
2E1	P05182	7 (25)	8 (28)
3A18	Q64581	3 (11)	nd
4A2	P20816	3 (12)	nd
4A3	P20817	4 (16)	3 (15)
4F1	P33274	3 (12)	3 (12)
CYP19	P11715	nd	5 (19)
CYP17	P22443	nd	4 (16)

nd = not detected

Note: the total number of peptides identified and the total sequence coverage given in the table is the result of three analyses.

The spectra with poor MS/MS patterns that is, spectra, which had fragment ion patterns too complex or too low in intensity to interpret, were named as unidentified peptides. The Sequest summary for each CYP isoform identified from one datafile searched using Sequest is given in Appendix i.

Although the subject of this study was the detection of CYPs, other proteins identified with high peptide numbers/sequence coverage from rat liver are shown in Table 3.2.

Table 3.2: Proteins identified with high sequence coverage from rat liver

Accession number	Protein identified	% Sequence coverage
P38659	Isomerase ER-60 precursor	42%
O88658	Kinesin	38%
P30839	Aldehyde dehydrogenase	28%
P10719	ATP synthase beta chain	26%
P10860	Glutamate dehydrogenase	24%
P07687	Epoxide hydrolase	22%
P36511	UDP-glucuronosyltranferase 2B12 precursor	21%
P08542	UDP-glucuronosyltranferase 2B3 precursor	18%
P10719	F1-ATPase beta chain	18%

3.2. Human Liver

Human liver tissue was obtained from redundant material procured from six different multi-organ donors. Consent from both the LREC authorities and from the families concerned was obtained. Hepatic microsomes were prepared from individual human livers as described in chapter 2. Microsomal proteins was subjected to SDS-PAGE, the region between 48 and 62 kDa cut into five bands, the proteins in the individual bands were subjected to trypsin digestion and the resultant peptides were analysed by nano-scale LC-ESI-MS/MS. The information about the starting mass of the tissue, which was used to isolate microsomes, was not available.

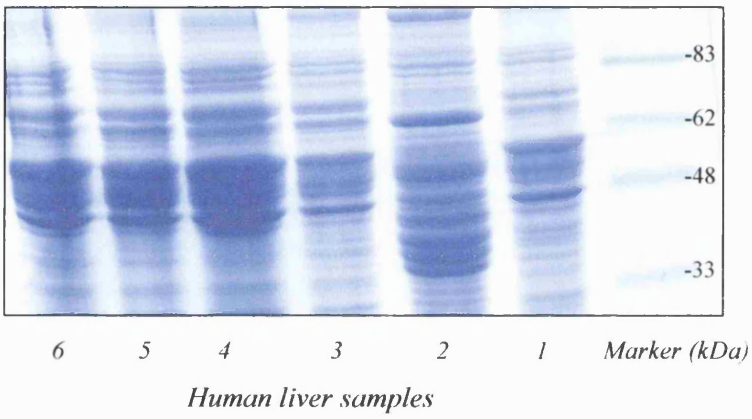


Figure 3.7: SDS-PAGE separation of human liver microsomes. The amount of the protein loaded onto the gel was approximately 25 μg.

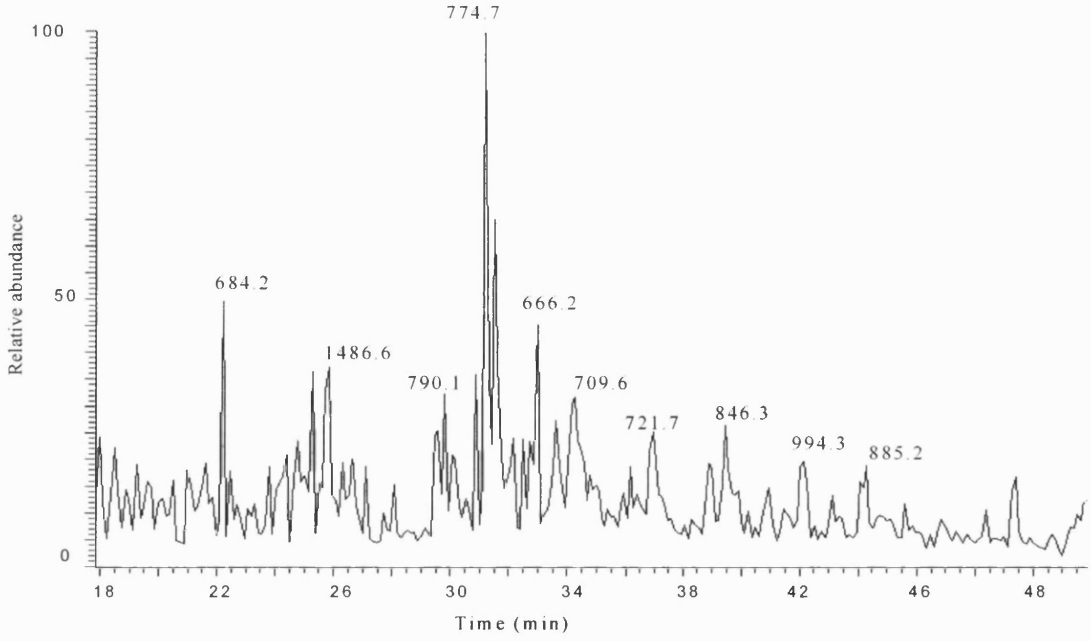


Fig 3.8. A. Base peak chromatogram of a tryptic digest obtained from human liver microsomes

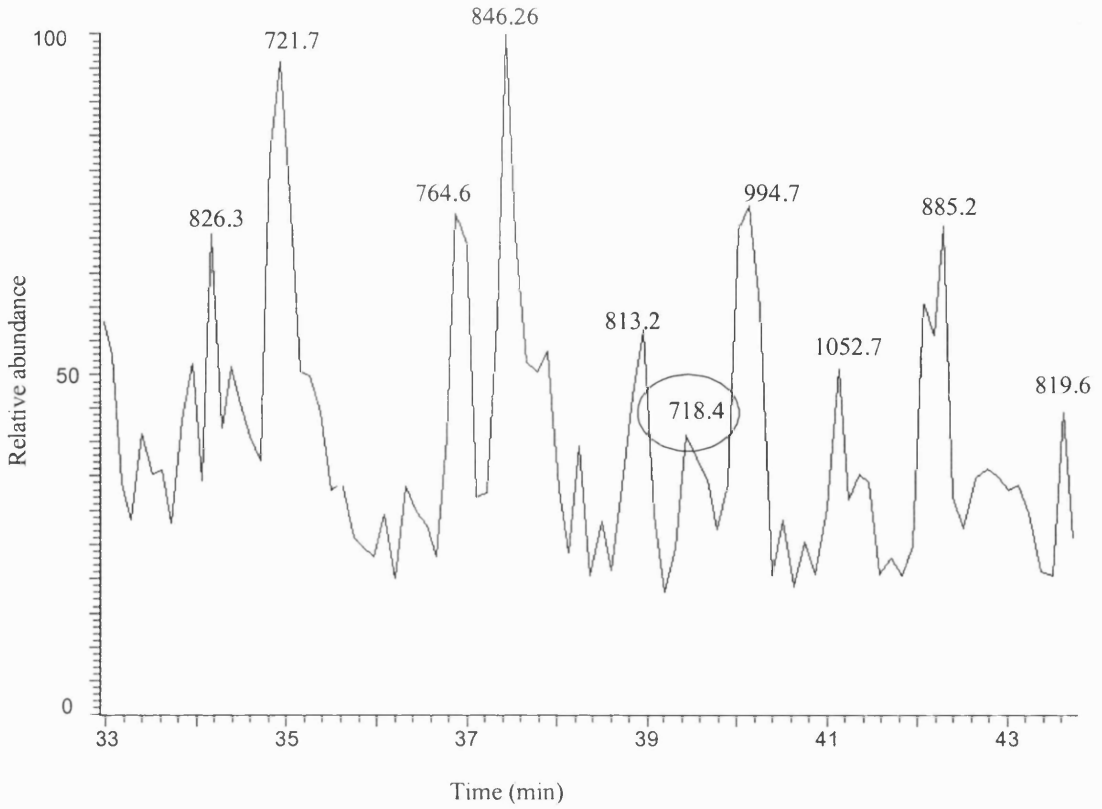


Fig 3.8 B. Expanded view showing where peptide ions of m/z 718.4 elute

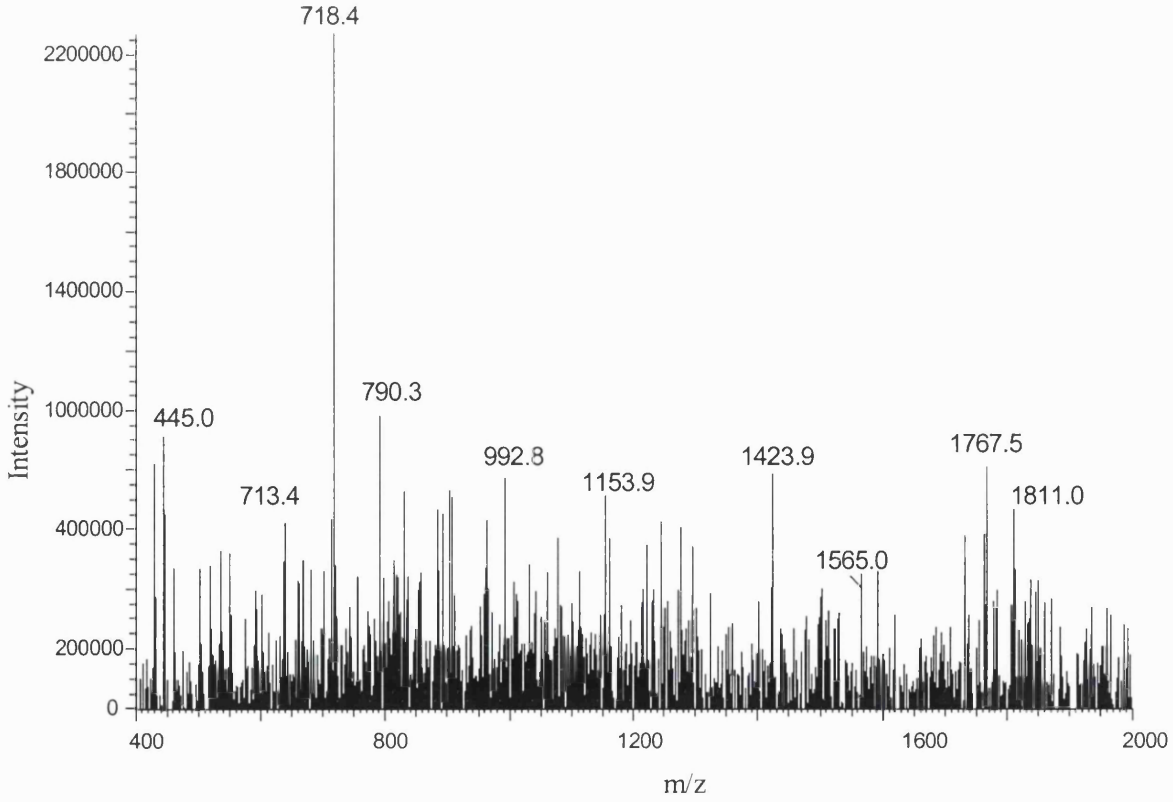
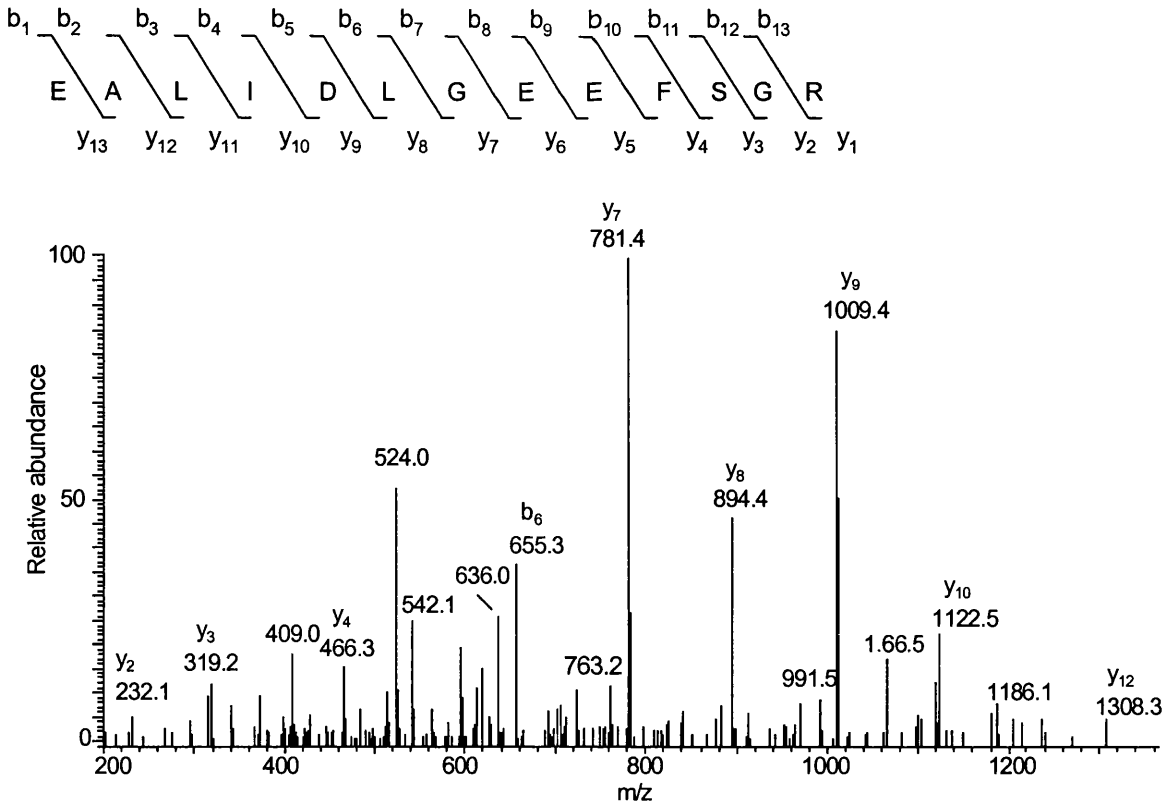


Fig 3.8 C. MS spectrum of the peak eluting at 39.2 min



D. MS/MS spectrum of a peptide fragment m/z 718.4

Fig. 3.8: A. LC-ESI-MS base peak chromatogram of tryptic peptides from the 52- 56 kDa band in figure 3.5, B. Expanded base peak chromatogram showing the elution of peptide ion of m/z 718.4, C. MS spectrum of the peptides eluting at 39.4 min, D. MS/MS spectrum m/z 718.4 $[M + 2H]^{2+}$ identified as ***EALIDLGEEFSGR****, to originate from CYP3A4.*

Each peptide was identified based on the cross correlation coefficient and Xcorr generated by the Sequest algorithm. See Table 3.3 for the number of peptides identified for each CYP.

Using this approach a total of 12 CYP proteins were identified in human liver as shown in Table 3.3.

Table 3.3: CYPs identified from human livers.

CYP isoform	Accession number	*Number of matched peptides (% Sequence coverage)*					
		Male Hu1	Female Hu2	Male Hu3	Female Hu4	Female Hu5	Female Hu6
1A2	P05177	11 (28)	10 (27)	8 (20)	8 (22)	5 (15)	4 (13)
2A6	P11509	-	5 (16)	-	7 (19)	-	5 (18)
2B6	P20813	-	-	-	-	-	3 (14)
2C8	P10632	-	-	-	-	-	5 (17)
2C9	P11712	12 (27)	10 (24)	11 (26)	4 (19)	4 (20)	4 (19)
2C17	P05093	3 (11)	-	-	-	-	-
2C18	P33260	4 (13)	-	5 (17)	-	-	-
2C19	P33261	3 (12)	2 (10)	-	3 (14)	-	-
2D6	P10635	5 (16)	3 (12)	-	-	-	3 (12)
2E1	P05181	13 (27)	16 (40)	6 (22)	7 (27)	4 (15)	4 (14)
3A4	P08684	12 (31)	10 (25)	15 (34)	4 (15)	11 (29)	17 (40)
3A5	P20815	-	4 (14)	-	-	3 (12)	-
4A11	Q02928	3 (14)	-	3 (12)	4 (15)	-	-
4F1	P78329	3 (13)	-	-	2 (9)	-	-
4F3	Q02928	5 (18)	4 (13)	4 (12)	-	3 (12)	-
Protein concentration (mg/ml of microsomal suspension)		2.3	2.4	2.6	2.7	2.6	2.7

- Not identified

Medical history:

Sex; age; ethnic origin and medical history

Hu1. Male; 35 yrs; Caucasian; social drinker, heavy smoker

Hu2. Female; 52yrs; Caucasian; hypertension, smoker

Hu3. Male; 62 yrs; Caucasian; colon surgery, morphine

Hu4. Female; 42 yrs; Caucasian; anaemia, asthmatic, vitamin b12, atenolol, salbutamol, aspirin

Hu5. Female; 55 yrs; Caucasian; no medication history

Hu6. Female; 39 yrs; Caucasian; no medication history

Note:
The total number of peptides identified and the total sequence coverage given in the table is the result of three separate data files generated on the LC-LCQ and then Sequest searched. The Sequest summary for each CYP isoform identified from one data file searched using Sequest is given in Appendix ii.

CYP1A2, 3A4, 2C9 and 2E1 were identified with high sequence coverage, which is consistent with the known high abundance of these CYPs in human liver (Shimada *et al.,* 1994). Further evidence to indicate that these CYPs are present in abundance in human liver is shown in the tryptic peptide chromatogram. The most intense peaks in the base peak chromatogram shown in Fig 3.9 correspond to *m/z* 685.5 (doubly charged peptide from CYP3A4) and 932.8 (doubly charged peptide from CYP1A2) (see Fig. 3.9).

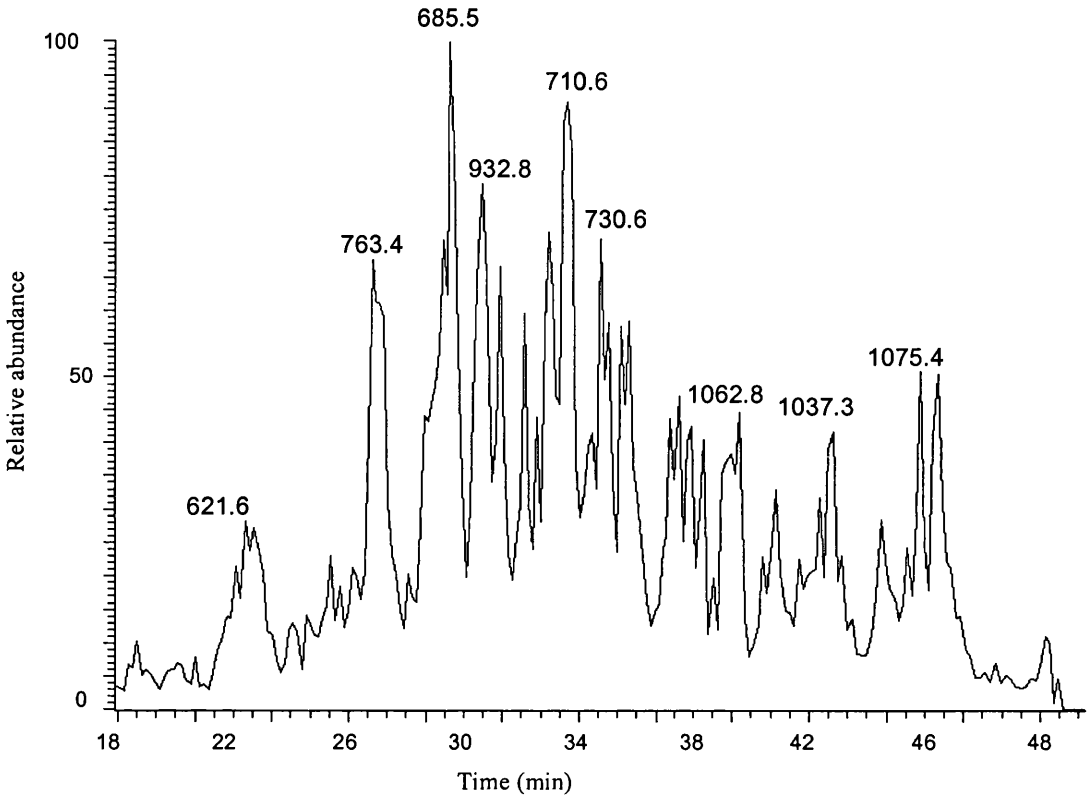


Fig 3.9: LC-MS/MS base peak chromatogram of tryptic peptides from a human liver sample from band 03 in Fig 3.5.

The MS spectra for peaks 932.8 and 685.5 are shown in Fig 3.10 and Fig 3.11 with normalized display mode values. In the normalize display mode, LCQ Duo always displays the largest peak in the spectrum at full scale. Therefore, the NL value gives the intensity of the signal of a mass of interest.

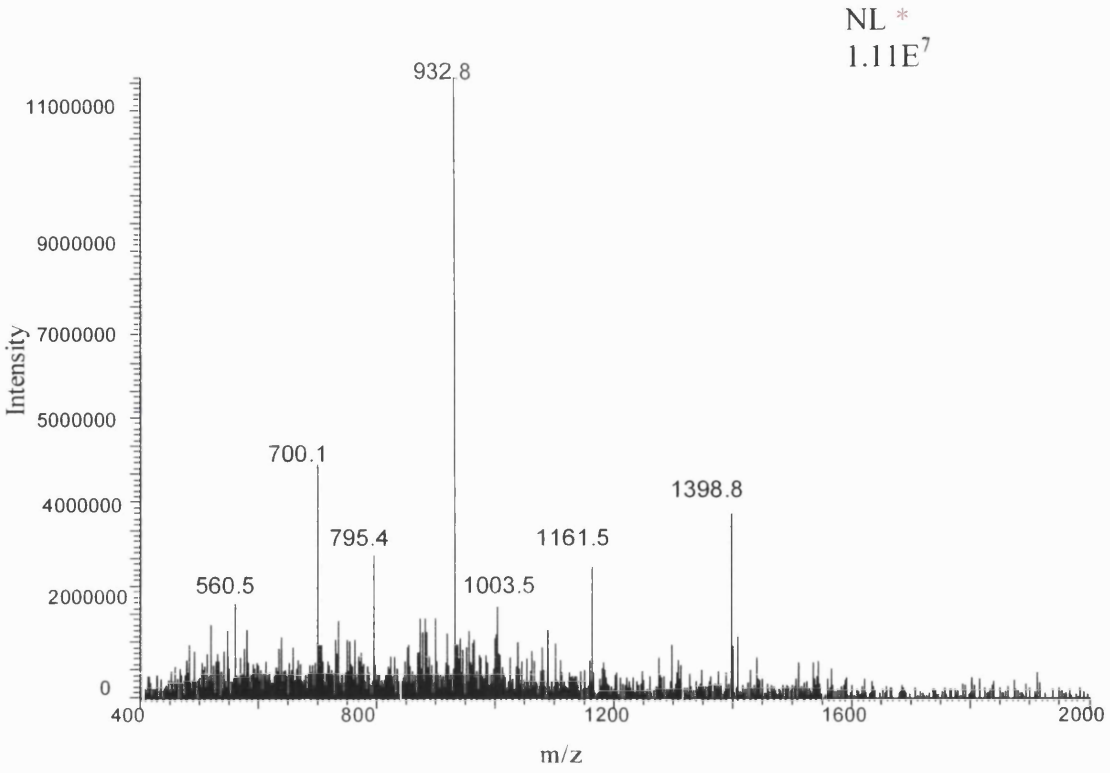


Fig 3.10: MS spectrum of the chromatographic peak eluting at 28.50, corresponding to a peptide of amino acid sequence, ***TVQEHYQDFDKNSVR*** *of m/z 932.8 $[M + 2H]^{2+}$, to originate from CYP1A2.*

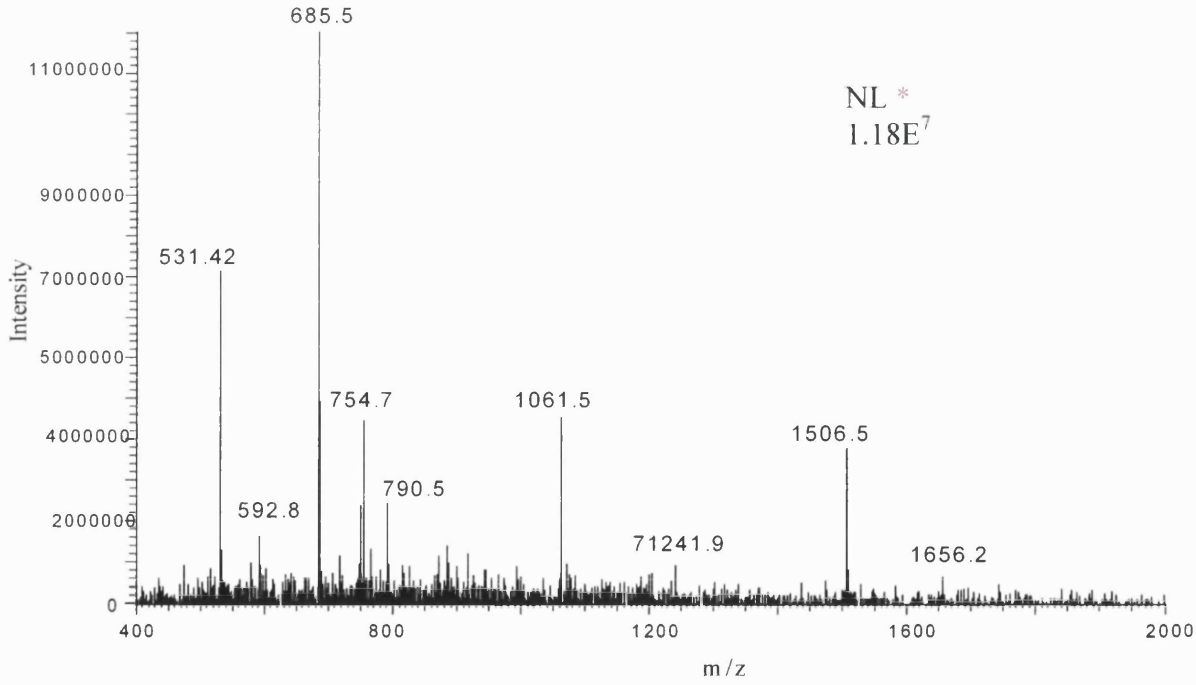


Fig 3.11: MS spectrum of the chromatographic peak eluting at 28.50, corresponding to a peptide of amino acid sequence, ***LQEEIDAVLPNK*** *of m/z 685.5 $[M + 2H]^{2+}$, to originate from CYP3A4*

* NL normalized display mode

It was possible to differentiate between closely related CYPs, for example the highly homologous CYP2C9 and 2C19 which have similarity of 91.8 % (see Fig 3.12). This differentiation was made based on peptides that were unique to CYP 2C9 and 2C19.

```
2C9    MDSLVVLVLCLSCLLLLSLWRQSSGRGKLPPGPTPLPVIGNILQIGIKDISKSLTNLSKV
2C19   MDPFVVLVLCLSCLLLLSIWRQSSGRGKLPPGPTPLPVIGNILQIDIKDVSKSLTNLSKI
          **                    *                    *          *

2C9    YGPVFTLYFGLKPIVVLHGYEAVKEALIDLGEEFSGRGIFPLAERANRGFGIVFSNGKKW
2C19   YGPVFTLYFGLERMVVLHGYEVVKEALIDLGEEFSGRGHFPLAERANRGFGIVFSNGKRW
                  ***       *                  *                   *

2C9    KEIRRFSLMTLRNFGMGKRSIEDRVQEEARCLVEELRKTKASPCDPTFILGCAPCNVICS
2C19   KEIRRFSLMTLRNFGMGKRSIEDRVQEEARCLVEELRKTKASPCDPTFILGCAPCNVICS

2C9    IIFHKRFDYKDQQFLNLMEKLNENIKILSSPWIQICNNFSPIIDYFPGTHNKLLKNVAFM
2C19   IIFQKRFDYKDQQFLNLMEKLNENIRIVSTPWIQICNNFPTIIDYFPGTHNKLLKNLAFM
          *                       * * *          **                *

2C9    KSYILEKVKEHQESMDMNNPQDFIDCFLMKMEKEKHNQPSEFTIESLENTAVDLFGAGTE
2C19   ESDILEKVKEHQESMDINNPRDFIDCFLIKMEKEKQNQQSEFTIENLVITAADLLGAGTE
       * *              *   *       *      *  *      * **  *  *

2C9    TTSTTLRYALLLLLKHPEVTAKVQEEIERVIGRNRSPCMQDRSHMPYTDAVVHEVQRYID
2C19   TTSTTLRYALLLLLKHPEVTAKVQEEIERVIGRNRSPCMQDRGHMPYTDAVVHEVQRYID
                                                  *

2C9    LLPTSLPHAVTCDIKFRNYLIPKGTTILISLTSVLHDNKEFPNPEMFDPHHFLDEGGNFK
2C19   LIPTSLPHAVTCDVKFRNYLIPKGTTILTSLTSVLHDNKEFPNPEMFDPRHFLDEGGNFK
         *           *              *                      *

2C9    KSKYFMPFSAGKRICVGEALAGMELFLFLTSILQNFNLKSLVDPKNLDTTPVVNGFASVP
2C19   KSNYFMPFSAGKRICVGEGLARMELFLFLTFILQNFNLKSLIDPKDLDTTPVVNGFASVP
          *                * *        *          *

2C9    PFYQLCFIPV
2C19   PFYQLCFIPV
```

**Different amino acid sequence*
Highlighted sequences were identified for CYP2C9 and 2C19 using LC-MS/MS

Figure 3.12: Sequence alignment and sequence coverage map for CYP2C9 and 2C19.

Other proteins, which were identified with good sequence coverage are shown in Table 3.4

Table 3.4: Proteins identified with high sequence coverage from human liver micrpsomes

Accession number	Protein identified	% Sequence coverage
P23141	Liver carboxylesterase 1 precursor	46%
P07099	Epoxide hydrolase 1	44%
P27797	Calreticulin precursor	43%
P07237	Protein disulfide isomerase precursor	41%
CAA15908	Flavin-containing monooxygenase 3	38%
P31513	UDP-glucuronosyltransferase 1 family	37%
P22309	ATP synthase	34%

3.3. CYP expression in normal human liver and colorectal metastases, normal colon and primary colon tumours

Microsomes were isolated from colorectal metastasis and livers of eight Caucasian donors using the method described in 2.2.4. Isolated microsomes were then separated by a SDS-PAGE as shown below.

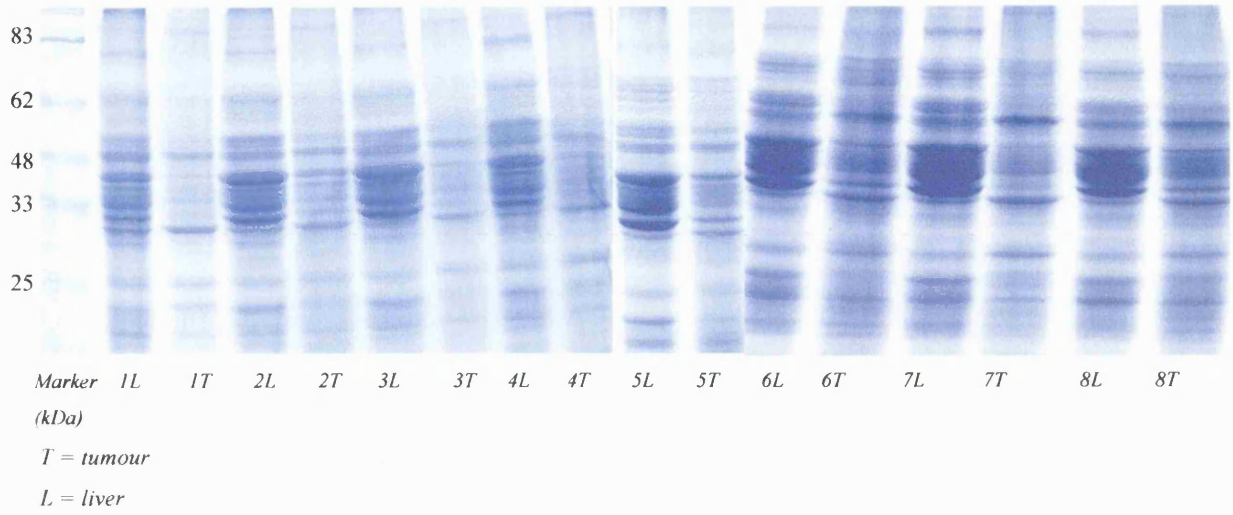


Figure 3.13: SDS-PAGE separation of human liver microsomes and colorectal metastases. The amount of the protein loaded onto the gel was approximately 25 μg except 1T (3.8 μg), 2T (10.8 μg), 3T (8.4 μg), 4T (10.8 μg) and 6T (14.4μg).

Peptides were extracted from the individual gel slices cut from 62-48 kDa region and each peptide mixture was subjected to LC-MS/MS. Individual peptide ions were automatically selected for MS/MS followed by correlation of the uninterpreted product ion spectra with predicted amino acid sequences present in a non redundant human protein sequence database using the Sequest search algorithm. A comparison of CYP expression by normal liver and colorectal metastases is shown the Table 3.6

Table 3.5: Microsomal protein concentrations for samples taken from patients with metastatic colorectal cancer of the liver

Sample	Mass of the tissue used to prepare microsomes (g)	Microsomal protein in tissue (mg/g)	Amount of protein loaded on the gel (μg)	mass of the tissue loaded on the gel (g)
1T	0.3	0.32	3.8	0.01
1L	0.68	2.0	24	0.012
2T	1.95	0.9	10.8	0.012
2L	1.82	5.6	28.0	0.005
3T	0.58	0.7	8.4	0.012
3L	0.58	2.6	26.0	0.01
4T	0.54	0.9	10.8	0.012
4L	0.86	4.4	26.4	0.006
5T	1.1	2.1	25.2	0.012
5L	2.2	3.5	24.5	0.007
6T	1.98	5.2	27.0	0.005
6L	2.0	5.4	27.0	0.005
7T	1.89	4.2	25.2	0.006
7L	2.3	8.5	25.4	0.0029
8T	2.40	3.5	27.2	0.007
8L	3.21	4.6	25.3	0.005

Table 3.6: CYPs identified in the microsomal fractions of liver and tumour samples from eight patients with metastatic colorectal cancer of the liver

Sample ID	# Protein content (mg/ml)	CYP isoform / % sequence coverage by amino acid count (number of peptides)												
		1A2	2A6	2B6	2C8	2C9	2C17	2C19	2D6	2E1	3A4	4A11	4F2	4F8
1T [a]	0.3	–	–	–	–	–	–	–	–	–	–	–	–	–
1L	2.1	6(1)	12(3)	–	12(4)	14(3)	–	–	10(2)	17(5)	19(5)	–	–	–
2T	0.9	14(3)	–	–	–	25(6)	–	–	–	25(6)	25(7)	–	–	–
2L	5.6	18(5)	15(4)	–	28(7)	37(12)	–	–	–	30(8)	31(10)	15(3)	–	–
3T	0.7	11(12)	–	–	–	6(1)	–	–	–	6(1)	12(3)	–	12(4)	8(3)
3L	2.7	21(5)	16(4)	–	–	30(8)	–	–	–	17(5)	22(6)	10(2)	–	–
4T	0.88	–	–	–	–	–	–	–	–	–	16(4)	–	–	–
4L	4.4	18(4)	14(4)	–	12(3)	28(6)	–	–	–	32(13)	33(9)	16(4)	–	–
5T	2.3	–	4(1)	–	–	–	–	–	–	–	11(2)	–	8(2)	8(2)
5L	3.5	12(3)	10(2)	–	–	6(1)	–	–	–	14(3)	13(3)	–	6(2)	3(1)
6T	1.2	4(1)	–	–	–	–	–	–	–	–	12(3)	–	–	–
6L	5.6	21(4)	26(7)	12(3)	15(3)	30(8)	16(4)	12(3)	5(1)	23(6)	31(10)	10(2)	4(1) [b]	4(1) [b]
7T	4.1	8(2)	18(4)	–	6(1)	31(8)	–	–	–	16(4)	10(2)	5(1)	6(2)	6(2)
7L	8.8	17(4)	48(16)	8(2)	22(5)	34(9)	–	–	3(1)	25(6)	31(9)	12(3)	3(1) [b]	3(1) [b]
8T	3.8	6(1)	–	–	16(4)	13(3)	–	–	3(1)	15(3)	5(1)	3(1)	7(2)	7(2)
8L	4.6	10(2)	15(3)	–	7(1)	27(6)	–	–	10(2)	18(4)	22(5)	–	–	–

 # Protein content mg/ml of microsomal solution determined using the Bradford assay (Bradford, 1976), [a] the microsomal protein concentration of sample 1T was too low for analysis, [b] peptides common to CYP4F2, 4F3, 4F11 and 4F12, - indicates that the CYP was not found

Medical history of patients with metastatic colorectal cancer of the liver

Hu1(1T/1L) 33 year old female, No drug history, 20 u/alcohol/wk, non-smoker, non drinker.

Hu2 (2T/2L) 73 year old male, prostatism. No drug history. 20u/wk alcohol. non-smoker. Previous liver resection. 5-Fluorouracil after excision colorectal cancer excision

Hu3 (3T/3L) 60 year old male, non drinker, Smoker 20/day, depression, hypertension. Lisinopril, bendrofluozide, olanzipine. 5-Fluorouracil after colorectal cancer excision

Hu4 (4T/4L) 75 year old male, alcoholic, non smoker, hypertension, non-insulin dependent diabetes mellitus (type II diabetes), chronic obstructive airway diseases. Perindopril, Gliglazide, omeprazole.

Hu5 (5T/5L) 64 year old female, hypercholesterolaemia, paroxysmal atrial fibrillation. atrovasstatin, aspirin, verapamil, nefidipine

Hu6 (6T/6L) 71 year old male, smoker, trype II diabetes, hypertension, bladder tumour (excised in 1998). Gliclazide, bumetanide, solpanol, candesartan, simvastatin

Hu7 (7T/7L) 59 year old female, non drinker, smoker, arthritis, drug history not available

Hu8 (8T/8L) 54 year old male, non drinker, non smoker, tonsillectomy and appendix as child, vasectomy 1977, hypertension. Losartan, bendrofluazide.

Using this method, thirteen CYP isoforms were identified from liver samples and ten from tumours (Table 3.6).

Two of the tumours (7T and 8T) were found to express the same pattern of CYPs as the liver, although generally fewer peptides were identified from the tumours. No CYP enzymes were identified in sample 1T, only CYP3A4 was identified in sample 4T, and only CYP1A2 and 3A4 were identified in sample 6T. The generally low level of CYP identification in tumour samples may be explained by the low tumour mass available for analysis and/or generally low protein content of the tumour (see Table 3.6). The chromatographic separations show that the number of proteins present in tumours located in the 48-62 kDa region was less than present in normal liver. Fig 3.14 is a representative example of the comparative chromatograms for liver and tumour obtained from the same individual.

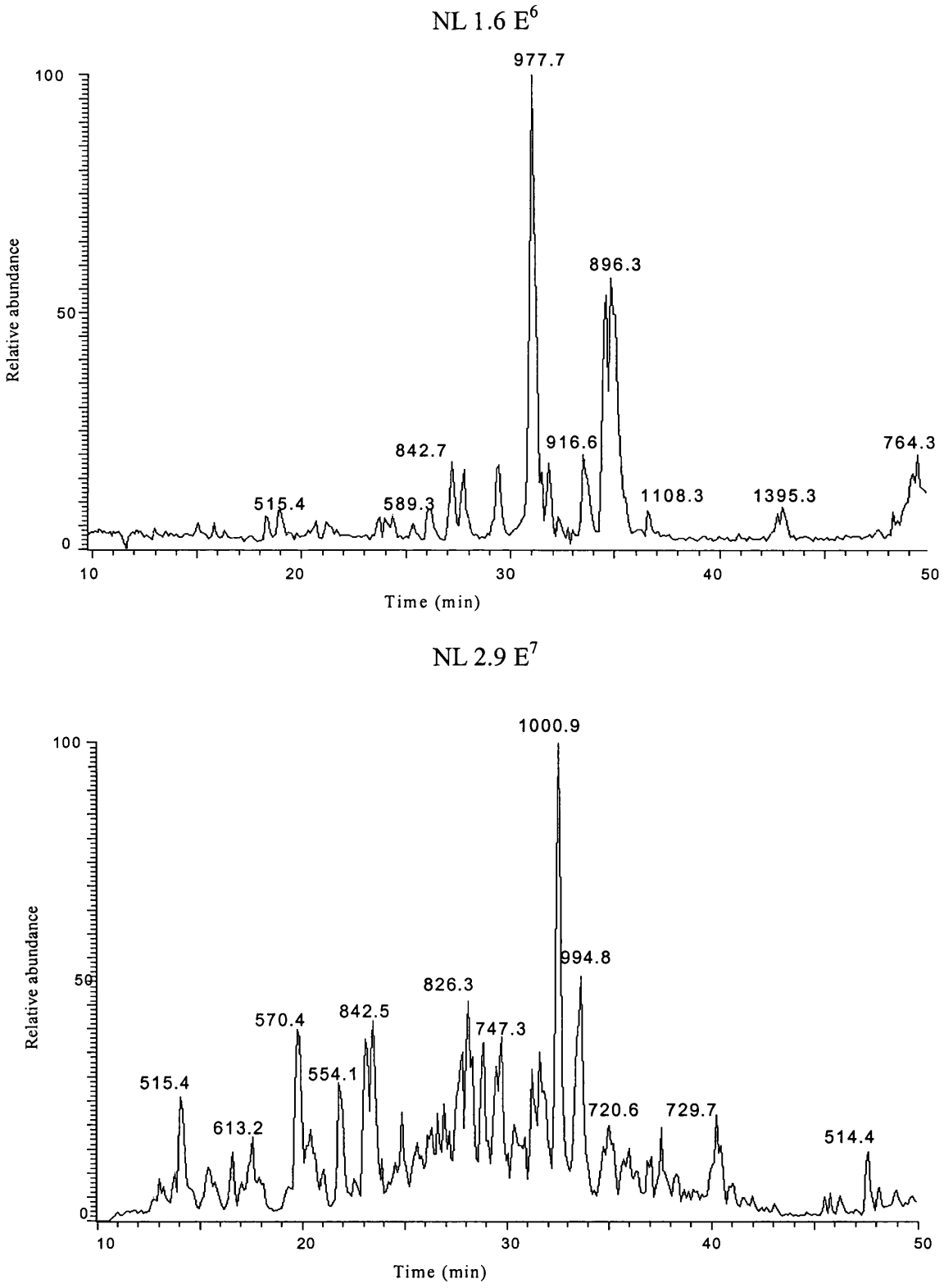


Fig 3.14: Chromatographic separation of peptides isolated from liver (lower panel) and tumour (upper panel) tissues obtained from the same individual.

Note: Equal amounts of the microsomal proteins were analysed for tumour and normal liver. However, visual observation of the gel separation of tumour sample showed less intense bands in 48-62 kDa.

Although the main purpose of this study was the detection of CYPs, other proteins identified with high peptide numbers/sequence coverage from tumours and livers are shown in Table 3.7.

Table 3.7: Proteins identified with high sequence coverage from tumour and liver samples

Accession number	Protein identified			
	Tumour	% Sequence coverage	Liver	% Sequence coverage
P30613	Pyruvate kinase	38	–	
P13674	Prolyl 4-hydroxylase alpha-1	28	Prolyl 4-hydroxylase alpha-1	35
P04264	Human keratin type II cytoskeletal 1	25	–	
BAA00016	F1-beta precursor	28	F1-beta precursor	32
P07099	Microsomal epoxide hydrolase	23	–	
P10719	ATPase beta	18	ATPase beta	29
P23141	–		Liver carboxylesterase	42
P27797	–		Calreticulin precursor	35

– *not identified*

Human pyruvate kinase, prolyl 4-hydroxylase, human keratin type II cytoskeletal 1 (Cytokeratin 1), F1-beta precursor, ATPase beta and microsomal epoxide hydrolase which were present in all samples analysed, except for the necrotic tumour sample 1T, 4T and 6T.

Normal colon and primary colon tumour

The results obtained from the colorectal metastasis led to a study of normal colon and primary colon tumour to enable a comparison with the CYP profile from the secondary (metastatic) colorectal tumours. Normal colon and primary colon tumour were obtained from the same individual and compared for CYP content. Microsomal proteins isolated from primary colon tumour and normal colon were separated by SDS-PAGE. Bands covering the molecular weight range 48-62 kDa were cut out and subjected to in-gel digestion with trypsin. The resultant peptides were extracted and analysed by LC-ESI-MS/MS. There were two primary colon tumours available for analysis. The mass of the tumour available for analysis, and the mass of tissue loaded on the gel is shown in Table 3.8.

Table 3.8: Microsomal protein concentrations for samples taken from patients with primary colon cancer

Sample	Mass of the tissue used to prepare microsomes (g)	Microsomal protein in tissue (mg/g)	Amount of protein loaded on the gel (μg)	mass of the tissue loaded on the gel (g)
[a]NC1	0.15	0.7	8.4	0.012
[b]TC1	0.95	2.1	25.2	0.012
[c]NC2	0.4	1.8	21.2	0.01
[d]TC2	0.15	1.5	18.0	0.012

[a] normal colon sample from human 01, [b] tumour colon sample from human 01

[c] normal colon sample from human 02, [d] tumour colon sample from human 02

Table 3.9 lists the proteins identified from colon tissues.

Table 3.9: Proteins identified from colorectal samples

Protein identified from paired samples			
Accession number	***Normal colon from human 01 and 02 Protein identified***	***% sequence coverage***	
		H01	***H02***
P13674	Prolyl 4-hydroxylase	40	33
P30613	Pyruvate kinase	31	36
P27797	Calreticulin precursor	39	40
NP_001272	Cytoskeleton-associated protein 4	–	28
P07237	Protein disulfide-isomerase	35	32
P10719	ATP synthase	40	42
AB50217	Protein disulfide isomerase-related protein	30	–
P08684	Cytochrome P450 3A	–	14
	Primary colon tumour from human 01 and 02 Protein identified	***% sequence coverage* (number of peptides)**	
		H01	***H02***
P01876	Ig alpha-1 chain C	22	–
P05783	Cytokeratin 18	38	42
P27797	Calreticulin precursor	16	27
P07237	Protein disulfide-isomerase	35	33
P13645	Keratin, type 1 cytoskeleton 10	26	22
BAA00016	F1-beta precursor	42	46
P10719	ATP synthase	28	15
CAA45026	Mutant beta-actin	25	–
P22695	Ubiquinol-cytochrome C reductase	49	–

– *Not identified*

The results indicate that enzymes of the CYP3A subfamily are present in colon tissue. However, the results are based on low sequence coverage, which also indicate that the protein is in low abundance. Previously, in liver it was shown that high protein abundance related to high sequence coverage. Four peptides were identified from the CYP 3A subfamily, all of which are present in CYP3A3, 3A4 and 3A5. Therefore it was not conclusively determined which specific isoform was present. The expression of CYP3A in colon is discussed in more detail in section 4.3.

MALIPDLAMETWLLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYHKGFCMFDMECHKKYG
MALIPDLAMETWLLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYHKGFCMFDMECHKKYG
MDLIPNLAVETWLLLAVSLVLLYLYGTRTHGLFKRLGIPGPTPLPLLGNVLSYRQGLWKFDTECYKKYG

KVWGFYDGQQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPVGFMKSAISIAEDEEWKRLRSLLSPTFT
KVWGFYDGQQPVLAITDPDMIKLVLVKECYSVFTNREPFGPVGFMKSAISIAEDEEWKRLRSLLSPTFT
KMWGTYEGQLPVLAITDPDVIRTVLVKECYSVFTNRRSLGPVGFMKSAISLAEDEEWKRIRSLLSPTFT

SGKLKEMVPIIAQYGDVLVRNLRREAETGKPVTLKDVFGAYSMDVITSTSFGVNIDSLNNPQDPFVENT
SGKLKEMVPIIAQYGDVLVRNLRRERETGKPVTLKDVFGAYSMDVITSSSFGVNVDSLNNPQDPLVENT
SGKLKEMFPIIAQYGDVLVRNLRREAEKGKPVTLKDIFGAYSMDVITGTSFGVNIDSLNNPQDPFVEST

KKLLRFDFLDPFFLSITVFPFLIPILEVLNICVFPREVTNFLRKSVKRMKESRLEDTQKHRVDFLQLMI
KKLLRFDFLDPFFLSITVFPFLIPILEVLNICVFPREVTNFLRKAVKRMKESRLEDTQKHRVDFLQLMI
KKFLKFGFLDPLFLSIILFPFLTPVFEALNVSLFPKDTINFLSKSVNRMKKSRLNDKQKHRLDFLQLMI

DSQNSKETESHKALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQQKLQEEIDAVLPNKAPP
DSHKNSKETESHKALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQQKLQEEIDAVLPNKAP
DSQNSKETESHKALSDLELAAQSIIFIFAGYETTSSVLSFTLYELATHPDVQQKLQKEIDAVLPNKAPP

TYDTVLQMEYLDMVVNETLRLFPIAMRLERVCKKDVEINGMFIPKGWVVMIPSYALHRDPKYWTEPEKF
PTYDTVLQMEYLDMVVNETLRLFPIAMRLERVCKKDVEINGMFIPKGWVVMIPSYALHRDPKYWTEPEK
TYDAVVQMEYLDMVVNETLRLFPVAIRLERTCKKDVEINGVFIPKGSMVVIPTYALHHDPKYWTEPEEF

LPERFSKKNKDNIDPYIYTPFGSGPRNCIGMRFALMNMKLALIRVLQNFSFKPCKETQIPLKLSLGGLL
FLPERFSKKNKDNIDPYIYTPFGSGPRNCIGMRFALMNMKLALIRVLQNFSFKPCKETQIPLKLSLGGL
RPERFSKKKDSIDPYIYTPFGTGPRNCIGMRFALMNMKLALIRVLQNFSFKPCKETQIPLKLDTQGLLQ

QPEKPVVLKVESRDGTVSGA
LQPEKPVVLKVESRDGTVSGA
PEKPIVLKVDSRDGTLSGE

CYP3A4
CYP3A3
CYP3A5

Highlighted sequences were identified for CYP3A4, 3A3 and 3A5 using LC-MS/MS

Fig 3.15: Sequence alignment for CYP3A4, 3A3 and 3A5.

Analysis of the colon sample using LC-MS/MS identified the peptides of amino acid sequence SLLSPTFTSGKLK and ETQIPLK. These peptides are both present in CYP3A3, 3A4 and 3A5 isoforms. Further analysis was carried out in an attempt to distinguish which of the three isoforms was present (see Fig 3.15).

3.4. Experimental tumours

Having investigated the colon tumour samples, the analysis was extended to human xenografts grown in mice. Four xenografts were selected for analysis for their CYP content namely, BE (colon), Colo205 (colon), IGROV1 (ovarian) and MDA-MB-435 (breast). Fig 3.16 shows SDS-PAGE separation of microsomal proteins isolated from xenografts.

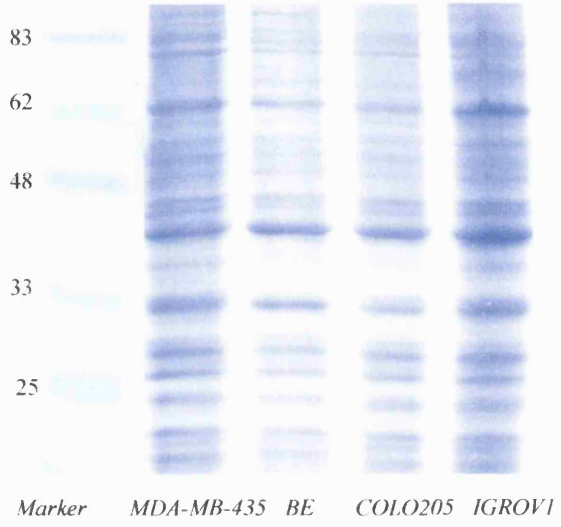


Fig 3.16: SDS-PAGE separation of microsomal protein isolated from human xenografts. The amount of the microsomal protein loaded onto the gel was 4.3 μg (MDA-MB-4.35), 5.8 μg (BE), 3 μg (COLO203) and 4.1 μg (IGROV1).

A number of proteins were identified, however no CYPs were found in MDA-MB-435, BE-, COLO and IGROV1 xenografts (Table 3.11).

Table 3.10: Microsomal protein concentrations for human xenograft samples

Sample	Mass of the tissue used to prepare microsomes (g)	Microsomal protein in tissue (mg/g)	Amount of protein loaded on the gel (μg)	Mass of the tissue loaded on the gel (g)
MDA-MB	0.17	0.36	4.32	0.012
BE	0.36	0.52	5.76	0.01
COLO205	0.04	0.25	3.0	0.012
IGROV1	0.36	0.34	4.08	0.012

Table 3.11: Proteins identified in the 48-62 kDa band from xenografts

Xenografts	Accession number	Proteins identified	% Sequence coverage
BE	P30613	Pyruvate kinase	32.5
	P08670	Vimentin	25.6
	P68104	Elongation factor 1-alpha	30.2
	P10719	ATP synthase	18.8
Colo205	P08670	Vimentin	69.2
	P30613	Pyruvate kinase	51.8
	P07237	Protein disulfide-isomerase	26.7
	P10809	Heat shock 60kDa protein	27.6
	P68104	Elongation factor 1 alpha	17.7
IGROV1	P30613	Pyruvate kinase	15.8
	P10809	Heat shock 60kDa protein	21.5
	BAA00016	F1-beta precursor	18.5
MDA-M-435	P10809	Heat shock 60kDa protein	18.6
	P30613	HUMAN Pyruvate kinase	18.2
	P05787	Keratin, type II cytoskeletal 8	49.8
	P08729	Keratin type II, cytoskeleton 7	62.3
	BAA00016	F1-beta precursor	32.7
	P10719	ATP synthase	36.6
	P05783	Cytokeratin 18	34.4

This study was further extended to the search for CYPs in cell lines. Human breast adenocarcinoma cells (MCF-7) were selected for this study, as previously the expression of CYP1A1 and 1B1 has been shown in MCF-7 cell lines induced with 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) (Spink *et al.,* 1998). MCF-7 cell with and without TCDD induction were grown, homogenised and centrifuged to isolate the S9 fraction and microsomes. These S9 fractions, microsomes and an E.Coli expressed CYP1B1 were then separated by a SDS-PAGE (Fig 3.19).

The tryptic digests of the MCF-7 microsomal proteins contained in the gel were analysed using LC-MS/MS. The peptide fragment information for the digested proteins was then subjected to Sequest search. The CYPs identified from MCF-7 cell

lines were CYP1A1 and CYP2A6. In the case of CYP1A1, a single peptide of 27 amino acids (XCorr 3.2) was identified. Two peptides were identified from CYP2A6 (see Fig 3.17).

>**gi|15147326**|ref|NP_000753.2| cytochrome P450, family 2, subfamily A, polypeptide 6; coumarin 7-hydroxylase; cytochrome P450, subfamily IIA (phenobarbital-inducible), polypeptide 3; cytochrome P450, subfamily IIA (phenobarbital-inducible), polypeptide 6; xenobiotic monooxygenase; flavoprotein-linked monooxygenase [Homo sapiens] [MASS=56501]

MLASGMLLVA LLVCLTVMVL MSVWQQRKSK GKLPPGPTPL PFIGNYLQLN TEQMYNSLMK
ISERYGPVFT IHLGPRRVVVLCGHDAVREA LVDQAEEFSG R**GEQATFDWV FK**GYGVVFSN
GERAKQLRRF SIATLRDFGV GKRGIEERIQ EEAGFLIDALRGTGGANIDP TFFLSRTVSN VISSIVFGDR
FDYKDKEFLS LLRMMLGIFQ FTSTSTGQLY EMFSSVMKHL PGPQQQAFQLLQGLEDFIAK
KVEHNQRTLD PNSPRDFIDS FLIRMQEEEK NPNTEFYLKN LVMTTLNLFI GGTETVSTTL
RYGFLLLMK**HPEVEAKVHEE IDR**VIGKNRQ PKFEDRAKMP YMEAVIHEIQ RFGDVIPMSL
ARRVKKDTKF RDFFLPKGTE VFPMLGSVLRDPSFFSNPQD FNPQHFLNEK GQFKKSDAFV
PFSIGKRNCF GEGLARMELF LFFTTVMQNF RLKSSQSPKD IDVSPKHVGFATIPRNYTMS FLPR

Mass (average): 56501.4 **Identifier:** gi|15147326 **Database:** C:/Xcalibur/database//human01July03edited.fasta
Protein Coverage: 25/494 = **5.0%** by amino acid count

GEQATFDWV FK	**102-112**
HPEVEAKVHEE IDR	**229-250**

>**gi|4503199**|ref|NP_000490.1| cytochrome P450, subfamily I (aromatic compound-inducible), polypeptide 1 [gi|117139|sp|P04798|CP11_HUMAN CYTOCHROME P450 1A1 (CYP1A1) (P450-P1) (P450 FORM 6) (P450-C) [gi|2144296|pir||O4HU6 aryl hydrocarbon (benzo[a]pyrene) hydroxylase (EC 1.14.14.-) cytochrome P450 1A1 - human [gi|30347|emb|CAA27843.1| (X04300) P-450 c [Homo sapiens] [MASS=58165]

MLFPISMSAT EFLLASVIFC LVFWVIRASR PQVPKGLKNP PGPWGWPLIG HMLTLGKNPH
LALSRMSQQY GDVLQIRIGSTPVVVLSGLD TIRQALVRQG DDFKGRPDLY TFTLISNGQS
MSFSPDSGPV WAARRRLAQN GLKSFSIASD PASSTSCYLEEHVSKEAEVL ISTLQELMAG
PGHFNPYRYVVVSVTNVICAICFGRRYDHNHQELLSLVNLNNNFGEVVGSGNPADFIPILRYLPNPSLNAF
KDLNEKFYSFMQKMVKEHYKTFEKGHIR**DITDSLIEHCQEKQLDENANVQLSDEK**IINIVLDLFGAGFDT
VTTAISWSLMYLVMNPRVQRKIQEELDTVIGRSRRPRLSDRSHLPYMEAFILETFRHSSFVPFTIPHSTTRDT
SLKGFYIPKGRCVFVNQWQINHDQKLWVNPSEFLPERFLTPDGAIDKVLSEKVIIFGMGKRKCIGETIARW
EVFLFLAILLQRVEFSVPLGVKVDM TPIYGLTMKH ACCEHFQMQL RS

Mass (average): 58165.3 **Identifier:** gi|4503199 **Database:** C:/Xcalibur/database//human.fasta
Protein Coverage: 27/512 = **5.3%** by amino acid count

DITDSLIEHCQEKQLDENANVQLSDEK	**280 – 306**

Fig 3.17: Sequest summary showing identification of CYP2A6 and 1A1

There were a number of other proteins, which were identified with good sequence coverage (Table 3.12).

Table 3.12: Proteins identified in the 48-62 kDa band from MCF-7 cell

Accession number	Proteins identified	% Sequence coverage
P05783	Cytokeratin 18	58.2
P08727	Cytokeratin 19	53.6
P68104	Elongation factor 1-alpha	23.6
P00558	Phosphoglycerate kinase1	21.0
P11509	CYP 2A6	5.0
P04798	CYP1A1	5.4

Absence of CYP1B1 from mass spectrometry analysis led to the western blot experiment of the MCF-7 samples for positive determination. The S9 fractions, microsomes and a recombinant expressed CYP1B1 were then separated by a SDS-PAGE and the proteins were blotted onto a nitrocellulose paper.

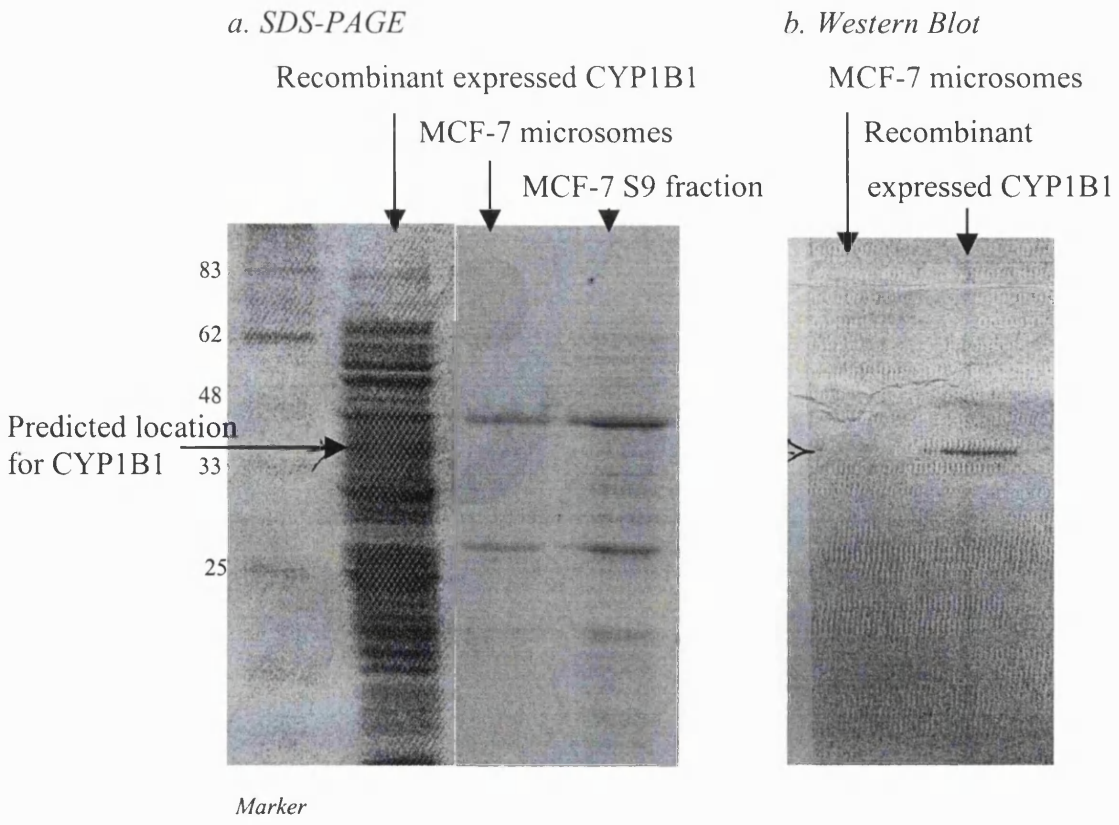


Fig 3.18: a. SDS-PAGE separation of recombinant expressed CYP1B1, MCF-7 micosomes and S9 fractions of MCF-7 cell lines. The amount of the protein loaded onto the gel was 25 μg except MCF-7 micosomes (6 μg) and S9 fractions (12 μg) b. Western blot of the gel separated proteins.

Western blot analysis showed that MCF-7 microsomes used for mass spectrometry analysis are negative for CYP1B1.

3.5. Optimisation of protein identification

An attempt was made to improve the identification method (with regards to amino acid sequence coverage and the number of proteins) by carrying out a number of repeat injections in a full scan range (*m/z* 400-2000) mode, or by narrowing the scan range. The later was done by splitting the full scan range into three *m/z* ranges i.e. 400-800, 800-1500 and 1500-2000. To test this experiment, normal liver and liver metastases from the same individual were separated on a SDS-PAGE and three bands were cut from the lanes corresponding to the normal and a metastases sample, in the 48-62 kDa region. The results are shown in tables 3.13-3.18. For each injection, 1st and the succeeding datafiles (run in full scan range), were then grouped under the same directory and searched with Sequest.

Table 3.13: Proteins identified from band 01 of normal liver

Accession number	Protein identified	Sequence coverage (%)					
		x1 datafile	x2 datafile	x3 datafile	x4 datafile	x5 datafile	Combined narrow scan range
P13674	Prolyl 4-hydroxylase	48.0	62.0	70.5	77.5	77.2	65.6
P23141	Human liver carboxylesterase	35.3	52.0	68.4	68.4	68.4	60.6
P27797	Calreticulin precursor	15.6	4.0	53.7	53.7	53.7	56.1
Q6NT91	UDP glycosyltransferase 1 family	—	8.3	12.4	12.4	21.0	17.6
P07237	Protein disulfide isomerase	33.2	52.3	66.7	66.7	66.7	66.7
P35504	UDP glycosyltransferase 1-5 precursor	—	—	—	14.4	14.4	8
P27338	Amine oxidase B	—	—	—	9.0	10.5	6.0
P49326	Flavin mono oxygenase 5	—	—	—	—	—	8.3

— Not identified

x1 means one datafile was searched using Sequest

x2, x3, x4 and x5 means two, three, four and five datafiles were combined respectively for Sequest analysis

Table 3.14: Proteins identified from band 02 of normal liver

Accession number	Protein identified	Sequence coverage (%)					
		x1 datafile	x2 datafile	x3 datafile	x4 datafile	x5 datafile	Combined narrow scan range
P07687	Epoxide hydrolase	50.0	51.4	51.4	51.4	56.3	55.4
P11509	Coumarin 7-hydroxylase, CYP2A6	26.3	28.5	28.5	31.2	31.2	28.5
P23141	Liver carboxylesterase precursor	23.0	27.9	30.5	30.5	30.5	33.2
P68133	Alpha skeletal muscle actin	15.2	33.2	33.2	33.2	33.2	29.6
P20813	Cytochrome P450-2B6	9.6	18.1	18.1	20.8	20.8	19.6
NP_001066	UDP glycosyltransferase 2 family, polypeptide B10	—	14.4	14.4	21.2	21.2	17.6
P08684	Cytochrome P450-3A4	17.5	20.4	24.1	26.4	28.6	20.1
NP_001065	UDP glycosyltransferase 2 family, polypeptide B7	5.3	12.9	12.9	15.1	15.1	13.6
NP_065392	Chromosome 20 open reading frame 3	—	—	—	—	—	15.5

Table 3.15: Proteins identified from band 03 of normal liver

Accession number	Protein identified	Sequence coverage (%)					
		x1 datafile	x2 datafile	x3 datafile	x4 datafile	x5 datafile	Combined narrow scan range
P06133	UDP glucuronosyltransferase 2B4 precursor	22.3	23.3	33.7	39.0	43.2	36.2
BAA00016	F1-beta precursor	19.7	52.7	56.0	56.0	56.0	55.1
P08684	Cytochrome P450 3A4	29.0	37.4	41.7	46.3	46.7	31.4
P07237	Protein disulfide-isomerase	21.4	30.7	38.0	38.0	41.5	42.2
NP_001065	UDP glycosyltransferase 2 family, polypeptide B7	21.6	31.9	31.9	32.1	32.1	26.5

P05181	Cytochrome P450 2E1	13.8	26.9	28.8	31.2	31.2	25.4
P11712	Cytochrome P450 2C9	8.9	28.4	28.4	28.4	30.1	20.6
Q14554	Protein disulfide isomerase-related protein 5	22.3	29.5	29.7	36.4	41.3	45.0
P23141	liver carboxylesterase	10.4	21.9	21.9	29.2	31.8	35.1
NP_009051	UDP glycosyltransferase 1 family, polypeptide A4	22.4	20.0	20.0	21.0	24.7	21.0
P16662	UDP glycosyltransferase 2B7	21.6	–	–	–	–	26.7
P36537	UDP glycosyltransferase 2B10	–	16.8	–	16.9	18.8	27.7

Proteins identified from Liver metastases

Table 3.16: Proteins identified from band 01 of liver metastases

Accession number	Protein identified	Sequence coverage (%)					
		x1 datafile	x2 datafile	x3 datafile	x4 datafile	x5 datafile	Combined narrow scan range
P13674	Prolyl 4-hydroxylase	11.0	48.2	56.5	57.5	59.0	41.7
P23141	Human liver carboxylesterase	7.0	25.5	28.2	50.4	50.4	35.9
P27797	Calreticulin precursor	–	7.4	22.0	22.5	22.5	16.5
P21397	Monoamine oxidase A	–	4.2	10.0	10.0	14.7	9.7
P10809	Heat shock 60 KDa protein 1	–	5.0	11.7	11.7	11.7	13.6
P27338	Amine oxidase B	–	–	–	8.8	12.2	15.0
P51648	Aldehyde dehydrogenase	–	–	–	5.8	8.5	4.9

Table 3.17: Proteins identified from band 02 of liver metastases

Accession number	Protein identified	Sequence coverage (%)					
		x1 datafile	x2 datafile	x3 datafile	x4 datafile	x5 datafile	Combined narrow scan range
P06576	ATP synthase, beta chain	28.0	28.0	33.3	338	33.8	28.6
P25705	ATP synthase, alpha subunit	35.0	39.2	41.9	41.9	41.9	42.9
Q14554	Protein disulfide isomerase-related protein 5	15.0	15.5	26.5	28.5	30.6	37.3
Q6QQX7	UDP glycosyltransferase 2 family, polypeptide B4	11.2	14.2	19.3	25.4	27.1	22.5
P16662	UDP glycosyltransferase 2 family, polypeptide B7	19.3	22.9	25.2	28.9	28.9	20.8
P51648	Fatty aldehyde dehydrogenase	8.1	8.1	15.6	15.6	15.6	18.7
P10632	Cytochrome P450 2C8	6.6	9.7	13.4	14.2	21.4	15.7
P36537	UDP glycosyltransferase 2 family, polypeptide B10	12.5	13.8	13.4	20.8	20.8	13.6
P36537	Cytochrome P450 3A4	10.5	10.5	13.7	16.3	16.3	13.7
NP_001924	Dihydrolipoamide S-succinyltransferase	–	–	–	9.3	9.3	8.2
Q02928	Cytochrome P450 4A11	–	–	–	9.6	11.3	6.0
NP_001067	UDP glycosyltransferase 2 family, polypeptide B15	–	–	–	16.4	16.4	14.6

Table 3.18: Proteins identified from band 03 of liver metastases

Accession number	Protein identified	Sequence coverage (%)					
		x1 datafile	x2 datafile	x3 datafile	x4 datafile	x5 datafile	Combined narrow scan range
NP_004484	Hydroxyacyl-Coenzyme A dehydrogenase	–	10.9	17.3	22.7	22.5	22.5
NP_065392	Chromosome 20 open reading frame 3	5.4	5.4	12.3	16.6	16.6	18.8
BAA04090	LTG9/MLLT3 protein	–	–	–	5.6	11.0	10.8
P39656	Glycosyltransferase	–	–	–	5.3	5.3	7.0
P07687	Microsomal epoxide hydrolase	–	–	27.5	31.9	31.9	29.0
Q14554	Protein disulfide isomerase-related protein 5	–	–	–	3.6	3.6	3.6
P11509	Cytochrome P450 2A6	–	–	–	–	–	3.0

The results obtained from the triplicate full scan range experiments were compared with the results obtained from the narrowing scan range experiments in terms of the percentage sequence coverage and the number of identified proteins, as the later was also performed using three data files run at different scan ranges. From the triplicate analysis of the liver sample of 400-2000 *m/z* range 22 proteins were identified and from the narrow scan range analysis, 29 proteins were identified. Similarly, from the triplicate analysis of the liver metastases sample of 400-2000 *m/z* range 17 proteins were identified and from the narrow scan range analysis, where 26 proteins were identified.

Notably for proteins identified with low sequence coverage that is <20%, the sequence coverage was improved by recording three narrow scan ranges and combining the files, than by simply combining files from three repeat injections over the wide mass range.

3.6. Validation of absolute quantification (AQUA) of proteins and applications to CYPs

In theory, in a chromatographic run the peak area of a peptide should correlate to its concentration; also, the peak area of a peptide from one protein should correlate to the concentration of that particular protein. It is now possible to use synthetic peptides with incorporated stable isotope labels to provide absolute quantification of proteins present in complex mixtures via the so called "AQUA" method (Gerber *et al.,* 2003).

In the present study, this method was adopted as a way of quantifying the proteins present in the microsomal fractions isolated. To validate the AQUA method, horse heart myoglobin was first used as this protein is not present in human liver microsomes. A peptide internal standard was selected from the myoglobin protein and then synthesised so as to contain a single amino acid labelled with heavy isotopes. The internal standard peptide containing stable isotope labels corresponding to one tryptic peptide from myoglobin was used for quantification.

The peptide LFTGHPETLEK from native myoglobin was analysed by MS/MS to examine its peptide fragmentation patterns. The MS/MS spectrum of the native peptide $[M + H]^{2+}$of *m/z* 636.4 revealed prominent y-type fragment ion which was suitable for monitoring (see Fig 3.19). The stable isotope labelled peptide internal standard LFTGHPETL*EK (L* is labelled with heavy isotopes) was subjected to MS/MS (see Fig 3.20).

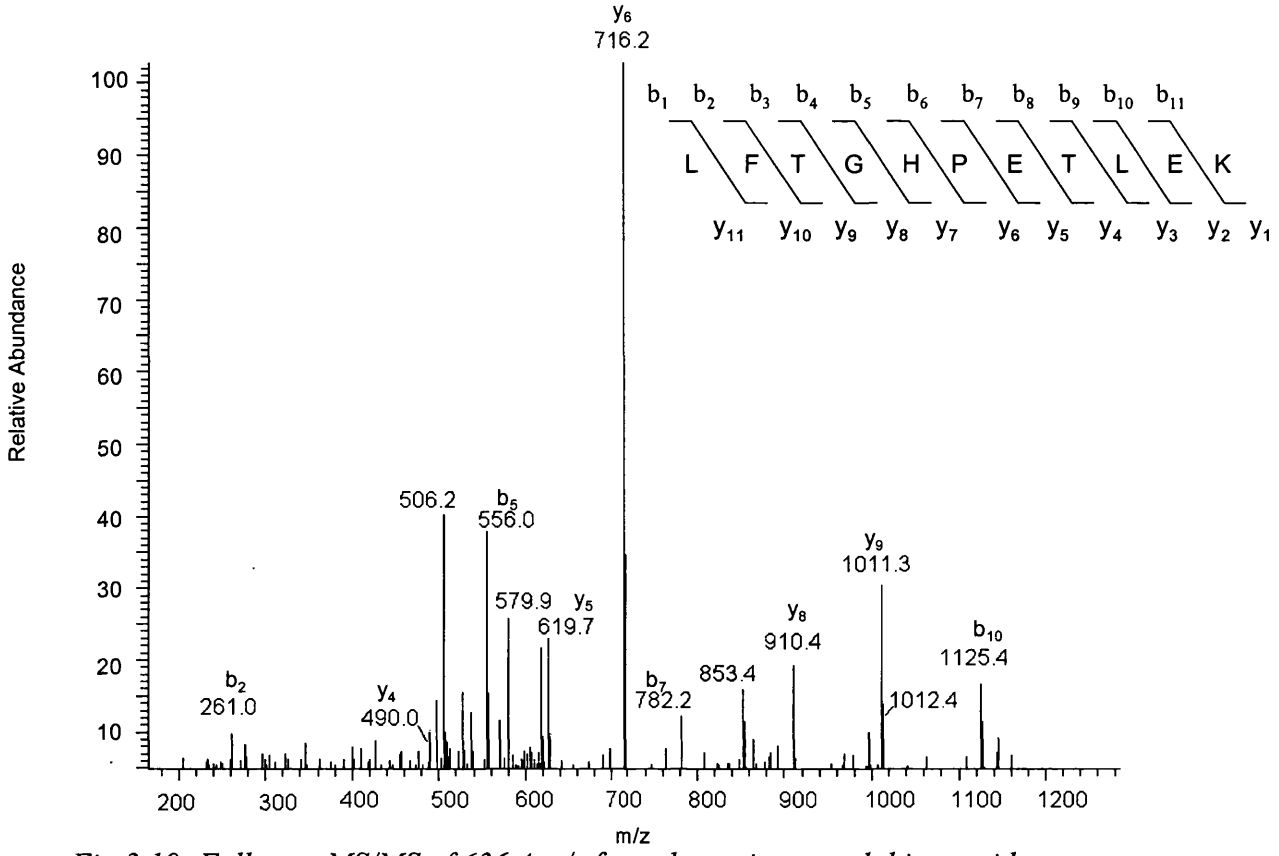


Fig 3.19: Full scan MS/MS of 636.4 m/z from the native myoglobin peptide LFTGHPETLEK

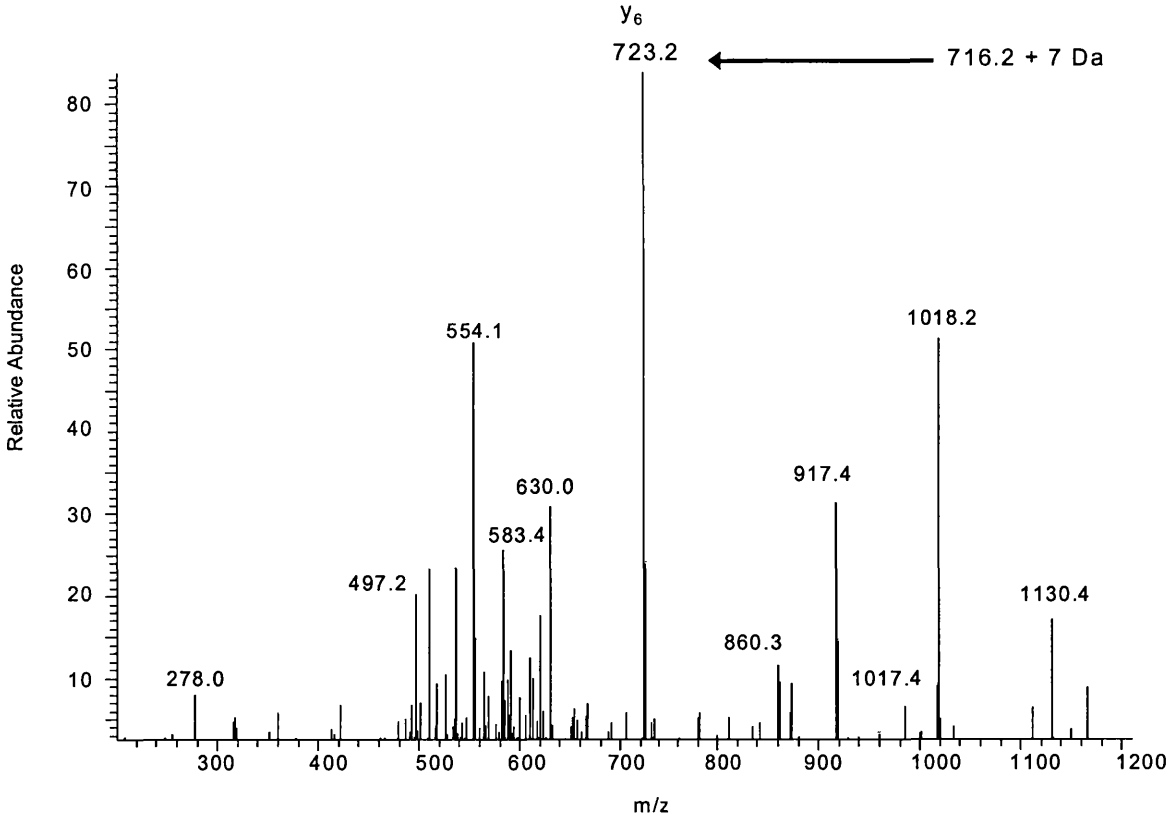


*Fig 3.20: Full scan MS/MS of 636.4 m/z from the isotope labelled myoglobin peptide LFTGHPETL*EK*

It can be seen that the fragmentation of *m/z* 636.4 to y_6 of *m/z* 716.2 are the most prominent fragmentation channel in the MS/MS spectrum.

The internal standard peptide was labeled with six ^{13}C and one ^{15}N atoms on the Leu residue. This results in its $[M + H]^{2+}$ ion appearing at *m/z* 639.9. MS/MS of this ion gives a prominent fragment at 723.4 corresponding to y_6.

The mass spectrometer was next set up to perform single reaction monitoring (SRM analysis. Two SRM transitions were monitored for the pair of native and isotope labelled peptide (in *m/z*: 636.4 → 716.4 and 639.9 → 723.4 for the LFTGHPETLEK pair).

A solution digest was then performed on myoglobin with added isotope labelled peptide to check that the myoglobin and the AQUA peptide were of the correct purity. A solution digest of a horse heart native myoglobin (500 fmol) was carried out with trypsin in the presence of an isotope labelled peptide LFTGHPETL*EK (where L* indicates a leucine with six ^{13}C and one ^{15}N atoms) (500 fmol). The solution digest was then analysed using LC-SRM and the abundance of the y_6 fragment ion from both the native peptide and the isotope labelled peptide was measured. The peak area ratio for myoglobin peptide to AQUA peptide was found to be ~ 0.8.

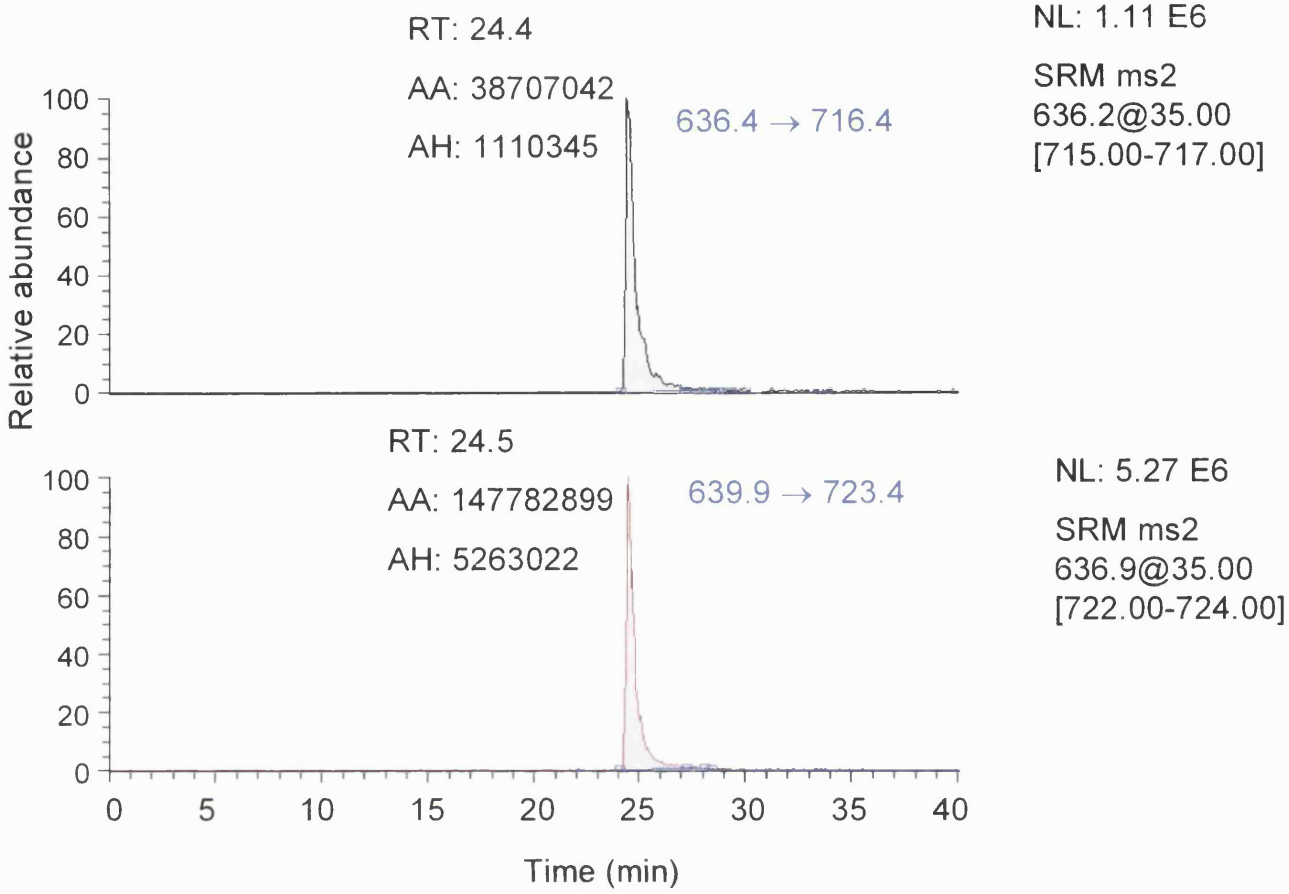


Fig 3.21: LC-SRM trace for the specific parent-to-product ion transitions of an AQUA internal standard and native myoglobin peptide.

RT: retention time, AA: area of the peak, AH: height, ms2: MS/MS

The number of moles of native myoglobin loaded onto the gel and the AQUA peptide added to the gel pieces was the same. Therefore, for 100% recovery one would expect the two peak areas to be the same. However, in an initial experiment when the recovery of native myoglobin peptide was calculated based on the peak area, it was far less than expected i.e. ~ 10%

The gel based experiment was then repeated by adding the internal standard peptide to the trypsin buffer solution, which was added to the dried gel pieces and then digested overnight. In the earlier experiment, the two were added separately to the gel pieces. The extraction procedure was also modified. This extraction involved the addition of 5% formic acid and 50% acetonitrile to the gel pieces after trypsin digestion, then dehydrating the gel pieces using 100 % acetonitrile and collecting the supernatant, rehydrating using 5 % formic acid and 50 % acetonitrile and dehydrating again. This process was repeated three times and the resultant solution was collected into a 1.5 ml

microfuge tube. These modifications increased the peptide recovery to 80 % from the gel. These results gave a confidence that LC/MS/MS can be used for quantification of proteins present in complex mixtures.

Chapter 4

Discussion

4.1. CYP expression in male and female rat liver

Sexual dimorphism in drug metabolism is due to the existence of multiple forms of hepatic CYPs whose gender-dependent expression is regulated by growth hormone (Shapiro *et al.,* 1995). The sex-dependent differences in xenobiotic metabolism are most pronounced in the rat. Consequently, this species is selected as the most popular animal model to study sexual dimorphisms in xenobiotic metabolism.

The use of MALDI-TOF-based PMF of 1D gel tryptic digests of liver microsomes uncovered four CYPs (2A1, 2C11, 2D2, 2D5) in uninduced male Sprague-Dawley rats (Galeva *et al.,* 2003). These isoforms were also found among the 24 CYPs identified in male and female rat livers using our LC-ESI-MS/MS methodology. MALDI-TOF based PMF of clofibrate-induced rat liver microsomes revealed the presence of CYP 2B1, 2B2, 4A1 and 4A3 (Galeva *et al.,* 2003). Of these, CYP 4A3 was the only isoform detected in our male and female livers suggesting that non-induced levels of CYP 4A1, 2B1 and 2B2 are low. However, we did detect CYP2B3 in female rat liver. It should be emphasized that while the MALDI-TOF PMF approach is excellent for the identification of separated proteins and simple protein mixtures (Jensen *et al.,* 1997; Zhang and McElvain, 2000), when proteins are present in low abundance or as part of complex mixtures, as may be the case with 1D gel bands, it is necessary to obtain more information (Fenyö, 2000). The most common method to achieve this is by MS/MS of isolated proteolytic peptides, as in our LC-ESI-MS/MS approach.

Although the results of the current study are of a qualitative nature, and the non-identification of a protein in a sample does not necessarily equate to its absence, at-least to a first approximation, the greater the concentration of a protein within a given sample the greater the probability of its identification. With this in mind some comments can be made in regard to the identification of CYP proteins based on rat gender (see Fig 4.1).

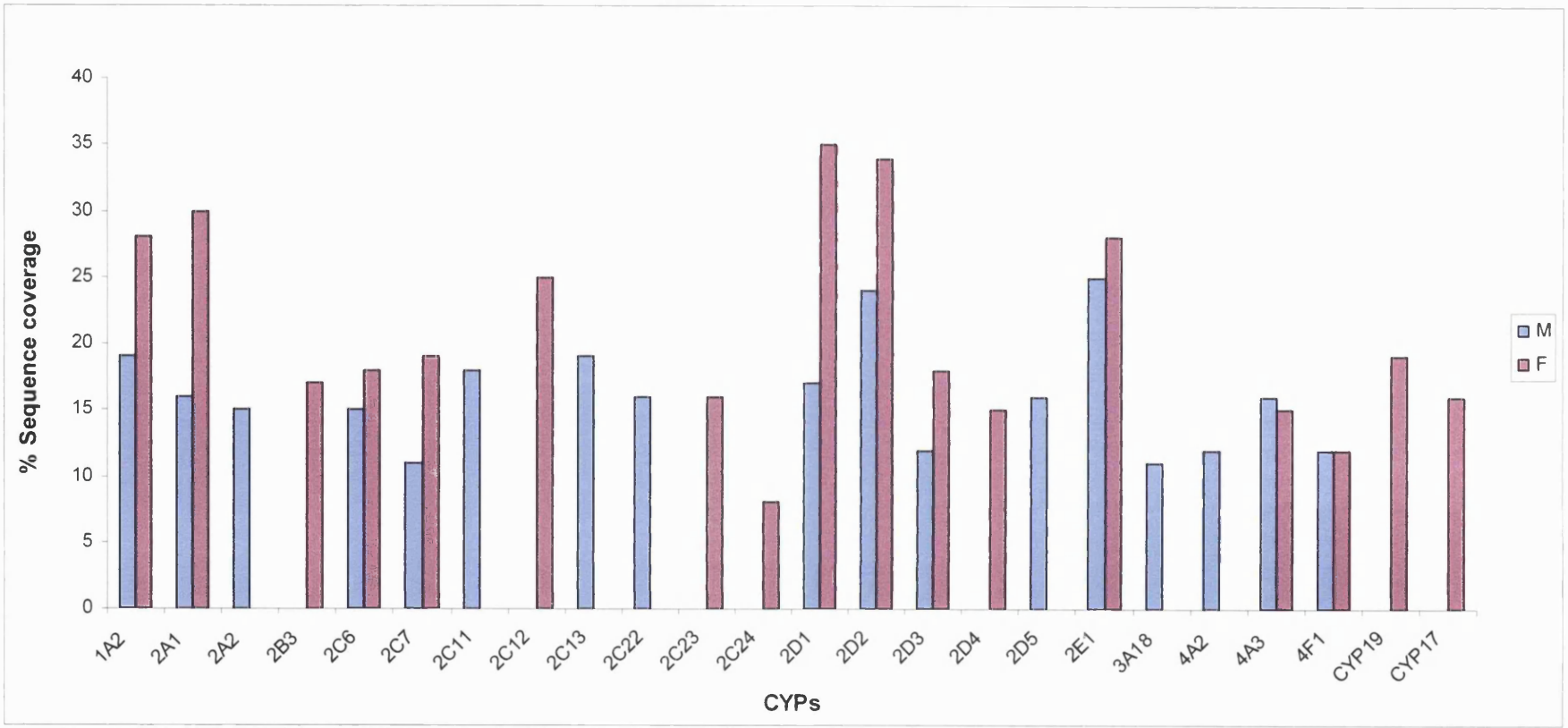


Figure 4.1: Comparison of CYP profile in male and female rat liver using LC-MS/MS.

M = Male rat, F = Female rat

Most of the known gender-related differences in compound toxicity in rats are due to gender-related differences in hepatic metabolism (Czerniak, 2001). Hormones and growth factors are strongly implicated in the gender-influenced expression of hepatic CYPs (Waxman, 1992; Pampori and Shapiro, 1999; Kalsotra *et al.,* 2002). Sexually dimorphic expression of hepatic CYP isoforms is regulated by the gender-dependent secretory GH profiles, *i.e.* episodic in males and continuous in females. In the case of the feminine GH profile, the continuous presence of the hormone in the circulation completely suppresses male-specific CYP2A2, 2C11, and 3A2, while stimulating full expression of female-dependent CYP2A1, 2C7 and 2C12. Similarly, females can reverse the gender-dependent expression of the CYPs to the episodic masculine GH profile. Under these conditions, females will now express the male-specific isoforms and suppress the female-dependent forms, whereas the opposite will occur in the males (Pampori and Shapiro, 1999). Nevertheless, it is not clear whether the levels of expression or suppression are comparable in male and female rats exposed to the same sex-dependent GH profiles. Whether CYP expression is gender specific rather than gender predominant or biased will depend on the absolute expression of protein in one sex compared to the other. It has been suggested that a CYP is gender specific only if the relative expression is 10-fold or higher in one sex compared with the other; lower than this is suggested to represent enriched CYP expression in one gender rather than sex specific (Kato and Yamazoe, 1993).

Our results show that CYP2C11 and 2C13 were found in male but not female rats whereas CYP2C12 was found in female only, supporting previous work that indicated that these three isoforms are gender-specific/predominant (Agrawal and Sharpiro, 2001).

In contrast to female specific CYP2C12, CYP2C7 is appropriately designated as female predominant because this isoform is expressed in male liver at significantly lower concentration than in female liver (Pampori and Shapiro, 1999). The current study suggests that other CYP2C members are gender predominant; specifically CYP2C22 was found only in male livers whilst CYP2C23 and 2C24 were found only in female livers. Regarding female-predominant CYP2A1, 2C6 and 2C7 (Agrawal and Sharpiro, 2001) these isoforms were identified in both genders by LC-ESI-MS/MS. However, they were identified with a higher percentage sequence coverage in female rats, which to a first

approximation indicates that a higher concentration of protein was present in female compared to male rats. CYP2C6 and 2A1 are female-predominant, with males expressing lower hepatic concentrations of the transcript (~60% less). CYP2C7 protein levels in females have been shown to be expressed at 3-4 times more that in males by Agrawal & Shapiro, 2001.

The rat has six members in the CYP2D subfamily: CYP2D1, 2D2, 2D3, 2D4, 2D5 and 2D18, although it is known that CYP2D18 is not expressed in liver. In our study we have identified all liver CYP2D isoforms supporting the fact that CYP2D is one of the major subfamilies expressed in rat liver. CYP2D1, 2D2 and 2D3 were found in both sexes whilst CYP2D4 was female specific and CYP2D5 was male specific. Gender selectivity of some 2D family members was previously identified (Schultz-Utermoehl *et al.*, 1999). CYP enzymes belonging to the CYP2D subfamily have been shown to be one of the determinants of polymorphic drug oxidations in the rat. Debrisoquine 4-hydroxylation is a typical reaction catalyzed by these enzymes.

In rats, the CYP3A subfamily consists of five related genes, CYP3A1, CYP3A2, CYP3A9, CYP3A18 and 3A23. CYP3As in rats show sex-, tissue-, species-, and age-dependent expression patterns (Waxman *et al.,* 1995). For instance, CYP3A1 and CYP3A18 have been shown to be male-dominant isoforms (Strotkamp *et al.* 1995). In the present study CYP 3A18 and 4A2 were observed in male but not female rat liver. This result is supported by previous studies showing that most members of the CYP 3A (including CYP3A18) and CYP4A families to be male gender specific or predominant (Anakk *et al.,* 2003; Holla *et al.,* 2001; Michell *et al.,* 2001; Robertson *et al.,* 1998; Waxman *et al.,* 1995). The RT-PCR approach has showed that CYP3A18 mRNA levels are 25-fold higher in male livers compared to females, while CYP3A9 showed a reverse pattern with 6-fold higher expression in the liver of females. Exposure of male rats to the female pattern of growth hormone secretion led to an increase in hepatic CYP3A9 mRNA expression and suppressed expression of CYP3A18 (Robertson *et al.,* 1998).
Our mass spectrometry approach showed that CYP17 and CYP19, two enzymes that contribute to sex steroid synthesis were present in female but not male liver. The apparent predominance of these enzymes in female rat liver is intriguing and warrants a more

detailed study. However it is likely to be of secondary importance to the complex interplay of glandular tissue with many peripheral tissues in the total provision of male and female sex steroids.

Western blot analysis clearly showed a higher expression of CYP4F in females than males in liver, kidneys and lungs. Furthermore, the results from real-time PCR clearly showed that CYP4F1 is highly expressed among all CYP4Fs in the liver, kidneys, and brain of either sex (Kalsotra *et al.,* 2002). The present results showed the identification of only CYP4F1 from CYP4F subfamily in male and female rat liver verifying that CYP4F1 is the highly expressed isoform in rat livers.

Progress has been made in identifying the CYP isoforms that are present in human liver (Nelson *et al.,* 1996), with 28 genes identified as coding for this superfamily of enzymes in the human genome. As in rodents, only gene families 1, 2, and 3 are involved in xenobiotic metabolism in humans. However, the major CYP isoform subfamily detected in human liver, CYP3A, is in relatively low concentration in rat liver (Table 4.1). Another key difference is that several CYP450 subfamilies have different substrate specificities in rodent as compared with human liver (Wrighton *et al.,* 1993). For example, human CYP3A has coumarin-7-hydroxylase activity, but none of the isoforms in the rat CYP3A subfamily show significant coumarin-7-hydroxylase activity (Guengerich, 1990). A comparison of major CYP isoforms in rodent and human is shown in table 4.1.

Table 4.1 - Comparison of Major Isoforms of Cytochrome P450 in Rodent and Human

Isoform	Rat	Human
CYP1A		
1A1	Present; induced by polycyclic aromatic hydrocarbons.	Present in liver and lung; induced by cigarette smoke.
1A2	Present; induced by polycyclic aromatic hydrocarbons	Present in liver and lung; induced by cigarette smoke.
CYP2A		
2A1	Rat testosterone 7 alpha-hydroxylase.	Not present
2A2	Present	Not present
2A3	Present in liver and lung	Not present
2A4	Present	Not present
2A5	Present	Coumarin 7-hydroxylase activity
CYP2B		
2B1	Phenobarbital-induced	Not present
2B2	?	Not present
2B6	?	Present
CYP2C	Major subfamily in rats;sex specific isoenzymes	
2C5	?	Not present
2C8	?	Retinol metabolism
2C9/10	?	Tolbutamide metabolism
2C18	?	Mepheytoin metabolism
CYP2D		
2D6	?	Debrisoquine metabolism
CYP2E		
2E1	Induced by ethanol, acetone	Induced by ethanol
CYP3A		Major subfamily in adult liver
3A1	Phenobarbital-inducible	?
3A2	Present in males only	?
3A3	present	See below
3A3/4	?	Major isoform in adult
3A5	?	Higher in adolescent liver
3A7	?	Major fetal form; not present in adults
CYP4A	?	Small role in metabolism of fatty acids

? not identified

Source: Mugford and Kedderis (1998)

Exaggerated sex-dependent variations in metabolism by rats may be the result of extensive inbreeding or differential evolution of CYP isoforms in mammals. Animal studies are used to help to determine the metabolism and toxicity of many chemical agents and in an attempt to extrapolate the risk to humans from exposure to these agents (Mugford and Kedderis, 1998). One of the most important concepts to consider in using rodent studies to identify sensitive individuals in the human population is that human CYPs differ from rodent CYPs in both isoform composition and catalytic activities. However, metabolism of xenobiotics by male rats can reflect human metabolism when the compound of interest is metabolized by CYP1A or CYP2E because there is strong regulatory conservation of these isoforms between rodents and humans (Mugford and Kedderis, 1998).

4.2. Human liver

Human liver microsomal samples were characterized for both CYP and other protein content. Fifteen CYP proteins were identified in human liver as shown in Table 3.3. All the major CYP involved in drug metabolism i.e. CYP1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, 3A4, were identified as were CYP4A11 or 4A22 (CYP4A22 is 95% homologous to 4A11), 4F2 and 4F11, which are involved in eicosanoid metabolism. The relatively high peptide sequence coverage for CYP 3A4, 2C8, 2C9, 2E1 and 2A6 generally coincided with the expected high expression profile of these CYPs identified previously using immunohistochemistry (Shimada *et al.,* 1994). Using immunochemistry, Shimada *et al.* (1994) have shown that 70% of liver CYPs could be accounted for by CYP1A2, 2A6, 2B6, 2C, 2D6, 2E1 and 3A. CYP3A (about 30% of total CYP) and 2C (about 20%) were the major CYP forms in human liver microsomes. Shimada *et al.* (1994) determined that CYP1A2 (about 13%), CYP2E1 (about 7%), and CYP2A6 (about 4%) are the major CYP isoforms, whereas the CYP2D6 (about 2%) and CYP2B6 (about 0.2%) were minor CYP forms (see Fig 4.2).

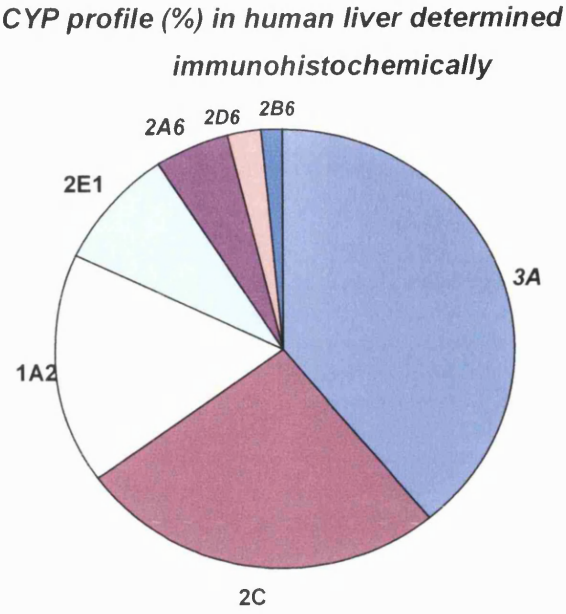


Fig 4.2: CYP profile in human liver

Source: Shimada et al. (1994)

The results obtained from the present study showed that CYP1A1, 2A7 and 3A7 proteins are not present or are present in low abundance in the human liver although these CYPs were previously detected in an adult human liver (Hakkola *et al.*, 1998) (see table 4.2).

Table 4.2: Expression of CYP mRNA or protein and the stage of upregulation in human

CYPs	*Expression*	*Principal stage of upregulation*
1A1 mRNA	+	
Protein	?	ND
1A2 mRNA	+++	
Protein	+++	During the first postnatal year
2A6 mRNA	++	
Protein	++	ND
2A7 mRNA	+	
Protein	?	ND
2B6 mRNA	+	
Protein	+	ND
2C mRNA	+++	Rapid postnatal rise followed by slow
Protein	+++	maturation during the 1st year
2D6 mRNA	+	Rapid postnatal rise followed by slow
Protein	+	maturation during the 1st year
2E1 mRNA	++	Rapid postnatal rise followed by slow
Protein	++	maturation during the 1st year
3A4 mRNA	+++	During the first postnatal months,
Protein	+++	possibly overlapping with CYP3A7
3A5 mRNA	++	
Protein	++	ND
3A7 mRNA	+	Down-regulated in late gestation and
Protein	-	postnatally

?: unknown, ND : not determined, - : not detected, + : low level, ++ : moderate level, +++ : high level.

Modified from Hakkola et al. (1998)

In adult human liver, CYP1A1 protein is considered a minor form but its contribution to liver metabolism is unknown (McKinnon *et al.,* 1991). In contrast, Ioannides & Parke (1993) showed that CYP1A1 is highly induced in many extrahepatic tissues due to cigarette smoking. There is some evidence suggesting that CYP1A1 could be expressed in fetal liver (Hakkola *et al.,* 1998). Recently, Stiborová *et al.* (2004) showed that an

enzymatically functional CYP1A1 appears to be expressed in human livers, however it was not identified using the LC-MS/MS technique in human liver. This could be due to the low levels of CYP1A1 expression in livers, about 0.7% of the total hepatic CYP (Stiborová *et al.*, 2004).

The amino acid sequence coverage determined for the identified proteins in the LC-ES-MS/MS experiment ranged from a maximum of 40% to a minimum of 9% (see Fig 4.3). This variation may be due to a number of factors, including protein abundance, the accessibility of the protein to complete trypsin digestion and mass spectrometry sensitivity. While the non-detection of a particular CYP does not mean that it is not present in a given sample, this does indicate that if present it is at a relatively low abundance. Similarly, at least to a first approximation, amino acid sequence coverage relates to relative protein abundance.

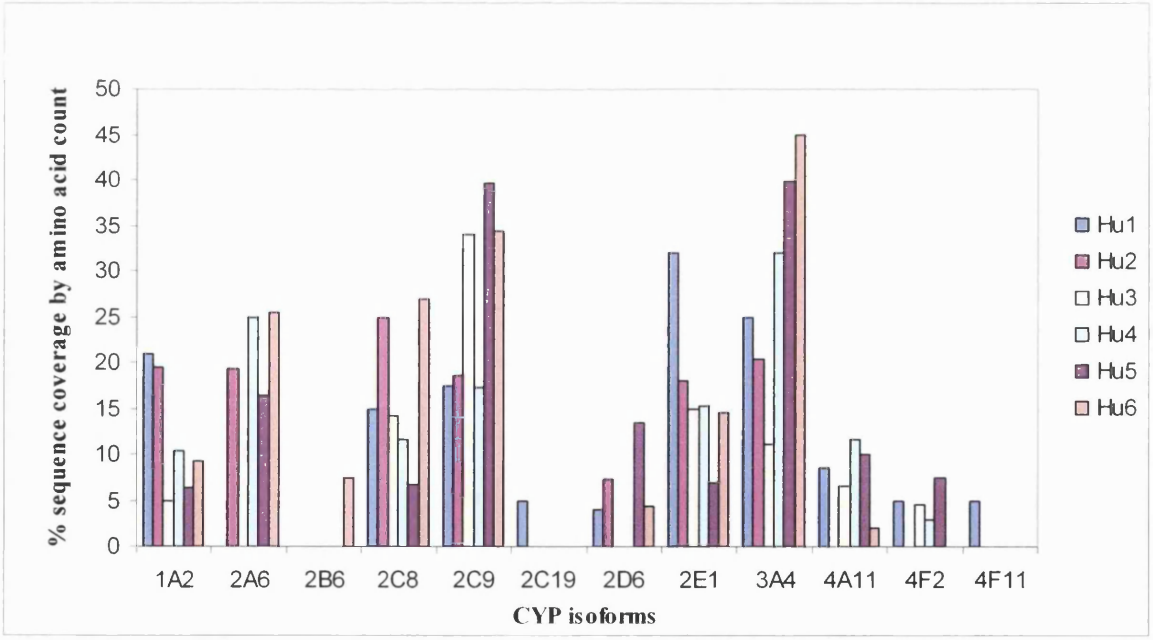


Fig 4.3: CYPs identified in 6-individual human livers using LC-MS/MS.

The present results also reveal that some CYPs are detected in some individuals but not in others. For example CYP2B6, which is a minor form of CYP in human liver, was seen in one individual (Hu6) (see table 3.3). The low abundance or absence of peptides may reflect

the low level of 2B6 and of 2C19, which are known to be quantitatively minor CYPs (Shimada *et al.,* 1994). Nicotine is known to be metabolized to its major metabolite cotinine by members of the CYP monooxygenase superfamily. Although CYP2A6 is identified as the principal enzyme which catalyses this biotransformation, CYP2D6 is also an active nicotine C-oxidase. Some 8% of the Caucasian populations have reduced or absent CYP2A6 activity (as seen in Hu1, Hu3 and Hu5); CYP2D6 may play a significant role in nicotine metabolism in these individuals. Hu1, a heavy smoker did not express CYP2A6, in this case CYP2D6 play an important role in nicotine metabolism. CYP2D6 is highly polymorphic - a number of studies linking CYP2D6 genotype to smoking behavior have now been published. CYP polymorphism could explain the absence of measurable CYP2A6 or 2D6 in certain individuals shown in Fig 4.2. For example *CYP*2A6*4 and *CYP2*D6*5 polymorphisms in Caucasians and Asians are known to result in no enzyme expression. The abundance of CYPs is also affected by both physiological and environmental factors. For example, CYP1A2 is highly inducible by tobacco smoke and was identified with high sequence coverage in individuals Hu1 and Hu2 who were smokers. CYP2E1 is induced by ethanol and was identified with high sequence coverage in Hu1 who acknowledged being a "social drinker". CYP4A11, a laurate and arachidonate omega-hydroxylase implicated in the control of systemic blood pressure, was identified in all individuals except Hu2 who intriguingly was known to suffer from hypertension.

The results obtained from this study show that there appears to be no obvious sex-specific isoforms of CYP in humans. Differences in xenobiotic metabolism among humans are more likely the consequence of intraindividual variations because of genetic or environmental exposure to chemicals rather than being due to sex-dependent differences in enzyme composition. Shimada *et al.* (1994) demonstrated that there were no clear sex differences in the levels of individual forms of CYPs in human liver microsomes. However, studies conducted using neonates showed that female neonates contain considerable levels of CYP2C and 3A proteins in liver microsomes; the levels were similar to those in adults (Shimada *et al.,* 1994). The main factors, which may affect difference in expression of CYP proteins among individuals and in turn lead to alteration in drug interactions, are age, gender, hormones, genetic polymorphism, hepatic disease, inflammation, nutrition, environmental factors and pregnancy.

4.3. Comparison of CYPs in normal liver, colorectal metastases, normal colon and primary colon tumours

The process of tumourigenesis is a complex event associated with the accumulation of multiple genetic changes (Jass, 1999). The problems involved with correlating changes in cancer cells to mRNA expression have lead investigators to study altered protein expression in cancer progression. The main concern is that the mRNA expression may be altered without a corresponding change observed in protein levels, while frequently protein levels are altered without a corresponding change observed in mRNA levels (Ozturk *et al.,* 1998). The genetic changes are importance in tumour CYP expression in influencing the outcome of chemotherapy through resistance or prodrug activation, which is poorly understood. Hence, there is a need to study the CYP expression in a way that does not rely on anticipating the CYPs present. Our LC-MS/MS experiments comprise the first study that investigates CYP enzymes in colorectal metastases without pre-selection of the proteins anticipated to be present.

Microsomal samples from livers of eight Caucasian donors were evaluated for expression of CYPs. Colorectal metastasis grown within these livers was also excised for comparative CYP expression study. The homogenization methods which were initially used for microsomal preparation from rat and human liver were slightly modified to prepare microsomes from colorectal metastases of human liver and matched liver samples. The tumours have a paucity of smooth endoplasmic reticulum and therefore were homogenized using a method which involved freezing the tissue in liquid nitrogen, transformation into a powdered form and homogenization, using a pestle and mortar (see 2.2.4 section for method). A visual examination of the samples during microsomal preparation showed a clear difference between microsomal samples isolated from tumours and those from liver. This difference was observed from the color of the samples, as hepatocytes present in the normal liver give a reddish yellow color, on the other hand, the tumour samples were white, fibrous and hard.

The results of the LC-MS/MS analysis showed that the CYPs expressed in the colon metastases were very similar in profile to the surrounding liver. The metastases were

shown to be free from liver tissue by histological examination; therefore this similarity is unlikely to be as a result of hepatocyte contamination. The data shows, for the first time, the direct and unequivocal detection of multiple CYPs in colorectal metastases (see Fig 4.4). All the major CYP isoforms involved in drug metabolism were identified in normal liver i.e. CYP1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, 3A4, were identified and CYP 4A11, 4F2 and 4F11 involved in ecosanoid metabolism. However, CYP isoforms were detected in both normal liver and colorectal metastasis and no obvious difference was found in CYP expression. It is intriguing that the generally close similarity of the tumour and the normal CYP expression may reflect the influence of liver tissue environment on the control of expression of CYPs in a tumour of extrahepatic origin. The influence of the liver environment on the CYP expression profile of colorectal metastases led to the investigation of the presence of CYPs in primary colon tumours and normal colon. Two individuals which had primary colon tumors with matched normal colon were obtained from the Royal Free Hospital.

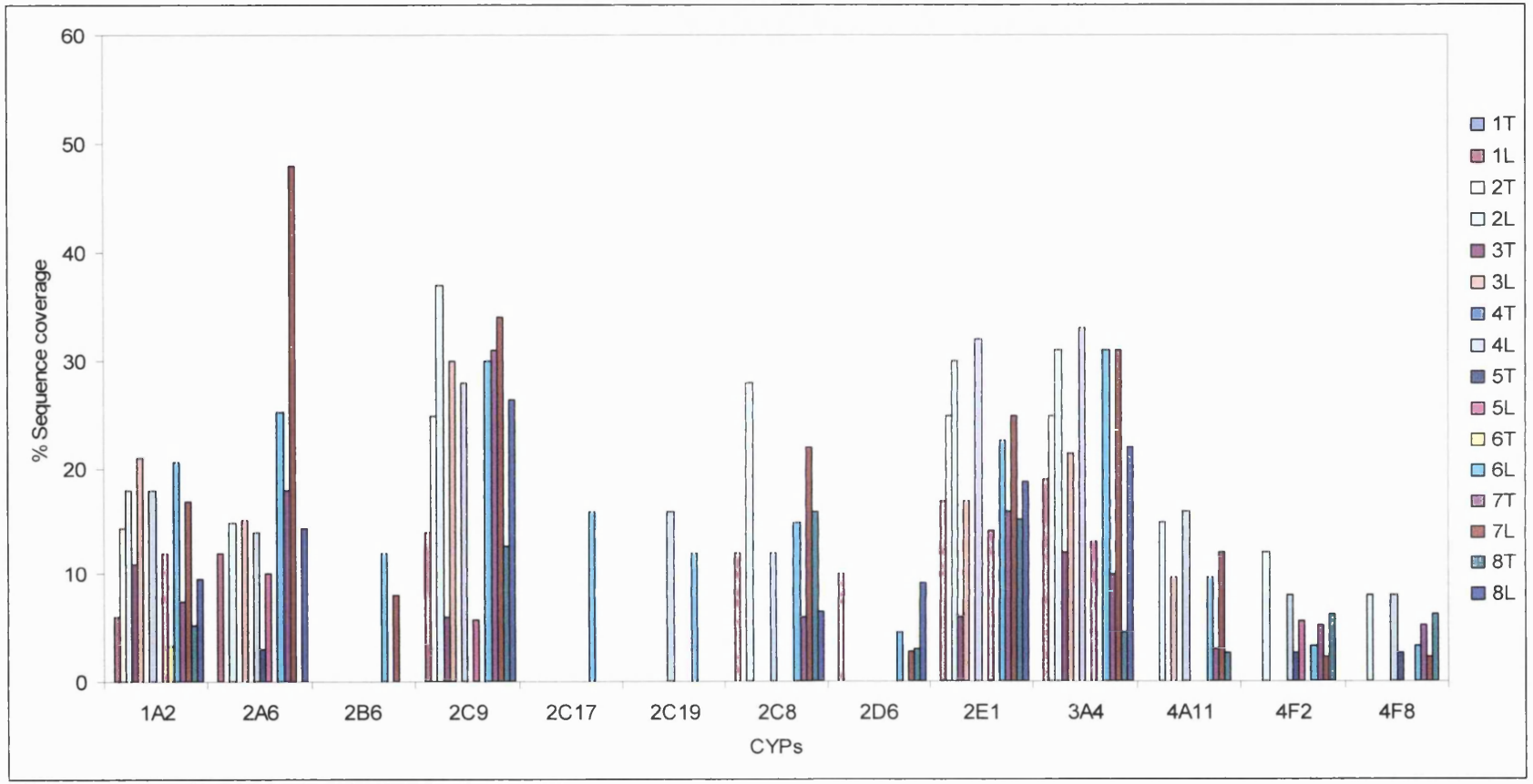


Fig 4.4: Comparison of Cytochrome P450 isoforms identified from normal human liver and colorectal metastasis.

Previously, western blot analysis of 23 samples isolated from histologically non-neoplastic human colon tissues showed that CYP2E1 and 3A proteins were present in each sample while CYP1A1, 2C9, 2C19 and 2D6 were present in some samples, but not in others (Bernauer *et al.,* 2002). A few studies have investigated the mRNA and protein expressions of CYP1A1, CYP1A2, CYP2C9/10, CYP2E1 and CYP3A in human colon (CaCo-2) cells (Lampen *et al.,* 1998).

The occurrence of CYP 3A-dependent metabolism in colorectal cancer tissue was previously demonstrated using the ability of CYP3A to metabolize the anticancer drug paclitaxel (Martinez *et al.*, 2002). Our LC-MS/MS results obtained from the normal colon identified CYP3A3 and/or CYP3A4. The peptides identified were the same in both CYPs therefore could not be used to distinguish whether one or both of the 3A isoforms were present. Previous studies have shown that CYP3A is the most abundant CYP present in small intestinal epithelium (Zhang *et al.*, 1999). Indeed, it has been suggested that vitamin D receptor-mediated induction of CYP3A4 may constitute a chemoprotective mechanism for detoxification of enteric xenobiotics and carcinogens and absence from the colon of certain of these enzymes, especially CYP3A4, may be involved in the comparatively high incidence rate of carcinogenesis in this organ (Thompson *et al.,* 2002).

The CYPs expressed in tumours may contribute to the resistance of these tumours to chemotherapy, for example taxol and irinotecan are deactivated by CYP3A4. Cyclophosphamide on the other hand is activated by CYP2B6 which was not detected in these samples of metastatic tumour.

Cytokeratin alterations have been proposed from different anatomical sites. Examples of proposed function of cytokeratins involve the participation in cell signaling and cholesterol transport. Glandular epithelia occurring in the intestine express mainly cytokeratin 8 and 18 (Liao and Omary, 1996). Recently, the increased transcription of cytokeratin 18 in human colon cancer line was observed (Prochasson *et al.,* 1999). In this study, cytokeratin 18 was identified in primary colon tumours with high sequence coverage (Table 3.9).

4.4. Experimental tumours and MCF-7 cell lines

Human tumours, grown as xenografts in immunodeficient mice are widely used models to study the activity of anticancer drugs. The xenograft model is considered to provide an accurate representation of the behavior of these compounds in human tumour tissues. In addition this approach has been successfully employed to study regulation of CYP expression in human colon and breast tumour xenografts (Smith *et al.,* 1993). The previous studies demonstrated that human breast and colon tumours constitutively expressed low levels of CYPs (notably CYP 2A, 2B, 2C, 3A and 4A). In the present study, there were no CYPs identified from the xenografts. This could be explained in terms of the amount of the sample available for the analysis. The xenografts were whitish and of very low protein content (see Table 3.10).

There could be several reasons for failure to identify a protein, or to match a peptide spectrum when searching a protein database; but predominantly because of the low abundance of a protein which may be present below the limit of detection of the method. Other explanations for failure to match a peptide are that the peptide sequence is not in the database searched, the presence of an unsuspected post-translational modification, the peptide is the result of non-specific cleavage or the spectrum is of poor quality.

MCF-7 is a stable cell line derived from a metastatic adenocarcinoma of the human breast that is a widely used model of estrogen-dependent tumours. Many primary tumours, including MCF-7 have been shown to express drug metabolizing enzymes (Murray *et al.,* 1997). However, the expression of CYPs in human tumour cell lines and xenografts is still not clear. McFadyen and coworkers (2001) have observed that MCF-7 cell lines do not express CYP1B1. In contrast Spink *et al.* (1998) showed that TCDD induce the expression of CYP1A1 and 1B1 in MCF-7 cell lines. The ability of TCDD to alter the expression of certain cytochrome P450 enzymes might play a role in the etiology of breast cancer. Previous studies have shown that some CYP enzymes are involved in the oxidative metabolism of the estrogens 17β-estradiol (E_2) and estrone (E_1). A major estrogen metabolizing pathway involves hydroxylation catalyzed by CYP1A1 and CYP1B1 (Badawi *et al.,* 2001). In addition, high mRNA

levels for CYP1B1, estrogen 4-hydroxylase, are found in malignant mammary tissues (Murray *et al.,* 1997 and Muskhelishvili *et al.,* 2001).

CYP1B1 and 1A1 expression in MCF-7 cell lines has previously been shown by Murray *et al.* (1997) using immunohistochemistry studies. In the present study CYP1B1 is not identified by Sequest searching of MS/MS data generated from TCDD induced MCF-7 cell lines. However, one of the peptides identified from MCF-7 cell lines is present in CYP 1A1, DITDSLIEHCQEKQLDENANVQLSDEK. This peptide has one miss cleavage and three basic amino acids, which are one H (His) and two K (lys). Although there was only one peptide identified from CYP1A1, a good XCorr (2.8) value and delta normalized correlation value suggested that CYP1A1 is present in microsomes isolated from MCF-7 cell lines induced with TCDD. The failure to identify CYP1B1 could be due to the fact that this CYPs is expressed at a very low level in cells or the mass spectrometer is not sensitive enough to pick out less abundant proteins from the mixture as there were high amounts of keratin in the sample. The high MS signals generated from keratin rich samples may have reduced the chances of less abundant protein to be subjected to MS/MS.

The results also showed that keratin was identified with high sequence coverage in all the tissues isolated from the breast tissues including MCF-7 cell lines and MDA-MB-435 (Table 3.11). It has been shown in many studies that serum concentrations of cytokeratin-18 and -19 correlated with disease progression in various malignancies and thus have been used as tumour markers (Nisman *et al.,* 1998; Doweck *et al.,* 2000). Cytokeratin-18 and -19 can be released into the extracellular space during epithelial cell apoptosis (Sheard *et al.,* 2002).

In ordered to enhance the probability of detecting peptide ions of low relative abundance (that is those co-elute with other peptide ions present at higher levels) MS/MS multiple injection and narrowing scan range experiments were carried out.

4.5. Optimisation of protein identification

When complex digest mixtures are analysed the use of a wide *m/z* range for ion selection of peptides results in low reproducibility of peptide ion selection between LC-ES-MS/MS runs. When mixtures of proteins digested with trypsin, it generates many different peptides which are thus available for ion selection; it is not surprising that the reproducibility for each set is low. This is due to the high number of components in a sample as low complexity samples show very good reproducibility. Another factor which could influence reproducibility is chromatographic conditions, as in such cases where chromatographic separation generates several overlapping peptides (that is co-eluting peptides), fewer analytical scans per unit time may prevent some peptides from being subjected to an MS/MS analysis.

CYPs can be upregulated or down regulated as a result of environmental factors. These factors may change the amino acid sequence of CYPs and in ordered to identify a modified amino acid in the sequence it is important to identify maximum possible peptides present in the sequence of a protein. Higher sequence coverage is very useful to identify polymorphic amino acid sites within the sequence and posttranslational modifications.

The narrow scan range experiments consisted of three LC-ESI-MS/MS analyses, each conducted over a unique *m/z* range, namely 400-1000, 1000-1500, 1500-2000. Results from these three analyses were compared to those from triplicate LC-ESI-MS/MS analysis over the full *m/z* range, 400-2000. As expected, fewer unique proteins were identified from the triplicate analysis of the liver and metastases samples using the wide or full *m/z* run (400-2000) than by analysis using three different narrow *m/z* ranges covering the same total *m/z* range.

The number of mass chromatographic peaks, and the proteins identified using narrow scan range analysis within 400-1000, 1000-1500, 1500-2000 *m/z* range were more than the proteins analysed using full scan range experiments. One of the limitations of the LC-MS/MS approach used for the identification of proteins present in complex mixtures is that only the most abundant ions are selected for sequence analysis by CID due to time constraints, and the co-elution of peptides from the column. Thus, peptide ions of low relative abundance (that is those that co-elute with other peptide ions

present at higher levels) are not amenable to identification. This indicates that the probability to omit a peptide for fragmentation is greater in the full scan range experiments. That is because when the data dependent MS/MS is performed on the first most intense ion from the first scan, then on second most intense ion and lastly on the third most intense ion from the same scan, by the time it comes to the third ion, the intensity of it may have fallen to prevent it from generating a good quality MS/MS selection.

Multiple data dependent LC-ESI-MS/MS analyses of the complex mixture of peptides produced by proteolysis of the liver sample had low reproducibility both at the level of individual peptides selected for CID and the identified parent protein. The obvious implication of this experiment is that failure to identify a protein from a complex mixture of peptides by data dependent LC-ESI-MS/MS analyses cannot be used as evidence that the protein is not in the sample. This is also true following multiple analysis of the same sample, as the results showed that the number of proteins identified from a single injection were less than the triplet analysis.

4.6. Absolute quantitation of proteins

As the goal of this study was to identify the CYP proteins in liver microsomes, it was decided not to pursue a stable-isotope dilution approach, such as the ICAT method (Gygi *et al.,* 1999), for relative quantification. Quantitation of peptide and protein mixtures by mass spectrometry has been a challenging analytical problem, largely because of ionization suppression among coeluting species (King *et al.,* 2000). The ICAT approach solves this problem by isotope labeling of peptides containing cysteine residues and is currently used for quantitative profiling of proteins in mixtures (Gygi *et al.,* 1999). With stable-isotope dilution methods proteins isolated from different sources are differentially stable-isotope labeled, combined, digested, and then analyzed by LC-MS/MS. This approach requires dilution of one sample by the other and also the analysis of selectively labeled peptides. Such a procedure would inherently limit differentiation between CYP isoforms with high sequence similarity, and may also elevate the CYP protein detection limit. In contrast to relative quantification strategies (ICAT), it is possible to use synthetic peptides with incorporated stable isotopes to provide absolute quantification (AQUA).

The AQUA strategy has two stages.

Stage 1 involves the selection and standard synthesis of a peptide from the protein of interest. The synthesised peptides were purchased and an amino acid residue was specified for 7 heavy atom labels (six ^{13}C and one ^{15}N) incorporation. In the case of myoglobin the peptide specified was LFTGHPETL*EK where L was selected for heavy isotope incorporation. The peptide labeled standards and native peptides were analyzed by MS/MS (see Fig 3.18) to examine the peptide fragmentation pattern. From the fragmentation patterns of a native peptide of *m/z* 636.40 (+2 charged), a fragment ion of *m/z* 716.40 (y_6) was selected for monitoring. For the heavy isotope labeled peptide the *m/z* of the parent peptide was 639.90 (as doubly charged therefore 7 Da / 2 = 3.5 Th addition to peptide mass) and fragment ion of *m/z* 723.4 (7 Da addition as singly charged ion) were selected to be monitored. The mass spectrometer was then set to perform SRM analysis in which a parent to product transition (636.40 → 716.40 for the native peptide and 639.90 → 723.40 for the isotope labeled peptide).

Stage 2 involves SDS-PAGE separation of purified horse heart myoglobin, proteolysis with trypsin in the presence of the AQUA heavy isotope labelled peptide. Then the abundance of a fragment ion from both the native and the labeled peptide were compared. As the molarity of native protein loaded onto the gel was exactly the same as the internal peptide concentration added to the gel pieces, it was expected that the peak areas for both would be the same. However, the results obtained indicated only 10% recovery of native peptide from the gel pieces. In ordered to make sure that the solutions were prepared correctly, a solution digest of myoglobin was carried out. To prepare a solution digest, horse heart myoglobin was digested by trypsin in the presence of organic solvents such as methanol, acetone, 2-propanol, and acetonitrile. The results showed that there were no major errors in preparations of solutions. There are several points which have to be considered to explain the discrepancy in the 1D gel experiment, including loss of protein while loading on the gel, the presence of salts in the myoglobin stock could lead to errors in the weight of the actual protein loaded.

After going through all the steps which were involved in preparing samples it was thought that a problem may arise at the stage when the trypsin and internal peptide were added to the dried gel pieces, initially the trypsin added 20 min after than internal peptide. As the labelled peptide was added first it was absorbed into the gel pieces efficiently, but trypsin was absorbed poorly into the gel pieces, and this in turn decreased the chances of digestion of the native myoglobin which was present in the gel pieces.

Another step where improvement was made was the extraction step which was initially carried out using 5 % formic acid and 50 % acetonitrile, followed by sonication, centrifugation and collection of supernatant in a tube. This was repeated two times. It was improved by extracting once using 5 % formic acid and 50 % acetonitrile, sonicating the gel pieces and pelleting down the gel pieces, collecting supernatant in the tube, then shrinking the gel pieces using 100 % acetonitrile and collecting the solution, rehydrating using 5% formic acid and 50 % acetonitrile and repeating the initial step. This was done three times. These two changes improved the recovery to 80 % of native peptide from the gel pieces. This experiment was performed 4 times to make sure that the results were reproducible. The method was further evaluated by adding different concentrations of horse heart myoglobin to a

human liver sample. The results confirmed that the quantitation method can also be used for quantitative profiling of proteins in complex mixtures such as liver microsomes.

In principle, this method can be extended to identify unique isotopically labelled peptides corresponding to all the CYPs present in human liver or any other complex tissue.

Summary and conclusions

Using a LC-ES-MS/MS method, male and female rat livers were studied for gender-specific CYPs. Sex-dependent differences in xenobiotic metabolism were found to be very pronounced in rats, as fourteen isoforms were shown to be gender biased. The study was extended to the identification of CYPs in human livers obtained from patients with different medical histories. The results showed the identification of fifteen CYP isoforms. The differences in the expression of xenobiotic metabolising CYPs among humans are likely to be a consequence of intra-individual variations as a result of genetics or environmental exposures rather than from sex-dependent differences in enzyme composition.

The CYPs were then identified from the colorectal metastasis in the liver (secondary colorectal tumours). The CYP profile of the colorectal metastases was compared to that of normal liver from the same individual and the results from these showed that CYP expression profiles are similar to the normal liver samples. From this it was concluded that it was the liver environment influencing the colorectal metastases to express a range of CYPs. The results obtained from the colorectal metastasis stimulated our interest to study the primary colon and primary tumour for CYP content and to compare their CYP profile to that from the secondary colorectal tumours. The results obtained identified only one CYP isoform (CYP3A). These results showed that the CYPs present in colon are less expressed than in liver.

It is also concluded that individuals investigated in this study have their own unique CYP profile probably due to a combination of genetic and environmental factors including the plethora of drugs used in treating their disease.

The successful identification of CYPs from complex mixtures necessitates developing further methods, which could improve the total sequence coverage for the identified proteins. This would not only enhance the chances of identification of less abundant CYPs present in extraheptic tissues but also help locate modifications in the amino acid sequences. From our results, it can be suggested that the triplicate analysis proved to be optimal.

The knowledge of the presence or absence of a protein in a tissue is important, but equally important is the measure of the amounts of protein present. An isotope labelled peptide strategy was successfully used to quantitate horse heart myoglobin.

In this study, the routinely applicable use of nano-scale liquid chromatography-electrospray-tandem mass spectrometry (LC-ES-MS/MS) in the identification of multiple CYPs from human liver was established.

References

Adams, D. J., Seilman, S., Amelizad, Z., Oesch, F. and Wolf, C. R. (1985) Identification of human cytochromes P450 analogous to forms induced by phenobarbital and3-methylcholanthrene in the rat. Biochem J. 232, 869-876.

Aebersold, R. and Goodlett, D. R. (2001) Mass spectrometry in proteomics. Chem Rev. 101, 269-95.

Aebersold, R., Gygi, S. P., Rist, B. and Gerber, S. A. (1999) Quantitative analysis of complex protein mixtures using ICATs. NAT Biotechnol. 17: 994-999.

Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. Nature. 422(6928), 198-207.

Agrawal, A. K. and Sharpiro, B. H. (2001) Intrinsic signals in sexually dimorphic circulating growth hormone profiles of the rat. Mol Cell Endocrinol. 173, 167-181.

Anakk, S., Ku, C. Y., Vore, M. and Strobel, H. W. (2003) Insights into gender bias: rat cytochrome P450 3A9. J Pharmacol Exp Ther. 305, 703-9.

Anderson, L. and Seilhamer, J. (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis. 18, 533-7.

Badawi, A. F., Cavalieri, E. L. and Rogan, E. G. (2001) Role of human cytochrome P450 1A1, 1A2, 1B1, and 3A4 in the 2-, 4-, and 16alpha-hydroxylation of 17beta-estradiol. Metabolism. 50, 1001–1003.

Bernauer, U., Ellrich, R., Heinrich-Hirsch, B., Teubner, W., Vieth, B. and Gundert-Remy, U. (2002) Expression of cytochrome P450 enzymes in human colon. IARC Sci Publ. 156, 487-9.

Bertz, R. J. & Granneman, G. R. (1997) Use of in vitro and in vivo data to estimate the likelihood of metabolic pharmacokinetic interactions. Clin Pharmacokinet. 32, 210-258.

Bradford, M. M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of dye-binding. Anal Biochem. 72, 248-54.

Cai, H. and Guengerich, F. P. (2001) Reaction of trichloroethylene and trichloroethylene oxide with Cytochrome P450 enzymes: inactivation and sites of modification. Chem Res Toxicol. 14, 451-458.

Carr, S. A. and Annanu, R. S. (1997) Overview of peptides and protein analysis by mass spectrometry. In: Current Protocols in Molecular Biology (Ausubel, F. M. *et al.*, ed.), pp. 10.21.1-10.21.27, Wiley, New York.

Chang, T. K., Chen, J., Pillay, V., Ho, J. Y. and Bandiera, S. M. (2003) Real-time polymerase chain reaction analysis of CYP1B1 gene expression in human liver. Toxicol Sci. 71, 11-9.

Chen, G., Gharib, T. G., Huang, C., Taylor, J. M. G., Misek, D. E., Kardia, S. L. R., Giordano, T. J., Iannettoni, M. D., Orringer, M. B., Hanash, S. M. and Beer, D. G. (2002) Disordant protein and mRNA expression in lung adenocarcinomas. Mol. Cell Proteomics. 1, 304-313.

Chow, T., Imaoka, S., Hiroi, T. and Funae, Y. (1999) developmental changes in the catalytic activity and expression of CYP2D isoforms in rat liver. Drug Metab Dispos. 27, 188-192.

Clauser, K. R., Baker, P. and Burlingame, A. L. (1999) Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal Chem. 71, 2871-82.

Czerniak, R. (2001) Gender-based differences in pharmacokinetics in laboratory animal models. Int J Toxicol. 20, 161-3.

Dalet-Beluche, I., Boulenc, X., Fabre, G., Maurel, P. & Bonfils, C. (1992) Purification of two cytochrome P450 isozymes related to CYP2A and CYP3A gene families from

monkey (baboon, Papio papio) liver microsomes. Cross reactivity with human forms. Eur J Biochem. 204, 641-648.

De Wildt, S. N., Kerans, G.L., Leeder, J.S. and van den Anker, J.N. (1999) Cytochrome P450 3A: ontogeny and drug disposition. Clin Pharmacokin. 37, 485-505.

Ding, X. and Kaminsky, L. S. (2003) Human extrahepatic cytochromes P450: function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts. Annu Rev Pharmacol Toxicol. 43, 149-73.

Doherty, M. M. and Michael, M. (2003) Tumoral drug metabolism: perspectives and therapeutic implications. Current drug metabolism. 4, 131-149.

Doweck, I., Barak, M., Uri, N. and Greenberg, E. (2000) The prognostic value of the tumour marker Cyfra 21-1 in carcinoma of head and neck and its role in early detection of recurrent disease. Br J Cancer. 83, 1696-701.

Ducret, A., Van Oostveen, I., Eng, J. K., Yates III, J. R. and Aebersold, R. (1998) High throughput protein characterisation by automated reverse-phase chromatography/electrospray tandem mass spectrometry. Protein Science. 7, 706-719.

Edwards, R. J., Adams, D. A., Watts, P. S., Davies, D. S. and Boobis, A. R. (1998) Development of a comprehensive panel of antibodies against the major xenobiotic metabolising forms of cytochrome P450 in humans. Biochem Pharmacol. 56, 377-387.

Eng, J.K., McCormack, A.L. and Yates III, J.R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom. 5, 976–989.

Eltom, S. E., Larsen, M. C. and Jefcoate, C. R. (1998) Expression of CYP1B1 but not CYP1A1 by primary cultured human mammary stromal fibroblasts constitutively and in response to dioxin exposure: role of the Ah receptor. Carcinogenesis. 19, 1437-1444.

Eng, J. K., McCormack, A. L. and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectom. 5, 976-989.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science. 246, 64–71.

Fenyö, D. (2000) Identifying the proteome: software tools. Curr Opinion Biotechnol. 11, 391-5.

Forbes, A. J., Mazur, M. T., Patel, H. M., Walsh, C. T., Kelleher, N. L. (2001) Towards efficient analysis of >70 kDa proteins with 100% sequence coverage Proteomics. 1, 927-33.

Forrester, L. M., Henderson, C. J., Glancey, M. J., Back, D. J., Park, B. K., Ball, S. E., Kitteringham, N. R., McLaren, A. W., Miles, J. S., Skett, P. (1992) Relative expression of cytochrome P450 isoenzymes in human liver and association with the metabolism of drugs and xenobiotics. Biochem J. 281, 359-68.

Galeva, N. and Alterman, M. (2002) Comparison of one-dimensional and two-dimensional gel electrophoresis as a separation tool for proteomic analysis of rat liver microsomes: Cytochrome P450 and other membrane proteins. Proteomics. 2, 713-722.

Galeva, N., Yakovlev, D., Koen, Y., Duzhak, T. and Alterman, M. (2003) Direct identification of Cytochrome P450 isoenzymes by matrix assisted laser desorption/ionisation time of flight-based proteomic approach. Drug Metab Dispos. 31, 351-355.

Garcia, M. C., Thangavel, C. and Shapiro, B. H. (2001) Epidermal growth factor regulation of female-dependent CYP2A1 and CYP2C12 in primary hepatocytes culture. Drug Metab Dispos. 29, 111-120.

Garrrels, J. I., McLaughlin, C. S., Warner, J. R., Futcher, B., Latter, G. I., Abdersen, D. S. and Payne, W. E. (1997) Proteome studies of Saccharomyces cerevisiae: identification and characterisation of abundant proteins. Electrophoresis. 18, 1347-1360.

Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W. and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A. 100, 6940-5.

Gervot, L., Carriere, V., Costet, P., Cugnenc, P. H., Berger, A., Beaune, P. and de Waziers, I. (1996) CYP3A5 is the major cytochrome P450 3A expressed in human colon and colonic cell lines. Environ Tox Pharm. 2, 381-388.

Gomez, A. and Tang, A. (1994) Charge and fission of droplets in electrostatic sprays. Phys Fluids. 6, 404-414.

Gonzalez, F. J. (1989) The molecular biology of cytochrome P450s. Pharmacol Rev. 40, 243-88.

Gonzalez, F. J. (1992) Human cytochromes P450: problems and prospects. Trends Pharmacol Sci. 13, 346-352.

Graham, S. E. and Peterson, J. A. (1999) How similar are P450s and what can their differences teach us? Arch Biochem Biophys. 369, 24-9.

Griffiths W. J., Andreas, P. J., Suya, L., Dilip, K. R. and Yuqin, W. (2001) Electrospray and tandem mass spectrometry in biochemistry. Biochem J. 355, 545–561.

Guengerich, F. P. (1990) Mechanism-based inactivation of human liver microsomal cytochrome P-450 IIIA4 by gestodene. Chem Res Toxicol. 3, 363-71.

Gygi, S. P., Rochon, Y., Franza, B. R. and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol. 19, 1720-1730.

Haehner, B. D., Gorski, J. C., VandenBranden, M., Wrighton, S. A., Janardan, S. K., Watkins, P. B. and Hall, S. D. (1996) Bimodal distribution of renal cytochrome P450 3A activity in humans. Mol Pharmacol. 50, 52-59.

Hakkola, J., Pasanen, M., Purkunen, R., Saarikoski, S., Pelkonen, O., Mäenpää, J., Rane, A. and Raunio, H. (1994) Expression of xenobiotic-metabolizing cytochrome P450 forms in human adult and fetal liver. Biochem Pharmacol. 48, 59-64.

Hakkola, J., Pasanen, M., Hukkanen, J., Pelkonen, O., Mäenpää, J., Edwards, R. J., Boobis, A. R. and Raunio, H. (1996) Expression of xenobiotic-metabolizing cytochrome P450 forms in human full-term placenta. Biochem Pharmacol. 51, 403-411.

Hakkola, J., Pasanen, M., Pelkonen, O., Hukkanen, J., Evisalmi, S., Anttila, S., Rane, A., Mäntylä, M., Purkunen, R., Saarikoski, S., Tooming, M. and Raunio, H. (1997) Expression of CYP1B1 in human adult and fetal tissues and differential inducibility of CYP1B1 and CYP1A1 by Ah receptor ligands in human placenta and cultured cells. Carcinogenesis. 18, 391-397.

Hames, B. D. (1998) Gel elctrophoresis of proteins: a practical approach (3rd ed.), Oxford University Press, Oxford.

Harry, J. L., Wilkins, M. R., Herbert, B. R., Packer, N. H., Gooley, A. A. and Williams, K. L. (2000) Proteomics: capacity versus utility. Electrophoresis. 21, 1071-81.

He, K., Bornheim, L. M., Falick, A. M., Maltby, D., Yin, H. and Correia, M. A. (1998) Identification of the heme-modified peptides from cumene hydroperoxide-inactivated Cytochrome P450 3A4. Biochemistry. 37, 17448-17457.

Hellmold, H., Lamb, J. G., Wyss, A., Gustafsson, J. A. and Warner, M. (1995) Developmental and endocrine regulation of P450 isoforms in rat breast. Mol Pharmacol. 48, 630-8.

Henzel, W. J., Billeci, T. M., Stultz, J. T., Wong, S. C., Grimley, C. and Wantanabe, C. (1993) Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc Natl Acad Sci USA. 90, 5011–5015.

Holla, V. R., Adas, F., Imig, J. D., Zhao, X., Price, E., Jr Olsen, N., Kovacs, W. J., Magnuson, M. A., Keeney, D. S., Breyer, M. D., Falck, J. R., Waterman, M. R. and Capdevila, J. H. (2001) Alterations in the regulation of androgen-sensitive CYP4A monooxygenases cause hypertension. Proc Natl Acad Sci USA. 98, 5211-6.

Huang, L., Baldwin, M. A., Maltby, D. A., Medzihradszky, K. F., Baker, P. R., Allen, N., Rexach, M., Edmondson, R. D., Campbell, J., Juhasz, P., Martin, S. A., Vestal, M. L. and Burlingame, A. L. (2002) The identification of protein-protein interactions of the nuclear pore complex of Saccharomyces cerevisiae using high throughput matrix-assisted laser desorption ionization time-of-flight tandem mass spectrometry. Mol Cell Proteomics. 1, 434-50.

Huang, Z., Fasco, M. J., Figge, H. L., Keyomarsi, K. and Kaminsky, L. S. (1996) Expression of cytochromes P450 in human breast tissue and tumors. Drug Metab Dispos. 24, 899-905.

Ingelman-Sundberg, M. (2004) Human drug metabolising cytochrome P450 enzymes: properties and polymorphisms. Nauyn Schmeidebergs Arch Pharmacol. 369, 89–104.

Jass, J. R. (1999) Broadsheet number 52: Molecular genetics of colorectal cancer. Pathology. 31, 354-64.

Jensen, O. N., Podtelejnikov, A. V. and Mann, M. (1997) Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. Anal Chem. 69, 4741-50.

Kalsotra, A., Anakk, S., Boehme, C. L. and Strobel, H. W. (2002) Sexual dimorphism and tissue specificity in the expression of CYP4F forms in Sprague Dawley rats. Drug Metab Dispos. 30, 1022-8.

Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 60, 2299-301.

Kashino Y, Koike H, Satoh K. (2001) An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. Electrophoresis. 22, 1004-7.

Kato, R. and Yamazoe, Y. (1993) Hormonal regulation of cytochrome P450 in rat liver, in Cytochrome P450. Springer-Verlag. Berlin.

King, R., Bonfiglio, R., Fernandez-Metzler, C., Miller-Stein, C. and Olah, T. (2000) Mechanistic investigation of ionization suppression in electrospray ionization. J Am Soc Mass Spectrom. 11, 942-50.

Kitada, M., Kamataki, T., Itahashi, K., Rikihisa, T., Kato, R. and Kanakubo, Y. (1985) Purification and properties of cytochrome P-450 from homogenates of human fetal livers. Arch Biochem Biophys. 241, 275-280.

Kitada, M. and Kamataki, T. (1994) Cytochrome P450 in human fetal liver: significance and fetal specific expression. Drug Metab Rev. 26, 305-323.

Kivistö, K. T., Griese, E-U., Fritz, P., Linder, A., Hakkola, J., Raunio, H., Beaune, P. and Kroemer, H. K. (1996) Expression of cytochrome P4503A enzymes in human lung: a combined RT-PCR and immunohistochemical analysis of normal tissue and lung tumors. Naunyn-Schmiedeberg Arch Pharmacol. 353, 207-212.

Koenigs, L. L., Peter, R. M., Hunter, A. P., Haining, R. L., Rettie, A. E., Friedberg, T., Pritchard, M. P., Shou, M., Rushmore, T. H. and Trager, W. F. (1999) Electrospray ionisation mass spectrometric analysis of intact Cytochrome P450: identification of tielenic acid adducts to P450 2C9. Biochemistry. 38, 2312-2319.

Korzekwa, K. R., Krishnamachary, N., Shou, M., Ogai, A., Parise, R. A., Rettie, A. E., Gonzalez, F. J. and Tracy, T. S. (1998) Evaluation of atypical cytochrome P450

kinetics with two-substrate models: evidence that multiple substrates can simultaneously bind to cytochrome P450 active sites. Biochemistry. 37, 4137-47.

Koskela, S., Hakkola, J., Hukkanen, J., Pelkonen, O., Sorri, M., Saranen, A., Anttila, S., Fernandez-Salguero, P., Gonzalez, F. J. and Raunio, H. (1999) Expression of CYP2A genes in human liver and extrahepatic tissues. Biochem Pharmacol. 57, 1407-1413.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227 (5259), 680-5.

Lane, C. S., Patterson, L. H. and Welham, K. J. (2002) Analysis of human cytochrome P450 isoforms using nanoelectrospray ionization tandem mass spectrometry. AAPS PharmSci Supplement. 4:4.

Lechevrel, M., Casson, A. G., Wolf, C. R., Hardie, L. J., Flinterman, M. B., Montesano, R. and Wild, C. P. (1999) Characterization of cytochrome P450 expression in human oesophageal mucosa. Carcinogenesis. 20, 243-248.

Lewis, D. F. V. (1992) Computer-assisted methods in the evaluation of chemical toxicity, Review in Computational Chemistry. 3, 173-222.

Lewis, S., Korsmeyer, K. K. and Correia, M. A. (1993) Matrix-assisted laser desorption mass spectrometry of cytochromes P450. Rapid Commun Mass Spectrom. 7, 16-9.

Li, D. N., Seidel, A., Pritchard, M. P., Wolf, C. R. and Friedberg, T. (2000) Polymorphisms in P450 CYP1B1 affect the conversion of estradiol to the potentially carcinogenic metabolite 4-hydroxyestradiol. Pharmacogenetics. 10, 343-353.

Liao, J. and Omary, M. B. (1996) 14-3-3 proteins associate with phosphorylated simple epithelial keratins during cell cycle progression and act as a solubility cofactor. J Cell Biol. 133, 345-57.

Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A. and Hemminki, K. (2000) Environmental and heritable factors in the causation of cancer analyses of cohorts of twins. N Engl J Med. 343, 78-85.

Lightning, L. K., Jones, J. P., Friedberg, T., Pritchard, M. P., Shou, M., Rushmore, T. H. and Trager, W. F. (2000) Mechanism-based inactivation of Cytochrome P450 3A4 by L-754, 394. Biochemistry. 39, 4276-4287.

Lieber, C. S. (1997) Cytochrome P-4502E1: Its physiological and pathological role. Physiol Rev. 77, 517-544.

Lewis, D. F. V. (1992) Computer-assisted methods in the evaluation of chemical toxicity. Review in Computational Chemistry. 3, 173-222.

Lewis, S., Korsmeyer, K. K. and Correia, M. A. (1993) Matrix-assisted laser desorption mass spectrometry of cytochromes P450. Rapid Commun Mass Spectrom. 7, 16-9.

Ioannides, C. and Parke, D. V. (1993) Induction of cytochrome P4501 as an indicator of potential chemical carcinogenesis. Drug Metab Rev. 25, 485-501.

Mahoney, W. C. and Hermodson, M. H. (1980) Separation of large denatured peptides by reverse phase high performance liquid chromatography. Trifluoroacetic acid as a peptide solvent. J of Biological Chemistry. 255, 11199-203.

Martinez, C., Garcia-Martin, E., Pizarro, R. M., Garcia-Gamito, F.J. and Agundez, J. A. (2002) Expression of paclitaxel-inactivating CYP3A activity in human colorectal cancer: implications for drug therapy. Br J Cancer. 87, 681-6.

Massaad, L., de Waziers, I., Ribrag, V., Janot, F., Beaune, P. H., Morizet, J., Gouyette, A. and Chabot, G. G. (1992) Comparison of mouse and human colon tumors with regard to phase I and phase II drug-metabolizing enzyme systems. Cancer Res. 52, 6567-75.

Masubuchi, Y., Iwasa, T., Hosokawa, S., Suzuki, T., Horie, T., Imaoka, S., Funae, Y. and Narimatsu, S. (1997) Selective deficiency of debrisoquine 4-hydroxylase activity in mouse liver microsomes. J Pharmacol Exp Ther. 282, 1435-41.

McFadyen, M. C., McLeod, H. L., Jackson, F. C., Melvin, W. T., Doehmer, J. and Murray, G. I. (2001) Cytochrome P450 CYP1B1 protein expression: a novel mechanism of anticancer drug resistance. Biochem Pharmacol. 62, 207-12.

McFadyen, C. E. M., Rooney, P. H., Melvin, W. T. and Murray, G. I. (2003) Quantitative analysis of the Ah receptor/Cytochrome P450 CYP1B1/CYP1A1 signalling pathway. Biochem Pharmacol. 65, 1663-1674.

McKay, J. A., Murray, G. I., Weaver, R. J., Ewen, S. W., Melvin, W. T. and Burke M. D. (1993) Xenobiotic metabolising enzyme expression in colonic neoplasia. Gut. 34, 1234-1239.

McKinnon, R. A., Burgess, W. M., Gonzalez, F. J., Gasser, R. and McManus, M. E. (1994) Species-specific expression of CYP4B1 in rabbit and human gastrointestinal tissues. Pharmacogenetics. 4, 260-270.

McKinnon, R. A., Hall, P. D., Quattrochi, L. C., Tukey, R. H. and McManus, M. E. (1991) Localization of CYP1A1 and CYP1A2 messenger RNA in normal human liver and in hepatocellular carcinoma by in situ hybridization. Hepatology. 14, 848-56.

McKinnon, R. A. and McManus, M. E. (1996) Localization of cytochromes P450 in human tissues: implications for chemical toxicity. Pathology. 28, 148-55.

Medzihradszky, K. F., Leffler, H., Baldwin, M. A. and Burlingame, A. L. (2001) Protein identification by in-gel digestion, high-performance liquid chromatography, and mass spectrometry: peptide analysis by complementary ionization techniques. J Am Soc Mass Spectrom. 12, 215-21.

Mitchell, S. R., Sewer, M. B., Kardar, S. S. and Morgan, E. T. (2001) Characterization of CYP4A induction in rat liver by inflammatory stimuli: dependence on sex, strain, and inflammation-evoked hypophagia. Drug Metab Dispos. 29, 17-22.

Mugford, C A. and Kedderis, G. L. (1998) Sex-dependent metabolism of xenobiotics. Drug Metab Rev. 30, 441-98.

Murray, G. I., Pritchard, S., Melvin, W. T. and Burke, M. D. (1995) Cytochrome P450 CYP3A5 in the human anterior pituitary gland. FEBS Lett. 364, 79-82.

Murray, G. I., Taylor, M. C., McFadyen, M. C. E., McKay, J. A., Greenlee, W. F., Burke, M. D. and Melvin, W. T. (1997) Tumor-specific expression of cytochrome P450 CYP1B1. Cancer Res. 57, 3026-3031.

Murray, G. I., Taylor, M. C., Burke, M. D. and Melvin, W. T. (1998) Enhanced expression of cytochrome P450 in stomach cancer. Br J Cancer. 77, 1040-4.

Murray, G. I., McFadyen, M. C., Mitchell, R. T., Cheung, Y. L., Kerr, A. C. and Melvin, W. T. (1999) Cytochrome P450 CYP3A in human renal cell cancer. Br J Cancer. 79, 1836-42.

Muskhelishvili, L., Thompson, P. A., Kusewitt, D. F., Wang, C. and Kadlubar, F. F. (2001) In situ hybridization and immunohistochemical analysis of cytochrome P450 1B1 expression in human normal tissues. J Histochem Cytochem. 49, 229–236.

Nebert, D. W., McKinnon, R. A. and Puga, A. (1996) Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. DNA Cell Biol. 15, 273-280.

Nakajima, M., Yamamoto, T., Nunoya, K. I., Yokoi, T., Nagashima, K., Inoue, K., Funae, Y., Shimada, N. and Kuroiwa, Y. (1996) Role of human cytochrome P4502A6 in C-oxidation of nicotine. Drug Metab Dispos. 24, 1212-1217.

Nelson, D. R. (1999) Cytochrome P450 and the individuality of species. Arch Biochem Biophys 369, 1-10.

Nelson, D. R. (2003) Comparison of P450s from human and fugu: 420 million years of vertebrate P450 evolution. Arch Biochem Biophys. 409, 18-24.

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

Newton, R. P., Brenton, A. G., Smith, C. J. and Dudley, E. (2004) Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments. Phytochemistry. 65, 1449-1485.

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

Nisar, S., Lane, C. S., Wilderspin, A. F., Welham, K. J., Griffiths, W. J. and Patterson, L. H. (2004) A proteomic approach to the identification of cytochrome P450 isoforms in male and female rat liver by nano-scale liquid chromatography electrospray ionisation tandem mass spectrometry. Drug Metab Dispos. 32, 382–386.

Nisman, B., Lafair, J., Heching, N., Lyass, O., Baras, M., Peretz, T. and Barak, V. (1998) Evaluation of tissue polypeptide specific antigen, CYFRA 21-1, and carcinoembryonic antigen in nonsmall cell lung carcinoma: does the combined use of cytokeratin markers give any additional information? Cancer. 82, 1850-9.

Omura, T. (1999) Forty years of Cytochrome P450. Biochem Biophys Res Commun. 266, 690-698.

Ozturk, M., Bolkent, S., Yilmazer, S., Kaner, G. and Unal, H. (1998) Detection of c-erbB-2 mRNAs using dig-labelled oligonucleotide probe with in situ hybridisation in human breast carcinoma: comparison with immunohistochemical results. Anal Cell Pathol. 16, 201-209.

Ortiz de Montellano, P. R. (1995) Cytochrome P450 structure, mechanism and biochemistry (1st ed.), Plenum press, New York.

Pampori, N. A. and Shapiro, B. H. (1999) Gender differences in the responsiveness of the sex-dependent isoforms of hepatic P450 to the feminine plasma growth hormone profile. Endocrinology. 140, 1245-54.

Patterson, L. H., McKeown, S. R., Robson, T., Gallagher, R., Raleigh, S. M. and Orr, S. (1999) Antitumour prodrug development using cytochrome P450 (CYP) mediated activation. Anticancer Drug Des. 14, 473-86.

Patterson LH, McKeown SR, Ruparelia K, Double JA, Bibby MC, Cole S, Stratford IJ. (2000) Enhancement of chemotherapy and radiotherapy of murine tumours by AQ4N, a bioreductively activated anti-tumour agent. Br J Cancer. 82, 1984-90.

Patterson, L. H. and Murray, G. I. (2002) Tumour Cytochrome P450 and drug activation. Curr Pharm Des. 8, 1335-47.

Pelkonen, O., Mäenpää, J., Taavitsainen, P., Rautio, A. and Raunio, H. (1998) Inhibition and induction of human cytochrome P450 (CYP) enzymes. Xenobiotica. 28, 1203-1253.

Pianezza, M. L., Sellers, E. M. and Tyndale, R. F. (1998) Nicotine metabolism defect reduces smoking. Nature. 393, 750-750.

Prochasson, P., Gunther, M., Laithier, M., Fossar, N., Lavialle, C. and Brison, O. (1999) Transcriptional mechanisms responsible for the overexpression of the keratin 18 gene in cells of a human colon carcinoma cell line. Exp Cell Res. 248, 243-59.

Rabilloud, T., Adessi, C., Giraudel, A. and Lunardi, J. (1997) Improvements of the solubilisation of proteins in 2DE with immobilized pH gradients, Electrophoresis. 18, 307-316.

Raleigh, S. M., Wanogho, E., Burke, M. D., McKeown, S. R. and Patterson, L. H. (1998) Involvement of human cytochromes P450 (CYP) in the reductive metabolism of AQ4N, a hypoxia activated anthraquinone di-N-oxide prodrug. Int J Radiat Oncol Biol Phys. 42, 763-7.

Raunio, H., Husgafvel-Pursiainen, K., Anttila, S., Hietanen, E., Hirvonen, A. and Pelkonen, O. (1995b) Diagnosis of polymorphisms in carcinogen-activating and inactivating enzymes and cancer susceptibility-review. Gene. 159, 113-121.

Raunio, H., Pasanen, M., Mäenpää, J., Hakkola, J. and Pelkonen, O. (1995a) Expression of extrahepatic cytochrome P450 in humans (Pacifici GM & Fracchia GN ed.). Advances in Drug Metabolism in Man: 234-287. European Commission, Office for Official Publications of the European Communities, Luxembourg.

Regal, K. A., Schrag, M. L., Kent, U. M., Wienkers, L. C. and Hollenberg, P. F. (2000) Mechanism-based inactivation of Cytochrome P450 2B1 by 7-ethynylcoumarin: verification of apo-P450 adduction by electrospray ion trap mass spectrometry. Chem Res Toxicol. 13, 262-270.

Robertson, G. R., Farrell, G. C. and Liddle, C. (1998) Sexually dimorphic expression of rat CYP3A9 and CYP3A18 genes is regulated by growth hormone. Biochem Biophys Res Commun. 242, 57-60.

Rodríguez-Antona, C., Jover, R., Gómez-Lechón, M. J. and Castell, J. V. (2000) Quantitative RT-PCR measurement of human cytochrome P-450s: application to drug induction studies. Arch Biochem Biophys. 376, 109-116.

Rodrigues, A. D. and Rushmore H. T. (2002) Cytochrome P450 Pharmacogenetics in Drug Development: In Vitro Studies and Clinical Consequences. Curr Drug Metab. 3, 559.

Roepstorff, P. and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom. 11, 601.

Rubin, E. and Farber, J. L. (1988) Pathology, 2nd edition. Published by J. B. Lippincott Company. Philadelphia.

Russell, W. K., Park, Z.-Y., Russell, D. H (2001) Proteolysis in Mixed Organic-Aqueous Solvent Systems: Applications for Peptide Mass Mapping Using Mass Spectrometry, Anal. Chem. 73, 2682-2685.

Santoni, V., Rabilloud, T., Doumas, P., Rouquie, D., Mansion, M., Kieffer, S., Garin, J. and Rossignal, M. (1999) Towards the recovery of hydrophobic proteins on 2-DE gels. Electrophoresis. 20, 705-711

Sarkar, M. A. and Jackson, B. J. (1994) Theophylline N-demethylations as probes for P4501A1 and P4501A2. Drug Metab Dispos. 22, 827-34.

Schmidt, J. V. and Bradfield, C. A. (1996) Ah receptor signalling pathways. Annu Rev Cell Dev Biol. 12, 55-89.

Schuetz, J. D., Beach, D. L. and Guzelian, P. S. (1994) Selective expression of cytochrome P450 CYP3A mRNAs in embryonic and adult human liver. Pharmacogenetics. 4, 11-20.

Schulz-Utermoehl, T., Bennett, A. J., Ellis, S. W., Tucker, G. T., Boobis, A. R. and Edwards RJ (1999) Polymorphic debrisoquine 4-hydroxylase activity in the rat is due to differences in CYP2D2 expression. Pharmacogenetics. 9, 357-66.

Schuetz, E. G., Schuetz, J. D., Grogan, W. M., Naray-Fejes-Toth, A., Fejes-Toth, G., Raucy, J., Guzelian, P., Gionela, K. and Watlington, C. O. (1992) Expression of cytochrome P450 3A in amphibian, rat, and human kidney. Arch Biochem Biophys. 294, 206-214.

Shan, X., Aw, T. Y., Smith, E. R., Ingelman-Sundberg, M., Mannervik, B., Iyanagi, T. and Jones, D. P. (1992) Effect of chronic hypoxia on detoxication enzymes in rat liver. Biochem Pharmacol. 43, 2421-6.

Shapiro, B. H., Agrawal, A. K., Pampori, N. A. (1995) Gender differences in drug metabolism regulated by growth hormone. Int J Biochem Cell Biol. 27, 9–20.

Sheard, M. A., Vojtesek, B., Simickova, M. and Valik, D. (2002) Release of cytokeratin-18 and -19 fragments (TPS and CYFRA 21-1) into the extracellular space during apoptosis. J Cell Biochem. 85, 670-7.

Shimada, T., Hayes, C. L., Yamazaki, H., Amin, S., Hecht, S. S., Guengerich, F. P. and Sutter, T. R. (1996) Activation of chemically diverse procarcinogens by human cytochrome P-450 1B1. Cancer Res. 56, 2979-2984.

Shimada, T., Watanabe, J., Kawajiri, K., Sutter, T. R., Guengerich, F. P., Gillam, E. M. J. and Inoue, K. (1999) Catalytic properties of polymorphic human cytochrome P450 1B1 variants. Carcinogenesis. 20, 1607-1613.

Shimada, T., Yamazaki, H., Mimura, M., Inui, Y. and Guengerich, F. P. (1994) Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. J Pharmacol Exp Ther. 270, 414-423.

Shou, M., Lu, T., Krausz, K. W., Sai, Y., Yang, T., Korzekwa, K. R., Gonzalez, F. J. and Gelboin, H. V. (2000) Use of inhibitory monoclonal antibodies to assess the contribution of cytochrome P450 to human drug metabolism. Eur J Pharmacol. 394, 199-209.

Simpson, A. E. (1997) The cytochrome P450 4 (CYP4) family. Gen Pharmacol. 28, 351-9.

Simpson, R. J. and Dorow, D. S. (2001) Cancer proteomics: from signaling networks to tumor markers. Trends Biotechnol. 19(10 Suppl), S40-8.

Smith, P. J., Blunt, N. J., Desnoyers, R., Giles, Y. and Patterson, L. H. (1997) DNA topoisomerase II-dependent cytotoxicity of alkylaminoanthraquinones and their N-oxides. Cancer Chemother Pharmacol. 39, 455-61.

Smith, G., Harrison, D. J., East, N., Rae, F., Wolf, H. and Wolf, C. R. (1993) Regulation of cytochrome P450 gene expression in human colon and breast tumour xenografts. Br J Cancer. 68, 57-63.

Sotaniemi, E. A., Rautio, A., Bäckström, M., Arvela, P. and Pelkonen, O. (1995) CYP3A4 and CYP2A6 activities marked by the metabolism of lignocaine and coumarin in patients with liver and kidney diseases and epileptic patients. Br J Clin Pharmacol. 39, 71-76.

Sotaniemi, E. A., Arranto, A. J., Pelkonen, O. and Pasanen, M. (1997) Age and cytochrome P450-linked drug metabolism in humans: an analysis of 226 subjects with equal histopathologic conditions. Clin Pharmacol Ther. 61, 331-9.

Spink, D. C., Spink, B. C., Cao, J. Q., DePasquale, J. A., Pentecost, B. T., Fasco, M. J., Li, Y. and Sutter, T. R. (1998) Differential expression of CYP1A1 and CYP1B1 in human breast epithelial cells and breast tumor cells. Carcinogenesis. 19, 291-8.

Srivastava, P. K., Yun, C. H., Beaune, P. H., Ged, C. and Guengerich, F. P. (1991) Separation of human liver microsomal tolbutamide hydroxylase and (S)-mephenytoin 4'-hydroxylase cytochrome P450 enzymes. Mol Pharmacol. 40, 69-79.

Stevens, T. J. and Arkin, I. T. (2000) Turning an opinion inside-out: Rees and Eisenberg's commentary on "Are membrane proteins 'inside-out' proteins?". Proteins. 40, 463-4.

Stiborová, M., Martinek, V., Rydlova, H., Koblas, T. and Hodek P. (2005) Expression of cytochrome P450 1A1 and its contribution to oxidation of a potential human carcinogen 1-phenylazo-2-naphthol (Sudan 1) in human livers. Cancer Lett. 220, 145-54.

Stoilov, I., Akarsu, A. N. and Sarfarazi, M. (1997) Identification of three different truncating mutations in cytochrome P4501B1 (CYP1B1) as the principal cause of primary congenital glaucoma (Buphthalmos) in families linked to the GLC3A locus on chromosome 2p21. Hum Mol Genet. 6, 641-647.

Sueyoshi, T., Kawamoto, T., Zelko, I., Honkakoski, P. and Negishi, M. (1999) The repressed nuclear receptor CAR responds to phenobarbital in activating the human CYP2B6 gene. J Biol Chem 274, 6043-6046.

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

Tanaka, E. (1983) Quantitative image reconstruction with weighted backprojection for single photon emission computed tomography. J Comput Assist Tomogr. 7, 692-700.

Tang, Y. M., Chen, G-F., Thompson, P. A., Lin, D-X., Lang, N. P. and Kadlubar, F. F. (1999) Development of an antipeptide antibody that binds to the C-terminal region of human CYP1B1. Drug Metab Dispos. 27, 274-280.

Thompson, P. D., Jurutka, P. W., Whitfield, G. K., Myskowski, S. M., Eichhorst, K. R., Dominguez, C. E., Haussler, C. A. and Haussler, M. R. (2002) Liganded VDR induces CYP3A4 in small intestinal and colon cancer cells via DR3 and ER6 vitamin D responsive elements. Biochem Biophys Res Commun. 299, 730-8.

Tomer, K. B. (2000) Separations combined with mass spectrometry. Chem Rev. 101, 297-328.

Waxman, D. J. (1992) Regulation of liver specific steroid metabolising cytochrome P450: Cholesterol 7alpha hydroxylases, bile acid 6beta-hydroxylase, and growth responsive steroid hormone hydroxylases. J Steroid Biochem Mol Biol. 43, 1055-1072.

Waxman, D. J., Ram, P. A., Pampori, N. A. and Shapiro, B. H. (1995) Growth hormone regulation of male-specific rat liver P450s 2A2 and 3A2: induction by intermittent growth hormone pulses in male but not female rats rendered growth hormone deficient by neonatal monosodium glutamate. Mol Pharmacol. 48, 790-7.

Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F. and McRee, D. E. (2000) Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. J Inorg Biochem. 81, 183-90.

Williams, P. A., Cosme, J., Ward, A., Angove, H. C., Vinkovic, D. M. and Jhoti, H. (2003) Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature. 424, 464-8.

Williams, P. A., Cosme, J., Vinkovic, D. M., Ward, A., Angove, H. C., Day, P. J., Vonrhein, C., Tickle, I. J. and Jhoti, H. (2004) Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. Science. 305, 683-6.

Willey, J. C., Coy, E. .L, Frampton, M. W., Torres, A., Apostolakos, M. J., Hoehn, G., Schuermann, W. H., Thilly, W. G., Olson, D. E., Hammersley, J. R., Crespi, C. L. and Utell, M. J. (1997) Quantitative RT-PCR measurement of cytochromes p450 1A1, 1B1, and 2B7, microsomal epoxide hydrolase, and NADPH oxidoreductase expression in lung cells of smokers and nonsmokers. Am J Respir Cell Mol Biol. 17, 114-124.

Willoughby, R., Sheehan, E. and Mitrovich, S. (1998) A Global view of LC/MS (First ed.), pp. 531, Global View Publishing, Pittsburgh, Pennsylvania.

Wrighton, S. A., Ring, B. J., Watkins, P. B. and VandenBranden, M. (1989) Identification of a polymorphically expressed member of the human cytochrome P-450III family. Mol Pharmacol. 86, 97-105.

Wrighton, S. A., Stevens, J. C., Becker, G. W. and VandenBranden, M. (1993) Isolation and characterization of human liver cytochrome P450 2C19: correlation between 2C19 and S-mephenytoin 4'-hydroxylation. Arch Biochem Biophys. 306, 240-5.

Yamano, S., Tatsuno, J. and Gonzalez, F. J. (1990) The CYP2A3 gene product catalyzes coumarin 7-hydroxylation in human liver microsomes. Biochemistry. 29, 1322-1329.

Yasuhiro Masubuchi, Takashi Iwasa, Shin Hosokawa, Tokuji Suzuki, Toshiharu Horie, Susumu Imaoka, Yoshihiko Funae and Shizuo Narimatsu (1997) Selective Deficiency of Debrisoquine 4-Hydroxylase Activity in Mouse Liver Microsomes. J Pharmacol Exp Ther. 282, 1435-1441.

Yates lll, J. R., Eng, J. K., McCormack, A. L. and Schieltz, D. (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal Chem. 67, 1426-1436.

Yates, J. R., Speicher, S., Griffin, P. R. and Hunkapiller, T. (1993) Peptide mass maps: a highly informative approach to protein identification. Anal. Biochem. 214, 397–408

Yokotani, N., Sogawa, K., Matsubara, S., Gotoh, O., Kusunose, E., Kusunose, M. and Fujii-Kuriyama, Y. (1990) cDNA cloning of cytochrome P-450 related to P-450p-2 from the cDNA library of human placenta: gene structure and expression. Eur J Biochem. 187, 23-29.

Young, P. M. and Wheat, T. E. (1990) Optimization of high performance liquid chromatographic peptide separations with alternative mobile and stationary phases. J of Chromatography. 512, 273-281.

Yun, C-H., Shimada, T. and Guengerich, F. P. (1991) Purification and characterization of human liver microsomal cytochrome P-450 2A6. Mol Pharmacol. 40, 679-685.

Zevin, S. and Benowitz, N. L. (1999) Drug interactions with tobacco smoking. An update. Clin Pharmacokinet. 36, 425-438.

Zhang, Q. Y., Dunbar, D., Ostrowska, A., Zeisloft, S., Yang, J. and Kaminsky, L. S. (1999) Characterization of human small intestinal cytochromes P-450. Drug Metab Dispos. 27, 804-9.

Zhang, Z. and McElvain, J. S. (2000) De novo peptide sequencing by two-dimensional fragment correlation mass spectrometry. Anal Chem. 72, 2337-50.

Appendices

Appendix i

Amino acid sequence summary for CYPs identified from the rat liver

>**gi|320081**|pir||A34272 testosterone 7alpha-hydroxylase (EC 1.14.14.-) cytochrome P450 2A1 - rat□gi|203744|gb|AAA41020.1| (M33312) hepatic steroid hydroxylase IIA1 (CYP2A1) [Rattus norvegicus] [MASS=55995]

MLDTGLLLVV ILASLSVMLL VSLWQQKIRG RLPPGPTPLP FIGNYLQLNT KDVYSSITQL SERYGPVFTI HLGPRRVVVLYGYDAVKEAL VDQAEEFSGR GEQATYNTLF KGYGVAFSSG ERAKQLRRLS IATLRDFGVG KRGVEERILE EAGYLIKMLQGTCGAPIDPT IYLSKTVSNV ISSIVFGERF DYEDTEFLSL LQMMGQMNRF AASPTGQLYD MFHSVMKYLP GPQQQIIKVTQKLEDFMIEK VRQNHSTLDP NSPRNFIDSF LIRMQEEKNG NSEFHMKNLV MTTLSLFFAG SETVSSTLRY GFLLLMKHPDVEAKVHEEIE QVIGRNRQPQ YEDHMKMPYT QAVINEIQRF SNLAPLGIPR RIIKNTTFRG FFLPKGTDVF PILGSLMTDPKFFPSPKDFD PQNFLDDKGQ LKKNAAFLPF STGKRFCLGD GLAKMELFLL LTTILQNFRF KFPMKLEDIN ESPKPLGFTRIIPKYTMSFM PI

Mass (average): 55995.1 **Identifier:** gi|320081 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 139/492 = **28.3%** by amino acid count, 15551.9/55995.1 = **27.8%** by mass

DVYSSITQLSER	52 - 63
EALVDQAEEFSGR	88 - 100
FSNLAPLGIPR	360 - 370
ILEEAGYLIK	148 - 157
LEDINESPKPLGFTR	466 - 480
LPPGPTPLPFIGNYLQLNTK	32 - 51
NFIDSFLIR	265 - 273
RVVVLYGYDAVK	76 - 87
TVSNVISSIVFGER	176 - 189
YGPVFTIHLGPR	64 - 75
YLPGPQQQIIK	228 - 238

>**gi|6978741**|ref|NP_036825.1|| cytochrome P450 IIA2□gi|117193|sp|P15149|CPA2_RAT CYTOCHROME P450 2A2 (CYPIIA2) (TESTOSTERONE 15-ALPHA-HYDROXYLASE) (P450-UT-4)□gi|92120|pir||A31887 testosterone 7alpha-monooxygenase (EC 1.14.14.-) cytochrome P450 2A2 - rat□gi|203749|gb|AAA41021.1| (M34392) hepatic steroid hydroxylase IIA2 (CYP2A2) [Rattus norvegicus]□gi|204902|gb|AAA41424.1| (J04187) IIA2 protein [Rattus norvegicus] [MASS=56345]
MLDTGLLLVV ILASLSVMFL VSLWQQKIRE RLPPGPTPLP FIGNYLQLNM KDVYSSITQL SERYGPVFTI HLGPRRIVVLYGYDAVKEAL VDQAEEFSGR GELPTFNILF KGYGFSLSNV EQAKRIRRFT IATLRDFGVG KRDVQECILE EAGYLIKTLQGTCGAPIDPS IYLSKTVSNV INSIVFGNRF DYEDKEFLSL LEMIDEMNIF AASATGQLYD MFHSVMKYLP GPQQQIIKVTQKLEDFMIEK VRQNHSTLDP NSPRNFIDSF LIRMQEEKYV NSEFHMNNLV MSSLGLLFAG TGSVSSTLYH GFLLLMKHPDVEAKVHEEIE RVIGRNRQPQ YEDHMKMPYT QAVINEIQRF SNLAPLGIPR RIIKNTTFRG FFLPKGTDVF PIIGSLMTEPKFFPNHKDFN PQHFLDDKGQ LKKNAAFLPF SIGKRFCLGD SLAKMELFLL LTTILQNFRF KFPMNLEDIN EYPSPIGFTRIIPNYTMSFM PI

Mass (average): 56345.5 **Identifier:** gi|6978741 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 93/492 = **18.9%** by amino acid count, 10440.0/56345.5 = **18.5%** by mass

DFGVGK	136 - 141
DVYSSITQLSER	52 - 63
GELPTFNILFK	101 - 111
GTDVFPIIGSLMTEPK	386 - 401
MPYTQAVINEIQR	347 - 359
NFIDSFLIR	265 - 273
TVSNVINSIVFGNR	176 - 189
YGPVFTIHLGPR	64 - 75

>**gi|117210**|sp|P13107|CPB3_RAT CYTOCHROME P450 2B3 (CYPIIB3)□gi|92125|pir||A29818 cytochrome P450 2B3, hepatic - rat□gi|203684|gb|AAA41006.1| (M20406) cytochrome P450IIB3 [Rattus norvegicus] [MASS=56384]
MDTSVLLLLA VLLSFLLFLV RGHAKVHGHL PPGPRPLPLL GNLLQMDRGG FRKSFIQLQE KHGDVFTVYF GPRPVVMLCGTQTIREALVD HAEAFSGRGI IAVLQPIMQE YGVSFVNEER WKILRRLFVA TMRDFGIGKQ SVEDQIKEEA KCLVEELKNHQGVSLDPTFL FQCVTGNIIC SIVFGERFDY RDRQFLRLLD LLYRTFSLIS SFSSQMFEVY

SDFLKYFPGV HREIYKNLKEVLDYIDHSVE NHRATLDPNA PRDFIDTFLL HMEKEKLNHY TEFHHWNLMI
SVLFLFLAGT ESTSNTLCYG FLLMLKYPHVAEKVQKEIDQ VIGSQRVPTL DDRSKMPYTE AVIHEIQRFS
DVSPMGLPCR ITKDTLFRGY LLPKNTEVYF ILSSALHDPQYFEQPDTFNP EHFLDANGAL KKCEAFMPFS
IGKRMCLGEG IARSELFLFF TTILQNYSVS SPVDPNTIDM TPKESGLAKVAPVYKICFVA R

Mass (average): 56384.3 **Identifier:** gi|117210 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 75/491 = **15.3%** by amino acid count, 8541.7/56384.3 = **15.1%** by mass

DFGIGK	134 - 139
EALVDHAEAFSGR	86 - 98
EIDQVIGSQR	327 - 336
EVLDYIDHSVENHR	240 - 253
MPYTEAVIHEIQR	346 - 358
PLPLLGNLLQMDR	36 - 48
VAPVYK	480 - 485

>**gi|117223**|sp|P05178|CPC6_RAT CYTOCHROME P450 2C6 (CYPIIC6) (P450 PB1) (PTF2)□gi|2117384|pir||A25954 cytochrome P450 2C6, hepatic - rat [MASS=56003]
MDLVMLLVLT LTCLILLSIW RQSSGRGKLP PGPIPLPIIG NIFQLNVKNI TQSLTSFSKV YGPVFTLYFG TKPTVILHGY
EAVKEALIDH GEEFAERGSF PVAEKINKDL GIVFSHGNRW KEIRRFTLTT LRNLGMGKRN IEDRVQEEAR
CLVEELRKTNGSPCDPTFIL GCAPCNVICS IIFQNRFDYK DQDFLNLMEK LNENMKILSS PWTQFCSFFP
VLIDYCPGSH TTLAKNVYHIRNYLLKKIKE HQESLDVTNP RDFIDYYLIK WKQENHNPHS EFTLENLSIT
VTDLFGAGTE TTSTTLRYAL LLLLKCPEVTAKVQEEIDRV VGKHRSPCMQ DRSRMPYTDA HDHEVQRFID
LIPTNLPHAV TCDIKFRNYL IPKGTTIITS LSSVLHDSKEFPDPEIFDPG HFLDGNGKFK KSDYFMPFSA
GKRMCAGEGL ARMELFLFLT TILQNFKLKS VLHPKDIDTT PVFNGFASLPPFYELCFIPL

Mass (average): 56002.9 **Identifier:** gi|117223 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 48/490 = **9.8%** by amino acid count, 5695.3/56002.9 = **10.2%** by mass

EFPDPEIFDPGHFLDGNGK	400 - 418
EHQESLDVTNPR	250 - 261
FDYKDQDFLNLMEK	187 - 200
IKEHQESLDVTNPR	248 - 261
KIKEHQESLDVTNPR	247 - 261

>**gi|117224**|sp|P05179|CPC7_RAT CYTOCHROME P450 2C7 (CYPIIC7) (P450F) (PTF1)□gi|203780|gb|AAA41036.1| (M18335) cytochrome P450 [Rattus norvegicus] [MASS=56187]
MDLVTFLVLT LSSLILLSLW RQSSRRRKLP PGPTPLPIIG NFLQIDVKNI SQSLTKFSKT
YGPVFTLYLGSQPTVILHGYEAIKEALIDN GEKFSGRGSY PMNENVTKGF GIVFSNGNRW KEMRRFTIMN
FRNLGIGKRN IEDRVQEEAQ CLVEELRKTKGSPCDPSLIL NCAPCNVICS ITFQNHFDYK DKEMLTFMEK
VNENLKIMSS PWMQVCNSFP SLIDYFPGTH HKIAKNINYMKSYLLKKIEE HQESLDVTNP RDFVDYYLIK
QKQANNIEQS EYSHENLTCS IMDLIGAGTE TMSTTLRYAL LLLMKYPHVTAKVQEEIDRV IGRHRSPCMQ
DRKHMPYTDA MIHEVQRFIN FVPTNLPHAV TCDIKFRNYL IPKGTKVLTS LTSVLHDSKEFPNPEMFDPG
HFLDENGNFK KSDYFLPFSA GKRACVGEGL ARMQLFLFLT TILQNFNLKS LVHPKDIDTM PVLNGFASLP
PTYQLCFIPS

Mass (average): 56187.1 **Identifier:** gi|117224 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 62/490 = **12.7%** by amino acid count, 6998.0/56187.1 = **12.5%** by mass

HMPYTDAMIHEVQR	344 - 357
IEEHQESLDVTNPR	248 - 261
KIEEHQESLDVTNPR	247 - 261
LPPGPTPLPIIGNFLQIDVK	29 - 48
VLTSLTSVLHDSK	387 - 399

>**gi|117228**|sp|P08683|CPCB_RAT CYTOCHROME P450 2C11 (CYPIIC11) (P-450(M-1)) (P450H) (P450-UT-A) (UT-2)□gi|92127|pir||A26685 cytochrome P450 2C11 - rat□gi|203868|gb|AAA41062.1| (J02657) cytochrome P-450(M-1) [Rattus norvegicus] [MASS=57181]

MDPVLVLVLT LSSLLLLSLW RQSFGRGKLP PGPTPLPIIG NTLQIYMKDI GQSIKKFSKV YGPIFTLYLG
MKPFVVLHGYEAVKEALVDL GEEFSGRGSF PVSERVNKGL GVIFSNGMQW KEIRRFSIMT LRTFGMGKRT
IEDRIQEEAQ CLVEELRKSKGAPFDPTFIL GCAPCNVICS IIFQNRFDYK DPTFLNLMHR FNENFRLFSS
PWLQVCNTFP AIIDYFPGSH NQVLKNFFYIKNYVLEKVKE HQESLDKDNP RDFIDCFLNK MEQEKHNPQS
EFTLESLVATVTDMFGAGTETTSTTLRYGLLLLLKHVDVTAKVQEEIERVIGRNRSPCMKDRSQMPYTDAVVHEIQR
YID LVPTNLPHLV TRDIKFRNYF IPKGTNVIVS LSSILHDDKEFPNPEKFDPG HFLDERGNFK KSDYFMPFSA
GKRICAGEAL ARTELFLFFT TILQNFNLKS LVDVKDIDTTPAISGFGHLPPFYEACFIPV QRADSLSSHL

Mass (average): 57181.1 **Identifier:** gi|117228 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 229/500 = **45.8%** by amino acid count, 26254.0/57181.1 = **45.9%** by mass

ADSLSSHL	493 - 500	HVDVTAK	316 - 322
DIGQSIK	49 - 55	LPPGPTPLPIIGNTLQIYMK	29 - 48
DPTFLNLMHR	191 - 200	NFFYIK	236 - 241
EALVDLGEEFSGR	85 - 97	NYFIPK	378 - 383
EHQESLDKDNPR	250 - 261	NYVLEK	242 - 247
FDPGHFLDER	407 - 416	SDYFMPFSAGK	422 - 432
FDYKDPTFLNLMHR	187 - 200	SLVDVK	460 - 465
FNENFR	201 - 206	SQMPYTDAVVHEIQR	343 - 357
GLGVIFSNGMQWK	109 - 121	TFGMGK	133 - 138
GSFPVSER	98 - 105	VKEHQESLDKDNPR	248 - 261
GTNVIVSLSSILHDDKEFPNPEK	384 - 406	VQEEIER	323 - 329
YIDLVPTNLPHLVTR	358 - 372	YGLLLLLK	308 - 315

>**gi|6978745**|ref|NP_036841.1|| cytochrom P450 15-beta gene□gi|205932|gb|AAA41784.1| (M33550) cytochrome P450 15-beta (CYP2C12) [Rattus norvegicus]□gi|226780|prf||1605163A cytochrome P450 15beta [Rattus norvegicus] [MASS=55894]

MDPFVVLVLS LSFLLLLYLW RPSPGRGKLP PGPTPLPIFG NFLQIDMKDI RQSISNFSKT YGPVFTLYFG
SQPTVVLHGYEAVKEALIDY GEEFSGRGRM PVFEKATKGL GISFSRGNVW RATRHFTVNT LRSLGMGKRT
IEIKVQEEAE WLVMELKKTKGSPCDPKFII GCAPCNVICS IIFQNRFDYK DKDFLSLIEN VNEYIKIVST PAFQVFNAFP
ILLDYCPGNH KTHSKHFAAIKSYLLKKIKE HEESLDVSNP RDFIDYFLIQ RCQENGNQQM NYTQEHLAIL
VTNLFIGGTE TSSLTLRFAL LLLMKYPHITDKVQEEIGQV IGRHRSPCML DRIHMPYTNA MIHEVQRYID
LAPNGLLHEV TCDTKFRDYF IPKGTAVLTS LTSVLHDSKEFPNPEMFDPG HFLDENGNFK KSDYFMPFSA
GKRKCVGEGL ASMELFLFLT TILQNFKLKS LSDPKDIDIN SIRSEFSSIPPTFQLCFIPV

Mass (average): 55893.8 **Identifier:** gi|6978745 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 69/490 = **14.1%** by amino acid count, 7964.8/55893.8 = **14.2%** by mass

DFIDYFLIQR	262 - 271
DKDFLSLIENVNEYIK	191 - 206
EALIDYGEEFSGR	85 - 97
GTAVLTSLTSVLHDSK	384 - 399
IKEHEESLDVSNPR	248 - 261

>**gi|117230**|sp|P20814|CPCD_RAT CYTOCHROME P450 2C13, MALE-SPECIFIC (CYPIIC13) (P450-G) (UT-5)□gi|205935|gb|AAA41785.1| (J02861) cytochrome P-450g [Rattus norvegicus] [MASS=55860]

MDPVVVLLLS LFFLLFLSLW RPSSGRGKLP PGPTPLPIIG NFFQVDMKDI RQSLTNFSKT YGPVYTLYVG
SQPTVVLHGYEALKEALVDH GEEFSGRGRL PICEKVAKGQ GIAFSHGNVW KATRHFTVKT LRNLGMGKGT
IEDKVQEEAK WLVKELKKTNGSPCDPQFIM GCAPGNVICS IILQNRFDYE DKDFLNLIEK VNEAVKIISS
PGIQVFNIFP ILLDYCPGNH NIYFKNHTWLKSYLLEKIKE HEESLDVSNP RDFIDYFLIE RNQENANQWM
NYTLEHLAIM VTDLFFAGIE TVSSTMRFAL LLLMKYPHVTAKVQEEIDHV IGRHRSPCMQ DRSHMPYTNA
MVHEVQRYID IGPNGLLHEV TCDTKFRNYF IPKGTAVLTS LTSVLHDSKEFPNPEMFDPG HFLDENGNFK
KSDYFIPFSA GKRMCLGESL ARMELFLFLT TILQNFKLKS LVDPKDINTT PICSSLSSVPPTFQMRFIPL

Mass (average): 55859.6 **Identifier:** gi|117230 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 49/490 = **10.0%** by amino acid count, 5590.2/55859.6 = **10.0%** by mass

EALVDHGEEFSGR	85 - 97
FDYEDKDFLNLIEK	187 - 200
GTAVLTSLTSVLHDSK	384 - 399
NYFIPK	378 - 383

>**gi|6978731**|ref|NP_036885.1|| cytochrome P450, subfamily XVII_gi|117286|sp|P11715|CPT7_RAT CYTOCHROME P450 17 (CYPXVII) (P450-C17) (STEROID 17-ALPHA-HYDROXYLASE/17,20 LYASE)_gi|92770|pir||A30828 steroid 17alpha-monooxygenase (EC 1.14.99.9) cytochrome P450 17 - rat_gi|56052|emb|CAA32248.1| (X14086) cytochrome P-450-(17-alpha) (AA 1 - 507) [Rattus norvegicus]_gi|205910|gb|AAA41777.1| (M31681) 17-alpha hydroxylase [Rattus norvegicus]_gi|205922|gb|AAA41783.1| (M22204) P-450 17-alpha protein [Rattus norvegicus]_gi|940818|emb|CAA49470.1| (X69816) 17-alpha hydroxylase [Rattus norvegicus] [MASS=57250]

MWELVGLLLLILAYFFWVKSKTPGAKLPRSLPSLPLVGSLPFLPR**RGHMHVNFFKLQEKYGPIYSLR**LGTTTTVIIG
HYQLAREVLIKKGKEFSGRPQMVTQSLLSDQGKGVAFADAGSSWHLHRKLVFSTFSLFKDGQKLEKLICQEAKSLCD
MMLAHDKESIDLSTPIFMSVTNIICAICFNISYEK**NDPKLTAIKTFTEGIVDATGDR**NLVDIFPWLTIFPNKGLEVI
KGYAKVRNEVLTGIFEKCREKFDSQSISSLTDILIQAKMNSDNNNSCEGRDPDVFSDRHILATVGDIFGAGIETTTT
VLKWILAFLVHNPEVKKKIQKEIDQYVGFSRTPTFNDRSHLLMLEATIREVLRIRPVAPMLIPHKANVDSSIGEFTV
PKDTHVVVNLWALHHDENEWDQPDQFMPERFLDPTGSHLITPTQSYLPFGAGPRSCIGEALARQELFVFTALLLQRF
DLDVSDDKQLPRLEGDPKVVFLIDPFKVKITVRQAWMDAQAEVST

Mass (average): 57250 **Identifier:** gi|117286 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 44/507 = **8.6%** by amino acid count

RGHMHVNFFKLQEKYGPIYSLR	46-67
NDPKLTAIKTFTEGIVDATGDR	189-211

>**gi|117295**|sp|P22443|CPV1_RAT CYTOCHROME P450 19 (AROMATASE) (CYPXIX) (ESTROGEN SYNTHETASE) (P-450AROM)_gi|92112|pir||A36121 aromatase (EC 1.14.14.-) cytochrome P450 19 - rat_gi|203805|gb|AAA41044.1| (M33986) aromatase cytochrome P450 [Rattus norvegicus] [MASS=58412]

MFLEMLNPMHYNVTIMVPETVPVSAMPLLLIMGLLLLIRNCESSSSIPGPGYCLGIGPLISHGRFLWMGIGSACNYYNKMYGEFMR
VWISGEETLIISKSSSMVHVMKHSNYISRFGSKRGLQCIGMHENGIIFNNNPSLWRTVR**PFFMKALTGPGLIR**MVEVCVESIKQHL
DRLGDVTDNSGYVDVVTLMRHIMLDTSNTLFLGIPLDESSIVKKIQGYFNAWQALLIKPNIFFKISWLYRKYERSVKDLKDEIEIL
VEKKRQKVSSAEKLEDCMDFATDLIFAERRGDLTKENVNQCILEMLIAAPDTMSVTLYVMLLLIAEYPEVETAILKEIHTVVGDRD
IRIGDVQNLKVVENFINESLRYQPVVDLVMRRALEDDVIDGYPVKKGTNIILNIGRMHRLEYFPKPNEFTLENFEKNVPYRYFQPF
GFGPRSCAGKYIAMVMMKVVLVTLLKRFHVKTLQKRCIENMPKNNDLSLHLDEDSPIVEIIFRHIFNTPFLQCLYISL

Mass (average): 57250 **Identifier:** gi|117295 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 44/507 = **5.3%** by amino acid count

VWISGEETLIISK	87-99
PFFMKALTGPGLIR	146-159

>**gi|117236**|sp|P19225|CPCM_RAT CYTOCHROME P450 2C22 (CYPIIC22) (P450 MD) (P450 P49)□gi|2117380|pir||A39257 cytochrome P450 2C22 - rat□gi|56827|emb|CAA37570.1| (X53477) cytochrome P-450Md (AA 1-489) [Rattus norvegicus] [MASS=56157]

MALFIFLGIW LSCLVFLFLW NQHHVRRK**LP PGPTPLPIFG NILQVGVK**NI SKSMCMLAKE YGPVFTMYLG
MKPTVVLYGYEVLKEALIDR GEEFSDKMHS SMLSKVSQGL GIVFSNGEIW KQTRRFSLMV LRSMGMGKRT
IENR**IQEEVV YLLEALR**KTNGSPCDPSFLL ACVPCNVISS VIFQHRFDYS DEKFQKFIEN FHTKIEILAS
PWAQLCSAYP VLYYLPGIHN KFLKDVTEQKKFILMEINRH RASLNLSNPQ DFIDYFLIKM EKEKHNEKSE
FTMDNLIVTI GDLFGAGTET TSSTIKYGLL LLLKYPEVTAKIQEEITRVI GRHRRPCMQD RNHMPYTDAV
LHEIQR**YIDF VPIPLPR**KTT QDVEFRGYHI PKGTSVMACL TSALHDDKEFPNPEKFDPGH FLDEKGNFKK
SDYFMAFSAG RRACIGEGLA RMEMFLILTS ILQHFTLKPL VNPEDIDTTP VQPGLLSVPPPFELCFIPV

Mass (average): 56156.6 **Identifier:** gi|117236 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 44/489 = **9.0%** by amino acid count, 4925.9/56156.6 = **8.8%** by mass

IQEEVVYLLEALR	145 - 157
LPPGPTPLPIFGNILQVGVK	29 - 48
YIDFVPIPLPR	357 - 367

>**gi|6166042**|sp|P24470|CPCN_RAT CYTOCHROME P450 2C23 (CYPIIC23) (ARACHIDONIC ACID EPOXYGENASE)□gi|2117398|pir||A46588 cytochrome P450 arachidonic acid epoxygenase - rat□gi|438419|gb|AAA03716.1| (U04733) cytochrome P450 arachidonic acid epoxygenase [Rattus norvegicus] [MASS=56433]

MELLGFTTLA LVVSVTCLSL LSVWTKLRTR GRLPPGPTPL PIIGNLLQLN LKDIPASLSK LAKEYGPVYT
LYFGTSPTVVLHGYDVVKEA LLQQGDEFLG RGPLPIIEDT HKGYGLIFSN GERWKVMRRF SLMTLRNFGM
GKRSLEERVQ EEARCLVEELQKTKAQPFDP TFILACAPCN VICSILFNDR FQYNDKTFLN LMDLLNKNFQ
QVNSVWCQMY NLWPTIIKYL PGKHIEFAKRIDDVKNFILE KVKEHQKSLD PANPRDYIDC FLSKIEEEKD
NLKSEFHLEN LAVCGSNLFT AGTETTSTTL RFGLLLLMKYPEVQAKVHEE LDRVIGRHQP
PSMKDKMKLPYTDAVLHEIQ RYITLLPSSL PHAVVQDTKF RDYVIPKGTT VLPMLSSVMLDQKEFANPEK
FDPGHFLDKN GCFKKTDYFV PFSLGKRACV GESLARMELF LFFTTLLQKF SLKTLVEPKD LDIKPITTGI
INLPPPYKLC LVPR

Mass (average): 56433.0 **Identifier:** gi|6166042 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 68/494 = **13.8%** by amino acid count, 7467.0/56433.0 = **13.2%** by mass

>**gi|117242**|sp|P10634|CPD2_RAT CYTOCHROME P450 2D2 (CYPIID2) (P450-DB2) (P450-CMF2) (DEBRISOQUINE 4-HYDROXYLASE)□gi|57812|emb|CAA36269.1| (X52027) cytochrome P450 IID protein [Rattus norvegicus]□gi|2575859|dbj|BAA23123.1| (AB008423) CYP2D2 [Rattus norvegicus] [MASS=56684]

MGLLIGDDLW AVVIFTAIFL LLVDLVHRHK FWTAHYPPGP VPLPGLGNLL QVDFENMPYS LYKLRSRYGD
VFSLQIAWKPVVVINGLKAV RELLVTYGED TADRPLLPIY NHLGYGNKSK GVVLAPYGPE WREQRRFSVS
TLRDFGVGKK SLEQWVTEEAGHLCDTFAKE AEHPFNPSIL LSKAVSNVIA SLVYARRFEY EDPFFNRMLK
TLKESFGEDT GFMAEVLNAI PILLQIPGLPGKVFPKLNSF IALVDKMLIE HKKSWDPAQP PRDMTDAFLA
EMQKAKGNPE SSFNDENLRL VVIDLFMAGM VTTSTTLSWALLLMILHPDV QRRVHEEIDE VIGQVRRPEM
ADQARMPFTN AVIHEVQRFA DIVPTNIPHM TSRDIKFQGF LIPKGTTLIPNLSSVLKDET VWEKPLRFHP
EHFLDAQGNF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQRFSFS VLAGRPRPSTHGVYALPVTP
QPYQLCAVAR

Mass (average): 56683.9 **Identifier:** gi|117242 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 170/500 = **34.0%** by amino acid count, 19527.0/56683.9 = **34.4%** by mass

AVSNVIASLVYAR	184 - 196
DMTDAFLAEMQK	273 - 284
EAEHPFNPSILLSK	170 - 183
ELLVTYGEDTADRPLLPIYNHLGYGNK	92 - 118
FADIVPTNIPHMTSR	369 - 383
FHPEHFLDAQGNFVK	418 - 432
GNPESSFNDENLR	287 - 299
GTTLIPNLSSVLKDETVWEKPLR	395 - 417
MPFTNAVIHEVQR	356 - 368
RFEYEDPFFNR	197 - 207
RVHEEIDEVIGQVR	333 - 346
VHEEIDEVIGQVR	334 - 346

>**gi|3915646**|sp|P12938|CPD3_RAT CYTOCHROME P450 2D3 (CYPIID3) (P450-DB3) (DEBRISOQUINE 4-HYDROXYLASE)□gi|2575861|dbj|BAA23124.1| (AB008424) CYP2D3 [Rattus norvegicus] [MASS=56642]

MELLAGTGLW PMAIFTVIFI LLVDLMHRRQ RWTSRYPPGP VPWPVLGNLL QVDLCNMPYS MYKLQNRYGD
VFSLQMGWKPVVVINGLKAV QELLVTCGED TADRPEMPIF QHIGYGHKAK GVVLAPYGPE WREQRRFSVS
TLRNFGVGKK SLEQWVTDEASHLCDALTAE AGRPLDPYTL LNKAVCNVIA SLIYARRFDY GDPDFIKVLK
ILKESMGEQT GLFPEVLNMF PVLLRIPGLADKVFPGQKTF LTMVDNLVTE HKKTWDPDQP PRDLTDAFLA
EIEKAKGNPE SSFNDANLRL VVNDLFGAGM VTTSITLTWALLLMILHPDV QCRVQQEIDE VIGQVRHPEM
ADQAHMPFTN AVIHEVQRFA DIVPMNLPHK TSRDIEVQGF LIPKGTTLIPNLSSVLKDET VWEKPLRFHP
EHFLDAQGNF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQRFSFS VPTGQPRPSDYGVFAFLLSP SP
YQLCAFKR

Mass (average): 56641.8 **Identifier:** gi|3915646 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 103/500 = **20.6%** by amino acid count, 11690.4/56641.8 = **20.6%** by mass

DETVWEKPLR	408 - 417
DIEVQGFLIPK	384 - 394
DLTDAFLAEIEK	273 - 284
FHPEHFLDAQGNFVK	418 - 432
FSVSTLR	137 - 143
GTTLIPNLSSVLK	395 - 407
IPGLADK	236 - 242
TFLTMVDNLVTEHKK	249 - 263
VQQEIDEVIGQVR	334 - 346

>**gi|117244**|sp|P13108|CPD4_RAT CYTOCHROME P450 2D4 (CYPIID4) (P450-DB4) (P450-CMF3) (DEBRISOQUINE 4-HYDROXYLASE)□gi|2117364|pir||D31579 cytochrome P450 2D4 - rat□gi|57816|emb|CAA36271.1| (X52029) cytochrome P450 IID3 protein [Rattus norvegicus] [MASS=56698]

MRMPTGSELW PIAIFTIIFL LLVDLMHRRQ RWTSRYPPGP VPWPVLGNLL QIDFQNMPAG FQKLRCRFGD LFSLQLAFESVVVLNGLPAL REALVKYSED TADRPPLHFN DQSGFGPRSQ GVVLARYGPA WRQQRRFSVS TFRHFGLGKK SLEQWVTEEARCLCAAFADH SGFPFSPNTL LDKAVCNVIA SLLFACRFEY NDPRFIRLLD LLKDTLEEES GFLPMLLNVF PMLLHIPGLLGKVFSGKKAF VAMLDELLTE HKVTWDPAQP PRDLTDAFLA EVEKAKGNPE SSFNDENLRV VVADLFMAGM VTTSTTLTWALLFMILHPDV QCRVQQEIDE VIGQVRRPEM ADQARMPFTN AVIHEVQRFA DILPLGVPHK TSRDIEVQGF LIPKGTTLITNLSSVLKDET VWEKPLRFHP EHFLDAQGNF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQRFSFS VPTGQPRPSDYGIFGALTTP RPYQLCASPR

Mass (average): 56697.8 **Identifier:** gi|117244 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 121/500 = **24.2%** by amino acid count, 13998.8/56697.8 = **24.7%** by mass

DETVWEKPLR	408 - 417
DIEVQGFLIPK	384 - 394
DLTDAFLAEVEK	273 - 284
FHPEHFLDAQGNFVK	418 - 432
GNPESSFNDENLR	287 - 299
MPFTNAVIHEVQR	356 - 368
VQQEIDEVIGQVR	334 - 346
WTSRYPPGPVPWPVLGNLLQIDFQNMPAGFQKLR	32 - 65

>**gi|117245**|sp|P12939|CPD5_RAT CYTOCHROME P450 2D5 (CYPIID5) (P450-DB5) (P450-CMF1B) (DEBRISOQUINE 4-HYDROXYLASE)□gi|65662|pir||O4RTD5 cytochrome P450 2D5 - rat□gi|57818|emb|CAA36272.1| (X52030) cytochrome P450 IID5 protein [Rattus norvegicus]□gi|203674|gb|AAA41003.1| (J02869) cytochrome P-450 IID5 [Rattus norvegicus]□gi|203776|gb|AAA41034.1| (M25143) cytochrome P450 [Rattus norvegicus] [MASS=57076]

MELLNGTGLW PMAIFTVIFI LLVDLMHRHQ RWTSRYPPGP VPWPVLGNLL QVDPSNMPYS MYKLQHRYGD VFSLQMGWKPMVIVNRLKAV QEVLVTHGED TADRPPVPIF KCLGVKPRSQ GVVFASYGPE WREQRRFSVSTLRTFGMGKK SLEEWVTKEAGHLCDAFTAQ NGRSINPKAM LNKALCNVIA SLIFARRFEY EDPYLIRMLT LVEESLIEVS GFIPEVLNTF PALLRIPGLADKVFQGQKTF MAFLDNLLAE NRTTWDPAQP PRNLTDAFLA EVEKAKGNPE SSFNDENLRM VVVDLFTAGM VTTATTLTWALLLMILYPDV QRRVQQEIDE VIGQVRCPEM TDQAHMPYTN AVIHEVQRFG DIAPLNLPRI TSCDIEVQDF VIPKGTTLIINLSSVLKDET VWEKPLRFHP EHFLDAQGNF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQHFSFS VPAGQPRPST LGNFAISVAP LPYQLCAAVR EQGH

Mass (average): 57076.2 **Identifier:** gi|117245 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 98/504 = **19.4%** by amino acid count, 11123.5/57076.2 = **19.5%** by mass

DETVWEKPLR	408 - 417
FHPEHFLDAQGNFVK	418 - 432
FSVSTLR	137 - 143
GNPESSFNDENLR	287 - 299

IPGLADK	236 - 242
ITSCDIEVQDFVIPK	380 - 394
NLTDAFLAEVEK	273 - 284
TFGMGK	144 - 149
VQQEIDEVIGQVR	334 - 346

>**gi|461810**|sp|P33274|CPF1_RAT CYTOCHROME P450 4F1 (CYPIVF1) (P450-A3)□gi|423662|pir||S29723 cytochrome P450 4F1 - rat□gi|203791|gb|AAA41040.1| (M94548) cytochrome P450 4F1 [Rattus norvegicus]□gi|6643935|gb|AAF20822.1|AF200361_1 (AF200361) cytochrome P450 4F1 [Rattus norvegicus] [MASS=59868]

MSQLSLSWLG LGPEVAFPWQ TLLLFGASWI LAQILTQIYA AYRNFRRLRG FPQPPKRNWL MGHVGMVTPT
EQGLKELTRL VGTYPQGFLM WIGPMVPVIT LCHSDIVRSI LNASAAVALK DVIFYTILKP WLGDGLLVSA
GDKWSRHRRM LTPAFHFNILKPYVKIFNDS TNIMHAKWKR LISEGSSRLD MFEHVSLMTL DSLQKCVFSF
DSNCQEKSSE YIAAILELSA LVAKRHQQPLLFMDLLYNLT PDGMRFHKAC NLVHEFTDAV IRERRRTLPD
QGLDEFLKSK AKSKTLDFID VLLLTKDEDG KELSDEDIRAEADTFMFEGH DTTASGLSWI LYNLANDPEY
QERCRQEVQE LLRDRDPEEI EWDDLAQLPF LTMCIKESLR LHPPVTVISRCCTQDILLPD GRTIPKGIIC LISIFGIHHN
PSVWPDPEVY NPFRFDPENI KDSSPLAFIP FSAGPRNCIG QTFAMSEMKVALALTLLRFR LLPDDKEPRR
QPELILRAEG GLWLRVEPLT AGAQ

Mass (average): 59868.4 **Identifier:** gi|461810 **Database:** C:/Xcalibur/database//rat2.fasta
Protein Coverage: 12/524 = **2.3%** by amino acid count, 1390.7/59868.4 = **2.3%** by mass

TLDFIDVLLLTK	295 - 306

>**gi|6691477**|dbj|BAA89312.1| (AB031863) cytochrome P450 2D27 [Mesocricetus auratus] [MASS=56489]

MALLIGDGLW SGVIFTALFL LLVDLMHRRK FWRARYPPGP MPLPGLGNLL QVDFEHMPYS LYKFRQRYGD
VFSLQMAWKPVVVINGLKAV REVLVNCGED TADRPPVPIF NHVGFGHNSQ GVAFARYGPQ WREQRRFCVS
TMRDFGVGKK SLEQWVTEEAGHLCDAFTQE AGHPFNPTTL LNKSVCNVIS SLIYAHRFDY EDPFFNSLLK
MLQESFGEDT GFIAEVLNAV PVLLRIPGLPGKAFPKLTAF MDSLYKMLIE HKTTWDPAQP PRGLTDAFLA
EVEKAKGRPE SSFNDENLRM VVADMFIAGM VTTSTTLSWALLLMILHPDV QSRVQQEIDD VIGQVRRPEM
ADQARMPYTN AVIHEVQRFG DIAPVNIPHM TSHDVEVQGF LIPKGTTLIPNLSSVLKDET VWEKPLHFHP
EHFLDAQGRF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQRFSFS VPAGQPRPSDQGIFALPVTP
TPYELCAVVR

Mass (average): 56489.4 **Identifier:** gi|6691477 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 26/500 = **5.2%** by amino acid count, 2882.4/56489.4 = **5.1%** by mass

GTTLIPNLSSVLK	395 - 407
MPYTNAVIHEVQR	356 - 368

>**gi|117241**|sp|P10633|CPD1_RAT CYTOCHROME P450 2D1 (CYPIID1) (P450-DB1) (P450-CMF1A) (P450-UT-7) (DEBRISOQUINE 4-HYDROXYLASE)□gi|92195|pir||A26822 debrisoquine 4-hydroxylase (EC 1.14.14.-) cytochrome P450 2D1 - rat□gi|203670|gb|AAA41001.1| (J02867) cytochrome P-450 IID1 [Rattus norvegicus]□gi|203834|gb|AAA41054.1| (M16654) debrisoquine 4-hydroxylase [Rattus norvegicus] [MASS=57175]

MELLNGTGLW SMAIFTVIFI LLVDLMHRRH RWTSRYPPGP VPWPVLGNLL QVDLSNMPYS LYKLQHRYGD
VFSLQKGWKPMVIVNRLKAV QEVLVTHGED TADRPPVPIF KCLGVKPRSQ GVILASYGPE WREQRRFSVS
TLRTFGMGKK SLEEWVTKEAGHLCDAFTAQ AGQSINPKAM LNKALCNVIA SLIFARRFEY EDPYLIRMVK
LVEESLTEVS GFIPEVLNTF PALLRIPGLADKVFQGQKTF MALLDNLLAE NRTTWDPAQP PRNLTDAFLA
EVEKAKGNPE SSFNDENLRM VVVDLFTAGM VTTATTLTWALLLMILYPDV QRRVQQEIDE VIGQVRCPEM
TDQAHMPYTN AVIHEVQRFG DIAPLNLPRF TSCDIEVQDF VIPKGTTLIINLSSVLKDET VWEKPHRFHP
EHFLDAQGNF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQRFSFS VPVGQPRPSTHGFFAFPVAP
LPYQLCAVVR EQGL

Mass (average): 57175.4 **Identifier:** gi|117241 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 64/504 = **12.7%** by amino acid count, 7242.0/57175.4 = **12.7%** by mass

AVQEVLVTHGEDTADRPPVPIFK	89 - 111
FHPEHFLDAQGNFVK	418 - 432
GNPESSFNDENLR	287 - 299
VQQEIDEVIGQVR	334 - 346

>**gi|117242**|sp|P10634|CPD2_RAT CYTOCHROME P450 2D2 (CYPIID2) (P450-DB2) (P450-CMF2) (DEBRISOQUINE 4-HYDROXYLASE)□gi|57812|emb|CAA36269.1| (X52027) cytochrome P450 IID protein [Rattus norvegicus]□gi|2575859|dbj|BAA23123.1| (AB008423) CYP2D2 [Rattus norvegicus] [MASS=56684]

MGLLIGDDLW AVVIFTAIFL LLVDLVHRHK FWTAHYPPGP VPLPGLGNLL QVDFENMPYS LYKLRSRYGD
VFSLQIAWKPVVVINGLKAV RELLVTYGED TADRPLLPIY NHLGYGNKSK GVVLAPYGPE WREQRRFSVS
TLRDFGVGKK SLEQWVTEEAGHLCDTFAKE AEHPFNPSIL LSKAVSNVIA SLVYARRFEY EDPFFNRMLK
TLKESFGEDT GFMAEVLNAI PILLQIPGLPGKVFPKLNSF IALVDKMLIE HKKSWDPAQP PRDMTDAFLA
EMQKAKGNPE SSFNDENLRL VVIDLFMAGM VTTSTTLSWALLLMILHPDV QRRVHEEIDE VIGQVRRPEM
ADQARMPFTN AVIHEVQRFA DIVPTNIPHM TSRDIKFQGF LIPKGTTLIPNLSSVLKDET VWEKPLRFHP
EHFLDAQGNF VKHEAFMPFS AGRRACLGEP LARMELFLFF TCLLQRFSFS VLAGRPRPSTHGVYALPVTP
QPYQLCAVAR

Mass (average): 56683.9 **Identifier:** gi|117242 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 103/500 = **20.6%** by amino acid count, 11815.2/56683.9 = **20.8%** by mass

DMTDAFLAEMQK	273 - 284
EAEHPFNPSILLSK	170 - 183
ELLVTYGEDTADRPLLPIYNHLGYGNK	92 - 118
FHPEHFLDAQGNFVK	418 - 432
GNPESSFNDENLR	287 - 299
MPFTNAVIHEVQR	356 - 368
RPEMADQAR	347 - 355

>**gi|1352193**|sp|P05182|CPE1_RAT CYTOCHROME P450 2E1 (CYPIIE1) (P450-J) (P450RLM6)□gi|203774|gb|AAA41033.1| (M20131) cytochrome P450IIE1 [Rattus norvegicus]□gi|3126851|gb|AAC15991.1| (AF061442) cytochrome P450 2E1 [Rattus norvegicus] [MASS=56627]
MAVLGITIAL LVWVATLLVI SIWKQIYNSW NLPPGPFPLP ILGNIFQLDL KDIPKSFTKL AKRFGPVFTL
HLGSRRIVVLHGYKAVKEVL LNHKNEFSGR GDIPVFQEYK NKGIIFNNGP TWKDVRRFSL SILRDWGMGK
QGNEARIQRE AQFLVEELKKTKGQPFDPTF LIGCAPCNVI ADILFNKRFD YNDKKCLRLM SLFNENFYLL
STPWIQLYNN FADYLRYLPG SHRKIMKNVSEIKQYTLEKA KEHLQSLDIN CARDVTDCLL IEMEKEKHSQ
EPMYTMENVS VTLADLFFAG TETTSTTLRY GLLILMKYPEIEEKLHEEID RVIGPSRVPA VRDRLDMPYM
DAVVHEIQRF INLVPSNLPH EATRDTVFQG YVIPKGTVVI PTLDSLLYDSHEFPDPEKFK PEHFLNENGK
FKYSDYFKAF SAGKRVCVGE GLARMELFLL LSAILQHFNL KSLVDPKDID LSPVTVGFGSIPPQFKLCVI PRS

Mass (average): 56626.9 **Identifier:** gi|1352193 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 86/493 = **17.4%** by amino acid count, 9774.1/56626.9 = **17.3%** by mass

DIDLSPVTVGFGSIPPQFK	468 - 486
DRLDMPYMDAVVHEIQR	343 - 359
FINLVPSNLPHEATR	360 - 374
FKPEHFLNENGK	409 - 420
GTVVIPTLDSLLYDSHEFPDPEK	386 - 408

>**gi|5921916**|sp|Q64581|CP3I_RAT CYTOCHROME P450 3A18 (CYPIIIA18) (P450(6)BETA-2)□gi|1083640|pir||S52097 cytochrome P450III - rat□gi|531374|emb|CAA56312.1| (X79991) cytochrome P450III [Rattus norvegicus]□gi|2463546|dbj|BAA22526.1| (D38381) P450 6beta-2 [Rattus norvegicus]□gi|1094401|prf||2106148A cytochrome P450 [Rattus norvegicus] [MASS=57307]

MEIIPNLSIE TWVLLATSLM LFYIYGTYSH GLFKKLGIPG PKPVPLFGTI FNYGDGMWKF DDDCYKKYGK
IWGFYEGPQPFLAIMDPEII KMVLVKECYS VFTNRRCFGP MGFMKKAITM SEDEEWKRLR TILSPTFTSG
KLKEMFPLMR QYGDTLLKNLRREEAKGEPI NMKDIFGAYS MDVITGTSFG VNVDSLNNPQ DPFVQKAKKI
LKFQIFDPFL LSVVLFPFLT PIYEMLNFSIFPRQSMNFFK KFVKTMKKNR LDSNQKNRVD FLQLMMNTQN
SKGQESQKAL SDLEMAAQAI IFIFGGYDAT STSISFIMYELATRPNVQKK LQNEIDRALP NKAPVTYDAL
MEMEYLDMVV NESLRLYPIA TRLDRVSKKD VEINGVFIPK GTVVTIPIYPLHRNPEYWLE PEEFNPERFS
KENKGSIDPY VYLPFGNGPR NCIGMRFALI SMKLAVIGVL QNFNIQPCEK TQIPLKISRQPIFQPEGPII LKLVSRD

Mass (average): 57307.1 **Identifier:** gi|5921916 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 38/497 = **7.6%** by amino acid count, 4251.0/57307.1 = **7.4%** by mass

GSIDPYVYLPFGNGPR	425 - 440
TQIPLKISRQPIFQPEGPIILK	471 - 492

>**gi|117164**|sp|P20816|CP42_RAT CYTOCHROME P450 4A2 PRECURSOR (CYPIVA2) (LAURIC ACID OMEGA-HYDROXYLASE) (P450-LA-OMEGA 2) (P450 K-5) (P-450 K-2)□gi|92156|pir||A32965 cytochrome P450 4A2 - rat□gi|203789|gb|AAA41039.1| (M57719) cytochrome P-450 IVA2 [Rattus norvegicus] [MASS=57969]

MGFSVFSPTR SLDGVSGFFQ GAFLLSLFLV LFKAVQFYLR RQWLLKALEK FPSTPSHWLW GHNLKDREFQ QVLTWVEKFPGACLQWLSGS TARVLLYDPD YVKVVLGRSD PKPYQSLAPW IGYGLLLLNG KKWFQHRRML TPAFHYDILK PYVKIMADSVSIMLDKWEKL DDQDHPLEIF HYVSLMTLDT VMKCAFSHQG SVQLDVNSRS YTKAVEDLNN LIFFRVRSAF YGNSIIYNMSSDGRLSRRAC QIAHEHTDGV IKTRKAQLQN EEELQKARKK RHLDFLDILL FAKMEDGKSL SDEDLRAEVD TFMFEGHDTTASGISWVFYA LATHPEHQER CREEVQSILG DGTSVTWDHL DQMPYTTMCI KEALRLYSPV PSVSRELSSP VTFPDGRSIPKGIRVTILIY GLHHNPSYWP NPKVFDPSRF SPDSPRHSHA YLPFSGGARN CIGKQFAMNE LKVAVALTLL RFELLPDPTRIPVPMPRLVL KSKNGIHLRL KKLR

Mass (average): 57968.9 **Identifier:** gi|117164 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 12/504 = **2.4%** by amino acid count, 1444.7/57968.9 = **2.5%** by mass

HLDFLDILLFAK	282 - 293

>**gi|1352184**|sp|P20817|CP43_RAT CYTOCHROME P450 4A3 (CYPIVA3) (LAURIC ACID OMEGA-HYDROXYLASE) (P450-LA-OMEGA 3)□gi|320080|pir||A32966 cytochrome P450 4A3 - rat□gi|204990|gb|AAA41458.1| (M33936) cytochrome P450 (IVA3) [Rattus norvegicus] [MASS=58232]

MGFSVFTPTR SLDGVSGFFQ GAFLLSLFLV LFKAVQFYLR RQWLLKALEK FPSTPSHWLW GHDLKDREFQ QVLTWVEKFPGACLQWLSGS KTRVLLYDPD YVKVVLGRSD PKASGIYQFL APWIGYGLLL LNGKKWFQHR RMLTPAFHYG ILKPYVKIMADSVNIMLDKW EKLDDQDHPL EIFHYVSLMT LDTVMKCAFS HQGSVQLDVN SRSYTKAVED LNNLTFFRVR SAFYGNSIIYNMSSDGRLSR RACQIAHEHT DGVIKMRKAQ LQNEEELQKA RKKRHLDFLD ILLFAKMEDG KSLSIDEDLRA EVDTFMFEGHDTTASGISWV FYALATHPEH QERCREEVQS ILGDGTSVTW DHLDQIPYTT MCIKEALRLY PPVPSVSREL SSPVTFPDGRSIPKGITTTI LIYGLHHNPS YWPNPKVFDP SRFSPDSPRH SHAYLPFSGG ARNCIGKQFA MNELKVAVAL TLLRFELLPDPTRIPVPMAR LVLKSKNGIH LRLKKLR

Mass (average): 58232.3 **Identifier:** gi|1352184 **Database:** C:/Xcalibur/database//rat.fasta
Protein Coverage: 23/507 = **4.5%** by amino acid count, 2495.8/58232.3 = **4.3%** by mass

HSHAYLPFSGGAR	440 - 452
LYPPVPSVSR	379 - 388

Appendix ii

Amino acid sequence summary for CYPs identified from the human liver

The total sequence coverage given here is the result of a single datafile searched. The total sequence coverage for each CYP isoforms given in table 3.1.1 was manually calculated from three independent datafiles run on Sequest. The peptide identification was confirmed on the basis of Xcorr and DelCn values.

>**gi|4503201**|ref|NP_000752.1|| cytochrome P450, subfamily I (aromatic compound-inducible), polypeptide 2□gi|117144|sp|P05177|CP12_HUMAN CYTOCHROME P450 1A2 (CYPIA2) (P450-P3) (P(3)450) (P450 4)□gi|65672|pir||O4HU4 cytochrome P450 1A2 - human□gi|30339|emb|CAA77335.1| (Z00036) cytochrome P3-450 (aa 1-515) [Homo sapiens]□gi|181308|gb|AAA52146.1| (M55053) cytochrome P3-450 [Homo sapiens]□gi|181382|gb|AAA52163.1| (M31667) cytochrome P450 [Homo sapiens]□gi|449342|prf||1918405A cytochrome P450 1A2 [Homo sapiens] [MASS=58294]

MALSQSVPFS ATELLLASAI FCLVFWVLKG LRPRVPKGLK SPPEPWGWPL LGHVLTLGKN PHLALSRMSQ
RYGDVLQIRIGSTPVLVLSR LDTIRQALVR QGDDFKGRPD LYTSTLITDG QSLTFSTDSG PVWAARRRLA
QNALNTFSIA SDPASSSSCYLEEHVSKEAK ALISRLQELM AGPGHFDPYN QVVVSVANVI GAMCFGQHFP
ESSDEMLSLV KNTHEFVETA SSGNPLDFFPILRYLPNPAL QRFKAFNQRF LWFLQKTVQE HYQDFDKNSV
RDITGALFKH SKKGPRASGN LIPQEKIVNL VNDIFGAGFDTVTTAISWSL MYLVTKPEIQ RKIQKELDTV
IGRERRPRLS DRPQLPYLEA FILETFRHSS FLPFTIPHST TRDTTLNGFYIPKKCCVFVN QWQVNHDPEL
WEDPSEFRPE RFLTADGTAI NKPLSEKMML FGMGKRRCIG EVLAKWEIFL FLAILLQQLEFSVPPGVKVD
LTPIYGLTMK HARCEHVQAR RFSIN

Mass (average): 58294.3 **Identifier:** gi|4503201 **Database:** C:/Xcalibur/database//human.fasta
Protein Coverage: 144/515 = **28.0%** by amino acid count, 15958.0/58294.3 = **27.4%** by mass

ASGNLIPQEK	297 - 306
DTTLNGFYIPK	393 - 403
FLTADGTAINKPLSEK	432 - 447
GRPDLYTSTLITDGQSLTFSTDSGPVWAAR	107 - 136
IGSTPVLVLSR	80 - 90
KIQKELDTVIGR	342 - 353
NPHLALSR	60 - 67
NTHEFVETASSGNPLDFFPILR	222 - 243
TVQEHYQDFDK	267 - 277
TVQEHYQDFDKNSVR	267 - 281
YLPNPALQR	244 - 252

>**gi|15147326**|ref|NP_000753.2| cytochrome P450, family 2, subfamily A, polypeptide 6; coumarin 7-hydroxylase; cytochrome P450, subfamily IIA (phenobarbital-inducible), polypeptide 3; cytochrome P450, subfamily IIA (phenobarbital-inducible), polypeptide 6; xenobiotic monooxygenase; flavoprotein-linked monooxygenase [Homo sapiens] [MASS=56501]
MLASGMLLVA LLVCLTVMVL MSVWQQRKSK GKLPPGPTPL PFIGNYLQLN TEQMYNSLMK ISERYGPVFT
IHLGPRRVVVLCGHDAVREA LVDQAEEFSG RGEQATFDWV FKGYGVVFSN GERAKQLRRF SIATLRDFGV
GKRGIEERIQ EEAGFLIDALRGTGGANIDP TFFLSRTVSN VISSIVFGDR FDYKDKEFLS LLRMMLGIFQ
FTSTSTGQLY EMFSSVMKHL PGPQQQAFQLLQGLEDFIAK KVEHNQRTLD PNSPRDFIDS FLIRMQEEEK
NPNTEFYLKN LVMTTLNLFI GGTETVSTTL RYGFLLLMKHPEVEAKVHEE IDRVIGKNRQ PKFEDRAKMP
YMEAVIHEIQ RFGDVIPMSL ARRVKKDTKF RDFFLPKGTE VFPMLGSVLRDPSFFSNPQD FNPQHFLNEK
GQFKKSDAFV PFSIGKRNCF GEGLARMELF LFFTTVMQNF RLKSSQSPKD IDVSPKHVGFATIPRNYTMS FLPR

Mass (average): 56501.4 **Identifier:** gi|15147326 **Database:** C:/Xcalibur/database//human01July03edited.fasta
Protein Coverage: 231/494 = **46.8%** by amino acid count, 26174.3/56501.4 = **46.3%** by mass

DFGVGK	137 - 142
DFIDSFLIR	266 - 274
DIDVSPK	470 - 476
DPSFFSNPQDFNPQHFLNEK	401 - 420
EALVDQAEEFSGR	89 - 101
FDYK	191 - 194
GEQATFDWVFK	102 - 112
GTGGANIDPTFFLSR	162 - 176
GYGVVFSNGER	113 - 123
HLPGPQQQAFQLLQGLEDFIAK	229 - 250
HPEVEAK	320 - 326
HVGFATIPR	477 - 485
IQEEAGFLIDALR	149 - 161
LPPGPTPLPFIGNYLQLNTEQMYNSLMK	33 - 60
MQEEEKNPNTEFYLK	275 - 289
TLDPNSPR	258 - 265
TVSNVISSIVFGDR	177 - 190
VHEEIDR	327 - 333
YGPVFTIHLGPR	65 - 76

>**gi|1168128**|gb|AAB35292.1| cytochrome P450 arachidonic acid epoxygenase isoform, Cyp 2C8 [human, kidney, Peptide Partial, 485 aa] [MASS=55197]
VLVLCLSFML LFSLWRQSCR RRKLPPGPTP LPIIGNMLQI DVKDICKSFT NFSKVYGPVF TVYFGMNPIV VFHGYEAVKEALIDNGEEFS GRGNSPISQR ITKGLGIISS NGKRWKEIRR FSLTTLRNFG MGKRSIEDRV QEEAHCLVEE LRKTKASPCDPTFILGCAPC NVICSVVFQK RFDYKDQNFL TLMKRFNENF RILNSPWIQV CNNFPLLIDC FPGTHNKVLK NVALTRSYIREKVKEHQASL DVNNPRDFID CFLIKMEQEK DNQKSEFNIE NLVGTVADLF VAGTETTSTT LRYGLLLLLK HPEVTAKVQEEIDHVIGRHR SPCMQDRSHM PYTDAVVHEI QRYSDLVPTG VPHAVTTDTK FRNYLIPKGT TIMALLTSVL HDDKEFPNPNIFDPGLFLDK NGNFKKSDYF MPFSAGKRIC AGEGLARMEL FLFLTTILQN FNLKSVDDLK NLNTTAVTKG IVSLPPSYQICFIPV

Mass (average): 55197.0 **Identifier:** gi|1168128 **Database:** C:/Xcalibur/database//human.fasta
Protein Coverage: 123/485 = **25.4%** by amino acid count, 13355.3/55197.0 = **24.2%** by mass

EALIDNGEEFSGR	80 - 92
HPEVTAK	311 - 317
LPPGPTPLPIIGNMLQIDVK	24 - 43
NFGMGK	128 - 133
NVALTR	231 - 236
NYLIPK	373 - 378
SEFNIENLVGTVADLFVAGTETTSTTLR	275 - 302
VQEEIDHVIGR	318 - 328
YGLLLLLK	303 - 310
YSDLVPTGVPHAVTTDTK	353 - 370

>**gi|6686268**|sp|P11712|CPC9_HUMAN CYTOCHROME P450 2C9 (CYPIIC9) (P450 PB-1) (P450 MP-4) (S-MEPHENYTOIN 4-HYDROXYLASE) (P-450MP)□gi|87271|pir||B38462 S-mephenytoin 4-hydroxylase (EC 1.14.14.-) cytochrome P450 2C9 - human□gi|359735|prf||1313295A cytochrome P450 [Homo sapiens] [MASS=55628]
MDSLVVLVLC LSCLLLLSLW RQSSGRGKLP PGPTPLPVIG NILQIGIKDI SKSLTNLSKV YGPVFTLYFG LKPIVVLHGYEAVKEALIDL GEEFSGRGIF PLAERANRGF GIVFSNGKKW KEIRRFSLMT LRNFGMGKRS IEDRVQEEAR CLVEELRKTKASPCDPTFIL GCAPCNVICS IIFHKRFDYK DQQFLNLMEK LNENIKILSS PWIQICNNFS PIIDYFPGTH NKLLKNVAFMKSYILEKVKE HQESMDMNNP QDFIDCFLMK MEKEKHNQPS EFTIESLENTAVDLFGAGTE TTSTTLRYAL LLLLKHPEVTAKVQEEIERV IGRNRSPCMQ DRSHMPYTDA VVHEVQRYID LLPTSLPHAV TCDIKFRNYL IPKGTTILIS LTSVLHDNKEFPNPEMFDPH HFLDEGGNFK KSKYFMPFSA GKRICVGEAL AGMELFLFLT SILQNFNLKS LVDPKNLDTT PVVNGFASVPPFYQLCFIPV

Mass (average): 55627.9 **Identifier:** gi|6686268 **Database:** C:/Xcalibur/database//human.fasta
Protein Coverage: 170/490 = **34.7%** by amino acid count, 18943.6/55627.9 = **34.1%** by mass

EALIDLGEEFSGR	85 - 97
EFPNPEMFDPHHFLDEGGNFK	400 - 420
GTTILISLTSVLHDNK	384 - 399
HNQPSEFTIESLENTAVDLFGAGTETTSTTLR	276 - 307
HPEVTAK	316 - 322
LPPGPTPLPVIGNILQIGIK	29 - 48
NFGMGK	133 - 138
NVAFMK	236 - 241
NYLIPK	378 - 383
SHMPYTDAVVHEVQR	343 - 357
SLTNLSK	53 - 59
SYILEK	242 - 247
VQEEIER	323 - 329
YALLLLLK	308 - 315

>**gi|117250**|sp|P05181|CPE1_HUMAN CYTOCHROME P450 2E1 (CYPIIE1) (P450-J)□gi|87272|pir||A31949 cytochrome P450 2E1 - human□gi|181356|gb|AAA52155.1| (J02843) cytochrome P450IIE1 [Homo sapiens]□gi|181360|gb|AAA35743.1| (J02625) cytochrome P450j [Homo sapiens] [MASS=56849]

MSALGVTVAL LVWAAFLLLV SMWRQVHSSW NLPPGPFPLP IIGNLFQLEL KNIPKSFTRL AQRFGPVFTL YVGSQRMVVMHGYKAVKEAL LDYKDEFSGR GDLPAFHAHR DRGIIFNNGP TWKDIRRFSL TTLRNYGMGK QGNESRIQRE AHFLLEALRKTQGQPFDPTF LIGCAPCNVI ADILFRKHFD YNDEKFLRLM YLFNENFHLL STPWLQLYNN FPSFLHYLPG SHRKVIKNVAEVKEYVSERV KEHHQSLDPN CPRDLTDCLL VEMEKEKHSA ERLYTMDGIT VTVADLFFAG TETTSTTLRY GLLILMKYPEIEEKLHEEID RVIGPSRIPA IKDRQEMPYM DAVVHEIQRF ITLVPSNLPH EATRDTIFRG YLIPKGTVVV PTLDSVLYDNQEFPDPEKFK PEHFLNENGK FKYSDYFKPF STGKRVCAGE GLARMELFLL LCAILQHFNL KPLVDPKDID LSPIHIGFGCIPPRYKLCVI PRS

Mass (average): 56848.9 **Identifier:** gi|117250 **Database:** C:/Xcalibur/database//human.fasta
Protein Coverage: 183/493 = **37.1%** by amino acid count, 21393.3/56848.9 = **37.6%** by mass

DRQEMPYMDAVVHEIQR	343 - 359
EAHFLLEALR	150 - 159
EALLDYKDEFSGR	88 - 100
FGPVFTLYVGSQR	64 - 76
FITLVPSNLPHEATR	360 - 374
FKPEHFLNENGK	409 - 420
GDLPAFHAHR	101 - 110
GTVVVPTLDSVLYDNQEFPDPEK	386 - 408
HFDYNDEKFLR	188 - 198
IPAIK	338 - 342
LHEEIDR	325 - 331
PEHFLNENGKFK	411 - 422
VCAGEGLARMELFLLLCAILQHFNLK	436 - 461
YPEIEEK	318 - 324
YSDYFKPFSTGK	423 - 434

>**gi|117156**|sp|P08684|CP34_HUMAN CYTOCHROME P450 3A4 (CYPIIIA4) (NIFEDIPINE OXIDASE) (NF-25) (P450-PCN1)□gi|35911|emb|CAA30944.1| (X12387) cytochrome P-450 (AA 1-503) [Homo sapiens] [MASS=57430]

MALIPDLAME TWLLLAVSLV LLYLYGTHSH GLFKKLGIPG PTPLPFLGNI LSYHKGFCMF DMECHKKYGK VWGFYDGQQPVLAITDPDMI KTVLVKECYS VFTNRRPFGP VGFMKSAISI AEDEEWKRLR SLLSPTFTSG KLKEMVPIIA QYGDVLVRNLRREAETGKPV TLKDVFGAYS MDVITSTSFG VNIDSLNNPQ DPFVENTKKL LRFDFLDPFF LSITVFPFLI PILEVLNICVFPREVTNFLR KSVKRMKESR LEDTQKHRVD FLQLMIDSQN SKETESHKAL SDLELVAQSI IFIFAGYETT SSVLSFIMYELATHPDVQQK LQEEIDAVLP NKAPPTYDTV LQMEYLDMVV NETLRLFPIA MRLERVCKKD VEINGMFIPK GWVVMIPSYALHRDPKYWTE PEKFLPERFS KKNKDNIDPY IYTPFGSGPR NCIGMRFALM NMKLALIRVL QNFSFKPCKE TQIPLKLSLGGLLQPEKPVV LKVESRDGTV SGA

Mass (average): 57430.3 **Identifier:** gi|117156 **Database:** C:/Xcalibur/database//human.fasta
Protein Coverage: 157/503 = **31.2%** by amino acid count, 17359.2/57430.3 = **30.2%** by mass

DNIDPYIYTPFGSGPR	425 - 440
EAETGKPVTLK	163 - 173
EMVPIIAQYGDVLVR	144 - 158
ETQIPLK	470 - 476
EVTNFLR	244 - 250
LGIPGPTPLPFLGNILSYHK	36 - 55
LQEEIDAVLPNK	331 - 342
LSLGGLLQPEKPVVLK	477 - 492
NKDNIDPYIYTPFGSGPR	423 - 440
SLLSPTFTSGK	131 - 141
TVLVK	92 - 96
VDFLQLMIDSQNSK	269 - 282
VWGFYDGQQPVLAITDPDMIK	71 - 91

>**gi|4503231**|ref|NP_000768.1|| cytochrome P450, subfamily IIIA (niphedipine oxidase), polypeptide 5□gi|117157|sp|P20815|CP35_HUMAN CYTOCHROME P450 3A5 (CYPIIIA5) (P450-PCN3)□gi|87283|pir||A34101 cytochrome P450 3A5 - human□gi|181346|gb|AAA02993.1| (J04813) cytochrome P450 PCN3 [Homo sapiens] [MASS=57109]

MDLIPNLAVE TWLLLAVSLV LLYLYGTRTH GLFKRLGIPG PTPLPLLGNV LSYRQGLWKF DTECYKKYGK MWGTYEGQLPVLAITDPDVI RTVLVKECYS VFTNRRSLGP VGFMKSAISL AEDEEWKRIR SLLSPTFTSG KLKEMFPIIA QYGDVLVRNLRREAEKGKPV TLKDIFGAYS MDVITGTSFG VNIDSLNNPQ DPFVESTKKF LKFGFLDPLF LSIILFPFLT PVFEALNVSLFPKDTINFLS KSVNRMKKSR LNDKQKHRLD FLQLMIDSQN SKETESHKAL SDLELAAQSI IFIFAGYETTSSVLSFTLYELATHPDVQQK LQKEIDAVLP NKAPPTYDAV VQMEYLDMVV NETLRLFPVA IRLERTCKKD VEINGVFIPK GSMVVIPTYALHHDPKYWTE PEEFRPERFS KKKDSIDPYI YTPFGTGPRN CIGMRFALMN MKLALIRVLQ NFSFKPCKET QIPLKLDTQGLLQPEKPIVL KVDSRDGTLS GE

Mass (average): 57108.6 **Identifier:** gi|4503231 **Database:** C:/Xcalibur/database//human.fasta

Protein Coverage: 107/502 = **21.3%** by amino acid count, 11909.0/57108.6 = **20.9%** by mass

ETQIPLK	469 - 475
ETQIPLKLDTQGLLQPEK	469 - 486
FLKFGFLDPLFLSIILFPFLTPVFEALNVSLFPK	210 - 243
LGIPGPTPLPLLGNVLSYR	36 - 54
LQKEIDAVLPNK	331 - 342
SAISLAEDEEWKR	116 - 128
SLLSPTFTSGK	131 - 141

>**gi|6166044**|sp|P78329|CPF2_HUMAN CYTOCHROME P450 4F2 (CYPIVF2) (LEUKOTRIENE-B4 OMEGA-HYDROXYLASE) (LEUKOTRIENE-B4 20-MONOOXYGENASE) (CYTOCHROME P450-LTB-OMEGA)□gi|7430699|pir||S45702 leukotriene-B4 20-monooxygenase (EC 1.14.13.30) cytochrome P450 4F3 - human□gi|1857022|dbj|BAA05490.1| (D26480) leukotriene B4 omega-hydroxylase [Homo sapiens]□gi|3347822|gb|AAC27730.1| (AC005336) CYP4F2; LEUKOTRIENE-B4 20-MONOOXYGENASE; YTOCHROME P450-LTB-OMEGA; LEUKOTRIENE-B4 OMEGA-HYDROXYLASE [Homo sapiens] [MASS=59853]

MSQLSLSWLG LWPVAASPWL LLLLVGASWL LAHVLAWTYA FYDNCRRLRC FPQPPRRNWF WGHQGMVNPT EEGMRVLTQLVATYPQGFKV WMGPISPLLS LCHPDIIRSV INASAAIAPK DKFFYSFLEP WLGDGLLLSA GDKWSRHRRM LTPAFHFNILKPYMKIFNES VNIMHAKWQL LASEGSACLD MFEHISLMTL DSLQKCVFSF DSHCQEKPSE YIAAILELSA LVSKRHHEILLHIDFLYYLT PDGQRFRRAC RLVHDFTDAV IQERRRTLPS QGVDDFLQAK AKSKTLDFID VLLLSKDEDG KKLSDEDIRAEADTFMFEGH DTTASGLSWV LYHLAKHPEY QERCRQEVQE LLKDREPKEI EWDDLAHLPF LTMCMKESLR LHPPVPVISRHVTQDIVLPD GRVIPKGIIC LISVFGTHHN PAVWPDPEVY DPFRFDPENI KERSPLAFIP FSAGPRNCIG QTFAMAEMKVVLALTLLRFR VLPDHTEPRR KPELVLRAEG GLWLRVEPLS

Mass (average): 59853.5 **Identifier:** gi|6166044 **Database:** C:/Xcalibur/database//human.fasta

Protein Coverage: 58/520 = **11.2%** by amino acid count, 6543.5/59853.5 = **10.9%** by mass

LHPPVPVISR	391 - 400
LVHDFTDAVIQER	262 - 274
TLDFIDVLLLSK	295 - 306
TLPSQGVDDFLQAK	277 - 290
VLPDHTEPR	491 - 499

>**gi|2117372**|pir||I65981 fatty acid omega-hydroxylase (EC 1.14.15.-) cytochrome P450 4A11 - human□gi|456998|gb|AAB29503.1| fatty acid omega-hydroxylase; CYP4A11v [Homo sapiens] [MASS=66678]

MSVSVLSPSR LLGDVSGILQ AASLLILLLL LIKAVQLYLH RQWLLKALQQ FPCPPSHWLF GHIQELQQDQ
ELQRIQKWVETFPSACPHWL WGGKVRVQLY DPDYMKVILG RSDPKSHGSY RFLAPWIGYG LLLLNGQTWF
QHRRMLTPAF HYDILKPYVGLMADSVRVML DKWEELLGQD SPLEVFQHVS LMTLDTIMKC AFSHQGSIQV
DRNSQSYIQA ISDLNNLVFS RVRNAFHQNDTIYSLTSAGR WTHRACQLAH QHTDQVIQLR KAQLQKEGEL
EKIKRKRHLD FLDILLLAKM ENGSILSDKD LRAEVDTFMFEGHDTTASGI SWILYALATH PKHQERCREE
IHSLLGDGAS ITWNHLDQMP YTTMCIKEAL RLYPPVPGIG RELSTPVTFPDGRSLPKGIM VLLSIYGLHH
NPKVWPNPEV FDPSRFAPGS AQHSHAFLPF SGGSRNCIGK QFAMNELKVA TALTLLRFELLPDPTRIPIP
IARLVLKSKM ESTCVSGGSL TLVKTRTSFE GLHLPSCLPD PRFCPLPVCP YPVFCLPTFP SSHLPAVPQS
ACPSLSHLSP GLPTCLSTCL LPTCISCWEK S

Mass (average): 66677.5 **Identifier:** gi|2117372 **Database:** C:/Xcalibur/database//human01July03edited.fasta

Protein Coverage: 82/591 = **13.9%** by amino acid count, 9214.4/66677.5 = **13.8%** by mass

ELSTPVTFPDGR	392 - 403
HLDFLDILLLAK	288 - 299
LYPPVPGIGR	382 - 391
NAFHQNDTIYSLTSAGR	234 - 250
NSQSYIQAISDLNNLVFSR	213 - 231
VWPNPEVFDPSR	424 - 435

List of Presentations and Publications

Poster Presentations

S. Nisar, C. Lane, A. Wilderspin, K. Welham, S. Orr, S. Kingston and L. H. Patterson (2002) Nanoelectrospray Ionisation Tandem Mass Spectrometric Identification of Cytochrome P450s in Human Liver. PC003
http://www.abdn.ac.uk/~bch196/HUG2003abstracts.doc

S. Nisar, C. S. Lane, W. J. Griffiths, K. J. Welham and L. H. Patterson (2004) Nanoelectrospray Ionisation Tandem mass Spectrometric Identification of Cytochromes P450 in Colorectal Metastases and Normal Human Liver. Presented at 52nd American Society of Mass Spectrometry, Nashville USA.

Oral Presentations

S. Nisar, C. S. Lane, A. F. Wilderspin, S. Orr, K. J. Welham and L. H. Patterson (2003) Nanoelectrospray Ionisation Tandem Mass Spectrometric Identification of Cytochrome P450s in Human Liver and Hepatocytes. A review of science at Harrogate, *Pharmaceutical Journal.* vol. 271, pp. 331.

Research articles

Nisar S, Lane CS, Wilderspin AF, Welham KJ, Griffiths WJ, Patterson LH. (2004) A proteomic approach to the identification of cytochrome P450 isoforms in male and female rat liver by nano-scale liquid chromatography electrospray ionisation tandem mass spectrometry. Drug Metab Dispos. 32, 382–386.

Lane,C. S., Nisar, S., Griffiths, W. J., Fuller, B. J., Davidson, B. R., Hewes, J., Welham, K. J. and Patterson, L. H. (2004) Identification of cytochrome P450 enzymes in human colorectal metastases and the surrounding liver: a proteomic approach. European Jounal of Cancer. 40, 2127-2134.

Nanoelectrospray Ionisation Tandem mass Spectrometric Identification of Cytochromes P450 in Colorectal Metastases and Normal Human Liver

S. Nisar, C. S. Lane, W. J. Griffiths, K. J. Welham and L. H. Patterson

The School of Pharmacy, University of London, 29-39 Brunswick Square, WC1N 1AX

Introduction:

Cytochromes P4450 (CYPs) have a central role in oxidative metabolism of carcinogens, drugs and steroid hormones. Detection of CYPs using mass spectrometry offers major advantages over traditional methods for the detection of CYP proteins which have relied either on techniques such as immunoblotting and activity assays, or on the detection of CYP mRNA. These techniques have significant limitations. Western blots, whilst being very sensitive rely on the availability of isoform-specific antibodies; it is necessary to pre-select which CYPs are expected to bee found and identify each isoform in turn. Activity assays invariably require multiple analysis techniques, and even so may not be totally isoform-specific. Measurement at the expression level is fraught with uncertainty since the presence and abundance of a particular type of mRNA does not infer similar presence and abundance of the corresponding protein. Mass spectrometry is increasingly used in proteomics and we have applied nano-ES-MS/MS in the identification of CYPs from human tissues, including liver and hepatocytes.

Methods:

Human colorectal metastases with corresponding normal liver samples (approximately 0.5-1gm) were received from eight consented patients at the UKHTB. These tissues were snap-frozen in liquid nitrogen and were ground in a percussion mortar and pestle. powdered tissue was homogenated to yield homogenate. Microsomes were prepared and total protein contents were determined using Bradford assay. Microsomal proteins were separated using SDS-PAGE. The gel region between 48.5 and 62 kDa was cut into five equal bands, digested with trypsin and the resultant peptides extracted as shown in Figure 1.

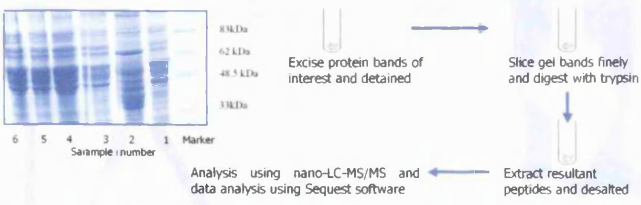


Figure 1: In gel tryptic digestion of proteins, and extraction of the resulting peptides for analysis by nano-LC-MS/MS.

MS/MS expeeriments were performed on a Thermofinnigan LCQduo with a nanospray ion source; onlinne nano-LC separations were performed using an LC Packings Ultimate with Famos autossampler. For dynamic analysis flow rates were maintained at 200nl/min.

Results and Discussion:

Human liveer and colorectal metastases microsomal preparations contained many proteins which were partially separated by 1D SDS PAGE. After proteolysis, further separation of the tryptic peptides was achieved using nano-LC prior to MS analysis as shown in Figure 2.

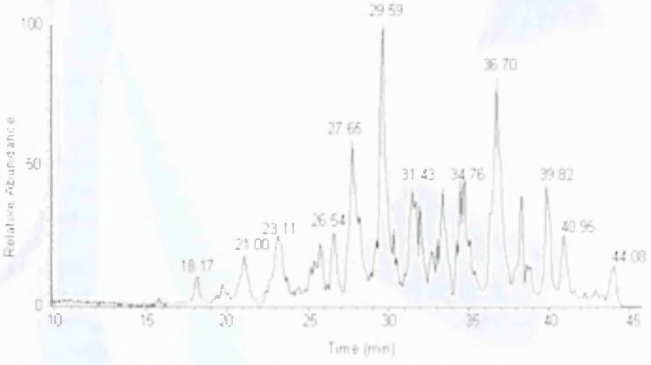


Figure 2: Nano-LC separation of a sample of human liver microsomes.

In ordered to obtain peptide sequence information peptide ions from full MS spectra were subjected to MS/MS. This generates fragment ions from which Sequest information can be obtained uusing a database search. An example is shown in Figure 3.

MS/MS spoectra were searched against a human protein database containing 56,106 entries using Sequest Browser software. Sequest has correlated tandem mass spectra from tryptic peptides of human liver samples with amino acid sequences from multiple CYPs in the human daatabase as shown in table 1.

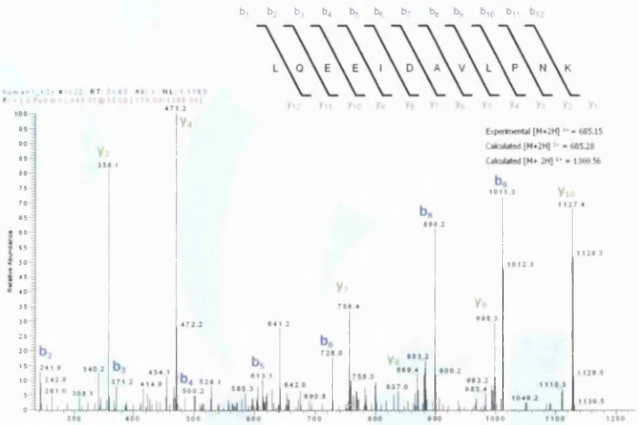


Figure 3: Collision induced dissociation generating b and y ion fragments from a CYP3A4 peptide (**LQEEIDAVLPNK**) of m/z ratio 685.15 in an ion trap.

Peptide identifications made were based on the cross correlation coefficient, Xcorr, generated by the Sequest algorithm. Proteins that were matched by two or more peptides with Xcorr value ≥ 2.5 were considered conclusively identified, provided that the peptides were unique to that protein in the database.

Table 1: CYPs identified from eight clinically derived paired samples: normal human liver vs colorectal metastases

Sample ID	# Protein content	CYP isoforms / % sequence coverage by amino acid count (number of peptides)												
		1A2	2A6	2B6	2C9	2C17	2C19	2C8	2D6	2E1	3A4	4A11	4F2	4F8
Tumour01	0.3 mg/ml	–	–	–	–	–	–	–	–	–	–	–	–	–
Liver 01	2.1 mg/ml	6 (1)	12 (3)	–	14 (3)	–	–	12 (4)	10 (2)	17 (5)	19 (5)	–	–	–
Tumour02	0.9 mg/ml	14 (3)	–	–	25 (6)	–	–	–	–	25 (6)	25 (7)	–	–	–
Liver 02	5.6 mg/ml	18 (5)	15 (4)	–	37 (12)	–	–	28 (7)	–	30 (8)	31 (10)	15 (3)	12 (4)	8 (3)
Tumour03	0.7 mg/ml	11 (2)	–	–	6 (1)	–	–	–	–	6 (1)	12 (3)	–	–	–
Liver 03	2.6 mg/ml	21 (5)	16 (4)	–	30 (8)	–	–	–	–	17 (5)	22 (6)	10 (2)	–	–
Tumour04	0.9 mg/ml	–	–	–	–	–	–	–	–	–	16 (4)	–	–	–
Liver 04	4.4 mg/ml	18 (4)	14 (4)	–	28 (6)	–	16 (4)	12 (3)	–	32 (13)	33 (9)	16 (4)	8 (2)	8 (3)
Tumour05	2.3 mg/ml	–	4 (1)	–	–	–	–	–	–	–	11 (2)	–	3 (1)	3 (1)
Liver 05	3.5 mg/ml	12 (3)	10 (2)	–	6 (1)	–	–	–	–	14 (3)	13 (3)	–	6 (2)	–
Tumour06	1.2 mg/ml	4 (1)	–	–	–	–	–	–	–	–	12 (3)	–	–	–
Liver 06	5.6 mg/ml	21 (4)	26 (7)	12 (3)	30 (8)	16 (4)	12 (3)	15 (3)	5 (1)	23 (6)	31 (10)	10 (2)	4 (1)	4 (1)
Tumour07	4.1 mg/ml	8 (2)	18 (4)	–	31 (9)	–	–	6 (1)	–	16 (4)	10 (2)	5 (1)	6 (2)	6 (2)
Liver 07	8.8 mg/ml	17 (4)	48 (16)	8 (2)	38 (9)	–	–	22 (5)	3 (1)	25 (6)	31 (9)	12 (3)	3 (1)	3 (1)
Tumour08	3.8 mg/ml	6 (1)	–	–	12.6 (3)	–	–	16 (4)	3 (1)	15 (3)	5 (1)	3 (1)	7(2)	7(2)
Liver 08	4.6 mg/ml	10 (2)	15 (3)	–	27 (6)	–	–	7 (1)	10 (2)	18 (4)	22 (5)	–	–	–

-Non identified, # Protein concentrations were determined using the Bradford assay.

Summary and Conclusions:

- CYP1A2, 3A4, 2E1 and 2C9 have been identified with high percentage sequence coverage in all individuals.
- Differences in the CYP expression in liver were observed in selected individuals which could reflect their medical history especially medications (e.g. CYP2D6, 3A5, 2C10).
- Nano-LC-MS/MS is a reliable method for the simultaneous identification of multiple CYP isoforms at the level of protein expression found in human liver and isolated hepatocytes.

References:

Srivastava P. K., Yun C. H., Beaune P. H., Ged C., and Guengerich F. P. (1991) *Mol. Pharmacol.* 40, 69-79.

Acknowledgement: Applied Biosystems (CASE) and EPSRC are gratefully acknowledged for support. BMSS

0090-9556/04//32004-382–386$20.00
DRUG METABOLISM AND DISPOSITION
Copyright © 2004 by The American Society for Pharmacology and Experimental Therapeutics
DMD 32:382–3866, 2004

Vol. 32, No. 4
1288/1140769
Printed in U.S.A.

Short Communication

A PROTEOMIC APPROACH TO THE IDENTIFICATION OF CYTOCHROME P450 ISOFORMS IN MALE AND FEMALE RAT LIVER BY NANOSCALE LIQUID CHROMATOGRAPHY-ELECTROSPRAY IONIZATION-TANDEM MASS SPECTROMETRY

(Received September 23, 2003; accepted January 13, 2004)

This article is available online at http://dmd.aspetjournals.org

ABSTRACT:

Nanoscale reeversed-phase liquid chromatography (LC) combined with electrosspray ionization-tandem mass spectrometry (ESI-MS/MS) has beeen used as a method for the direct identification of multiple cytoochrome P450 (P450) isoforms found in male and female rat liverr. In this targeted proteomic approach, rat liver microsomes were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by in-gel tryptic digestion of the proteins present in the 48- to 62-kDa bands. The resultant peptides were extracted and analyzed by LC-ESI-MS/MS. P450 identifications were made by searching the MS/MS data against a rat protein database containing 21,576 entries including 47 P450s using Sequest software (Thermo Electron, Hemel Hempstead, UK). Twenty-four P450 isoforms from the subfamilies 1A, 2A, 2B, 2C, 2D, 2E, 3A, 4A, 4F, CYP17, and CYP19 were positively identified in rat liver.

The cytochromes P450 (P450s[1]) are one of the largest known gene families and carry out a wide range of biological oxidation and reduction processes important in the metabolism of a large number of drugs, xenobiotics, and endogenous compounds. P450 enzymes are characterized into families and subfamilies by their sequence similarities; there are over 280 different families of P450s and currently more than 1925 sequenced and named isoforms (drnelson.utmem.edu/Cytochrome P4550.html). In the rat, a species traditionally important in drug development studies, there are about 50 P450 genes as identified from the UniGene database (drnelson.utmem.edu/UNIGENE.RAT.html; Wheeler et al., 2004). Traditional methods for the detection of P450 proteins have relied on immunodetection of protein, activity assays, or on the detection of P450 mRNA (Patterson and Murray, 2002). These techniques have significant limitations. Immunoblotting, although very sensitive, relies on the availability of isoform-specific antibodies, and it is necessary to preselect which P450s are to be analyzed and to identify each isoform in turn. Activity assays that are geared to investigate the activity of a P450 isoform invariably require multiple analysis techniques, and different assays must be developed for different target substrates; even then, they may not be totally isoform-specific. Measurements at the mRNA level are fraught with uncertainty, since the presence and abundance of a particular type of mRNA do not necessarily indicate a similar presence and abundance of the corresponding protein (Anderson and Seilhamer, 1997; Chen et al., 2002; McFadyen et al., 2003). Mass spectrometry is an alternative method for the analysis of expressed proteins, uniquely offering the ability to detect low levels of multiple proteins in a single analytical run. To date, there have been relatively few reports of the analysis of P450 proteins by mass spectrometry; the majority of these have been activation-based, concentrating on the observation of substrates and metabolites involved in P450 reactions rather than the P450s themselves. Matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry has been used to determine the molecular masses of closely related CYP2B1 and 2B2 (Lewis et al., 1993). MALDI-TOF analysis of peptides generated by cyanogen bromide cleavage of CYP3A4 (He et al., 1998; Lightning et al., 2000) and trypsin proteolysis of CYP2E1 (Cai and Guengerich, 2001) have been described. With regard to multiple protein identification, two-dimensional gel electrophoresis followed by MALDI-TOF and peptide mass fingerprinting (PMF) has become a principal approach for the proteomic profiling of various in vitro and in vivo biological systems (Henzel et al., 1993). A major drawback of the two-dimensional gel electrophoresis approach is its low performance in the separation of membrane proteins including P450s (Galeva and Alterman, 2002). In recognition of this, SDS-PAGE (one-dimensional gel) separation of microsomal proteins followed by MALDI-TOF and PMF was used to separate and analyze the endoplasmic reticulum proteins from rat and rabbit livers. Using this one-dimensional gel separation, up to eight P450s were positively identified (Galeva et al., 2003). However, PMF is most reliable when samples containing only a few proteins are analyzed because the presence of multiple proteins within an SDS-PAGE band may lead to spurious results. This is particularly the case when proteins are present in low abundance, and their tryptic peptide sequence coverage is low (Clauser et al., 1999; Fenyö, 2000; Huang et al., 2002). Liquid chromatography (LC)-electrospray ionization (ESI)-tandem mass spectrometry (MS/MS) has also been applied to the analysis of P450s

This work iis ssupported by a grant from the Engineering and Physical Sciences Research Coiuncil, UK.

Current address for K.J.W.: University of Hull, Cottingham Road, Hull HU16 7RX, UK.

[1] Abbreviations used are: P450, cytochrome P450; MALDI, matrix-assisted laser desorption/ionization; TOF, time-of-flight; PMF, peptide mass fingerprinting; PAGE, polyacrylamide gel electrophoresis; LC, liquid chromatography; ESI, electrospray ionization; MS/MS, tandem mass spectrometry; TFA, trifluoroacetic acid.

Address correspondence to: Prof. Laurence H. Patterson, Department of Pharmaceutical & Biological Chemistry, The School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX. E-mail: medchem@ullsopp.ac.uk

(Koenigs et al., 1999; Regal et al., 2000). Very recently, Kislinger et al. (2003) described a multidimensional LC-ESI-MS/MS method for the identification of mammalian proteins. They applied this method to analyze liver microsomes from female mice and were able to detect 20 P450 isoforms. The multidimensional LC-ESI-MS/MS or multidimensional protein identification technology method (Link et al., 1999) does not require prefractionation of the protein sample but rather relies on chromatographic separation of the very complex mixture of peptides generated by proteolysis of unseparated protein mixtures. Although the multidimensional protein identification technology approach may hold advantages in the unbiased analysis of protein mixtures, in the current study we wish to target our analysis to the identification of P450 isoforms. In this regard we have used SDS-PAGE to separate endoplasmic reticulum proteins into discrete bands according to molecular weight and then selected the bands in the P450 molecular weight range for further analysis. Proteins within bands in the 48- to 62-kDa range were digested with trypsin, and the resulting peptides were separated and analyzed by nanoscale LC-ESI-MS/MS. The MS/MS spectra provide amino acid sequence information that, in combination with peptide mass data, provides a secure identification of proteins. Male and female rats were investigated to illustrate the use of this technique in the identification of gender differences in multiple P450 protein expression.

Materials and Methods

Preparation of Rat Liver Microsomes. Fresh liver (9–10 g) from male and female outbred 10-week-old Wistar rats (Bantin and Kingman, Hull, UK) was washed with cold isotonic saline (0.9% sodium chloride, 4°C) to remove blood. Connective tissue was excised. Liver was homogenized using an Ultra Turrax T25 (Janke and Kunkel, IKA Labortecnik, Staufen, Germany) in 0.01 M Tris-HCl buffer, pH 7.4, containing 0.25 M sucrose, 15% glycerol, and 0.67 mM phenylmethanesulfonyl fluoride. Microsomes were prepared using differential centrifugation as follows: an initial centrifugation at 2400*g* for 10 min was used to sediment the cell debris, nuclei, and unbroken cells. The supernatant was centrifuged at 12,000*g* for 20 min at 4°C. Supernatant from this step were centrifuged at 180,000*g* for 1 h at 4°C. The resultant microsomal pellets were suspended in 0.1 M Tris-HCl, containing 15% glycerol and 1 mM EDTA, pH 7.4, and then recentrifuged at 180,000*g* for 1 h. The final pellet was resuspended in 0.1 M Tris-HCl, containing 15% glycerol and 1 mM EDTA, pH 7.4, and stored at −80°C. Microsomal protein was determined using the Bradford assay (Bradford, 1976).

SDS-PAGE. One-dimensional SDS-PAGE was performed using standard methods on the Hoefer Mighty Small gel system (Amersham Biosciences UK, Ltd., Little Chalfont, Buckinghamshire, UK). Microsome samples were boiled at 95°C for 8 min in a solubilization buffer (2% SDS, 20% glycerol, 0.01% bromphenol blue, and 50 mM Tris-HCl, pH 6.8) and 0.1 M dithiothreitol. Microsomal protein (20 μg) was resolved on a 10% acrylamide gel. Gels were stained with 0.2% Coomassie Brilliant Blue R-250 in 30% methanol and 10% acetic acid (1 h), and destained overnight with the same solvent.

In-Gel Tryptic Digestion and Peptide Extraction. The molecular weight region on the SDS-PAGE gel between 48 and 62 kDa was divided into five approximately equal bands, and each band was excised with a scalpel. Bands were washed in distilled water until the pH was neutral and completely destained using 50 mM NH_4HCO_3 in 40% ethanol. Bands were cut into fine pieces to increase the surface area then dried with acetonitrile and then in a SpeedVac (Thermo Savant, Holbrook, NY) for 30 min. Digestion was carried out using sequencing-grade modified trypsin (approximately 75 ng/μl) (Promega, Southampton, UK) in 25 mM NH_4HCO_3. Sufficient trypsin solution was added to swell the gel pieces, which were kept on ice for 30 min and then covered with 25 mM NH_4HCO_3 and incubated at 37°C overnight. Peptides

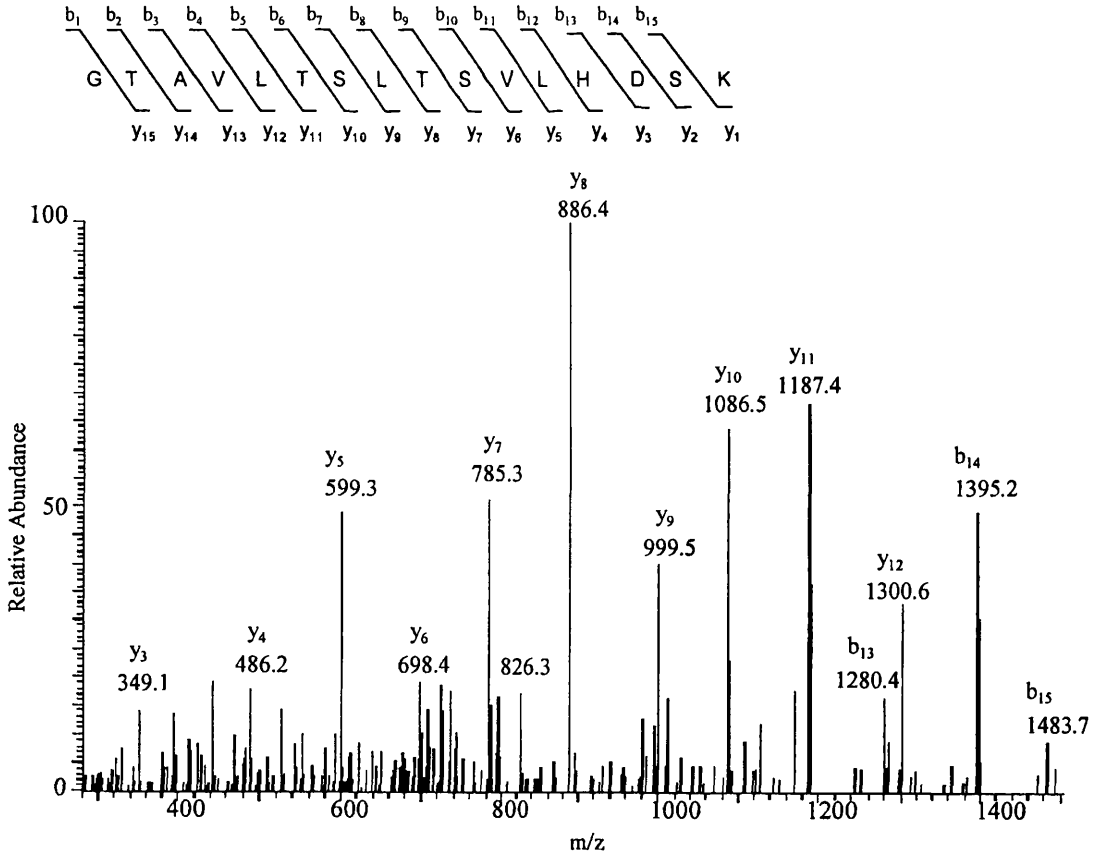


FIG. 1. *MS/MS spectrum of the tryptic peptide GTAVLTSLTSVLHDSK [M + 2H]$^{2+}$ ion of* m/z *815.4, identified to originate from CYP2C12.*

Shown in the inset is the peptide's amino acid sequence. y and b ions are formed by peptide bond cleavage with charge retention on the C terminus and N terminus, respectively.

were extracted from the gel pieces with ultrasonication, using sequential washings with a solution of 5% trifluoroacetic acid (TFA) in 50% acetonitrile. The extracts were combined and dried in a SpeedVac to complete dryness. Samples were stored at −80°C and reconstituted in 0.1% TFA before analysis.

LC-ESI–MS/MS. Nanoscale LC was performed using an LC Packings UltiMate capillary high-performance liquid chromatography system with FAMOS autosampler (Dionex, Camberley, Surrey, UK). A separate UltiMate Micro Pump (Dionex) was used as a loading pump. The sample (1 μl) was injected via a sample loop (using 0.1% TFA in water as carrier solvent) onto a 1 mm × 300 μm PepMap C18 guard column (5 μm) (LC Packings, Sunnyvale, CA). The sample was washed with 0.1% TFA for 3.5 min on the guard column before being switched onto a 15 cm × 75 μm PepMap C18 column (3 μm) (LC Packings) equilibrated with 95% mobile phase A (5% acetonitrile containing 0.1% formic acid) and 5% mobile phase B (80% acetonitrile containing 0.1% formic acid), at a flow rate of 200 nl/min. Five minutes after sample injection the proportion of mobile phase B was increased linearly to 50% over 30 min and then stepped to 95% and maintained at this level for 10 min (wash phase). The column was then re-equilibrated for 20 min with 95% mobille phase A, 5% mobile phase B. The column effluent was continuously directed into an LCQ^{duo} mass spectrometer fitted with a nano-ESI source (Thermo Electron, Hemel Hempstead, UK), and spectra were recorded. Mass spectrometer conditions were optimized using in-solution tryptic digests of purified recombinant P450 isoforms 1A2, 2E1, and 3A4, obtained from PanVera Corp. (Madison, WI).

ESI was performed under the following conditions: positive ionization mode; spray voltage, 1.8 kV; capillary voltage, 28 V; capillary temperature, 180°C; and no sheath or auxiliary gas was used. Data were collected in the full-scan and data-dependent MS/MS modes; three microscans were performed, with the maximum ion injection time of 200 ms. In the full-scan mode, ions were collected in the *m/z* range of 400 to 2000. The MS/MS collision energy was set to 35%.

Protein Identification. MS/MS spectra were searched using Sequest Browser software (Thermo Electron) (Eng et al., 1994; Yates et al., 1995), against a rat protein database containing 21,576 entries including 47 P450s (Finnigan Xcalibur; Thermo Finnigan, San Jose, CA; revision 1.0 P/N XCALI-64012, July 2000). The aim of the Sequest approach is to find the peptide sequence in a database that best explains the fragment ions present in a spectrum. Candidate sequences are found in the database on the basis of intact peptide masses, and the complete or partial spectra expected to result from the fragmentation of these candidate peptides are generated and compared with the experimental spectrum. The final score assigned to each candidate peptide sequence is the Xcorr, a measure of how well the theoretical spectrum cross-correlates to the observed spectrum. Proteins that were matched by two or more peptides with Xcorr values ≥2.5 were considered conclusively identified, provided that the peptides were unique to that protein in the database (Ducret et al., 1998).

Results and Discussion

Initially, it was considered necessary to evaluate the performance of the SDS-PAGE LC-ESI-MS/MS procedure for the identification of P450s. This was achieved by loading known amounts of purified recombinant P450 isoforms 1A2, 2E1, and 3A4 onto the SDS-PAGE gel and proceeding through the analytical cycle to identify the proteins. It was found that P450s could be successfully identified down to approximately 1 pmol loaded onto the gel (results not shown). A successfull identification was defined as one in which two or more unique peptides were found with Sequest Xcorr values ≥2.5.

Identification of P450s in Rat Liver Microsomes. Five bands of approximately equal size covering the molecular weight range of 48 to 62 kDa were cut out from SDS-PAGE and subjected to in-gel digestion with trypsin. The resultant peptides were extracted and analyzed by LC-ESI–MS/MS. To identify the proteins present, the MS/MS spectra were submitted to the Sequest algorithm. Sequest then identified the tryptic peptides by matching their MS/MS spectra against insilico generated theoretical spectra from the database. This is illustrated in Fig. 1, which shows the MS/MS spectrum of the doubly charged peptide of *m/z* 815.4. Sequest determined the amino acid sequence of the peptide to be GTAVLTSLTSVLHDSK, from the CYP2C12 protein. Using this approach, a total of 24 P450 proteins, about half of all the known rat P450s, were identified in male and/or female Wistar rat liver as shown in Table 1. Sequence coverage was found to vary from a maximum of 35% to a minimum of 8%. This variation may be due to a number of factors, including the level of protein abundance, the accessibility to trypsin digestion, and peptide ion mass spectrometric sensitivity. Nevertheless, it was possible to differentiate between closely related P450s, e.g., the highly homologous CYP2D2 and 2D3 (see Fig. 2).

Previously, the use of MALDI-TOF-based PMF of one-dimensional gel tryptic digests of liver microsomes uncovered four P450s (2A1, 2C11, 2D2, 2D5) in uninduced male Sprague-Dawley rats (Galeva et al., 2003). These isoforms were also found among the 24 P450s identified in male and female rat livers using our LC-ESI-MS/MS methodology. MALDI-TOF-based PMF of clofibrate-induced rat liver microsomes revealed the presence of CYP2B1, 2B2, 4A1, and 4A3 (Galeva et al., 2003). Of these, CYP4A3 was the only isoform detected in our male and female livers, suggesting that non-induced levels of CYP4A1, 2B1, and 2B2 are low. However, we did detect CYP2B3 in female rat liver. It should be emphasized that although the MALDI-TOF PMF approach is excellent for the identification of separated proteins and simple protein mixtures (Jensen et al., 1997; Zhang and McElvain, 2000), when proteins are present in low abundance or as part of complex mixtures, as may be the case with one-dimensional gel bands, it is necessary to obtain more information (Fenyö, 2000). The most common method to achieve this is by

TABLE 1

CYPs identified by MS/MS to be present in rat liver microsomes

Other proteins, which were identified with good sequence coverage (>15%), are: probable protein disulfide isomerase ER-60 precursor (42%), kinesin (38%), aldehyde dehydrogenase, microsomal (28%) rat ATP synthase β chain (26%), glutamate dehydrogenase precursor (24%), mitochondrial precursor (22%), microsomal epoxide hydrolase (21%), UDP-glucuronosyltransferase 2B12 precursor (16%), UDP-glucuronosyltransferase 2B3 precursor (18%), UDP-glucuronosyltransferase 2B2 precursor (19%), and F1-ATPase β subunit.

Cytochromes P450 Identified	SWISS-PROT Accession Numbers	Number of Matched Peptides (percentage of sequence coverage by amino acid count)	
		Male	Female
1A2	P04799	8 (19)	11 (28)
2A1	P11711	5 (16)	11 (30)
2A2	P15149	8 (15)	N.D.
2B3	P13107	N.D.	6 (17)
2C6	P05178	4 (15)	6 (18)
2C7	P05179	2 (11)	8 (19)
2C11	P08683	6 (18)	N.D.
2C12	P11510	N.D.	7 (25)
2C13	P20814	6 (18)	N.D.
2C22	P19225	4 (16)	N.D.
2C23	P24470	N.D.	6 (16)
2C24	P33273	N.D.	2 (8)
2D1	P10633	6 (17)	13 (35)
2D2	P10634	9 (24)	12 (34)
2D3	P12938	4 (12)	6 (18)
2D4	P13108	N.D.	5 (15)
2D5	P12939	7 (16)	N.D.
2E1	P05182	7 (25)	8 (28)
3A18	Q64581	3 (11)	N.D.
4A2	P20816	3 (12)	N.D.
4A3	P20817	4 (16)	3 (15)
4F1	P33274	3 (12)	3 (12)
CYP19	P11715	N.D.	5 (19)
CYP17	P22443	N.D.	4 (16)

N.D., not detected.

```
2D2     MGLLIGDDLWAVVIFTAIFLLLVDLVHRHKFWTAHYPPGPVPLPGLGNLLQVDFENMPYS 60
2D3     MELLAGTGLWPMAIFTVIFILLVDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLCNMPYS 60

2D2     LYKLRSRYGDVFSLQIAWKPVVVINGLKAVRELLVTYGEDTADRPLLPIYNHLGYGNKSK 120
2D3     MYKLQNRYGDVFSLQMGWKPVVVINGLKAVQELLVTCGEDTADRPEMPIFQHIGYGHKAK 120

2D2     GVVLAPYGPEWREQRRFSVSTLRDFGVGKKSLEQWVTEEAGHLCDTFAEAEHPFNPSIL 180
2D3     GVVLAPYGPEWREQRRFSVSTLRNFGVGKKSLEQWVTDEASHLCDALTAEAGRPLDPYTL 180

2D2     LSKAVSNVIASLVYARRFEYEDPFFNRMLKTLKESFGEDTGFMAEVLNAIPILLQIPGLP 240
2D3     LNKAVCNVIASLIYARRFDYGDPDFIKVLKILKESMGEQTGLFPEVLNMFPVLLRIPGLA 240

2D2     GKVFPKLNSFIALVDKMLIEHKKSWDPAQPPRDMTDAFLAEMQKAKGNPESSFNDENLRL 300
2D3     DKVFPGQKTFLTMVDNLVTEHKKTWDPDQPPRDLTDAFLAEIEKAKGNPESSFNDANLRL 300

2D2     VVIDLFMAGMVTTSTTLSWALLLMILHPDVQRRVHEEIDEVIGQVRRPEMADQARMPFTN 360
2D3     VVNDLFGAGMVTTSITLTWALLLMILHPDVQCRVQQEIDEVIGQVRHPEMADQAHMPFTN 360

2D2     AVIHEVQRFADIVPTNIPHMTSRDIKFQGFLIPKGTTLIPNLSSVLKDETVWEKPLRFHP 420
2D3     AVIHEVQRFADIVPMNLPHKTSRDIEVQGFLIPKGTTLIPNLSSVLKDETVWEKPLRFHP 420

2D2     EHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARMELFLFFTCLLQRFSFSVLAGRPRPST 480
2D3     EHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARMELFLFFTCLLQRFSFSVPTGQPRPSD 480

2D2     HGVYALPVTPQPYQLCAVAR 500
2D3     YGVFAFLLSPSPYQLCAFKR 500
```

**Different amino acid sequence*
Highlighted sequences were identified for CYP2D2 and 2D3 using LC-MS/MS

FIG. 2. *Sequence alignment and sequence coverage map for CYP2D2 and 2D3.*

The tryptic peptides ^{92}ELLVTYGEDTADRPLLPIYNHLGYGNK^{118}, ^{170}EAEHPFNPSILLSK^{183}, ^{184}AVSNVIASLVYAR^{196}, ^{197}RFEYEDPFFNR^{207}, and ^{273}DMTDAFLAEMQK^{284} are unique to CYP2D2, whereas the tryptic peptides ^{273}DLTDAFLAEIEK^{284} and ^{287}GNPESSFNDANLR^{299} are unique to CYP2D3.

MS/MS of isolated proteolytic peptides, as in our LC-ESI-MS/MS approach.

As the goal of this preliminary study was to identify the P450 proteins in liver microsomes, it was decided not to pursue a stable-isotope dilution approach, such as the isotope-coded affinity tag method (Gygi et al., 1999), for relative quantification. With stable-isotope dilution methods, proteins isolated from different sources are differentially stable-isotope labeled, combined, digested, and then analyzed by LC-MS/MS. This approach requires dilution of one sample by the other and also the analysis of selectively labeled peptides. Such a procedure would inherently limit differentiation between P450 isoforms with high sequence similarity and may also elevate the P450 protein detection limit.

Although the results of the current study are of a qualitative nature, and the nonidentification of a protein in a sample does not necessarily equate its absence, at least to a first approximation, the greater the concentration of a protein within a given sample, the greater the probability of its identification. With this in mind, some comments can be made with regard to the identification of P450 proteins based on rat gender (see Table 1). Most of the known gender-related differences in compound toxicity in rats are due to gender-related differences in hepatic metabolism (Czerniak, 2001). Hormones and growth factors are strongly implicated in the gender-influenced expression of hepatic P450s (Waxman, 1992; Pampori and Shapiro, 1999; Kalsotra et al., 2002). Whether P450 expression is gender-specific rather than gender-predominant or -biased will depend on the absolute expression of protein in one sex compared with the other. It has been suggested that a P450 is gender-specific only if the relative expression is 10-fold or higher in one sex compared with the other; lower than this has been suggested to represent enriched P450 expression in one gender rather than sex specificity (Kato and Yamazoe, 1993). Our results show that CYP2C11 and 2C13 were found in male but not female rats, whereas CYP2C12 was found in females only, supporting previous work that indicated that these three isoforms are gender-specific/predominant (Agrawal and Shapiro, 2001). Our study shows also that other CYP2C members are gender-predominant; specifically, CYP2C22 was found only in male livers, whereas CYP2C23 and 2C24 were found only in female livers. Regarding female-predominant CYP2A1, 2C6, and 2C7 (Agrawal and Shapiro, 2001), these isoforms were identified in both genders by LC-ESI-MS/MS. CYP2D1, 2D2, and 2D3 were found in both sexes whereas CYP2D4 was female-specific and CYP2D5 was male-specific. Gender selectivity of some 2D family members was previously identified (Schulz-Utermoehl et al., 1999). CYP3A18 and 4A2 were observed in male but not female rat liver, which is supported by previous studies showing that most members of the CYP3A (including CYP3A18) and

CYP4A families to be male gender-specific or -predominant (Waxman et al., 1995; Robertson et al., 1998; Holla et al., 2001; Mitchell et al., 2001; Anakk et al., 2003). CYP17 and CYP19, two enzymes that contribute to sex steroid synthesis, were identified in female but not male liver.

The sex-related differences indicated above suggest that the next step in the proteomic study of rat liver microsomal P450s should involve relative protein quantification between the sexes. Unfortunately, the straightforward use of stable-isotope dilution methods may not be applicable, as such methods will require isoform-specific labeling, which will not be trivial for P450 isoforms with high sequence similarity.

In summary, in the current study 24 P450 isoforms have been conclusively identified by LC-ESI-MS/MS. The results for the majority of P450s found are consistent with previously published studies describing expression profiles of selected P450s in rat liver. Clearly, there is a considerable advantage in using LC-ESI-MS/MS in the identification of multiple P450s from complex tissues.

Acknowledgments. We thank Mike Cocksedge and Emmanuel Samuel for technical assistance.

The School of Pharmacy, University of London, London, United Kingdom

S. NISAR
C. S. LANE
A. F. WILDERSPIN
K. J. WELHAM
W. J. GRIFFITHS
L. H. PATTERSON

References

Agrawal AK and Shapiro BH (2001) Intrinsic signals in sexually dimorphic circulating growth hormone profiles of the rat. *Mol Cell Endocrinol* **173**:167–181.

Anakk S, Kim CY, Vore M, and Strobel HW (2003) Insights into gender bias: rat cytochrome P450 3A9. *J Pharmacol Exp Ther* **305**:703–709.

Anderson L and Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18**:533–537.

Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of dye-binding. *Anal Biochem* **72**:248–254.

Cai H and Guengerich FP (2001) Reaction of trichloroethylene and trichloroethylene oxide with cytochrome P450 enzymes: inactivation and sites of modification. *Chem Res Toxicol* **14**:451–458.

Chen G, Gharib TG, Huang C, Taylor JMG, Misek DE, Kardia SLR, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM, and Beer DG (2002) Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* **1**:304–313.

Clauser KR, Baker P, and Burlingame AL (1999) Role of accurate mass measurement (±10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* **71**:2871–2882.

Czerniak R (2001) Gender-based differences in pharmacokinetics in laboratory animal models. *Int J Toxicol* **20**:161–163.

Ducret A, Van Oostveen I, Eng JK, Yates JR 3rd, and Aebersold R (1998) High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein Sci* **7**:706–719.

Eng JK, McCormack AL, and Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **5**:976–989.

Fenyö D (2000) Identifying the proteome: software tools. *Curr Opin Biotechnol* **11**:391–395.

Galeva N and Alterman M (2002) Comparison of one-dimensional and two-dimensional gel electrophoresis as a separation tool for proteomic analysis of rat liver microsomes: cytochrome P450 and other membrane proteins. *Proteomics* **2**:713–722.

Galeva N, Yakovlev D, Koen Y, Duzhak T, and Alterman M (2003) Direct identification of cytochrome P4560 isoenzymes by matrix assisted laser desorption/ionization time of flight-based proteomic approach. *Drug Metab Dispos* **31**:351–355.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, and Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**:994–999.

He K, Bornheim LM, Falick AM, Maltby D, Yin H, and Correia MA (1998) Identification of the heme-modified peptides from cumene hydroperoxide-inactivated Cytochrome P450 3A4. *Biochemistry* **37**:17448–17457.

Henzel WJ, Billeci TM, Stultz JT, Wong SC, Grimley C, and Wantanabe C (1993) Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci USA* **90**:5011–5015.

Holla VR, Adas F, Imig JD, Zhao X, Price E Jr, Olsen N, Kovacs WJ, Magnuson MA, Keeney DS, Breyer MD, et al. (2001) Alterations in the regulation of androgen-sensitive CYP4A monooxygenases cause hypertension. *Proc Natl Acad Sci USA* **98**:5211–5216.

Huang L, Baldwin MA, Maltby DA, Medzihradszky KF, Baker PR, Allen N, Rexach M, Edmondson RD, Campbell J, Juhasz P, et al. (2002) The identification of protein-protein interactions of the nuclear pore complex of *Saccharomyces cerevisiae* using high throughput matrix-assisted laser desorption ionization time-of-flight tandem mass spectrometry. *Mol Cell Proteomics* **1**:434–450.

Jensen ON, Podtelejnikov AV, and Mann M (1997) Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. *Anal Chem* **69**:4741–4750.

Kalsotra A, Anakk S, Boehme CL, and Strobel HW (2002) Sexual dimorphism and tissue specificity in the expression of CYP4F forms in Sprague Dawley rats. *Drug Metab Dispos* **30**:1022–1028.

Kato R and Yamazoe Y (1993) Hormonal regulation of cytochrome P450 in rat liver, in *Cytochrome P450*, pp 447–456, Springer-Verlag, Berlin.

Kislinger T, Rahman K, Radulovic D, Cox B, Rossant J, and Emili A (2003) PRISM, a generic large scale proteomic investigation strategy for mammals. *Mol Cell Proteomics* **2**:96–106.

Koenigs LL, Peter RM, Hunter AP, Haining RL, Rettie AE, Friedberg T, Pritchard MP, Shou M, Rushmore TH, and Trager WF (1999) Electrospray ionization mass spectrometric analysis of intact cytochrome P450: identification of tienilic acid adducts to P450 2C9. *Biochemistry* **38**:2312–2319.

Lewis S, Korsmeyer KK, and Correia MA (1993) Matrix-assisted laser desorption mass spectrometry of cytochromes P450. *Rapid Commun Mass Spectrom* **7**:16–19.

Lightning LK, Jones JP, Friedberg T, Pritchard MP, Shou M, Rushmore TH, and Trager WF (2000) Mechanism-based inactivation of cytochrome P450 3A4 by L-754, 394. *Biochemistry* **39**:4276–4287.

Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, and Yates JR (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* **17**:676–682.

McFadyen CEM, Rooney PH, Melvin WT, and Murray GI (2003) Quantitative analysis of the Ah receptor/cytochrome P450 CYP1B1/CYP1A1 signalling pathway. *Biochem Pharmacol* **65**: 1663–1674.

Mitchell SR, Sewer MB, Kardar SS, and Morgan ET (2001) Characterization of CYP4A induction in rat liver by inflammatory stimuli: dependence on sex, strain, and inflammation-evoked hypophagia. *Drug Metab Dispos* **29**:17–22.

Pampori NA and Shapiro BH (1999) Gender differences in the responsiveness of the sex dependent isoforms of hepatic P450 to the feminine plasma growth hormone profile. *Endocrinology* **140**:1245–1253.

Patterson LH and Murray GI (2002) Tumour cytochrome P450 and drug activation. *Curr Pharm Des* **8**:1335–1347.

Regal KA, Schrag ML, Kent UM, Wienkers LC, and Hollenberg PF (2000) Mechanism-based inactivation of cytochrome P450 2B1 by 7-ethynylcoumarin: verification of apo-P450 adduction by electrospray ion trap mass spectrometry. *Chem Res Toxicol* **13**:262–270.

Robertson GR, Farrell GC, and Liddle C (1998) Sexually dimorphic expression of rat CYP3A9 and CYP3A18 genes is regulated by growth hormone. *Biochem Biophys Res Commun* **242**:57–60.

Schulz-Utermoehl T, Bennett AJ, Ellis SW, Tucker GT, Boobis AR, and Edwards RJ (1999) Polymorphic debrisoquine 4-hydroxylase activity in the rat is due to differences in CYP2D2 expression. *Pharmacogenetics* **9**:357–366.

Waxman DJ (1992) Regulation of liver specific steroid metabolizing cytochrome P450: cholesterol 7α hydroxylases, bile acid 6β-hydroxylase and growth responsive steroid hormone hydroxylases. *J Steroid Biochem Mol Biol* **43**:1055–1072.

Waxman DJ, Ram PA, Pampori NA, and Shapiro BH (1995) Growth hormone regulation of male-specific rat liver P450s 2A2 and 3A2: induction by intermittent growth hormone pulses in male but not female rats rendered growth hormone deficient by neonatal monosodium glutamate. *Mol Pharmacol* **48**:790–797.

Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, et al. (2004) Database resources of the National Center for Biotechnology Information: update. *Nucleic Acids Res* **32**: **Database issue**:D35–D40.

Yates JR III, Eng JK, McCormack AL, and Schieltz D (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* **67**:1426–1436.

Zhang Z and McElvain JS (2000) De novo peptide sequencing by two-dimensional fragment correlation mass spectrometry. *Anal Chem* **72**:2337–2350.