A CHARACTERIZATION OF HUMAN MU CLASS GLUTATHIONE S-TRANSFERASES

by

Kenine Elaine Comstock

A DISSERTATION

presented to the Department of Biochemistry and Molecular Biology
and the Oregon Health Sciences University
School of Medicine
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
April 1993

APPROVED:

(Dr. W. David Henner, Associate Professor, Dept. of Biochemistry and Molecular Biology)

(Dr. John Resko, Chairman, Graduate Council)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The difficult path to the completion of my graduate thesis research and writing of this dissertation was made a little less rugged because of the help, guidance and support of numerous people. Firstly, I would like to thank David Henner for his expert advice and his willingness to help me, even in the midst of demanding clinical responsibilities. He has always displayed a great deal of confidence in me, which helped at times when I didn't feel confident in myself.

I would also like to thank members of the Henner laboratory, both past and present, for their assistance with experiments, advice on difficult problems, and creating a pleasant working environment. I especially want to thank Kara Johnson for her assistance with the sequencing of the GSTM4 gene. I also want to express my gratitude to Danica Hargett for all her help with administrative work and travel arrangements.

Dr. Bengt Mannervik, University of Uppsala, was very helpful to the completion of this project by allowing me to come to his laboratory in Sweden, arranging for my accommodations for four months and letting me use Mu class expression clones constructed in his lab. I greatly appreciated the assistance of all the members of the Mannervik group, who gave technical advice, translated Swedish, and made my stay there very enjoyable by inviting me to their homes and helping me with the difficulties of living in a foreign country.

Finally I would like to recognize Scott Schuetze who has been both inspirational and extremely supportive throughout my graduate education. I look forward to sharing many future adventures.

ABSTRACT

The glutathione *S*-transferases (GSTs) are a family of related enzymes which catalyze the conjugation of glutathione to a broad range of reactive electrophilic substrates. GSTs are involved in the detoxification of potential carcinogens and drug resistance to anticancer drugs. There are four major classes of cytosolic GSTs: Alpha, Mu, Pi and the recently discovered Theta class. Within a class, the GST proteins are immunologically cross-reactive, have similarities in substrate and inhibitor specificities and share 70-90% amino acid sequence homology. The focus of this dissertation is the human Mu class glutathione transferases which are of interest for several reasons: 1) because all genes and enzymes encoded by them which may be important toxicologically have not been isolated, 2) they may be involved in drug resistance and, 3) there is a correlation between deletion of a Mu class gene and lung and other cancers. To further characterize the Mu class GSTs, several studies were initiated. A PCR-based assay was designed which could be used in epidemiological studies to investigate the correlation between a genetic predisposition to lung cancer and a Mu class gene deletion. The genomic DNA and the corresponding cDNA sequences were determined for a new member of the human GST gene superfamily, GSTM4. The nucleotide sequence of the presumed coding region of GSTM4 is highly homologous to the coding region of the cDNA sequence of GSTM1 (93%), GSTM2 (88%), GSTM3 (69%), and rat Yb1 (77%), which established GSTM4 as a Mu class GST gene. The genomic sequence consists of eight exons, with splice junctions identical to the rat Mu class gene Yb2, and the upstream region

contains the consensus sequence for an antioxidant responsive element (ARE). The cDNA for GSTM4 was isolated from a library derived from the cervical carcinoma cell line, HeLa. Northern blots of human RNA, probed with a unique 5' region of the GSTM4 cDNA, demonstrated the presence of GSTM4 mRNA in human heart, placenta, lung, brain, liver, skeletal muscle, pancreas, testis, cerebral cortex, uterus, ovary, a lymphoblastoid cell line, and four carcinoma cell lines. After GSTM4 cDNA sequences were cloned into an *E. coli* expression vector, recombinant GSTM4-4 protein was purified to yield a homogenous band of 27.5 kd on SDS-PAGE. This enzyme has a relatively low specific activity with the standard GST substrate CDNB (1.39 $\pm$ 0.21 U/mg protein). A detailed comparison of the activity with various substrates and inhibitors was performed between the recombinant GSTM4-4 and the following other recombinant human Mu class GSTs: GSTM1a-1a, GSTM2-2 and GSTM3-3. Altogether these studies have aided in the characterization of the human Mu class GSTs and contributed to an understanding of their role in detoxification.

I. INTRODUCTION

A. Introduction to detoxification systems

1. The need for detoxification

Organisms are constantly bombarded by toxic molecules from both exogenous and endogenous sources. Toxic foreign compounds, defined as xenobiotics, arise from sources such as industrial wastes, agricultural chemicals, drugs, and cigarette smoke. The products of cellular metabolism, especially from the utilization of oxygen in energy metabolism, can also be toxic. If these reactive chemicals are not removed from the cell, they can inflict damage upon proteins, membrane lipids and DNA.

Organisms have evolved several different lines of defense against foreign toxic molecules. The first obvious one is the physical barrier of the cell membrane and, in higher organisms, the skin, mucus cilia of the respiratory tract, and epithelial lining of the gastrointestinal tract. However, as most xenobiotics are lipophilic, they can penetrate these barriers quite readily. Therefore defense against xenobiotics, as well as reactive endogenous compounds, depends on metabolism to prepare them for inactivation and removal from the cell by normal means of excretion. This metabolism of toxic molecules is catalyzed by detoxification enzymes (Jakoby, 1980).

The number and variety of toxic molecules encountered by an organism is tremendous. These molecules have all possible descriptions: inert, reactive, electrophilic, nucleophilic, lipophilic, hydrophilic. The challenge of detoxification enzyme systems is to

be able to metabolize this almost endless number of substrates. Obviously, it would be inefficient for an organism to evolve a separate enzyme for each possible substrate. Not surprisingly therefore, organisms have evolved families of closely related detoxification enzymes, each of which can catalyze the inactivation of several distinct, yet overlapping substrates. This is an evolutionarily conservative way of accounting for a broad range of substrates.

2. Metabolism of toxic molecules

Xenobiotic metabolism usually consists of two steps. In the first step, the addition of a polar reactive group to the xenobiotic is catalyzed by phase I enzymes. These enzymes catalyze reactions including microsomal monooxygenations, cytosomal oxidations, reductions, and hydrolyses with the net result of an addition of a polar group to their substrates. Some representative phase I enzyme families are the cytochrome P450s, alcohol dehydrogenases and epoxide hydratases.

The most well known phase I enzymes, cytochrome P450s, are a large family of membrane-bound, heme-containing isozymes involved in the monooxygenation of foreign compounds. They are the only detoxification enzymes able to catalyze the initial step required for the metabolism of many inert compounds such as hydrocarbons and are also important in the metabolism of therapeutic drugs (Armstrong, 1987). Reactions catalyzed by cytochrome P450s sometimes yield products which are more toxic than the unmetabolized substrate and can be substrates for other detoxification enzymes (Anders et al., 1988).

In the second step of xenobiotic metabolism, phase II enzymes catalyze the conjugation of an endogenous molecule, such as an amino acid, sugar or glutathione, to the modified xenobiotic molecule. These xenobiotics have reactive groups such as hydroxyl, amino, carboxyl, epoxide or halogen. The resulting products are usually more polar, less toxic and more easily eliminated from the body by renal or hepatic routes. In some cases, the phase II enzymes can catalyze these conjugations without prior preparation by the phase I enzymes. The phase II enzymes include the sulfotransferases, methyltransferases, UDP-glucuronosyltransferases, and glutathione *S*-transferases.

B. Introduction to glutathione *S*-transferases

1. General description

The glutathione *S*-transferases (GSTs) are a gene family of related enzymes which catalyze the conjugation of glutathione (GSH), the tripeptide γ-glutamylcysteinyl glycine, on sulfur atom of cysteine to a broad range of electrophilic substrates (Mannervik & Danielson, 1988; Pickett, 1989). GSTs are involved in detoxification and it is this function which has been most well studied. A GSH conjugate is generally less hydrophobic and may be more easily excreted as such or may be further metabolized to a mercapturic acid which is the classical excretion product of xenobiotics (Figure 1).

GSTs have also been implicated in anticancer drug resistance. In addition, GSTs may play a role as intracellular binding proteins as they have been shown to bind some hydrophobic compounds such as heme, bilirubin, bile acids and steroid hormones. GSTs are also involved in the biosynthesis of certain leukotrienes and prostaglandins.

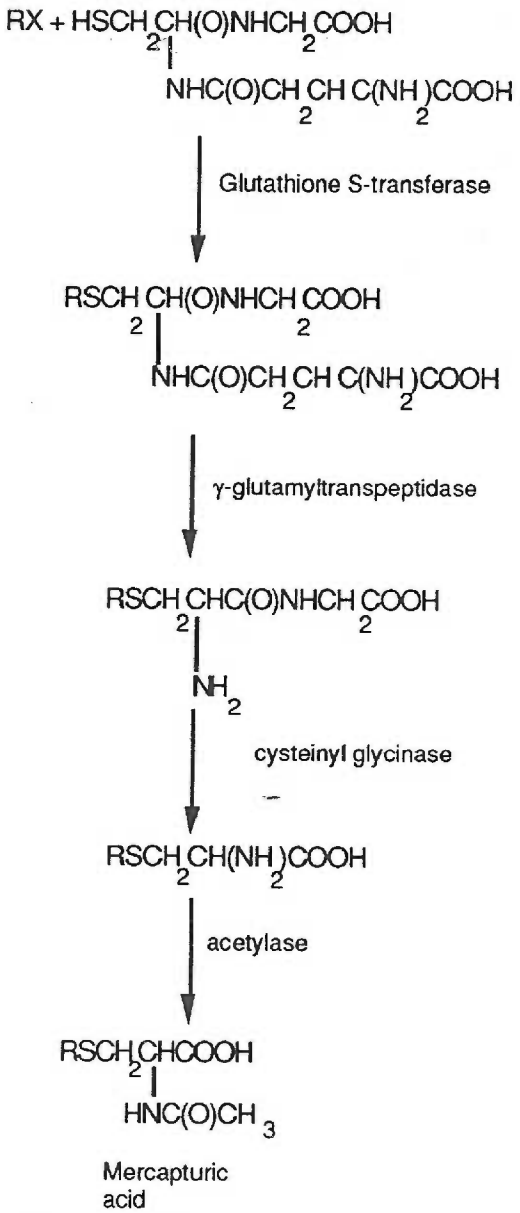


Figure 1. The mercapturic acid pathway.

There are four major classes of cytosolic GSTs, including the Alpha, Mu, Pi (Mannervik, 1985) and the recently discovered Theta class (Meyer et al., 1991) (Figure 2). A microsomal enzyme which is more distantly related to the other classes has also been isolated (McLellen et al., 1989). Within a class, the GST proteins are immunologically cross-reactive, have similarities in substrate and inhibitor specificities, and share 70-90% amino acid sequence homology, while only 30% homology is found between different classes (Alin et al., 1985). The presence of multiple GST isozymes is a characteristic of all species studied including humans, rats, chickens, trout, cockroaches, corn

and peas (Mannervik, 1985). So far, it has not been possible to definitively assign human homologues to GSTs from other species either by nucleotide sequence identity or by comparing similarities in substrate specificities or specific activities for particular substrates.

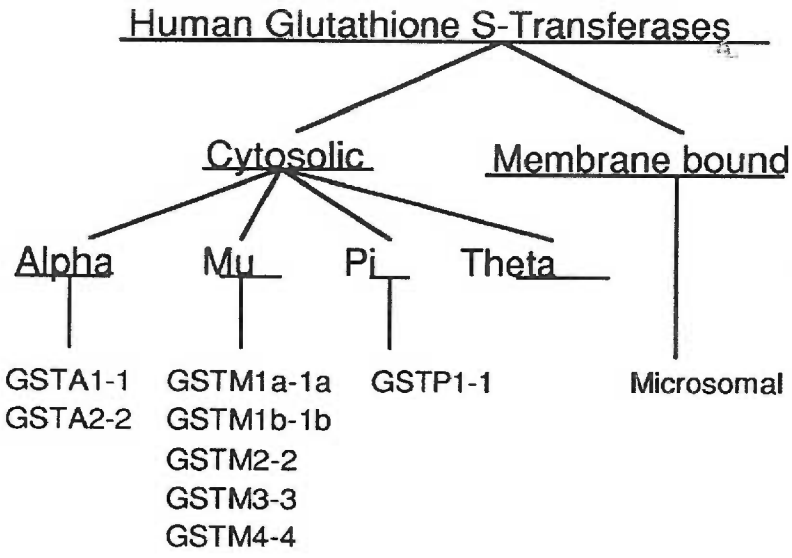


Figure 2. Evolutionary relationships between the Human GSTs.

The cytosolic GSTs are dimers of 22-25 kd subunits, while the microsomal enzyme is a trimer of identical 17 kd subunits (Morgenstern et al., 1985). The cytosolic enzymes have two active sites per dimer, each of which functions independently of the other (Danielson & Mannervik, 1985); however, there is no evidence for a catalytically-active monomer. Although heterodimers have been isolated and may form between subunits of two cytosolic GST

isozymes from the same class, subunits from different classes do not appear to dimerize (Reddy et al., 1984; Tu et al., 1985; Mannervik, 1985). The ability of the enzymes to form heterodimers further increases the number of enzyme species found in cells.

GSTs are not known to have posttranslational modifications other than a modification of the amino terminus in some isozymes which causes blocking when protein sequencing is attempted. In most cases the nature of this modification is not known. However in the rat Alpha class GST 8-8 it has been identified as acetylation of the initiator methionine (Alin et al., 1989). The initiating methionine is removed in all GSTs examined with unmodified amino termini (Widersten et al., 1992).

2. Nomenclature

The nomenclature for the human GSTs has been extremely confusing since many different isozymes have been isolated in a relatively short time period and there have been no consistent rules for naming them to which all investigators in the field adhere. In some cases, the same isozyme subunit has several different names, depending upon the investigator studying it.

Recently a new system for naming human GSTs has been proposed by a group of 14 of the most prolific workers in the field (Mannervik et al., 1992). This system assigns the soluble human human GSTs to each of the four classes, based on the criteria that enzymes within a class must have greater than 50% amino acid sequence homology (Table 1).

Table 1. Nomenclature for the cytosolic human glutathione transferases.

Previous designation	Class	New designation	Gene
ε, B_1B_1, GST2-type 1, H_a(subunit 1), $\alpha_x\alpha_x$	Alpha	GSTA1-1	GSTA1
γ, B_2B_2, GST2-type 2, H_a(subunit 2), $\alpha_y\alpha_y$	Alpha	GSTA2-2	GSTA2
μ, GST1-type2, H_b(subunit 4)	Mu	GSTM1a-1a	GSTM1
ψ, GST1-type1	Mu	GSTM1b-1b	GSTM1
muscle, GST4	Mu	GSTM2-2	GSTM2
brain, GST5	Mu	GSTM3-3	GSTM3
π, GST3	Pi	GSTP1-1	GSTP1
GSTθ	Theta	*	*

* N-terminal amino acid sequences prove the exsistance of the Theta class; individual isozymes will by named once full length sequences are available.
This table was modified from Mannervik et al. (1992) Biochem. J. **282**, 305-308.

The letter designation of the enzymes (A, M, P, and T) indicates to which class it has been assigned. Arabic numerals directly following the letter designating class represent the subunit composition. The subunits are numbered according to the order in which they were described. Allelic variants are represented by a lower case letter directly following the subunit designation, as in GSTM1a. The gene encoding a subunit is named in the same way as the protein subunit. For example, the GSTP1 protein subunit is encoded by the GSTP1 gene. The dimer composition is represented by

the number designating the subunit components separated by a hyphen, as in GSTP1-1.

C. Glutathione *S*-Transferase catalysis

1. Structure of the active site

The GSTs have a binding site for glutathione (G-site) and a binding site for electrophilic substrates (H-site) on each subunit of the dimer. The determination of the three-dimensional structure of a Pi class enzyme isolated from pig lung and, recently, the human GSTP1-1 have provided some structural details of the active site (Reinemer et al., 1991 Reinemer et al., 1992).

It is likely that the basic mechanism by which GSTs activate glutathione, the physiological substrate, at the G-site is the same for all enzymes. Although the enzymes' specificity for GSH is high, the GSH analog homoglutathione (wherein an alanine residue replaces the glycine) can function as a substrate with reduced efficiency (Adang et al., 1988). Two models have been proposed to explain the catalytic mechanism (Jakoby et al., 1978 Mannervik & Danielson, 1988). The first one involves general acid-base catalysis where the GST enhances the nucleophilicity of GSH by providing a base of an appropriate pK_a in the active site for deprotonating the thiol group. Alternatively, the GST's active site could stabilize a thiolate anion (GS-) of bound GSH by lowering the pK_a of its thiol moiety. Direct evidence for the existence of the (GS-) at the active site has been provided by UV absorption spectroscopy studies (Graminski et al., 1989).

The binding site of a competitive inhibitor, S-hexylglutathione, was analyzed making use of the three-dimensional

structure of class Pi GST from human placenta and the location of the G-site was determined. Residues in contact with the GSH part of the inhibitor were determined, most notably Tyr7 which is strictly conserved among all GSTs. The replacement of this residue by any other amino acid results in a marked reduction in enzyme activity (Stenberg et al., 1991b; Kolm et al., 1992; Kong et al., 1992; Lui et al., 1992). Reinemer and coworkers postulate that Tyr7 plays a role in the catalytic mechanism (Reinemer et al., 1992). Interestingly, the three-dimensional structure also reveals a role in the catalytic mechanism for Asp96, the only amino acid from the other subunit to interact with GSH, and substantiates the hypothesis of Dirr and coworkers that an active monomer of GST does not exist (Dirr et al., 1991). There is evidence that the binding in the G-site involves ionic bonds between arginine residues and the carboxyl group of GSH (Schasteen et al., 1983). Site-directed mutagenesis studies on a human Alpha class GST have also shown the importance of conserved arginine residues in the enzyme's affinity for GSH (Stenberg et al., 1991a).

The second substrate binding site, the H-site, exists next to the G-site, but unfortunately the three-dimensional structures so far give little information about which residues are important in binding at this site. The H-site is believed to be hydrophobic because of the nature of the electrophilic substrates and it is assumed that activation of the electrophilic substrate at this site is necessary. An amazing variety of chemical groups can be accommodated by the H-site, however this site can have fairly rigorous stereochemical requirements. For example, the human

GSTP1-1 displays an enantiopreference for (7*S*)-styrene oxide compared to (7*R*)-styrene oxide (Dostal et al., 1988)

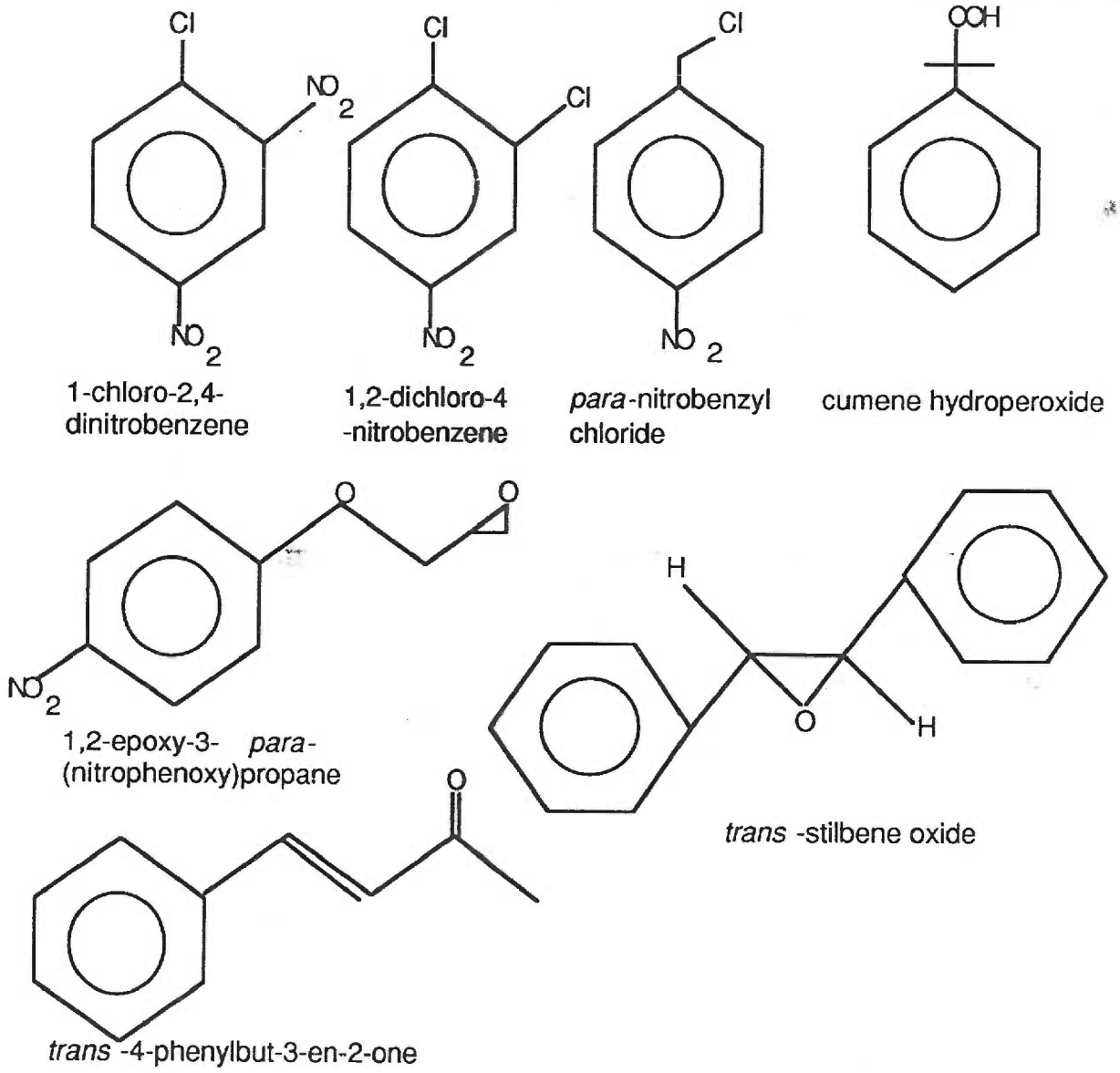


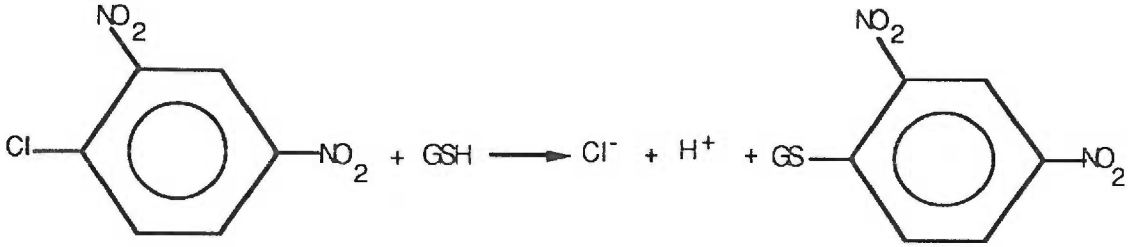
Table 2: Substrates for GSTs *in vitro*

2. Substrates used *in vitro*

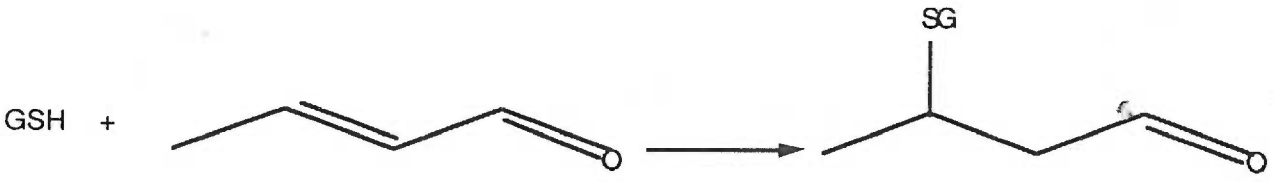
Glutathione *S*-transferases catalyze the conjugation of GSH to substrates with several different types of reactive centers including electrophilic carbon, as well as electrophilic nitrogen, sulfur, and oxygen. To characterize these enzymes, investigators have frequently used substrates which are products of the chemical industry and may have no real biological relevance. Nevertheless, these substrates are readily available and GSH conjugation is usually easily detectable, so they have proven very useful to the study of GSTs. Some examples of these substrates and the types of reactions catalyzed by the GSTs are provided in Table 2 and below.

Perhaps the most useful substrate for glutathione transferases is 1-chloro-2,4 dinitrobenzene (CDNB) which has been used in the demonstration of multiple forms of GSTs (Habig et al., 1974) It has been called a "universal" substrate, but in fact many enzymes have relatively low activity toward CDNB, and the Theta class enzymes lack any activity toward this substrate (Mannervik and Danielson, 1988; Meyer et al., 1991).

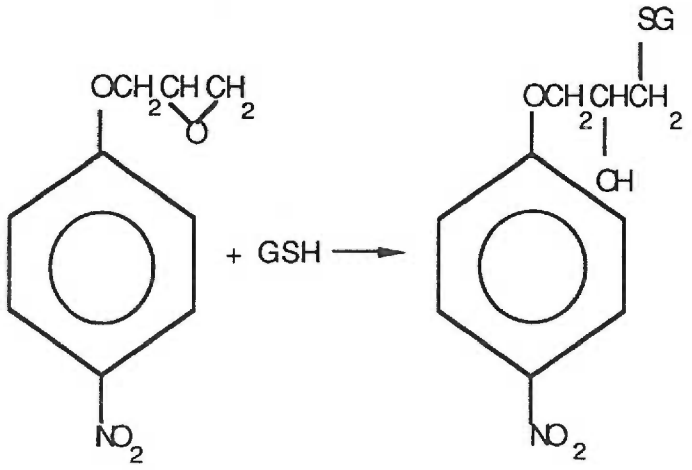
GSTs can catalyze substitutions on aryl and alkylhalides, as demonstrated by the conjugation of CDNB with GSH shown below. Other substrates which demonstrate this type of reaction are 1,2-dichloro-4-nitrobenzene and *para*-nitrobenzylchloride (Habig et al., 1974).



GSTs catalyze the addition to an activated double bond as demonstrated by reactions with ethacrynic acid, *trans*-4-phenylbut-3-en-2-one (Habig et al., 1974) and acrolein (Berhane et al., 1989) shown below.



GSTs also catalyze the addition to arene and alkyl epoxides. Good examples of this type of reaction are provided by reactions with the substrates *trans*-stilbene oxide and 1,2-epoxy-3-(*para*-nitrophenoxy)propane, shown below (Fjellstedt et al., 1973).



3. Biologically relevant substrates

The primary function of the GSTs *in vivo* is believed to be detoxification; however, these enzymes may also have a role in the biosynthesis of certain hormones and as binding proteins. A wide variety of compounds of both endogenous and exogenous origins are biologically relevant substrates of the GSTs.

Epoxides are a group of substrates that are formed *in vivo* by the oxidation of the carbon-carbon double bonds of xenobiotics (Levi, 1987). These compounds can arise from sources such as cigarette smoke and industrial chemicals and are known to be mutagenic and carcinogenic. Many GST isozymes show activity toward these substrates and are thought to be involved in their detoxification. The rat Theta class GST 5-5 was the first pure enzyme shown to have activity toward epoxides (Fjellstedt et al., 1973). The human GSTs GSTM1a-1a, the class Alpha enzymes and GSTP1-1 show activity toward styrene 7,8-oxide, an epoxide which is a metabolite of styrene to which workers in the chemical industry may be exposed (Warholm et al., 1983). GSTA1-1, GSTA2-2 and GSTM1-1 detoxify AFB_1-*exo*-8,9-epoxide, which is formed by the cytochrome P450-catalyzed epoxidation of aflatoxin B_1 (AFB_1), a well known hepatocarcinogen produced by the fungi *Aspergillus flavus* and *A. parasiticus* (Levi, 1987; Raney et al., 1992).

Oxygen free radicals can have a number of deleterious effects on the cell, including the peroxidation of lipids and damage to DNA. The GSTs may play a role in protecting cells from these highly reactive products of oxygen metabolism. For example, the rat Alpha class GST 8-8 has been shown to have activity toward 4-

hydroxyalkenals and acrolein, which are produced during lipid peroxidation (Stenberg et al., 1992). There is also some evidence for a role in the detoxification nucleic acid hydroperoxides and repair of radical damage to DNA. The human GSTP1-1, GSTM1a-1a and rat Alpha class GST 3-3, GST 4-4, and Theta class GST 5-5 show glutathione peroxidase activity toward nucleic acid hydroperoxides (Tan et al., 1985). Tan and coworkers also claim that rat GST 5-5, which shows the highest level of activity toward DNA hydroperoxides of those tested, can be found in the nucleus where it would be assumed to be most effective in the detoxification of these compounds formed *in vivo* .

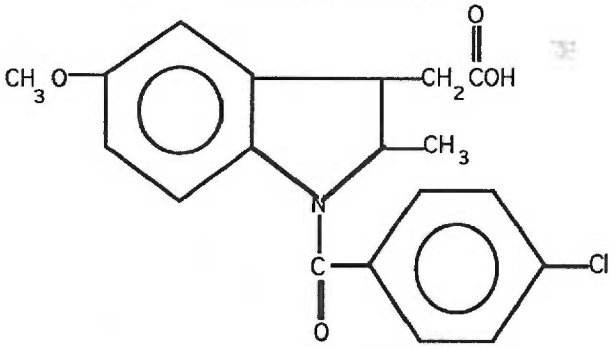
Some GSTs are known to be involved in the biosynthesis of leukotriene C_4 (Samuelsson et al., 1983; Tsuchida et al., 1987) and prostaglandin $F_{2\alpha}$ (Chang et al., 1987; Burgess et al., 1989). The steroid isomerase activity demonstrated by the Δ^5-androstendione assay also suggests a role for GSTs in androgen and estrogen biosynthesis (Warholm et al., 1986; Benson et al., 1977).

The GSTs may also have a role as intracellular binding proteins. Some forms act as binding or carrier proteins for glucocorticoid (Homma et al., 1986) and thyroid (Ishigaki et al., 1989) hormones, neurotransmitters (Abramovitz et al., 1988), hematin (Senjo et al., 1985), bilirubin (Kamisaka et al., 1975; Vander Jagt et al., 1982) and leukotriene C_4 (Sun et al., 1986).

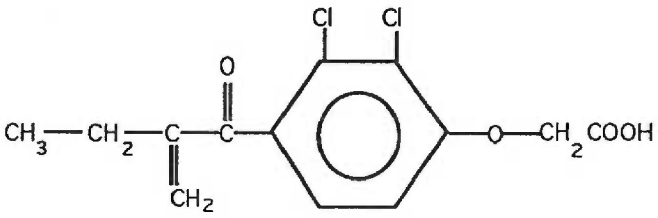
4. Inhibitors

Inhibitors of GSTs are important tools used to study the structure of the active site and the mechanism of catalysis of the GSTs. A number of chemicals have been shown to inhibit GSTs

including GSH derivatives, bile acids, steroid hormone derivatives, anti-inflammatory drugs, diuretic drugs and dyes. Glutathione derivatives which are substrate analogs are useful to probe the structure of the G-site, while product analogs such as glutathionyl-derivatives of electrophilic substrates are useful to probe the H-site. Other inhibitors such as bile acids and hematin bind at a second site, which is distinct from the active site, and may elicit some conformational change which affects catalysis (Hayes et al., 1986). Inhibitors can be used to distinguish individual GST isozymes from mixtures of isozymes which have very similar substrate specificities (Tahir et al., 1985). Simple inhibition studies can also be used to distinguish between homodimeric and heterodimeric isozymes of GSTs (Tahir et al., 1986).



Indomethacin



Ethacrynic acid

In addition, inhibitors may eventually prove useful in reversing GST mediated anticancer drug resistance in human patients. Some known inhibitors of GSTs are the pharmacologically active drugs

ethacrynic acid, a diuretic, and indomethacin, a nonsteroidal anti-inflammatory agent. Several groups have shown that treatment of drug resistant cell lines with known GST inhibitors increases their sensitivity to the drugs (Hall et al., 1989; Hansson et al., 1991). A recent study showed that patients treated with ethacrynic acid had decreased levels of GST activity in peripheral mononuclear cells, which demonstrates the effectiveness of this inhibitor *in vivo* and suggests that further clinical studies are warranted (O'Dwyer, 1991).

D. Role in carcinogenesis

Ultimate carcinogens are strong electrophiles which can form adducts nonenzymatically with nucleophilic oxygen as well as with nitrogen atoms of the DNA bases (Levi, 1987). The modification of DNA by reactive compounds may irreversibly alter it, resulting in a heritable change. While these mutations are not always harmful, they may result in the growth and development of tumor cells. GSTs are thought to be involved primarily in cancer prevention by conjugating GSH to reactive compounds resulting in their ultimate removal from the body (Ketterer, 1988). In addition to their role in detoxification, GSTs may be involved in cancer in several other ways: as markers for preneoplasia, in the activation of procarcinogens and as genetic markers for decreased risk for certain types of cancer.

1. GSTs as markers for preneoplasia in human tissues

During the early stages of the carcinogenic process, groups of preneoplastic cells which show alterations in specific enzymes are observed in organs prior to the appearance of malignant cancers.

Many detoxification enzymes, including the GSTs, may be useful as markers for these preneoplastic cells (Sato et al., 1989). By searching for the presence of marker enzymes, which are expressed at either a higher or lower level in preneoplastic cells than in normal cells, it is possible to detect the presence of these cells at an earlier stage than might otherwise be possible and begin treatment sooner. Numerous reports claim that the presence of increased levels of GSTs, mostly in the Pi class, may be used as a marker for neoplasia and preneoplasia in rats and humans. For example, an anti-GSTP1 antibody might be useful in the early detection of cervical cancer as one study showed that normal cervical mucosa did not bind antibody while severe dysplasia and squamous cell carcinoma of the cervix showed strong staining (Shiratori et al., 1987). Other groups have reported increased levels of GSTP1-1 in colon, breast, lung, uterine, and gastric adenocarcinomas and in melanomas (Kodate et al., 1986; Shea et al., 1988). Another group reported that serum levels of GSTP1-1 were elevated in patients with gastric and esophageal carcinomas, but decreased when tumors were surgically removed. Determining serum levels of GSTP1-1 may be useful in monitoring patients with cancers throughout treatment (Sato et al., 1989).

2. Metabolic activation of procarcinogens

In some cases, GSTs are involved in the activation of relatively innocuous compounds to products which can be carcinogenic or nephrotoxic. Several types of glutathione *S*-conjugates, including *S*-haloakyl and *S*-haloalkenyl glutathiones, are metabolized to the corresponding cysteine *S*-conjugates. When

these conjugates are further metabolized by cysteine conjugate ß-lyase, a 175,000 dalton, pyroxidal phosphate-requiring protein, the product is an unstable reactive thiol. For example, rat Alpha class GST 2-2 and Mu class GST 3-3 catalyze the conjugation of GSH with 1,2-dibromoethane and 1,2-dichloroethane, which are both widely used pesticides. When further metabolized, these compounds bind DNA and have been shown to be carcinogenic in rats (Cmarik et al., 1990) Trichloroethylene, a common industrial chemical used as a degreasing solvent and in drycleaning, may also be conjugated with GSH and metabolized further to the corresponding cysteine conjugate which has been shown to be mutagenic by the Ames test (Anders et al., 1988).

3. GSTM1 deletion as a marker for susceptibility to lung cancer in humans

A deletion of the human Mu class gene GSTM1 is a possible marker for greater susceptibility to lung cancer and certain other cancers among heavy smokers. This gene deletion and associated loss of enzyme activity occurs at a frequency of approximately 50% in Caucasian and Asian populations (Seidegard et al., 1985; Seidegard et al., 1988; Brockmoller et al., 1992). The hereditary deficiency of GSTM1 is associated with increased sensitivity of cells to cytogenetic damage (Wiencke et al., 1990; Van Poppel et al., 1992) and may be associated with a predisposition to lung squamous cell carcinoma, adenocarcinoma of the stomach and colon and bladder carcinoma (Seidegard et al., 1986; Seidegard et al., 1990; Strange et al., 1991; Bell et al., 1992), but not to breast cancer, certain types of colon cancer or polyposis coli (Peters et al., 1990;

Shea et al., 1990). Other studies, however, have failed to confirm the predisposition to lung cancer (Zhong et al., 1991; Heckbert et al., 1992). GSTM1-1 has relatively high activity with certain epoxides, some of which are present in cigarette smoke (Warholm et al., 1983). It is presumed that lack of ability to detoxify these reactive compounds leads to an increased risk of cancer in individuals exposed to them.

E. Relationship to drug resistance

1. Metabolism of anticancer drugs by GSTs

One of the biggest obstacles to the success of cancer therapy is the resistance of malignant cells to anticancer drugs. The overexpression of GSTs have been implicated as a mechanism of resistance to a wide variety of drugs which are used in the treatment of cancer. Specific GST isozymes have been shown to catalyze the inactivation *in vitro* of the commonly used anticancer drugs chlorambucil, cyclophosphamide (Tsuchida, 1992), L-phenylalanine mustard (Bolton et al., 1991; Dulik et al., 1986), and 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU) (Smith et al., 1989).

2. Expression of GSTs in drug resistance

Studies have shown GST isozymes to be elevated in some human and rodent tumor cell lines resistant to certain anticancer drugs and in tumors from patients who show clinical resistance. While the evidence is correlative, the case is stronger given data that GSTs catalyze the inactivation of these drugs *in vitro*. For example, the rat Mu class enzyme GST 4-4 is elevated in rat gliosarcoma cells resistant to BCNU (Evans et al., 1987) and has also been shown to catalyze the inactivation of BCNU *in vitro* by a

denitosation reaction (Smith et al., 1989). Numerous other studies have implicated GSTs in drug resistance to anticancer drugs. Increased expression of a Pi class gene was correlated with the acquisition of resistance to *cis*-platin in a Chinese hamster ovary cell line after exposure to this drug (Saburi et al., 1989). In one report, the amplification of an Alpha class gene in a Chinese hamster ovary cell line was associated with resistance to nitrogen mustards (Lewis et al., 1988). Not surprisingly, in some human tumor cell lines selected for resistance to anticancer drugs, increased levels of specific GST mRNAs have been observed (Wang et al., 1989; Saburi et al., 1989; Kuzmich et al., 1992; Waxman et al., 1992). It has been proposed that resistance to anticancer drugs can be due to induction of specific GST genes by the drug itself.

3. GSTs in multidrug resistance

Some cells have the ability to survive exposure to a wide range of anticancer drugs. This phenomenon is called multidrug resistance (mdr) and has usually been associated with overexpression of the membrane-associated glycoprotein, P-glycoprotein (Moscow et al., 1988). Recently, however, additional mechanisms have been implicated in mdr, such as changes in the levels of phase I and phase II drug metabolizing enzymes including the GSTs. Some investigators claim that mdr can be due to overexpression of GSTs alone. For example, transfection studies in COS cells have shown that the expression of recombinant GSTP1-1 confers resistance to the alkylating agents chlorambucil, melphalan, cisplatin and doxorubicin. Reversal of expression resulted in loss of the mdr phenotype (Puchalski et al., 1990). In contrast, other investigators

believe that elevations of GSTs alone are not enough to confer multidrug resistance. For instance, in a rat model of hepatocarcinogenesis, resistance to anticancer drugs is associated only with both the increased expression of P-glycoprotein and rat Pi class GST 7-7 (Fairchild et al., 1987). In addition, work by another group has shown that overexpression of a Pi class GST by itself does not render a cell resistant to a wide variety of anticancer drugs (Wang et al., 1989). These reports illustrate that mdr is much more complex than originally realized; however the role GSTs play in this phenomenon remains to be resolved.

4. Induction of drug resistance

A tumor which is resistant to cytotoxic drugs can be catagorized as being either inherently resistant, defined as resistance at the time of its malignant transformation, or having acquired resistance after exposure to the drug. Many tumors which are inherently resistant are associated with increased exposure to carcinogens, as is usually seen in lung and colon cancers. It is possible that increased levels of GSTs are induced by these compounds to protect cells which later, unfortunately, have the effect of increased inactivation of anticancer drugs necessary to treat the cancer. The phenomenon of induction will be discussed in detail below.

F. Gene expression

1. Induction by xenobiotics

Numerous studies in rodents have shown that GST enzyme levels can be induced by treatment with a wide variety of chemicals such as butylated hydroxyanisole, phenobarbital, 3-

methylcholanthrene, ß-naphthoflavone, and even coffee beans, green tea and broccoli extracts. Several of these chemicals have been shown to function as "anticarcinogens", that is, treatment imparts a protective effect against subsequent assault with carcinogens (Sparnins et al., 1982; Khan et al., 1992). In several cases, the increased levels of GST enzymes following exposure to certain chemicals are due to increases in transcription and/or mRNA stability (Ding et al., 1985; Davidson et al., 1990; Vandenberghe et al., 1991). Two cis-acting elements have been identified in the rat GST Ya (subunit 1) gene which are essential for induction (Rushmore et al., 1990). One of these, the xenobiotic responsive element (XRE), has been found in multiple copies in the cytochrome P-450la1 gene which is also induced by many chemicals. The other element, the antioxidant responsive element (ARE), is responsive to phenolic antioxidants, planar aromatic compounds and hydrogen peroxide (Rushmore et al., 1991). A functional ARE is also found in the mouse GST Ya gene (Friling et al., 1990) and the rat NAD(P)H:quinone reductase gene (Favreau et al., 1991). The presence of a transcription factor(s) which binds to the ARE in the rat GST Ya gene has been shown recently by DNAse I footprinting and gel mobility shift experiments by Nguyen and coworkers (Nguyen et al., 1992). Although the induction of GSTs in response to chemical stimuli has been studied in detail in rodents, similar studies of human GSTs were lacking until recently when a report claimed that Alpha class transcripts were preferentially induced by several chemicals in cultured human hepatocytes (Morel, 1993).

2. Tissue specific expression

Glutathione *S*-transferases are expressed in a highly tissue specific manner in all species. Several human enzymes of the Alpha class and GSTM1-1 have been observed in liver at relatively high concentrations (Tsuchida et al., 1987) while GSTP1-1 is expressed as a major form in many organs other than liver (Kano et al., 1987). Various levels of Pi, Mu and Alpha class GSTs have been observed in kidney, lung, brain, skin, intestine, adrenal gland, testis, prostate, uterus, heart, blood vessels, and/or skeletal muscle (Mannervik et al., 1985). The expression of GSTs is not only tissue specific, but also differentially regulated during development in the liver (Hiley et al., 1988; Strange et al., 1989). The mechanism of tissue specific regulation of GST genes is not yet known. The elucidation of normal means of regulation of expression of these genes will aid in the understanding of altered expression in drug resistant tumors.

G. Human Mu class GSTs

The human Mu class GSTs, which are the focus of this dissertation, consist of at least four isozyme subunits, GSTM1, GSTM2, GSTM3 and the recently discovered GSTM4 (see Manuscript #2) each of which is coded for by a distinct gene. It is likely there are at least two more genes and/or pseudogenes predicted by Southern blotting analysis using a probe which would be expected to hybridize to genes in this class (DeJong, 1991). There is some evidence for heterodimer formation among the members of this class which further increases the number of enzyme species found in tissues (Tsuchida et al., 1990 Hussey et al., 1991). These isozymes have highly homologous amino acid sequences (70-87% identity), but

distinct substrate specificities and tissue distributions. Because of their high degree of nucleotide sequence homology, it has been proposed that gene conversion occurred in the evolution of these genes (Taylor et al., 1991).

The known human Mu class enzymes also show a high degree of identity at the amino acid and nucleotide level with other mammalian Mu class enzymes isolated from rats, mice, pigs, hamsters and guinea pigs. Of the fifteen mammalian Mu class enzymes known, all but one have the same number of amino acids (218). All the gene and partial gene sequences known for rats, hamsters and humans in this class have identical intron-exon boundaries (Taylor et al., 1991; Lai et al., 1988; Fan et al., 1992).

GST M1a-1a is the most well characterized of the human Mu class enzymes. The GSTM1 gene locus is polymorphic and displays an allelic form, GSTM1b-1b, which differs from GSTM1a-1a by one amino acid (Singh et al., 1987). Both forms appear to be identical in their catalytic abilities and substrate specificities, but have differences in their isoelectric points. The GSTM1 gene locus shows an additional form, the null allele. The GSTM1 cDNA has been isolated and was used to probe a genomic Southern blot of DNA from several individuals to show that the null allele is due to a gene deletion (Seidegard et al., 1988). GSTM1a-1a and GSTM1b-1b have a relatively high activity with certain epoxides, some of which are present in cigarette smoke (Warholm et al., 1983). It is the deletion of the GSTM1 gene and resulting lack of enzymatic activity which has been correlated with an increased risk of certain cancers among smokers (see above). There is also some evidence for the

involvement of this enzyme in drug resistance as GSTM1a-1a has been reported to be elevated in glioma cells resistant to (2-chloroethyl)-3-sarcosinamide-1-nitrosourea (Skalski et al., 1990)

GSTM2-2 was first isolated from skeletal muscle and the 24 N-terminal amino acids have been determined (Suzuki et al., 1987 Board et al., 1988). The corresponding cDNA sequence for this enzyme has been recently reported and shares 94% nucleotide sequence identity with GSTM1 cDNA (Vorachek et al., 1991). An interesting feature of GSTM2 is its extremely high degree of sequence identity with GSTM1 in certain regions of the mRNA: in the first 152 nucleotides of the coding region there is only one silent substitution and there are only three differences in 450 nucleotides of the 3' untranslated region.

GSTM3-3 was recently isolated from human brain and testis (Campbell et al., 1990). The cDNA for this enzyme was also reported and examination of the sequence revealed that GSTM3 is the least structurally conserved among the human Mu class. While the nucleotide sequence identity between GSTM3 and the other Mu class cDNAs in the coding region is as much as 69%, the identity between these cDNAs and GSTM3 in the 3' untranslated region is negligible. Also, GSTM3 has some unique features in its primary structure not found in other mammalian Mu class enzymes: a blocked and extended amino terminus and three additional residues at the carboxyl terminus.

A new member of the human GST gene superfamily, GSTM4, has recently been discovered. We have determined both the genomic DNA and the corresponding cDNA sequences and also characterized the

tissue specificity of expression and properties of the recombinant enzyme. This data will be discussed in detail in manuscripts #2 and #3.

H. Thesis objectives

Humans are exposed to an increasing variety of toxic compounds. Many of these compounds are reactive against DNA which can cause mutations and may ultimately lead to cancer, the second leading cause of death among people in the industrialized world. Fortunately enzyme detoxification systems exist which can help protect us against this fate. Identification and study of the enzymes involved in detoxification of potential carcinogens can aid in determining safe exposure limits to toxic compounds. The glutathione *S*-transferases are one such detoxification system, involved primarily in the inactivation and removal of reactive electrophiles. Although many enzymes from this multigene family have been isolated and characterized in humans, it is likely there are still enzymes left undiscovered which may be important toxicologically and in drug resistance.

The focus of my thesis project has been the characterization of human Mu class GSTs. This class of enzymes was of interest to me in particular because of their possible involvement in drug resistance, the correlation between a deletion of the Mu class gene GSTM1 and lung and other cancers, and because it was suspected that there were members of this family which had not yet been identified.

It had been previously determined that a possible genetic predisposition to lung cancer is due to the deletion of the GSTM1 gene. Epidemiological studies had been done by testing individuals for GSTM1-1 activity with the substrate *trans*-stilbene oxide.

However this assay is not trivial and required the isolation of a fairly large number of lymphocytes from the subjects. Testing for the lack of GSTM1-1 could be done by Western blotting or ELISA, however these are relatively difficult assays, require a relatively large number of cells and could be complicated by the cross-reaction of anti-GSTM1 antisera with the other Mu class enzymes known to exist in tissues other than lymphocytes. What was needed was an assay which could be quickly and easily performed, had a high degree of accuracy and could be done with relatively small amounts of blood samples. This was accomplished by designing a PCR-based assay which could be used to study the hereditary predisposition to cancer.

At the time when I began working in this area it was suspected that several members of the Mu class existed, but only one enzyme had been characterized in any detail. Nothing much was known about the other enzymes which also might play important roles in detoxification and drug resistance. Only one human Mu class cDNA had been reported, the GSTM1a cDNA, but when this cDNA was used as a probe on genomic Southern blots, several bands hybridized which suggested that other closely related genes were present in the human genome. A rat Mu class gene Yb2 had been reported, but no human genomic sequences were available. As well as to identify new members of the human GST Mu class, one of my objectives was to study the regulation of the Mu class gene(s) by analyzing the promoter region. An understanding of the control of expression of these genes could lead to a better understanding of mechanisms of increased expression during drug resistance and ways of manipulating the levels of Mu class enzymes in tumors to overcome

drug resistance. To do this it is necessary to isolate and sequence clones containing these gene(s).

The major portion of my thesis project has been the isolation and characterization of a new Mu class GST gene which has been named GSTM4. I have studied GSTM4 at the level of the gene, mRNA and protein including its enzymatic properties. The following manuscripts describe investigations which were conducted to further characterize the human Mu class glutathione *S*-transferases and contribute to an understanding of their role in detoxification.

II. COLLECTION OF PAPERS

Manuscript #1:

GST1 gene deletion determined by polymerase chain reaction.

Kenine E. Comstock, Barbara J. S. Sanderson, Ginger Claflin, and W. David Henner

Division of Hematology and Medical Oncology, Oregon Health Sciences University, Portland OR 97201, USA

Submitted February 9, 1990
Published in Nucleic Acids Research, vol. 18, no. 12, 3670.

Results and Discussion

The human glutathione transferase (GST) isozyme GSTμ (renamed GSTM1-1 in 1992) is frequently deficient (31-61%) in individuals (1). The absence of GSTμ has been ascribed to a homozygous deletion of the gene GSTM1 (2) and correlates with an increased risk of lung cancer among smokers (3). We developed a simple assay based on PCR technology to determine the presence or absence of GSTM1, which may be useful in studies of the role of GSTμ in lung cancer risk and drug resistance. Primers were designed based on the cDNA sequence for human GSTμ (2) and on the intron/exon boundaries of the related rat gene Yb2 (4). The primers hybridize to the 5' region of exon 4 (5'-CTGCCCTACTTGATTGATGGG-3') and the 3' region of exon 5 (5'-CTGGATTGTAGCAGATCATGC-3') of GST1. Genomic DNA was prepared from human blood lymphocytes by standard procedures. A commercial PCR kit containing TAQ DNA polymerase (GeneAmp, US Biochemical Corp.) was used with 1 μg DNA. Reactions were heated for 2 min at 94°C, 1 min at 55°C, 1.5 min at 72°C for 35 cycles in an Ericomp Thermal Cycler. PCR products were electrophoresed on a 2.1% agarose gel. The 273 bp product, present in positive reactions, was identified by sequence analysis as GST1 intron 4, a previously unsequenced region (Fig. 3), flanked by exon 4 and exon 5 (EMBL Accession no. X51451). The presence or absence of GSTμ was determined in 10 cell lysates by Western blotting and ELISA (Mukit, Medlabs, Dublin, Ireland). All Western blot/Mukit positive samples contained the GST1 gene when analyzed by PCR assay. Western blot/Mukit negative samples lacked

the gene (6/6) (Fig. 4). Eighteen samples of human breast carcinoma also had a one to one correspondence between the presence of the 273 bp PCR product and the presence of GSTμ by Western blotting (not shown). This assay is an efficient alternative to other methods for determining GSTμ status.

References

1. Seidegard, J., DePierre, J., Pero, R. (1985) Carcinogenesis **6**, 1211-1216.

2. Seidegard, J., Vorachek, W., Pero, R., Pearson, W. (1988) Proc. Natl. Acad.,Sci. USA **85**, 7293-7297.

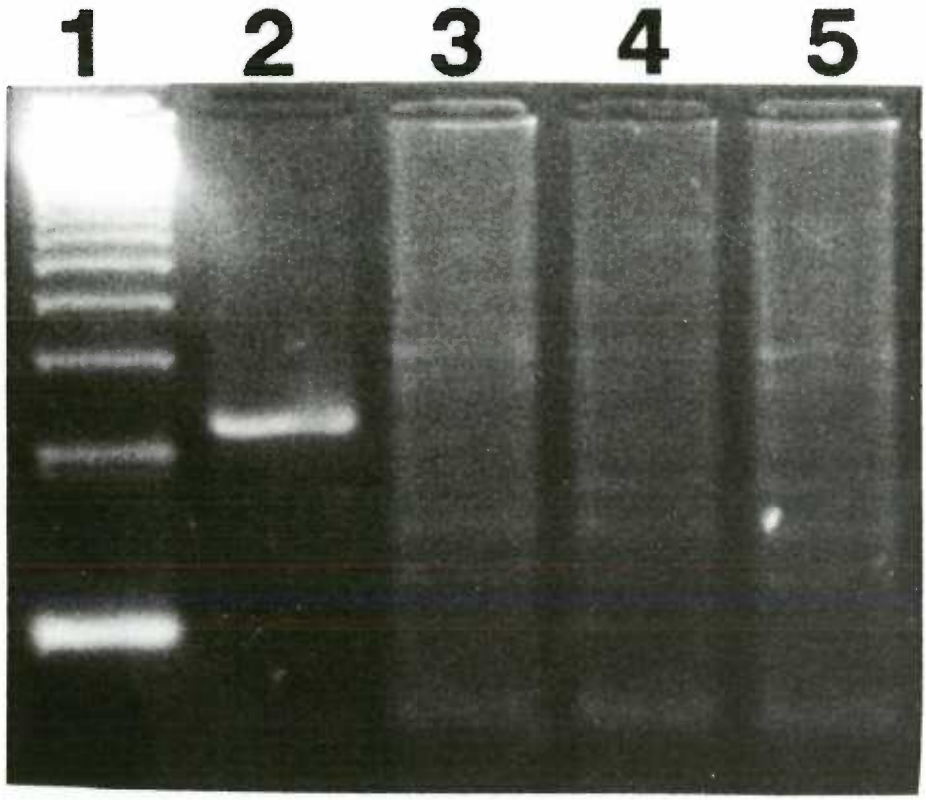
3. Seidegard, J., Pero, R., Miller, D., Beattie, E. (1986) Carcinogenesis **7**, 751-753.

4. Lai, H.-C., Qian, B., Grove, G., Tu, C.-P. (1988) J. Biol. Chem. **263**, 11389-11395.

```
1     GTG AGT GTG GGT GGC TGC AAT GTG TAG GGG GAA GGT     35
37    GGC CTC CTC CTT GGC TGG GCT GTG ATG CTG AGA TTG     72
73    AGT CTG TGT TTT GTG GGC AG
```

Figure 3. Nucleotide sequence of GSTM1 intron 4 (splice site at nucleotide 276 of the human cDNA (2)).

Figure 4. Lane 1: 123 bp ladder. Lane 2: Positive sample, Lanes 3-5: Negative samples.


1
2
3
4
5

Manuscript #2:

Isolation and Analysis of the Gene and cDNA for a Human Mu Class Glutathione S-Transferase, GSTM4.

Kenine E. Comstock[1], Kara J. Johnson[2], Dean Rifenbery[2], and W. David Henner[2]

From the [1]Department of Biochemistry and Molecular Biology and [2]Division of Hematology and Medical Oncology, Oregon Health Sciences University, Portland, OR.

Submitted July 7, 1992, to the Journal of Biological Chemistry.

*This research was supported by a grant (BM 100-R153) from Bristol-Myers-Squibb. K.E.C. was supported in part by a Tatar fellowship from the Medical Research Foundation of Oregon. The nucleotides sequences reported in this paper have been submitted to the GenBank data bank and assigned accession numbers M96233 and M96234.

Abstract

We have isolated the human gene and cDNA that encode a new member of the glutathione S-transferase (GST) multigene family. The complete gene sequence (GenBank Accession No. M96233) and the cDNA sequence (No. M96234) for this fourth member of the Mu class of human GSTs are reported and are named GSTM4. The gene is comprised of eight exons with an organization similar to that of the rat Mu class GST, Yb2. The cDNA for GSTM4 was isolated from a library derived from the cervical carcinoma cell line, HeLa. Northern blots of human RNA, probed with a unique 5' region of the GSTM4 cDNA, demonstrate the presence of GSTM4 message in human heart, placenta, lung, brain, liver, skeletal muscle, pancreas, testis, cerebral cortex, uterus, ovary, a lymphoblastoid cell line, and four carcinoma cell lines, including HeLa. The deduced amino acid sequence of GSTM4 is 87% (GSTM1), 83% (GSTM2) and 70% (GSTM3) identical to the previously described human Mu class GSTs. The polypeptide encoded by the GSTM4 cDNA, purified from overexpressing E. coli, exhibits the characteristic activity of many GSTs in conjugating 1-chloro,2,4-dinitrobenzene to GSH. The GSTM4 gene encodes a human GST expressed in many human carcinoma cell lines and in many human tissues.

Introduction

Many detoxification enzymes exist as isozymes with overlapping substrate specificities that are encoded by a family of multiple genes. The presence of multiple, closely related gene products provides a broad substrate specificity and thereby allows detoxification of a wide variety of xenobiotics. In addition, the multiplicity of gene promoters allows the organism great flexibility in the regulation of expression in specific tissues, during development, and in response to inducing agents. The GSTs[3] are one such multigene family of related enzymes that detoxify a variety of electrophilic xenobiotics by conjugating them to GSH (reviewed in Ketterer, 1988, Mannervik, 1985). In all animal species studied, multiple GST isozymes have been identified (Mannervik, 1985). The human and other mammalian GSTs have been grouped into four classes of cytosolic isozymes (Alpha, Pi, Theta and Mu) and a microsomal isozyme, based on their amino acid sequence homologies, immunocrossreactivity and substrate specificities (Mannervik, 1985, Meyer et al., 1991). The recently proposed nomenclature for the human GSTs, which is based on the above classes of Alpha, Pi, Mu and Theta, as well as information regarding nucleotide sequences of the cDNAs or genes that encode individual GSTs, has been adopted in this report (Mannervik et al., 1992).

---

[3] Abbreviations used are: GST, glutathione S-transferase; PCR, polymerase chain reaction; RT-PCR, reverse transcriptase-PCR; IPTG, isopropyl-β-D-thiogalactopyranoside; CDNB, 1-chloro-2,4-dinitrobenzene; ARE, antioxidant response element; GRE, glucocorticoid response element, kb, kilobase; SSC, 0.15M NaCl, 15 mM $Na_3citrate \cdot 2H_2O$ (pH 7.0).

The previously described human GSTs of the Mu class, GSTM1, GSTM2 and GSTM3 are the product of three gene loci, but share a high level of amino acid sequence homology. The GSTM1 locus displays three allelic forms. The gene product of one of these alleles, GSTM1a-1a, was first isolated from liver (Warholm et al., 1980) and has a high level of activity with epoxides as a substrate (Warholm et al., 1983). Another GSTM1 allele is GSTM1b, whose gene product differs by only one amino acid from GSTM1a. The third allelic form of GSTM1 is a null allele, whose phenotype is a lack of the enzyme activity characteristic of GSTM1a-1a or GSTM1b-1b. A substantial portion of human livers and lymphocytes display this phenotype of deficiency of GSTM1 activity (Seidegard et al., 1985, Warholm et al., 1980). When the cDNA sequence for GSTM1a was isolated (DeJong et al., 1988, Seidegard et al., 1988), it was shown that the GSTM1 null phenotype was due to a homozygous deletion of the GSTM1 gene (previously termed GST1) in about 50% of the population (Seidegard et al, 1988, Comstock et al., 1990, Zhong et al., 1991). Lymphocytes from patients with GSTM1 deficiency are more sensitive to epoxide-induced cytogenetic damage (Wiencke et al., 1990) and lymphocytes from smokers with GSTM1 deficiency have higher levels of cytogenetic damage (van Poppel al., 1992). There have been reports that homozygous deletion of GSTM1 may predispose cigarette smokers to lung (Seidegard et al., 1986) and bladder[4] cancer, but other studies have failed to confirm the predisposition to lung cancer (Zhong et al., 1991, Heckbert et al., 1992). Additional human

---

[4] Bell, D. A., Taylor, J. A., Paulson, D. F., Robertson, C. N., Mohler, J. L., Miller, C. R., and Lucier G. W. (1992) Printed in abstract form in the *Proceedings of the American Association for Cancer Research* **33,** 291.

Mu class GSTs for which cDNA sequences are published include GSTM2-2, whose deduced amino acid sequence matches the 24 N-terminal amino acids of a Mu class GST expressed in muscle (Board et al, 1988, Vorachek et al., 1991) and GSTM3-3, identified in human testis and brain (Campbell, et al.,1990).

On the basis of the multiplicity of hybridizing bands on Southern blots probed with Mu class cDNAs (Seidegard et al., 1988, DeJong et al., 1988, DeJong et al., 1991), we anticipated the existence of one or more additional human Mu class GST isozymes. In this communication, we report the isolation and characterization of a human Mu class gene and its mRNA and we show that it encodes a catalytically active GST.

Materials and Methods

Molecular cloning- A human genomic library in λFIX (Stratagene) was a gift of Dr. J. Adelman, Oregon Health Sciences Univ. A λZAPII (Short et al., 1988) HeLa cDNA library was purchased from Stratagene. A plasmid containing the insert from clone λGTH (coding for GSTM1 cDNA) in a pEMBL-19 vector was a gift of Dr. W. Pearson, Univ. of Virginia. A plasmid containing the region of the GSTM1 gene from exon 4 through exon 5 in a pTZ18U vector has been previously described (Comstock et al., 1990). Hybridization probes were uniformly labeled with [α-32P]dCTP by random hexanucleotide priming and bacteriophage plaques were screened by hybridization to labeled probe, both by standard techniques (Sambrook et al., 1989). The final post-hybridization wash was at 0.5 X SSC, 0.1% SDS at 25°C for one h for genomic screening and 0.1 X SSC, 0.1% SDS at 60°C for one h for cDNA screening. Following plaque purification by secondary and tertiary rounds of screening, purified λ DNA was restriction-digested and the insert DNA subcloned into pBluescript (Stratagene) or pTZ18U (Mead et al., 1985). In the case of λZAPII phage, inserts were subcloned into pBluescript by *in vivo* recombination (Short et al., 1988). DNA sequencing was performed by the dideoxy method (Sanger et al., 1977) using double-stranded plasmid DNA and Sequenase 2.0 (US Biochemical Corp.). DNA sequence analysis was performed using the MacVector program (IBI).

Chromosome assignment- Chromosomal location was determined by PCR amplification of GSTM4 DNA from a human/rodent hybrid cell panel, using a method analogous to that described by

Dionne and coworkers (1990). 10 ng each of the DNAs from from the NIGMS human/rodent somatic cell hybrid mapping panel no. 1 was added to a 50 µl reaction mix containing a forward primer (5'-GATCTGGCTTTACTCTCACG-3') and reverse primer (5'-AACAGGCGCGGGAGGCACGGAGCAG-3') at 1 µM each, 10 mM Tris-Cl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 1 mM deoxynucleoside triphosphates, and 1.25 U Taq DNA polymerase (U.S. Biochemical Corp.). DNA was amplified by incubating 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C for 35 cycles in a Coy thermal cycler. DNA products were resolved on a 2% agarose gel containing ethidium bromide. The NIGMS human/rodent somatic cell hybrid mapping panel no. 1 was a gift from Dr. Michael Litt, O.H.S.U.

Reverse transcriptase-PCR- Total RNA (1 µg), isolated by standard techniques (Ausubel et al, 1987), was added to a 20 µl reaction mixture containing 10 mM Tris-Cl (pH 8.3), 50 mM KCl, 4 mM $MgCl_2$, 0.01% gelatin, 1 mM deoxynucleoside triphosphates, 42.5 pmol random hexanucleotides, 40 U RNasin (Promega), 200 U M-MLV reverse transcriptase (Gibco BRL) and incubated at 37° C for 30 min. A portion (5 µl) of the reaction was then added to a 50 ml reaction mixture containing the forward (CTGCCCTACTTGATTGATGGG) and reverse (CAGGGCTGTAGCAGACTCTGG) oligonucleotide primers (1 µM each), 10 mM Tris-Cl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, and 1.25 U Taq DNA polymerase (U.S. Biochemical Corp.). DNA was amplified by incubating 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C for 35 cycles in a Coy thermal cycler. DNA products were resolved on a 2% agarose gel containing ethidium bromide.

Restriction Enzyme Digestion of PCR products- Human genomic DNA was isolated by standard techniques (Ausubel et al., 1987). 1 μg of genomic DNA was added to a 50 μl reaction mix containing a forward primer (5'- GGTGCAGTGCAGTCTAGACT-3') and reverse primer (5'-GGATCTAGAATCTGACCACTC-3') at 1 μM each, 10 mM Tris-Cl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 1mM deoxynucleoside triphosphates, and 1.25 U Taq DNA polymerase (U.S. Biochemical Corp.). DNA was amplified by incubating 1 min at 94°C, 1 min at 52°C, and 1 min at 72°C for 35 cycles in a Coy thermal cycler. DNA products were ethanol precipitated, then taken up in an appropriate buffer and digested overnight with 10U HhaI. The resulting DNA fragments were resolved on a 2% agarose gel containing ethidium bromide.

Genomic Southern blotting- High molecular weight human genomic DNA was isolated by standard techniques (Ausubel et al., 1987). Restriction digests were set up containing 10 μg DNA and an excess of restriction enzyme. Samples were resolved on a 0.7% agarose gel containing ethidium bromide. A 1 kb ladder and HindIII-digested λ DNA (Gibco BRL) were used as size markers. DNA was transferred to a nylon membrane, immobilized and probed with a labeled DNA fragment. The final post hybridization wash was in 0.5 X SSC, 0.1% SDS at 50°C for one h. The blot was exposed to autoradiography film one week at -70° C.

Northern blotting- Total RNA or polyadenylated RNA, isolated by standard techniques (Ausubel et al, 1987), was resolved by agarose-formaldehyde gel electrophoresis, transferred to a nylon filter and immobilized by a standard procedure (Sambrook et al.,

1989). A human multiple tissue Northern blot was purchased from Clontech Laboratories. The blot is a charge modified nylon membrane to which polyadenylated mRNA isolated from eight different human tissues have been transferred from a denaturing formaldehyde agarose gel. Blots were probed by hybridization to a radiolabeled DNA fragment by a standard technique. A 2 kb fragment of the β-actin cDNA was obtained from Clontech (Cleveland et al., 1980). The final post-hybridization wash was in 0.5 X SSC, 0.1% SDS at 60°C for one h. The blot was exposed to autoradiography film for one week at -70° C.

Cell Culture- HeLa, MCF-7 and WIL2-NS cells were obtained from American Tissue Culture Collection. HepG2 and SK-Hep-1 cells were gifts from Dr. Thomas Deloughery and Dr. John Ansel, Oregon Health Sciences Univ. respectively. All cell lines were maintained in monolayer culture except WIL2-NS, which was maintained in suspension. HepG2, SK-Hep-1 and MCF-7 were cultured in DMEM with 10% calf serum. HeLa and WIL2-NS cells were cultured in RPMI-1640 media with 10% calf serum (JRH Biosciences, Lenexa, KS).

GST Purification- The following procedures were performed at 4°C unless noted. HeLa cells were grown to confluency and harvested with a rubber cell scraper in Dulbecco's phosphate buffered saline, washed once in the same, and stored at -70°C. A total of 1.3 x 10^8 cells were thawed, homogenized in a Dounce homogenizer with 7 ml 22 mM sodium phosphate, (pH 7.0), and sonicated for 20 s using a Branson Model 200 sonifier to ensure complete lysis. The volume was brought to 60 ml with 22 mM sodium phosphate (pH 7.0), and the samples were centrifuged for 30

min at 15,600 x *g*. The supernatant was applied to a GSH-Sepharose affinity column (0.5 ml bed volume) prepared from GSH and epoxy-activated Sepharose beads by the method of Simons and Van der Jagt (1981) and the column washed with 5 ml of 22 mM sodium phosphate, (pH 7.0). The GSH binding proteins were eluted from the column with 10 bed volumes 5 mM reduced gluthathione (Sigma) in 100 mM Tris-HCl, (pH 9.6). Fractions (0.5 ml) were collected and assayed for protein using the BioRad protein assay (Bradford, 1976). The GSH affinity-purified proteins were analyzed by SDS-PAGE (Laemmli, 1970) and by Western blot analysis (Towbin et al., 1979). Antisera to GSTP1 (Shea et al., 1988) and GSTM1 (Shea et al., 1990) were prepared as previously described. Proteins were transferred to Immobilon-P membrane (Millipore) for sequence analysis (Matsudaira, 1987).

Expression in E. Coli- The coding region of the GSTM4 cDNA sequences was amplified by PCR using a reaction mixture that contained a forward primer containing an NdeI site (5'- ACA CCA ACC AGC CAT ATG TCC ATG -3'), a reverse primer containing a BamHI site (5'- CTC ACT CCC GGA TCC TGG CCT TCA -3'), plasmid pKC101 (250 ng) and XmnI-digested plasmid pKC104 (250 ng). The DNA product was resolved by agarose gel electrophoresis, digested with NdeI and BamHI and ligated into NdeI- and BamH1-digested pET11c plasmid vector (Studier et al., 1990). Constructs were propagated in the E. coli strain DH5α then transferred into the IPTG-inducible strain BL21(DE3)pLysS. For production of GSTM4 in E. coli, 250 ml of cells were grown in Luria-Bertani medium. GSTM4 synthesis was induced when cells reached an OD_{600} of 0.6 - 1.0 by adding 1 mM

IPTG. Cells were grown 3 h, then pelleted by centrifugation and resuspended in 20 ml of 22 mM sodium phosphate pH 7.0. Cells were sonicated on ice in 2 - 3 second bursts and cellular debris was pelleted by centrifugation at 4000 x *g* for 20 min. GSTM4 was purified from the supernatant by GSH-Sepharose chromatography as described above. GST activity with CDNB as a substrate was determined as described by Habig and coworkers (1974).

Results

Cloning and sequence of a human mu class genomic DNA- Approximately 45,000 plaques from a λFIX human genomic library were screened by hybridization to a previously described 275 base pair fragment of the GSTM1 gene extending from exon 4 through exon 5 (Comstock et al., 1990). Five positive clones were identified, one of which, λQ7, was purified and analyzed in detail. Preliminary analysis, obtained by sequencing of PCR products generated with GSTM1 specific primers, established that λQ7 contained sequences similar to, but distinct from, GSTM1, GSTM2 or GSTM3. Insert DNA from λQ7 was restriction-digested, subcloned and the sequence of both strands obtained by the dideoxy method using a combination of strategies of staggered deletions and sequence "walking" with multiple synthetic oligonucleotide primers. The one exception to these strategies was the region of 147 base pairs between the two EcoRI sites in intron 5, which was obtained by sequencing of two cloned PCR products that spanned this region. The sequences of both clones were identical. The sequence of the gene, which we have termed GSTM4, consists of 6082 base pairs, and is shown in Figure 5. This sequence has been deposited in the GenBank DNA sequence data bank as Accession number M96233.

Analysis of GSTM4 gene sequence - The exon/intron boundaries of GSTM4 were first deduced by comparison to the sequence of the rat mu class gene Yb2 (Lai et al., 1988), then confirmed by the cDNA sequence (see below). The organization of the GSTM4 gene is shown in Figure 6. The sequence obtained

includes 477 nucleotides upstream from the most 5' nucleotide of the cDNA sequence and eight exons. The location of the exon/intron boundaries and the size of the introns are very similar in the GSTM4 and rat Yb2 genes. The nucleotide sequence of the presumed coding region of GSTM4 is highly homologous to the coding region of the cDNA sequence of GSTM1 (93%), GSTM2 (88%), GSTM3 (69%), and rat Yb1 (77%), thus establishing GSTM4 as a Mu class GST gene. Comparison of the sequence of the GSTM4 introns and upstream sequences with the corresponding regions of the rat Yb2 gene revealed no significant homology between the intron sequences of GSTM4 and Yb2, except for the 99 bp intron 4, where there is 71% identity between the GSTM4 and Yb2, and a region of approximately ten nucleotides near each splice junction (Figure 7).

In the course of analysis of the λQ7 clone, a report of a partial gene sequence for a human Mu class GST gene appeared (Taylor et al., 1991). This partial gene sequence, which Taylor and coworkers termed GSTmu2, differs from the sequence of GSTM4 in exons 3 through 5 by only 11 nucleotides within introns. Thus, it is highly likely that the GSTmu2 and GSTM4 sequences are from the same gene.

The exact initiation site (or sites) for transcription of the GSTM4 gene have not yet been determined. However, based on the size of the mRNA for GSTM4 (see below), transcription is likely to be initiated at or near to the 5' end of the cDNA sequence isolated. An analysis of the GSTM4 gene sequences upstream from the cDNA sequence (see below) identified several features of potential significance. The sequences upstream from the cDNA sequence have

many of the characteristics of promoters usually associated with "housekeeping" genes including the absence of a TATA box, a region of high (80%) G+C content, and several consensus Sp1 binding sites (Boyer et al., 1989 , Boyer et al., 1990). There is also a sequence at positions 115 to 126 that matches the consensus sequence (5'-PuGTGACNNNGC-3') for an ARE (Rushmore et al., 1991). Analysis of GSTM4 intron sequences identified the presence of a $(GT)_{23}$ microsatellite in intron 1. The consensus sequence for a GRE (5'-TGTTCT-3') was identified in intron 3 at nucleotide position 1775-1780, and was also present in the GSTmu2 sequence as reported by Taylor and coworkers (1991).

Chromosomal location of the GSTM4 gene- Using PCR amplification of a unique region of the GSTM4 gene (see below) from a panel of rodent-human somatic cell hybrid DNAs, we have determined the chromosomal location of the GSTM4 gene. A primer pair was selected which specifically amplified a 217 bp GSTM4 DNA product from human DNA, but did not amplify a similarly sized product when mouse or hamster DNAs were used as a target. All of the four hybrid DNAs that contained chromosome 1 and none of the fourteen hybrid DNAs that lacked it produced a PCR product of the expected size. None of the other chromosomes were concordant. Thus, the GSTM4 gene is located on chromosome 1.

Identification of HeLa cells as containing GSTM4 mRNA and a Mu class GST protein - In order to further investigate the properties of the GSTM4 gene, its expression and putative protein product, we wished to obtain the cDNA for this gene. To facilitate the cloning of the GSTM4 cDNA, we wished to

identify a human cell line or tissue that contained the mRNA for GSTM4. RT-PCR was used to identify the presence of GSTM4 mRNA in the human carcinoma cell line HeLa. Following cDNA synthesis by reverse transcription of randomly primed HeLa cell total RNA, PCR was used to specifically amplify GSTM4 sequences. A reverse primer which is complementary to a region in the putative exon 5 of GSTM4 and whose sequence differs from that of the other known human mu class GST cDNAs was paired with a forward primer complementary to a region in exon 4 that is identical among GSTM1, GSTM2 and GSTM4. The product of the PCR with these primers is of the expected 175 bp and was identified as the product of RT-PCR of GSTM4 mRNA by restriction enzyme digestion and sequencing. The sequence of GSTM4 but not of GSTM1, GSTM2 or GSTM3, contains a BstNI restriction site. BstNI completely cleaved the 175 bp product to the expected 122 and 53 bp fragments. In addition, when the 175 bp product was subcloned and sequenced, the nucleotide sequence of two independent clones matched exactly the nucleotide sequence of the putative exon 4 and and exon 5 of GSTM4 (data not shown).

Because the presence of a Mu class GST in HeLa cells had not been previously reported, and because HeLa cell extracts contain very low levels of total GST activity using CDNB as a substrate (< 5 mU/mg protein), we wished to confirm that HeLa cells did indeed express a Mu class GST. Therefore, we isolated GSTs from HeLa cells by GSH affinity chromatography and analyzed them by SDS-PAGE and by Western blotting (Figure 8). Analysis of HeLa proteins, purified by GSH-affinity chromatography, shows two bands on SDS-PAGE, one at ~25 kilodaltons which reacts on Western blotting with

anti-GST Pi class antisera, raised against GSTP1-1, and a second, much less prominent, band at ~27.5 kilodaltons which reacts with anti-GST Mu class antisera, raised against GSTM1-1. Attempts to microsequence the small amount of Mu class GST protein following blotting onto PVDF were unsuccessful, most likely due to a blocked amino terminus. Therefore, on the basis of these qualitative results, we can conclude that HeLa cells contain GSTM4 mRNA and that HeLa cells express a Mu class GST protein. However, whether GSTM4 is the only Mu class GST protein in HeLa cells can not be determined.

Cloning and sequence analysis of the GSTM4 cDNA - On the basis of the above results, we chose to screen a HeLa cell cDNA library for the GSTM4 cDNA. One million phage of a HeLa cell cDNA library in λZAPII were screened by filter hybridization to the GSTM1 cDNA. Two phage were isolated and subcloned into pBluescript to yield plasmids pKC101 and pKC104. Sequencing of these plasmids revealed that each is a partial but overlapping clone of the GSTM4 cDNA (Figure 6). The nucleotide sequence of the cDNA has been deposited in the GenBank as Accession number M96234. The combined sequences of pKC101 and pKC104 contain a single large open reading frame of 654 base pairs coding for a deduced amino acid sequence of 218 amino acids (Figure 9). The nucleotide sequence of the GSTM4 cDNA clones differed at only 4 positions relative to the corresponding positions in the GSTM4 gene. Of the nucleotide differences, only one led to a change in deduced amino acid sequence, a change at amino acid position 17 from methionine in the gene to isoleucine in the cDNA. This difference in sequence

was confirmed by demonstrating the expected presence and absence of a HhaI restriction site in the gene and cDNA clones, respectively. To determine whether this difference represents a genetic polymorphism or a cloning artifact, we used PCR to amplify this region of genomic DNA from six individuals, including the individual whose DNA had been used as the source for the λFIX library from which Q7 was isolated. While the PCR product from Q7 contained the HhaI restriction site, none of the PCR products from the six donors contained a HhaI site. Therefore, it appears that the sequence coding for Met in the Q7 isolate represents a mutation that occurred during cloning and that the correct amino acid at this position is isoleucine.

Further analysis of the cDNA sequence revealed several interesting features. The 5' untranslated region of the GSTM4 cDNA is considerably larger (246 base pairs) than the 5' untranslated regions of the GSTM1, GSTM2 or GSTM3 cDNA sequences and contains an exact 28 base pair direct repeat. The most 3' region of the GSTM4 mRNA sequence is presumably missing from the isolated clones because neither a poly-A sequence nor a polyadenylation signal were present. The likely polyadenylation signal and polyadenylation site for this message are indicated on the gene sequence and correspond quite closely to the position of the GSTM1 and GSTM2 sites.

Comparison of deduced amino acid sequences of several Mu class cDNAs- The amino acid sequence of GSTM4, deduced from the cDNA and the deduced amino acid sequences of the other human Mu class GST cDNAs are shown in Figure 10. The amino acid sequence of GSTM4 is highly homologous to the deduced amino

acid sequences of the other human Mu class GSTs but is clearly distinct from them and the amino terminal amino acid sequences determined directly from purified human Mu class GST proteins (Suzuki et al., 1991, Tsuchida et al., 1990, Hussey et al., 1991, Board et al., 1988). However, several isolated Mu class proteins have also been reported to be blocked at the amino terminus (Hussey et al., 1991, Campbell et al.,1990). The human Mu class glutathione transferases show a high degree of amino acid sequence homology. There is an overall amino acid sequence identity of 87% between GSTM4 and GSTM1, 83% between GSTM4 and GSTM2, and 70% between GSTM4 and GSTM3. Also, there are long stretches of complete amino acid sequence identity among the four proteins.

GSTM4 mRNA levels in human tumor cell lines and tissues- We wished to determine the relative level of GSTM4 mRNA in various human cell lines and tissues. Because of the high degree of nucleotide sequence homology among the Mu class cDNAs, hybridization probes consisting of one cDNA sequence will also hybridize to mRNA of the other Mu class GSTs, even when washes are performed under very stringent conditions. The 5' untranslated region of the GSTM4 cDNA, however, was not homologous to any region in the other known Mu class GST cDNAs and therefore was likely to serve as specific probe. To ensure that this sequence could not be present in any other message, we determined that it is present as a single copy in the human genome by Southern blotting. A 260 bp HindIII-PstI fragment of pKC101 (including positions 478 to 719 on the gene sequence) was subcloned into pBluescript (plasmid pKC122), and the insert used to probe a human genomic

Southern blot (Figure 11). Only a single hybridizing band was observed in the EcoRI and PstI lanes. Two bands were observed in the BamHI lane, which was expected since there is a BamHI restriction site within the probe. Thus this region of GSTM4 is present as only a single copy in the genome with no highly homologous regions in other genes.

Since this 5' region of GSTM4 is unique, it could be used as a specific probe for GSTM4 mRNA. A Northern blot of total RNA hybridized with this probe is shown in Figure 12 and demonstrates the presence of an approximately 1.5 to 1.6 kb GSTM4 mRNA in the human adenocarcinoma cell line SK-Hep-1, the hepatoma line HepG2, the cervical carcinoma line HeLa, the breast carcinoma line MCF-7, and the human lymphoblastoid cell line WIL2-NS. Northern blots of polyadenylated mRNA from human tissues revealed expression of a 1.5-1.6 kb GSTM4 mRNA in human heart, brain, liver, skeletal muscle, pancreas, placenta, lung and kidney (Figure 13). In other experiments, GSTM4 mRNA was also detected on Northern blots of polyadenylated mRNA from testis, cerebral cortex, uterus and ovary (data not shown).

Expression of GSTM4 in E. coli- GSTM4 cDNA sequences were cloned into the pET11c expression vector as described in Methods to yield plasmid pKC114. After confirming that the sequence of the open reading frame was the same as that of the cDNA clones, bacterial cultures containing pKC114 were IPTG-induced, lysed and the GSH-binding proteins purified by affinity chromatography. A single protein band on SDS-PAGE chromatography of approximately 27.5 kilodaltons was obtained which reacted with

anti-Mu class GST anti-sera on Western blotting (data not shown). This protein, GSTM4-4, catalyzes conjugation of the standard GST substrate CDNB to GSH with a specific activity of 1.39 $\pm$ 0.21 U/mg protein. Therefore the GSTM4 cDNA sequence encodes a polypeptide which has a characteristic activity of a GST and the immunoreactivity of a Mu class GST. Further biochemical characterization of GSTM4-4 will appear in a subsequent report.

Discussion

This report provides the first complete gene sequence for a human Mu class GST and identifies a new member of this class, GSTM4. Although GSTM4 mRNA, and presumably the protein, are present in a variety of human cells and tissues, GSTM4 does not appear to have been previously characterized. Perhaps this is because the message and the protein are present at only very low levels in the cell lines and tissues examined. Even in HeLa cells, which have the highest relative amount of GSTM4 mRNA observed, the GSTM4 protein can be estimated to comprise less than 0.04% of the cytosolic protein. In addition, the relatively low specific activity of the GSTM4 protein for CDNB, the most commonly used substrate for the GSTs, may have contributed to this enzyme having previously escaped detection.

While the GSTM4 cDNA and amino acid sequence are clearly distinct from those of other Mu class GSTs, the degree of amino acid sequence homology with other human Mu class GSTs is quite striking. In view of the high degree of sequence conservation for the GSTM1, GSTM2, GSTM3 and GSTM4 isozymes, it is interesting to speculate as to the evolutionary pressures that select for conservation of these multiple, very similar forms. Perhaps there are important, as yet unrecognized, differences in substrate specificity among these isozymes, such that a particular form is required for a crucial detoxification of a xenobiotic or for an important step in an endogenous metabolic pathway. Alternatively, the differences in substrate specificity among the Mu class GSTs may be trivial and relatively unimportant, and it may be that the selection pressure is

for maintenance of multiple genes with multiple promoters, thereby allowing the organism to express the Mu class GSTs in a manner that is highly tissue-specific, developmentally-specific and responsive to inducing agents.

The identification of a portion of the 5' untranslated region of GSTM4 as a unique sequence has allowed us to determine the tissue-specific mRNA levels for GSTM4. The pattern of GSTM4 mRNA levels, while distinct, overlaps that of other Mu class GSTs. GSTM4 is expressed in liver, lung and kidney, as is GSTM1 (Warholm et al., 1980, Beckett et al., 1990). GSTM4 is expressed in skeletal muscle, as is GSTM2, and in brain and testis, as is GSTM3 (Board et al., 1988, Vorachek et al., 1991, Campbell et al., 1990). Factors which determine the tissue-specific expression of GSTM4 and other human mu class GSTs have not yet been determined.

In addition to tissue-specific expression, some of the rat and mouse GSTs are subject to induction of increased transcription in response to xenobiotic exposure (Ding and Pickett, 1985, Pearson et al., 1988, Davidson et al., 1990, Vandenberghe et al., 1991). In some cases, induction renders a rodent less susceptible to chemical carcinogens (Sparnins et al., 1982, McLellen et al., 1991). While human GSTs have not yet been demonstrated to be inducible by xenobiotics, we note the presence of an ARE in the upstream sequences of the GSTM4 gene. An ARE is a cis-acting, regulatory element shown previously to be inducible by metabolites of planar aromatic compounds and phenolic antioxidants. Identified in the rat glutathione transferase Ya subunit gene and NADP(H):quinone reductase gene, this element responds to stimuli even in the absence

of a functional Ah receptor and is active in human HepG2 cells (Rushmore et al., 1991). The presence of an ARE in the 5' flanking region of the GSTM4 gene raises the possibility that the GSTM4 gene may also be inducible by compounds such as β-naphthoflavone and 3-methylcholanthrene as is the rat Ya gene. Experiments to determine the functionality of this ARE sequence, as well as the GRE, in the GSTM4 gene are underway.

The multiplicity of Mu class GST genes and their high degree of nucleotide sequence homology have complicated previous attempts to determine the chromosomal location of the human Mu class GSTs (DeJong et al, 1988, Islam et al., 1989). As discussed above, the nucleotide sequence of GSTM4 matches the partial sequence of a cosmid clone (GSTmu2) reported by Taylor and coworkers and is likely to be the same gene locus (Taylor et al., 1991). Taylor and coworkers also reported the partial sequence of an additional Mu class GST gene from the same cosmid clone which they designated GSTmu3. Based on the sequence identity between GSTmu3 and GSTM2 cDNA sequence, these genes are likely to be identical. If this is correct, GSTM4 is closely linked to GSTM2, and since our data indicate GSTM4 is located on chromosome 1, GSTM2 is also on chromosome 1. This could indicate that at least two Mu class genes are clustered on chromosome 1. Experiments to further define the map position of the GSTM4 locus are underway.

Acknowledgment

We thank J. Hagerup for excellent technical assistance and Drs. Mike Litt and John Scott for helpful advice.

References

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith J. A., Seidman, J. G., and Struhl, K. (1987) *Current Protocols in Molecular Biology.* 1st ed. Wiley Interscience, New York, NY.

Beckett, G. J., Howie, A. F., Hume, R., Matharoo, B., Hiley, C., Jones, P., and Strange, R. C. (1990) *Biochem. Biophys. Acta* **1036**, 176-182.

Board, P. G., Suzuki, T., and Shaw, D. C. (1988) *Biochem. Biophys. Acta* **953**, 214-217.

Boyer, T. G., Krug, J. R., and Maquat, L. E. (1989) *J. Biol. Chem.* **264**, 5177-5187.

Boyer, T. G., and Maquat, L. E. (1990) *J. Biol. Chem.* **265**, 20524-20532.

Bradford, M. A. (1976) *Anal. Biochem.* **72**, 248-254.

Campbell, E., Takahashi, Y., Abramovitz, M., Peretz, M., and Listowsky, I. (1990) *J. Biol. Chem.* **265**, 9188-9193.

Cleveland D. W., Lopata, M. A., MacDonald, R. J., Cowan, N. J., Rutter, W. J., and Kirschner, M. W. (1980) *Cell* **20**, 95-105.

Comstock, K. E., Sanderson, B. J. S., Claflin, G., and Henner, W. D. (1990) *Nucleic Acids Res.* **18**, 3670.

Davidson, N. E., Egner, P.A., and Kensler, T. W. (1990) *Cancer Res.* **50**, 2251-2255.

DeJong, J. L., Chang, C.-M., Whang-Peng, J., Knutsen, T., and Tu, C.-H. D. (1988) *Nucleic Acids Res.* **16**, 8541-8554.

DeJong, J. L., Mohandas, T., and Tu, C.-P. D. (1991) *Biochem. Biophys. Res. Commun.* **180**, 15-22.

Dionne, C. A., Kaplan, R., Seuanez, H., O'Brien S. J., and Jaye, M. (1990) *Biotechniques* **8**, 190-194.

Ding, V. D.-H. and Pickett, C. B., (1985) *Arch. Biochem. Biophys.* **240**, 553-559.

Habig, W. H., Pabst, M. J., and Jakoby, W. B. (1974) *J. Biol. Chem.* **249**, 7130-7139.

Heckbert, S. R., Weiss, N. S., Hornung, S. K., Eaton, D. L., and Motulsky, A. G. (1992) *JNCI* **84**, 414-422.

Hussey, A. J., Kerr, I. A., Cronshaw, A. D., Harrison, D. J., Hayes, J. D., (1991) *Biochem. J.* **273**, 323-332.

Islam, Q. M., Platz, A., Szpirer, J., Szpirer, C., Levan, G., and Mannervik, B. (1989) *Hum. Genet.* **82**, 338-342.

Ketterer, B. (1988) *Mutation Res.* **202**, 343-361.

Lai, H.-C. J., Qian, B., Grove, G., and Tu, C.-P. D. (1988) *J. Biol. Chem.* **263**, 11389-11395.

Laemmli, U. K., (1970) *Nature* **227**, 680-685.

Mannervik, B., (1985) *Adv. Enzymol. Relat. Areas Mol. Biol.* **57**, 357-417.

Mannervik, B., Awasthi, Y. C., Board, P.G., Hayes, J. D., Dillio C., Ketterer, B., Listowski, I., Morgenstern, R., Muramatsu, M., Pearson, W. R., Pickett, C. B., Sato, K., Widersten, M., and Wolf, C. R. (1992) *Biochem. J.* **282**, 305-308.

Matsudaira, P. (1987) *J. Biol. Chem.* **262**, 10035-10038.

McLellen, L. I., Kerr, L. A., Cronshaw, A. D., and Hayes, J. D. (1991) *Biochem J.* **276**, 461-469.

Mead, D. A., Skorupa, E. S., and Kemper, B. (1985) *Nucl. Acids Res.* **13**, 1103-1118.

Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M., and Ketterer, B. (1991) *Biochem. J.* **274**, 409-414.

Pearson, W. R., Reinhart, J., Sisk, S. C., Anderson, K., and Adler, P. N. (1988) *J. Biol. Chem.* **263**, 13324-13332.

Rushmore T. H., Morton, M. R., and Pickett, C. B. (1991) *J. Biol. Chem.* **266**, 11632-11639.

Sambrook, J., Fritsch, E. F., and Maniatis, T., (1989) *Molecular Cloning: A Laboratory Manual*, 2nd E., Cold Spring Harbor Laboratory, Cold Spring NY.

Sanger, F., Nicklen, S., and Coulson, A. R., (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Seidegard, J., and Pero, R. W. (1985) *Hum. Genet.* **69**, 66-68.

Seidegard, J., Pero, R. W., Miller, D. G., and Beattie, E. J. (1986) *Carcinogenesis* **7**, 751-753.

Seidegard, J., Vorachek, W. R., Pero, R. W., and Pearson, W. R., (1988) *Proc. Natl. Acad. Sci.* **85**, 7293-7297.

Shea, T. C., Kelley, S. L., and Henner, W. D., (1988) *Cancer Res.* **48**, 527-533.

Shea, T. C., Claflin, G., Comstock, K. E., Sanderson, B. J. S., Burstein, N. A., Keenan, E. J., Mannervik, B., and Henner, W.D. (1990) *Cancer Res.* **50**, 6848-6853.

Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. (1988) *Nucleic Acids Res.* **16**, 7583-7600.

Simons, P. C., and Van der Jagt, D. L. (1981) *Methods Enzymol.* **77**, 235-237.

Sparnins, V. L., Venegas, P. L., Wattenberg, L. W. (1982) *JNCI* **68**, 493-496.

Studier, W. F., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60-89.

Suzuki, T., Shaw, D. C., and Board, P. (1991) *Biochem. J.* **274**, 405-408.

Taylor, J. B., Oliver, J., Sherrington, R., and Pemble, S. E. (1991) *Biochem. J.* **274**, 587-593.

Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350-4354.

Tsuchida, S., Maki, T., and Sato, K. (1990) *J. Biol. Chem.* **265**, 7150-7157

Vandenberghe, Y., Tee, L., Morel, F., Rogiers, V., Guillouzo, A., and Yeoh, G. (1991) *FEBS Letters* **284**, 103-108.

van Poppel, G., de Vogel, N., van Balderen, P. J., and Kok, F. J. (1992) *Carcinogenesis* **13**, 303-305.

Vorachek, W. R., Pearson, W. R., and Rule, G. S. (1991) *Proc. Natl. Acad. Sci.* **88**, 4443-4447.

Warholm, M., Guthenberg, C., Mannervik, B., von Bahr, C., Glaumann, H. (1980) *Acta Chem. Scand.,Ser. B* **B 34**, 607-621.

Warholm M., Guthenberg, C., and Mannervik, B. (1983) *Biochemistry* **22**, 3610-3617.

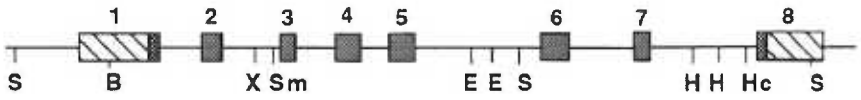
Wiencke, J. K., Kelsey, K. T., Lamela, R. A., and Toscano, W. A. (1990) *Cancer Res.* **50**, 1585-1590.

Zhong, S., Howie, A. F., Ketterer, B., Taylor, J., Hayes, J. D., Beckett, G. J., Wathen, C. G., Wolf, C. R., and Spurr N. K. (1991) *Carcinogenesis* **12**, 1533-1537.

Figure 5. Nucleotide sequence of the GSTM4 gene. Introns are indicated by lowercase letters, exons are indicated by uppercase letters. The translation initiation codon ATG is located at nucleotides 741-743. The amino acids encoded by the exons are listed below the nucleotide sequence. An ARE consensus sequence is underlined. Two overlapping Sp1 consensus sequences are shown in bold. The stop codon for translation is indicated by asterisks. The polyadenylation signal sequence is italicized and underlined.

Figure 6. Map of the GSTM4 gene (A) and cDNA (B). Exons are indicated by boxed regions that are either stippled (coding regions) or crosshatched (untranslated regions). Letters indicate the position of restriction sites for SacI (S), BamHI (B), XbaI (X), SmaI (Sm), EcoRI (E), HindIII (H), or HincII (Hc).

A.



B.

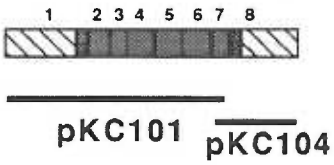


Figure 7. Comparison of the intron 4 nucleotide sequences of GSTM4 and the rat Mu class gene Yb2.

```
                2028        2038        2048        2058        2068
                  *           *           *           *           *
GSTM4     GTGAGTGTGGTTGGCTGCAGTGTGTGGGGGGAAGGTGGCATCCTCCTTGG
Yb2       .......G..C..A......G...G..ACA--..AA.C.....CT.....

                2078        2088        2098        2108        2118
                  *           *           *           *           *
GSTM4     CTGGATTGGGGTGCTATGCTCAGAGTGAGTCTGTGTTTTGTGGGTGGCAG
Yb2       ..T.GC...A.CAGG.....G......G.........G....T.CT....
```

Figure 8. Analysis of GSH-affinity purified HeLa protein extracts. (A) Coomassie blue stained SDS-PAGE of HeLa protein extract after GSH-affinity chromatography, lane 1; and total rat liver GST (Sigma), lane 2. An arrow indicates the position of the protein recognized by anti-Mu class GST antisera. A 12% polyacrylamide gel was used. (B) Western blots of GSH-affinity purified HeLa protein extract probed with anti-Pi class GST antisera, lane 1; or probed with anti-Mu class GST antisera, lane 2; and total rat liver GST probed with anti-Mu class GST antisera, lane 3.

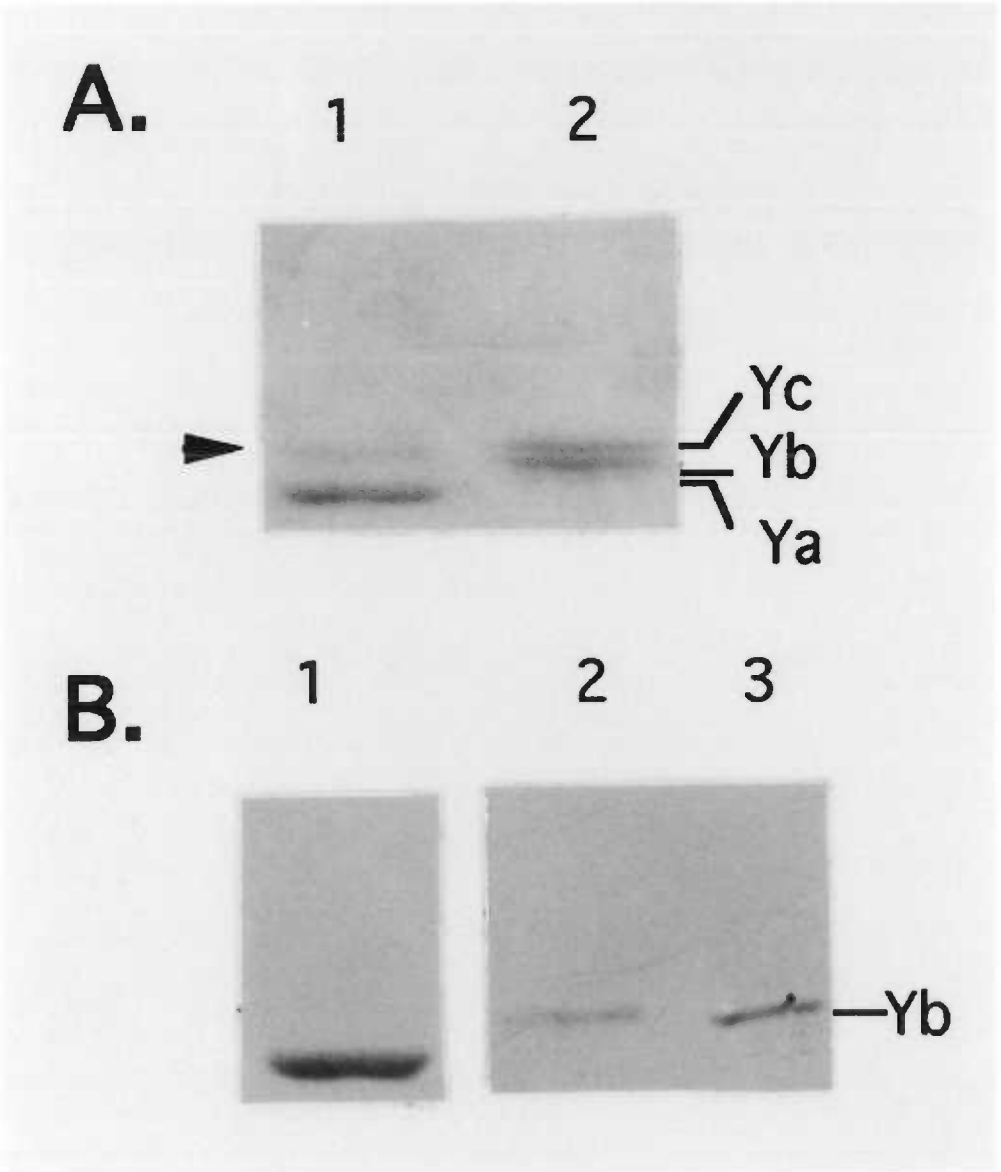

A.
1
2
Yc
Yb
Ya
B.
1
2
3
Yb

Figure 9. Nucleotide sequence of the overlapping GSTM4 cDNA clones pKC101 and pKC104. Coding regions are indicated by uppercase letters. The translation initiation codon ATG is position 264 to 266. The stop codon is indicated by asterisks. Nucleotide sequence differences between the gene and cDNA clones are shown in bold.

```
0001   cactaggaggtggcggtggatcttactccttccagccagtgaggatccag
0051   caacctgctccgtgcctcccgcgcctgttggttggaagtgacgaccttga
0101   agatcggccggttggaagtgacgaccttgaagatcggcgggcgcagcggg
0151   gccgaggggggcgggtctggcgctaggtccagcccctgcgtgccgggaacc
0201   ccagaggaggtcgcagttcagcccagctgaggcctgtctgcagaatcgac
0251   accaaccagcatcATGTCCATGACACTGGGGTACTGGGACATCCGCGGGC
                        MetSerMetThrLeuGlyTyrTrpAspIleArgGlyL

0301   TGGCCCACGCCATCCGCCTGCTCCTGGAATACACAGACTCAAGCTACGAG
       euAlaHisAlaIleArgLeuLeuLeuGluTyrThrAspSerSerTyrGlu

0351   GAAAAGAAGTATACGATGGGGGACGCTCCTGACTATGACAGAAGCCAGTG
       GluLysLysTyrThrMetGlyAspAlaProAspTyrAspArgSerGlnTr

0401   GCTGAATGAAAAATTCAAGCTGGGCCTAGACTTTCCCAATCTGCCCTACT
       pLeuAsnGluLysPheLysLeuGlyLeuAspPheProAsnLeuProTyrL

0451   TGATTGATGGGGCTCACAAGATCACCCAGAGCAACGCCATCCTGTGCTAC
       euIleAspGlyAlaHisLysIleThrGlnSerAsnAlaIleLeuCysTyr

0501   ATTGCCCGCAAGCACAACCTGTGTGGGGAGACAGAAGAGGAGAAGATTCG
       IleAlaArgLysHisAsnLeuCysGlyGluThrGluGluGluLysIleAr

0551   TGTGGACATTTTGGAGAACCAGGCTATGGACGTCTCCAATCAGCTGGCCA
       gValAspIleLeuGluAsnGlnAlaMetAspValSerAsnGlnLeuAlaA

0601   GAGTCTGCTACAGCCCTGACTTTGAGAAACTGAAGCCAGAATACTTGGAG
       rgValCysTyrSerProAspPheGluLysLeuLysProGluTyrLeuGlu

0651   GAACTTCCTACAATGATGCAGCACTTCTCACAGTTCCTGGGGAAGAGGCC
       GluLeuProThrMetMetGlnHisPheSerGlnPheLeuGlyLysArgPr

0701   ATGGTTTGTTGGAGACAAGATCACCTTTGTAGATTTCCTCGCCTATGATG
       oTrpPheValGlyAspLysIleThrPheValAspPheLeuAlaTyrAspV

0751   TCCTTGACCTCCACCGTATATTTGAGCCCAACTGCTTGGACGCCTTTCCA
       alLeuAspLeuHisArgIlePheGluProAsnCysLeuAspAlaPhePro

0801   AATCTGAAGGACTTCATCTCCCGCTTTGAGGGCTTGGAGAAGATCTCTGC
       AsnLeuLysAspPheIleSerArgPheGluGlyLeuGluLysIleSerAl

0851   CTACATGAAGTCCAGCCGCTTCCTCCCAAAACCTCTGTACACAAGGGTGG
       aTyrMetLysSerSerArgPheLeuProLysProLeuTyrThrArgValA

0901   CTGTCTGGGGCAACAAGTAAtgccttgaaggccaggaggtgggagtgagg
       laValTrpGlyAsnLys***                                .

0951   agcccatactcagcctgctgcccaggctgtgcagcgcagctggactctgc
1001   atcccagcacctgcctcctcgttcctttctcctgtttattcccatcttta
1051   cccccaagactttattgggcctcttcacttc
```

Figure 10. Comparison of the deduced amino acid sequences of human Mu class cDNAs.

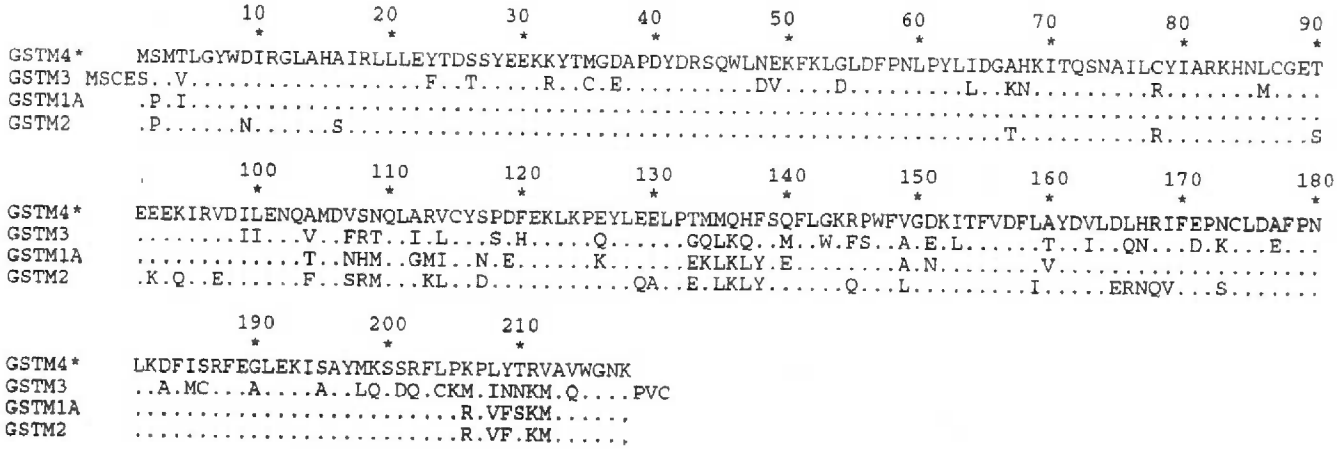

10 20 30 40 50 60 70 80 90
GSTM4* MSMTLGYWDIRGLAHAIRLLLEYTDSSYEEKKYTMGDAPDYDRSQWLNEKFKLGLDFPNLPYLIDGAHKITQSNAILCYIARKHNLCGET
GSTM3 MSCES..V.....F..T.....R..C.E.....DV....D.....L..KN.....R.....M....
GSTM1A .P.I.....
GSTM2 .P.....N.....S.....T.....R.....S
100 110 120 130 140 150 160 170 180
GSTM4* EEEKIRVDILENQAMDVSNQLARVCYSPDFEKLKPEYLEELPTMMQHFSQFLGKRPWFVGDKITFVDFLAYDVLDLHRIFEPNCLDAFPN
GSTM3 .....II...V..FRT..I.L...S.H.....Q.....GQLKQ..M..W.FS..A.E.L.....T..I..QN...D.K...E...
GSTM1A .....T..NHM..GMI..N.E.....K.....EKLKLY.E.....A.N.....V.....
GSTM2 .K.Q..E.....F..SRM...KL..D.....QA..E.LKLY.....Q...L.....I.....ERNQV...S.....
190 200 210
GSTM4* LKDFISRFEGLEKISAYMKSSRFLPKPLYTRVAVWGNK
GSTM3 ..A.MC...A....A..LQ.DQ.CKM.INNKM.Q....PVC
GSTM1A .....R.VFSKM.....,
GSTM2 .....R.VF.KM.....,

Figure 11. Human genomic Southern blot hybridized to radiolabeled pKC122, which contains a portion of the GSTM4 5' untranslated region. DNA was digested with BamHI, lane 1; EcoRI, lane 2; and PstI, lane 3. The units of the ordinate are kb.

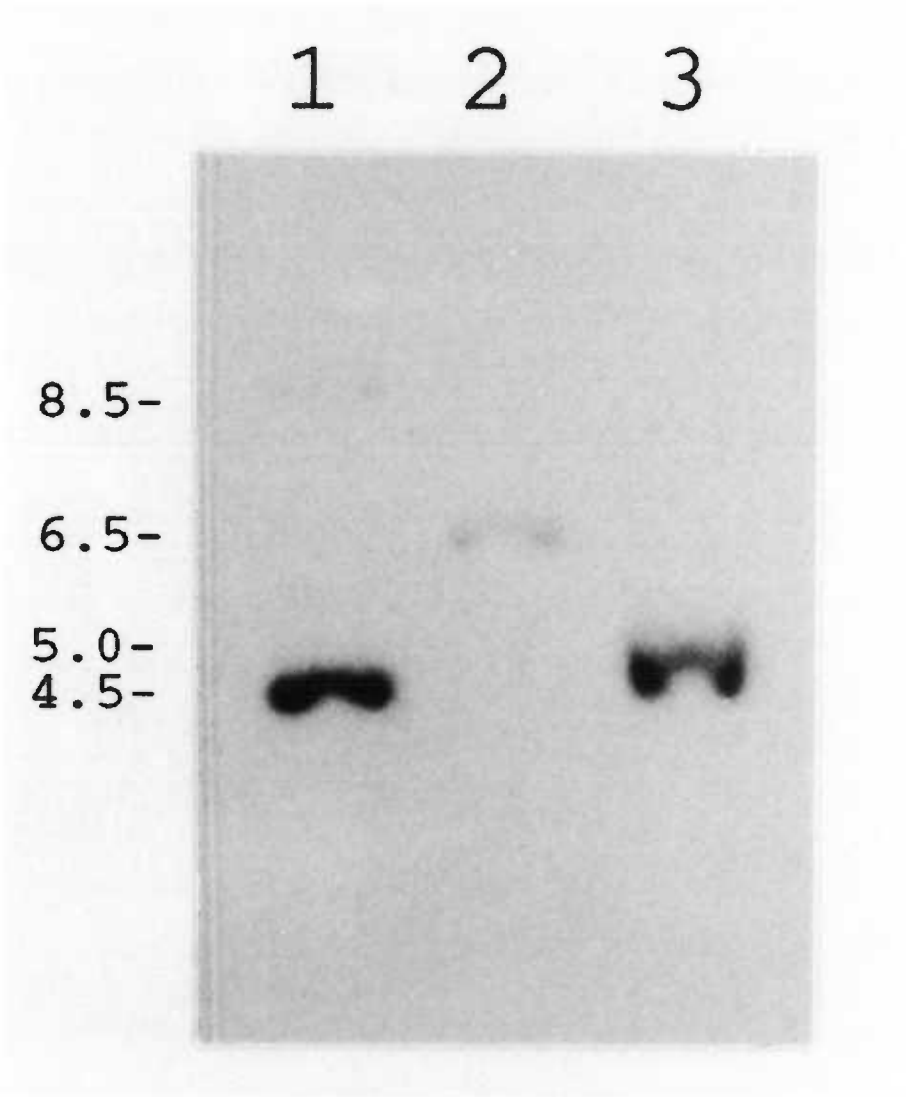

1
2
3
8.5-
6.5-
5.0-
4.5-

Figure 12. Northern blot of total RNA from human cell lines. The blot was hybridized with radiolabeled pKC122 (A), and equivalency of loading was determined by RNA visualization under ultraviolet light after intercalation of ethidium bromide (B). Each lane was loaded with 10 μg RNA from SK-Hep-1, lane 1; HepG2, lane 2; HeLa, lane 3; MCF-7, lane 4; WIL2-NS, lane 5.

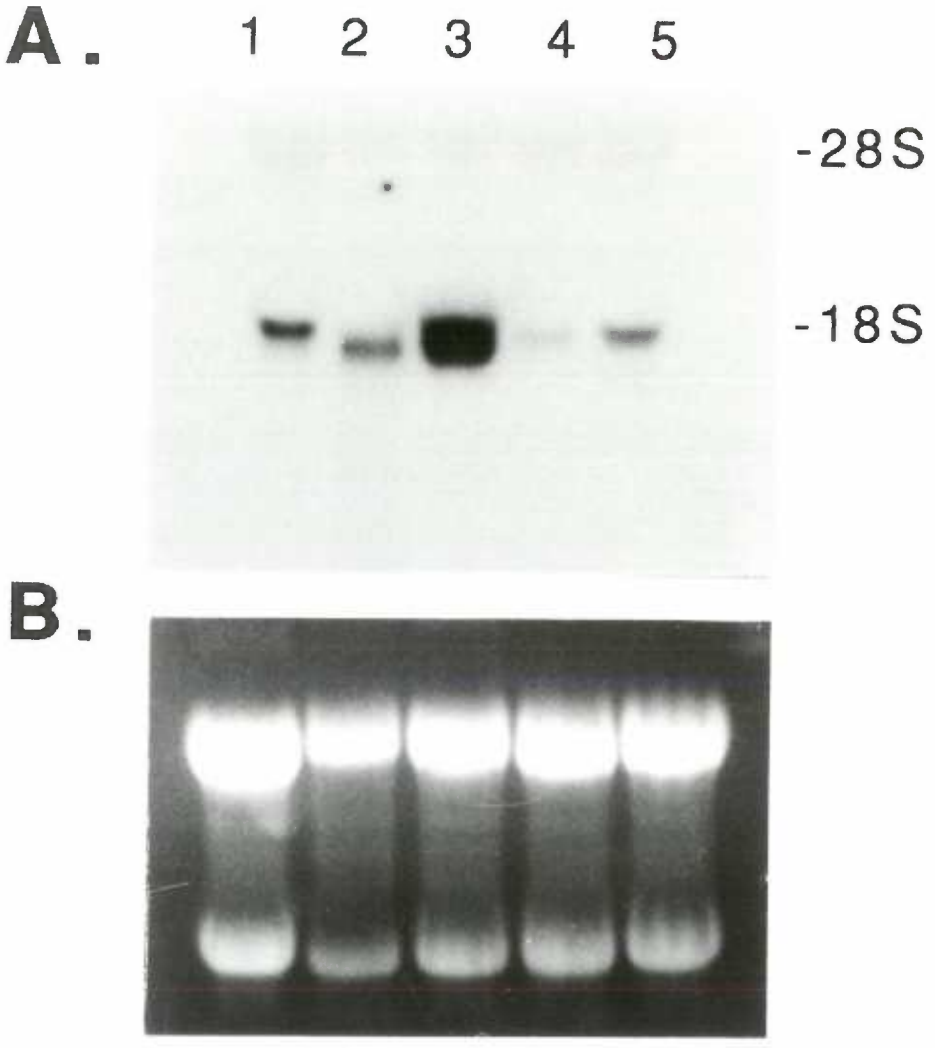
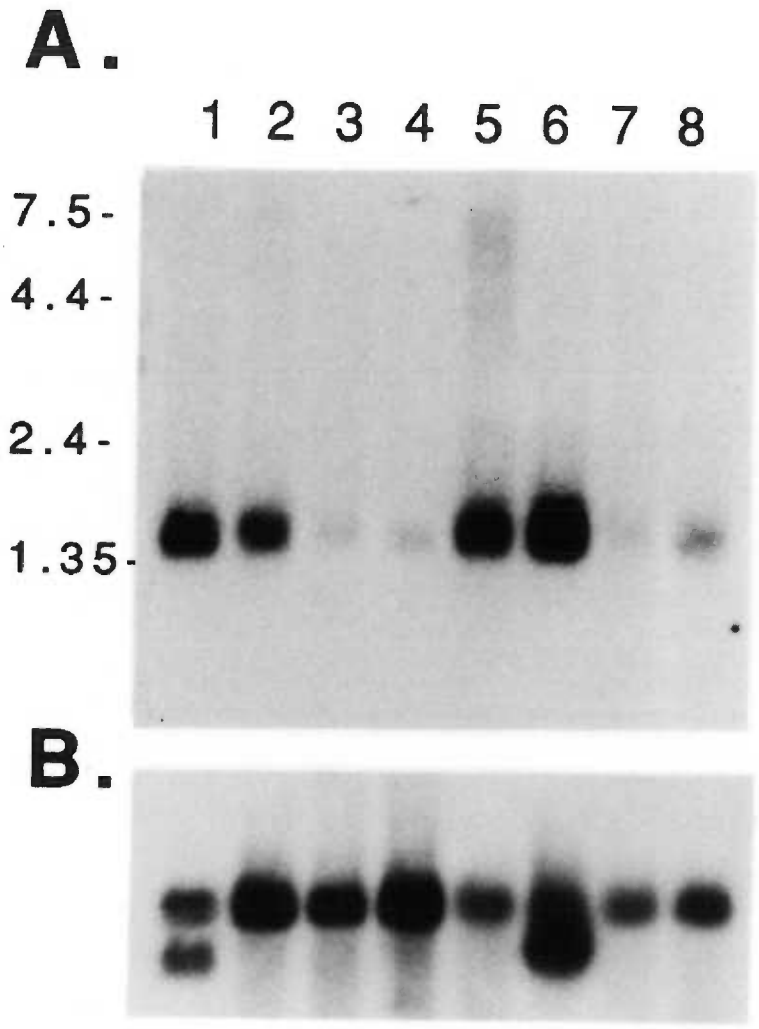

A.
1
2
3
4
5
-28S
-18S
B.

Figure 13. Northern blot of polyadenylated RNA from human tissues. The blot was hybridized with radiolabeled pKC122 (A), or a β-actin cDNA (B). Each lane was loaded with 2 μg RNA from heart, lane 1; brain, lane 2; placenta, lane 3; lung, lane 4; liver, lane 5; skeletal muscle, lane 6; kidney, lane 7; pancreas, lane 8. The units of the ordinate are kb.


A.
1 2 3 4 5 6 7 8
7.5-
4.4-
2.4-
1.35-
B.

Manuscript #3:

A Comparison of the Enzymatic and Physicochemical Properties of a Recombinant Glutathione S-Transferase GSTM4-4 with Three other Recombinant Human Mu Class Enzymes.

Kenine E. Comstock[1], Mikael Widersten[2], Xiao-Yong Hao[2], W. David Henner[1,3] and Bengt Mannervik[2]

From the [1]Department of Biochemistry and Molecular Biology and [3]Division of Hematology and Medical Oncology, Oregon Health Sciences University, Portland OR, U.S.A. and the [2]Department of Biochemistry, Uppsala University, Biomedical Center, Uppsala, Sweden.

Abstract

The multigene family of cytosolic glutathione *S*-transferases (GSTs) consists of four classes (Alpha, Mu, Pi and Theta), all involved in the detoxification of reactive electrophiles. The human Mu class GSTs consist of at least four isozyme subunits, GST M1, GST M2, GST M3 and GST M4, which have 70-90% amino acid sequence identity. The gene and cDNA sequences for GST M4 have been determined recently by some of us. When GST M4 cDNA sequences were cloned into an *E. coli* expression vector and induced, an enzyme was purified which has a relatively low specific activity with the standard GST substrate CDNB (1.39 $\pm$ 0.21 U/mg protein), but an activity equivalent to other Mu class enzymes with some substrates. This protein yields a homogenous band of approximately 26.7 kd on SDS-PAGE . A detailed comparison of the activity with various substrates and inhibitors was performed between the recombinant GST M4-4 and the other recombinant human Mu class GSTs, GST M1a-1a, GST M2-2 and GST M3-3. Despite the high level of amino acid sequence identity, the enzymatic properties of these enzymes were quite different. These findings support the hypothesis that most of the amino acid differences among these enzymes occur in the substrate-binding region and suggest that diversification of these genes occurs primarily in the substrate binding regions to cope with an increasing variety of foreign compounds.

Introduction

Human GSTs are a large multigene family of isozymes which catalyze the conjugation of glutathione to electrophilic substrates (for reviews see Mannervik, 1985, Ketterer, 1988, Mannervik et al., 1992). These enzymes are involved in the detoxification of both endogenous and exogenous electrophiles which are reactive toward cellular components such as DNA. The modification of DNA by reactive compounds can initiate carcinogenesis and the GSTs are thought to be involved in cancer prevention by inactivating carcinogens.

There are four major classes of cytosolic GSTs, including the Alpha, Mu, Pi and the recently discovered Theta class. Particular enzymes are placed in a class based upon their amino acid sequence identity. Within a class, the GST proteins share a high degree of amino acid sequence homology (70-90%), are immunologically cross-reactive, and have similarities in substrate and inhibitor specificities. The cytosolic GSTs are dimers of approximately 22-27 kd subunits. While heterodimers may form between subunits of two GST isozymes from the same class, subunits from different classes do not appear to dimerize (Reddy et al., 1984; Tu et al., 1985 Mannervik, 1985).

The human Mu class GSTs, which are the focus of this investigation, consists of at least four isozyme subunits, GST M1, GST M2, GST M3 and the recently discovered GST M4, each of which is coded for by a distinct gene. It is likely there are at least two more genes and/or pseudogenes predicted by Southern blotting

analysis (DeJong et al. 1991). These isozymes have a high level of amino acid sequence identity, but distinct physicochemical properties and tissue distributions.

The GST M1 gene locus is polymorphic and displays an allelic form, GST M1b, differing from GST M1a by one nucleotide that results in a single amino acid substitution (Singh et al., 1987). Both enzymes appear to be identical in their catalytic abilities and substrate specificities. The GST M1 gene locus shows an additional form, the null allele, which has been shown to be due to a gene deletion. Individuals homozygous for the null allele occur at a frequency of about 50% in the populations tested (Seidegard et al., 1988). Both GST M1a-1a and GST M1b-1b have a relatively high activity with certain epoxides, some of which are present in cigarette smoke (Warholm et al., 1983). The common deletion of the GST M1 gene and resulting lack of enzymatic activity has been correlated with an increased risk of lung and other cancers among smokers (Seidegard et al., 1990, 1986) but other studies have failed to confirm the predisposition to lung cancer (Zhong et al., 1991, Heckbert et al., 1992).

GST M2-2 has been isolated from skeletal muscle, the cDNA sequence has been reported and a recombinant expression clone was constructed. The recombinant enzyme was expressed and shown to have specific activities distinct from those of GST M1a-1a and GST M1b-1b (Vorachek et al, 1992). GST M3-3 was originally isolated from brain and testis. The cDNA sequence was reported, revealing that this enzyme is the least structurally conserved among the

human Mu class. A limited study of the substrate specificities of GST M3-3 has been reported (Campbell et al., 1991).

Recently, the gene sequence and the cDNA sequence for a Mu class GST isozyme, GST M4, has been determined in our laboratory (Comstock et al., in press). Northern blotting shows the presence of GST M4 mRNA in liver, heart, placenta, lung, brain, liver, skeletal muscle, pancreas, testis, cerebral cortex, uterus, and ovary. No GST bearing amino-terminal amino acid sequence matching the deduced amino acid sequence for this enzyme has been reported.

The previously described human Mu class isozyme subunits show a high degree of identity with GST M4: the deduced amino acid sequence of GST M4 is 87%, 83%, and 70% identical to GST M1a, GST M2, and GST M3 respectively. This high degree of homology might suggest a high degree of similarity in substrate specificities and catalytic ability. However the differences in amino acid sequence may occur mostly in the substrate binding regions resulting in marked differences in function. It has been postulated that diversification of cytochrome P450 genes occurs primarily in the substrate binding regions to cope with an increasing variety of foreign compounds (Gotoh, 1992) and this may also have occurred during the evolution of the Mu class GSTs . To test the hypothesis that there may be marked differences in the substrate and inhibitor specificities of GSTM4-4 and the other isozymes despite their high degree of identity, we have constructed or obtained recombinant expression clones for the human Mu class glutathione transferases GST M4-4, GST M1a-1a, GST M2-2, and GST M3-3 for the production and characterization of the enzymatic and physicochemical

properties these enzymes. This work constitutes the first comparative study of the known human Mu class GSTs.

Materials and Methods

Materials: *S*-hexylglutathione was synthesized by the method of Vince and coworkers (Vince et al., 1971). Oligonucleotides were purchased from Operon Technologies (Alameda, CA, U.S.A.) or were synthesized on an Applied Biosystems PCRmate oligonucleotide synthesizer.

Growth media was prepared with tryptone and yeast extract from Difco Laboratories (Detroit MI, U.S.A.). Antibiotics and isopropyl ß-D-thiogalactopyranoside were from Sigma Chemical Co. (St. Louis MO, U.S.A.).

Δ^5-androstene-3,17-dione and 2-cyano-1,3-dimethyl-1-nitrosoguanidine were gifts from Dr. Paul Talalay, Johns Hopkins University, Baltimore MD, U.S.A. and Dr. David Jenssen, Temple University, Philidelphia PA, U.S.A., respectively. Rose bengal, 1-chloro-2,4-dinitrobenzene, hydrogen peroxide and 1,2-epoxy-3-(p-nitrophenoxy)-propane were obtained from Sigma Chemical Co. (St. Louis MO, U.S.A.). *para*-nitrobenzyl chloride, (2*S*, 3*S*)-3-(4-nitrophenyl)-glycidol, (2*R*, 3*R*)-3-(4-nitrophenyl)-glycidol and *trans*-stilbene oxide were from Aldrich Chemical Co. (Milwaukee WI U.S.A.). Cumene hydroperoxide was from Merck (Munich, Germany). Cibacron Blue was from Ciba Ltd. (Basel, Switzerland).

Plasmid constructions: Construction of the plasmids pKHM1a, pKHM1b (Widersten et al., 1991) and pKC114 (Comstock et al., in press) have been described previously. The plasmid pKK-GST4 was a gift of William Pearson, University of Virginia, Charlottesville VA, U.S.A. The construction of the plasmid pKHM3

was done as follows: Total RNA was isolated from HeLa cells by standard techniques. 40 pmoles of oligonucleotide primer (5'-AAGAATTCTGCAGT-3') was annealed to 15 µg total RNA. First strand synthesis was carried out with 200 units of Superscript murine mammary leukemia virus H-reverse transcriptase (Bethesda Research Laboratories Life Technologies, Inc., Gaithersburg MD, U.S.A.) according to manufacturer's protocol. The cDNA was added to a PCR reaction containing 1µM each of the primers (5'-CCGAATTCATGTCGTGCGAGTCGTCTAGT-3') and (5'-GGCTGCAGCACCAGTAACATAAGTGCTAT-3'), 10 mM Tris-CL (pH 8.3), 1.5 mM MgCl2, 50 mM KCl, 0.1 mg/ml gelatin, 0.2 mM deoxynucleoside triphosphates, and 2.5 U Taq DNA polymerase (Boeringer Mannheim, Mannheim, Germany) in a 100 µl reaction. DNA was amplified by incubating 1 min at 95 ^{o}C, 2 min at 55 ^{o}C, and 1 min at 72 ^{o}C for a total of 30 cycles in a Perkin Elmer Cetus DNA Thermal Cycler.

The amplified fragment was digested with *Eco* RI and *Pst* I and subsequently subcloned into the polylinker of pGEM3Zf (Promega, Madison WI, U.S.A.). Plasmid DNA was sequenced by the dideoxy method (Sanger et al., 1977). After verification of the DNA sequence, the insert was subcloned into the plasmid pKK233-3 (Clonetech, Palo Alto CA, U.S.A.). This plasmid was denoted pKHM3.

Purification of GST M1a-1a, GST M1b-1b, GST M2-2 and GSTM3-3: Plasmids pKHM1a, pKHM1b and pKK-GST4 were transfected into competent JM103 (Bethesda Research Laboratories Life Technologies, Inc., Gaithersburg MD, U.S.A.). Plasmid pKHM3 was transfected into competent XL1-Blue (Stratagene, LaJolla CA, U.S.A.).

A 35 ml portion of LB broth (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% NaCl, pH 7.0), containing 50 µg/ml ampicillin, was inoculated with a loopful of one of the above. The culture was incubated approximately 3 h at 37 °C on a rotary shaker. Six 2000 ml Erlenmeyer flasks each containing 500 ml of LB broth and 50 µg/ml ampicillin were inoculated with 5 ml of the above culture. When the cell suspension reached a density of approximately A_{600}= 0.3, isopropyl-ß-D-thiogalactopyranoside was added to a final concentration of 0.2 mM and incubation was continued overnight. After being harvested at 5000 g in a Beckman JA14 rotor, the cells were resuspended in 20 ml of 50 mM Tris-HCl (pH 8.0), 50 mM EDTA, 15% (w/v) sucrose, 1 mg/ml lysozyme. The suspension was incubated for 30 min at room temperature, then cells were disrupted by sonication six times in a Branson Sonifier at setting 7 for 20 s, on ice. Cellular debris was removed by centrifugation at 25,000 g for 25 min at 4 °C. The lysate was added to approximately 20 ml *S*-hexylglutathione-Sepharose gel and incubated on a rocking platform for 1 h at 4 °C. After adsorption of the enzyme, the gel was packed into a column and washed with 22 mM sodium phosphate buffer (pH 7.0) 1 mM EDTA, 0.2 mM dithiothreitol, 50 µM phenylmethanesulphonyl fluoride. The GSTs were eluted with 10 mM Tris-HCl (pH 7.8), 5 mM *S*-hexylglutathione. Fractions were pooled and dialyzed against 2 x 4 liters of 0.1 M sodium phosphate buffer (pH 7.0), containing 0.2 mM dithiothreitol and 0.02% sodium azide.

Purification of GST M4-4: Plasmid pKC114 was transfected into competent BL21(DE3)pLysS (Novagen, Madison WI, U.S.A.). A 20 ml culture of LB containing 50 µg/ml ampicillin and 25

µg/ml chloramphenicol was inoculated with a loopful of the above and incubated overnight at 37 °C with shaking. A 2000 ml Erlenmeyer flask containing 1 liter of LB broth with antibiotics was inoculated with 10 ml of the above culture. When the cell suspension reached a density of approximately A_{600}= 0.5, isopropyl-ß-D-thiogalactopyranoside was added to a final concentration of 1.0 mM and incubation was continued for 5 hours. After being harvested at 5000 g in a Beckman JA14 rotor, the cells were resuspended in 20 ml 22 mM sodium phosphate buffer (pH 7.0), 1 mM EDTA, 0.2 mM dithiothreitol, 50 µM phenylmethanesulphonyl fluoride at 4 °C. Cells were disrupted by sonication six times for 20 s on ice and debris was removed by centrifugation at 25,000 g for 25 min at 4 °C. The lysate was added to approximately 20 ml *S* -glutathione-Sepharose gel and incubated on a rocking platform for 1 h at 4 °C. After adsorption of the enzyme, the gel was packed into a column and washed with 22 mM sodium phosphate buffer (pH 7.0) 1 mM EDTA, 0.2 mM dithiothreitol, 50 µM phenylmethanesulphonyl fluoride. GSTs were eluted with 10 mM Tris-HCl (pH 7.8), 5 mM GSH, fractions were pooled and then dialyzed against 2 x 4 liters of 0.1 M sodium phosphate buffer (pH 7.0), containing 0.2 mM dithiothreitol and 0.02% sodium azide.

Protein concentration was determined by the method of Peterson (1977) and/or the method of Bradford (1976). Also used was an extinction coefficient, A_{280} = 1 corresponds to 0.7 mg/ml of GST Mu, estimated by comparing concentrations determined by the first two methods to the absorbance at 280 nm.

Physicochemical Procedures: Native M_r was estimated by gel filtration on a Superose 12 column connected to an F.P.L.C. system (Pharmacia-LKB Biotechnology, Uppsala, Sweden). The chromatography was run at 0.5 ml/min in 0.05 M sodium phosphate buffer pH 7.0, 0.15 M NaCl. Marker proteins were monitored photometrically at 280 nm and were ovalbumin (43,000 Da) and trypsin inhibitor (20,100 Da).

Subunit M_r and purity was estimated by SDS/PAGE (Laemmli, 1970). GST M4, GST M1, GST M3 and GST M2, and molecular weight standard proteins were resolved on a 9-15% linear gradient polyacrylamide gel. Bands were visualized by silver staining.

The isoelectric points were determined by PAGE/isoelectric focusing on precast gel plates run under the conditions prescribed by the manufacturer (Pharmacia-LKB Biotechnology, Uppsala, Sweden), with a parallel run of standard pI marker proteins.

The amino-terminal amino acid sequence was determined by protein sequencing with Applied Biosystems equipment.

Kinetic Measurements: Enzyme activities were determined spectrophotometrically at a constant temperature of 30 ^{o}C with alternative electrophilic substrates as described in Table 3 except for *trans*-stilbene oxide which was performed as follows: Activity was assayed radiometrically with 0.25 mM [^{3}H]*trans*-stilbene oxide and 5 mM GSH in 0.228 M sodium phosphate, pH 7.2, at 37 ^{o}C with 0.1 M sodium phosphate used as a blank (Gill et al., 1983). After incubation with enzyme, the reaction mixture was extracted with 2 volumes of *n*-hexanol and the radioactivity remaining in the aqueous

phase was determined in a liquid scintillation counter (LKB Rackbeta 1209).

The inhibition studies were performed by addition of various concentrations of the inhibitors to the standard reaction mixture containing 1 mM 1-chloro-2,4-dinitrobenzene and 1 mM GSH. This data was used to generate an inhibition curve. All measurements were performed in quintuplicate in microtiter plates at a constant temperature of 30 ^{o}C, using a Molecular Devices microtiter plate reader.

Experiments to determine the kinetic constants were performed as follows: The V_{max} and K_m (substrate concentration at $V_{max}/2$) values were determined from v versus [S] plots, in which concentrations of 1-chloro-2,4-dinitrobenzene and glutathione were varied independently in the ranges of 62.5 to 1000 µM and 312.5 to 5000 µM, respectively. Measurements were made in triplicate in microtiter plates at a constant temperature of 30 ^{o}C, using a Molecular Devices microtiter plate reader. The kinetic data was analyzed by nonlinear regression analysis by using the computer program package SIMFIT (Bardsley et al., 1989).

The effect of pH on initial velocities was determined by examining the effect of varying the pH of the 0.1 M sodium phosphate buffer from pH 4.0 to pH 8.0 in the standard reaction containing 1 mM CDNB and 1 mM GSH and a constant amount of recombinant GST M4-4. The velocity of the spontaneous reactions at each pH was subtracted from the enzymatic reactions. All measurement were made in triplicate at a constant temperature of 30 ^{o}C.

Binding Studies: Measurements were performed at ambient temperature in phosphate buffered saline on an Aminco spectrofluorometer: λ_{ex} = 273 nm, and λ_{em} = 334 nm. The dissociation constant for the binding of GSH was determined by analysing the data by nonlinear regression analysis. The program SIMFIT (Bardsley et al., 1989) was used for all data analysis.

Results

Expression and purification of recombinant enzymes: After induction of pKC114 cultures, followed by a 5 h incubation and purification, 15 - 20 mg of pure GST M4-4 enzyme was routinely recovered from a 1 liter culture. 10 - 50 mg of pure GST M1a-1a, GST M1b-1b GST M2-2 and GST M3-3 were routinely recovered after induction, overnight incubation and purification of 6 liter cultures of pKHM1a, pKHM1b, pKK-GST4 and pKHM3 cultures, respectively. Homogeneity was established by SDS/PAGE. Amino acid sequence analysis of recombinant GST M4-4 showed that the amino terminus began with Ser-2, indicating that the initiating methionine had been removed as has been shown for other recombinant GSTs (Widersten et al. 1991).

Molecular Properties: The apparent native molecular weight of recombinant GST M4-4 was estimated as 45,000 Da by gel exclusion F.P.L.C. The apparent molecular weights of recombinant GST M1a-1a and GST M2-2 also have been estimated at 45,000 Da (Table 4). We were unable to obtain an accurate estimate of the native molecular weight of GST M3-3 as the enzyme appeared to be retained on the column.

In the presence of sodium dodecyl sulfate, GST M4-4 dissociated into two identical subunits of apparent Mr 26,700, as determined by polyacrylamide gel electrophoresis (Table 4 and Figure 14). Recombinant GST M1a-1a, GST M2-2, and GST M3-3 dissociated into subunits of apparent Mr 26,700, 26,600, and 26,800, respectively. The isoelectric point of recombinant GST M4-

4, determined by isoelectric focusing, was pH 5.20, compared to pH 6.2 for GST M1a-1a, pH 5.6 for GST M1b-1b, and pH 5.25 for GST M2-2 (Table 4). We were unable to determine an isoelectric point for GST M3-3 due to the inability of the protein to enter the gel for reasons which are unclear.

Kinetic Properties: A comparison of the specific activities of GST M4-4 and the other recombinant Mu class enzymes with some potential substrates was made. GST M4-4 had the lowest specific activity with CDNB (approximately 1 μmol min^{-1} mg^{-1}) of all the Mu class enzymes studied (Table 5), however CDNB was still the best substrate for GST M4-4 among those tested. The activity with CDNB was checked daily and remained constant throughout the period of time in which the assays for other activities were performed. GST M3-3 also has comparatively low CDNB activity, with a specific activity of approximately 7 μmol min^{-1} mg^{-1}. Different preparations of GST M1a-1a were found to vary in their CDNB activity between 60 and 90 μmol min^{-1} mg^{-1}, but the activity of a single preparation appeared to remain constant. The explanation for this is unclear. The specific activities of GST M1a-1a for other substrates did not vary among preparations.

GST M4-4 showed measurable activity with the substrates ethacrynic acid, the alkylhalide *para*-nitrobenzyl chloride, and the epoxides (2*R*, 3*R*)-3-(4-nitrophenyl)-glycidol and *trans*-stilbene oxide. The highest activity among the Mu class for the substrates *para*-nitrobenzyl chloride and cumene hydroperoxide was shown by GST M1a-1a. It was the only Mu class enzyme tested to show activity with 1,2-epoxy-3-(*para*-nitrophenoxy)-propane and Δ^5-

androstene-3,17-dione. GST M2-2 showed the highest activity among the Mu class enzymes for the substrates CDNB, 1,2-dichloro-4-nitrobenzene, (2*R*, 3*R*)-3-(4-nitrophenyl)-glycidol and was the only enzyme tested to show activity with 2-cyano-1,3-dimethyl-1-nitrosoguanidine. GST M3-3 showed the highest activity among the Mu class enzymes for ethacrynic acid and was the only enzyme tested to show activity with hydrogen peroxide and both 3-(4-nitrophenyl)-glycidol isomers.

Kinetic constants:

The k_{cat} (sometimes referred to as the turnover number), K_m[GSH], and K_m[CDNB], for the substrate CDNB were determined for GST M4-4 (Table 6).

Binding studies:

The binding of glutathione to GST M4-4 was determined by measuring the quenching of the intrinsic fluorescence upon titration of the protein with glutathione. The decrease in fluorescence upon binding followed a hyperbolic curve, allowing the dissociation constant for the binding of glutathione to GST M4-4 to be calculated as 39 μM ± 6 μM.

pH Dependence of Initial Velocities:

The pH optimum for the enzymatic reaction between glutathione and CDNB was at approximately pH 8 (Figure 15), compared to the pH optimum of approximately pH 7.5 for the same reaction catalyzed by GST M1a-1a, determined by Warholm and coworkers (Warholm et al., 1983).

Inhibition Studies:

A comparison of the inhibitory effects of a wide variety of compounds on catalysis by the recombinant Mu class GSTS was made (Table 7). There were some striking differences in the potency of the inhibitors against the Mu class enzymes studied. Triethyltin bromide was a more potent inhibitor of GST M4-4 than the other enzymes by at least a factor of 5. Cibacron blue was a more potent inhibitor of GST M2-2 by at least a factor of 5. Hematin was a more potent inhibitor of GST M3-3 than the other enzymes by a factor of 3 and *S*-(*p*-bromobenzyl)glutathione was a more potent inhibitor of GST M1a-1a than the other enzymes by at least a factor of 5.

Discussion

Numerous electrophilic compounds can serve as substrates for glutathione transferases. Substrates which are synthesized by the modern chemical industry are readily available and the products of conjugation to GSH are usually easily detectable, so they have proven very useful to the study of GSTs. Although the substrates we have tested are not necessarily biological substrates, some can give clues about the function of these enzymes *in vivo*. Epoxides are a group of substrates that can be formed by the oxidation of the carbon-carbon double bonds of xenobiotics (Levi, 1987). These compounds, known to be mutagenic and carcinogenic, can arise from sources such as cigarette smoke and industrial chemicals. Some of the Mu class enzymes show activity with the epoxides 1,2-epoxy-3-(p-nitrophenoxy)-propane and *trans*-stilbene oxide. GSTs are thought to be involved in cancer prevention by inactivating these potential carcinogens by the conjugation of glutathione. There is some data to suggest that the absence of epoxide-inactivating activity of GST M1a-1a and GST M1b-1b can result in a higher risk of lung and other cancers among smokers. GSTM4-4 shows activity toward *trans*-stilbene oxide (at approximately the same level as shown for the Alpha class enzyme GSTA1-1 (Mannervik & Danielson 1988)), so it is possible that GSTM4-4 may be involved in inactivating epoxides *in vivo* to a certain extent. Because there are differences in tissue distribution of the Mu class enzymes, the frequent deficiency of GSTM1 may only render some tissues susceptible to epoxide damage. GSTM4-4 may act as a backup

protector to some degree against these compounds in tissues where it is expressed.

The Δ^5-3-ketosteroid isomerase activity demonstrated by the Δ^5-androstendione assay also suggests a role for GST M1a-1a in the androgen and estrogen biosynthetic pathway (Benson et al. 1977). While the frequent deletion of the GSTM1 gene and the existence of other enzymes with this activity suggests that GST M1a-1a's role is probably not essential, androgen and estrogen levels in deleted individuals may be influenced in certain tissues. The activity toward hydroperoxides by GST M1a-1a and GST M3-3 suggests that these enzymes may play a role in protecting cells from the highly reactive products of oxygen metabolism. Oxygen free radicals can have a number of deleterious effects on the cell, including the peroxidation of lipids and damage to DNA.

The so-called "universal" substrate for GSTs, 1-chloro-2,4 dinitrobenzene (CDNB) (Habig et al., 1981) has been extremely useful in the demonstration of multiple forms of GSTs. However, many enzymes have relatively low activity toward CDNB and Theta class enzymes lack any activity toward this substrate (Danielson and Mannervik, 1988; Meyer et al., 1991). GST M4-4 is a good example of a GST for which CDNB is a poor substrate. Because investigators frequently test crude fractions with CDNB to aid in the isolation of new GSTs, this may explain why an enzyme whose sequence matches the deduced amino acid sequence of GST M4-4 has not yet been identified in tissues. Although GST M4-4 exhibited relatively low catalytic activity toward CDNB, it showed comparable activity to one or more of the other Mu class isozymes with the substrates

ethacrynic acid and *para*-nitrobenzyl chloride. Despite being a poor catalyst for the conjugation of GSH to CDNB, GST M4-4 may still be quite active with an unidentified environmental or physiological substrate.

The steady state kinetics of GST M4-4 with CDNB as a substrate was evaluated. The k_{cat} value for this substrate was approximately 7 times lower than the k_{cat} value previously determined for GST M1a-1a (Warholm et al., 1983). This lower k_{cat} value is reflected in the lower specific activity of GST M4-4 with CDNB as compared to GST M1a-1a. GST M4-4 does not appear to catalyze a simple reaction mechanism where the noncovalent enzyme-substrate complex rapidly and reversibly forms from the substrates and the enzyme and then decomposes into products. The K_m [GSH] for GST M4-4 is greater than the dissociation constant K_d, which suggests that the dissociation to enzyme and substrates is slower than the forward reaction and that there are several kinetically significant intermediate steps in the reaction. GSTs typically show significant deviations from Michaelis-Menten kinetics (Mannervik & Danielson, 1988).

The rates of enzymatic reactions commonly show a substantial pH dependence. A variation in velocity as seen in the case of GST M4-4, implies that the enzyme has ionizable group(s) which have an effect on activity, either directly as in the case of an ionizable catalytic amino acid or indirectly by the influence of ionizable group(s) on the microenvironment at the active site. While it is not possible to determine definitively which amino acids are involved, it appears that the residues may be slightly different between GST M4-

4 and GST M1a-1a based upon the differences seen in their pH dependence curves.

Inhibitors of GSTs are important tools which can be useful when studying the structure of the active site and the mechanism of catalysis of the GSTs. Glutathione derivatives which are substrate analogs can be useful to probe the structure of the GSH binding site. Bulkier *S* -alkylglutathiones seem to be more potent inhibitors of GST M4-4. Since these compounds are competitive with glutathione and probably act at the active site, these results indicate the importance of hydrophobic interactions.

Inhibitors can also be used to distinguish individual GST isozymes from mixtures of isozymes which have very similar substrate specificities (Tahir et al., 1985). Since the Mu class enzymes studied have at least one inhibitor each which can be used to easily distinguish it from the other enzymes, the results of these inhibitor studies can be used to determine identity of a particular Mu class isozyme and have the advantage that relatively small amounts of enzyme are needed.

Simple inhibition studies can also be used to distinguish between homodimeric and heterodimeric isozymes of GSTs (Tahir et al., 1986). The inhibition curves for heterodimeric enzymes can be predicted from those of the corresponding homodimeric isozymes and can therefore aid in the discrimination between homo- and heterodimers of GST M4-4 and other Mu class GSTs. It has been reported that the Mu class enzymes form heterodimers, however the identities of the monomers were not determined unambiguously (Tsuchida et al., 1990; Hussey et al., 1991). More studies are needed

to determine which monomers associate *in vivo* and how this affects catalytic function.

In addition, inhibitors may eventually prove useful in reversing GST mediated anticancer drug resistance in humans. GST M1a-1a has been reported to be elevated in glioma cells resistant to (2-chloroethyl)-3-sarcosinamide-1-nitrosourea (SCNU) which suggests possible involvement in drug resistance (Skalski et al., 1990). Because the pharmacologically active drugs, indomethacin, an anti-inflammatory agent and ethacrynic acid, a diuretic drug, are inhibitors of GST M1a-1a, it would be reasonable to determine whether these drugs have any effect on some patients with drug resistance to SCNU. Several reports have shown that treatment of drug resistant cell lines with these inhibitors increases their sensitivity to the drugs (Hall et al., 1989; Hansson et al., 1991). Also, a recent study showed that patients treated with ethacrynic acid had decreased levels of GST activity in peripheral mononuclear cells, which demonstrates the effectiveness of this inhibitor *in vivo* and suggests that further clinical studies are warranted (O'Dwyer et al., 1991).

We have found that there are significant differences in the substrate and inhibitor specificities between GST M4-4 and the other Mu class enzymes despite their high degree of amino acid sequence identity. This in turn suggests that, of the differences found in the primary structure, many must occur in the hydrophobic substrate binding region. It has been suggested by Gotoh that diversification of duplicate genes of cytochrome P450s occurs primarily in the substrate binding regions to cope with an increasing

number of foreign compounds (Gotoh, 1992). Our data suggests this may also occur in the evolution of the Mu class GSTs.

Acknowledgements

Dr. William G. Bardsley, University of Manchester, Manchester, U. K. generously provided the SIMFIT computer program package.

References

Bardsley, W. G., McGinlay, P. B., & Roig, M. G. (1989) J. Theor. Biol. **139**, 85-102

Bradford, M. A. (1976) Anal. Biochem. **72**, 248-254

Benson, A. M., Talalay, P., Keen, J. H., & Jakoby, W. B. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 158-162

Campbell, E., Takahashi, Y., Abramovitz, M., Peretz, M., & Listowsky, I. (1990) J. Biol. Chem. **265**, 9188-9193

Comstock, in press (Manuscript #2 of this dissertation)

Danielson, U. H., & Mannervik, B. (1988) CRC Crit. Rev. Biochem **23**, 283-337

DeJong, J. L. Mohandas, T. & Tu, C.-P. D. (1991) Biochem. Biophys. Res. Comm. **180**, 15-22.

Gill, S. S., Ota, K., & Hammock, B. D. (1983) Anal. Biochem. **131**, 273-281.

Gotoh, O. (1992) J. Biol. Chem. **267**, 83-90

Habig, W. H., & Jakoby, W. B. (1981) Methods Enzymol. **77**, 398-405

Hall, A., Robson, C. N., Hickson, I. D., Harris, A. L., Proctor, S. J., & Catten, A. R. (1989) Cancer Res. **49**, 6265-6268

Hansson, J., Berhane, K., Castro, V. M., Jungnelius, U., Mannervik, B., & Ringborg, U. (1991) Cancer Res. **51**, 94-98

Heckbert, S. R., Weiss, N. S., Hornung, S. K., Eaton, D. L., & Motulsky, A. G. (1992) JNCI **84**, 414-422

Hussey, A. J., Kerr, L. A., Cronshaw, A. D., Harrison, D. J., & Hayes, J. D., (1991) Biochem. J. **273**, 323-332

Ketterer, B. (1988) Mutation Res. **202**, 343-361

Laemmli, U. K., (1970) Nature **227**, 680-685

Levi P. E., (1987) in A Textbook of Modern Toxicology (Hodgson, E., & Levi, P. E. Eds.), pp 133-185, Elsevier New York

Mannervik, B. (1985) Adv. Enzymol. Rel. Areas Mol. Biol. **57**, 357-417

Mannervik, B. & Danielson, U. H. (1988) CRC Crit. Rev. Biochem. **23**, 283-337

Mannervik, B., Awasti, Y. C., Board, P. G., Hayes, J. D., Dillio, C., Ketterer, B., Listowsky, I., Morgenstern, R., Muramatsu, M., Pearson, W. R., Pickett, C. B., Sato, K., Widersten, M., & Wolf, C. R. (1992) Biochem. J. **282**, 305-306

Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M., & Ketterer, B. (1991) Biochem. J. **274**, 409-414

O'Dwyer, P. J., LaCreta, F., Nash, S., Tinsely, P. W., Schilder, R., Clapper, M. L., Tew, K. D., Panting, L., Litwin, S., Comis, R. L., & Ozols, R. F. (1991) Cancer Res. **51**, 6059-6065

Peterson, G. L. (1977) Anal. Biochem. **83**, 346-356.

Reddy, C. C., Li, N., & Tu, C.-P. D. (1984) Biochem. Biophys. Res. Comm. **121**, 1014-1020

Sanger, F., Nicklen, S., & Coulson, A. R., (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5467.

Seidegard, J., Pero, R. W., Miller, D. G., & Beattie, E. J. (1986) Carcinogenesis **7**, 751-753

Seidegard, J., Vorachek, W. R., Pero, R. W., & Pearson, W. R., (1988) Proc. Natl. Acad. Sci. **85**, 7293-7297

Seidegard, J., Pero, W. R., Markowitz, M. M., Roush, G., Miller, D. G., & Beattie, E. J. (1990) Carcinogenesis **11**, 33

Singh, S. V., Kurosky, A., & Awasthi, Y. C. (1987) Biochem. J. **243**, 61-67

Skalski, V., Yarosh, D. B., Batist, G., Gros, P., Feindel, W., Kopriva, D., & Panasci, L. C. (1990) Mol. Pharmacol. **38**, 299

Tahir, M. K., Guthenberg, C., & Mannervik, B. (1985) FEBS Lett. **181**, 249-252

Tahir, M. K., & Mannervik, B. (1986) J. Biol. Chem. **261**, 1048-1051

Tsuchida, S., Maki, T., & Sato, K. (1990) J. Biol. Chem. **265**, 7150-7157

Tu, C.-P. D., & Reddy, C. C. (1985) J. Biol. Chem. **260**, 9961-9964

Vince, R., Daluge, S., & Wadd, W. B. (1971) J. Med. Chem. **14,** 402-404

Vorachek, W. R., Pearson, W. R., & Rule, G. S. (1991) Proc. Natl. Acad. Sci. U.S.A. **88,** 4443-4447.

Warholm M., Guthenberg, C., & Mannervik, B. (1983) Biochemistry **22**, 3610-3617

Widersten, M., Pearson, W. R., Engstrom, A., & Mannervik, B. (1991) Biochem. J. **276**, 519-524

Zhong, S., Howie, A. F., Ketterer, B., Taylor, J., Hayes, J. D., Beckett, G. J., Wathen, C. G., Wolf, C. R., & Spurr, N. K. (1991) Carcinogenesis **12**, 1533-1537

Figure 14. SDS-PAGE of recombinant Mu class GSTs. Lanes 1 and 8, low molecular weight marker (Phamacia); lanes 2 and 7, mixture of recombinant Mu class GSTs; lane 3, GSTM4; lane 4, GSTM3; lane 5, GSTM1; lane 6, GSTM2. The values indicated for the marker proteins are apparent M_r.

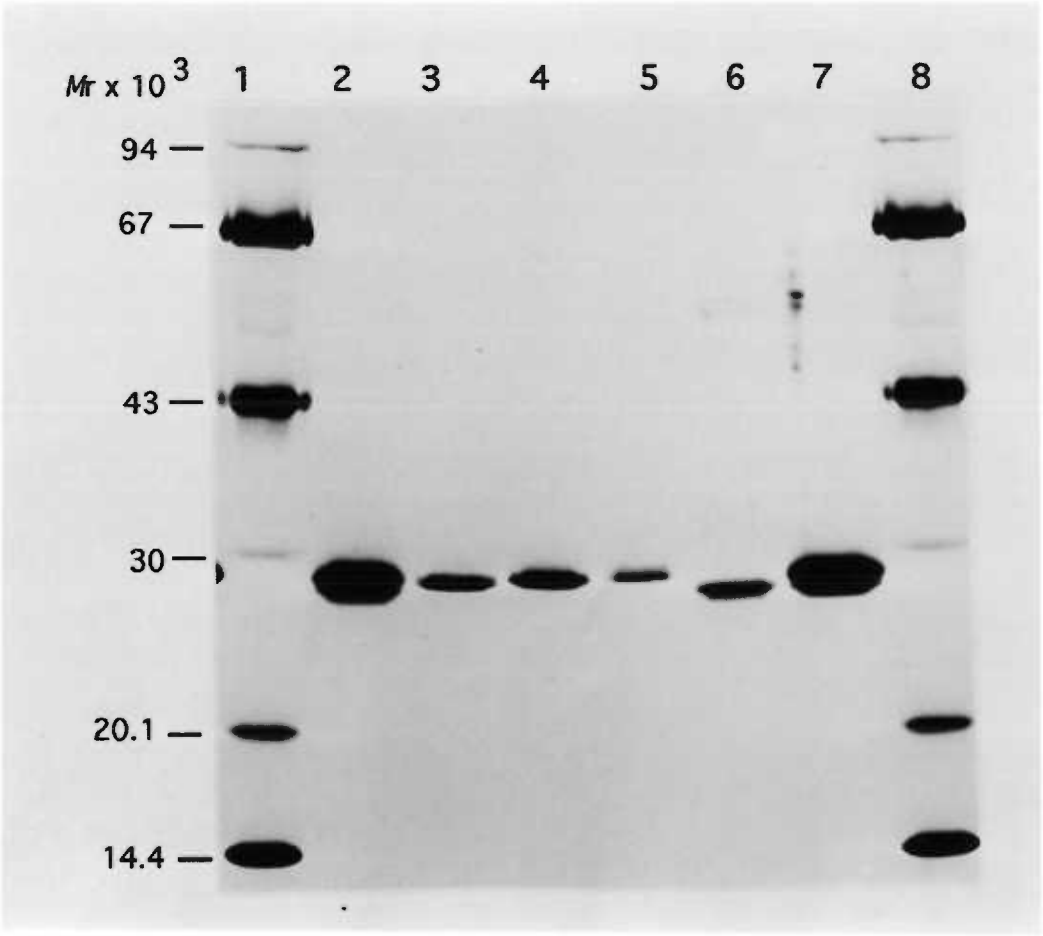

Mr x 10 3
1
2
3
4
5
6
7
8
94
67
43
30
20.1
14.4

Figure 15. pH dependence of initial velocities obtained with 1-chloro-2,4-dinitrobenzene. The experimental details are provided under Materials and Methods. (hollow squares) GSTM4-4; (filled squares) GSTM1a-1a.

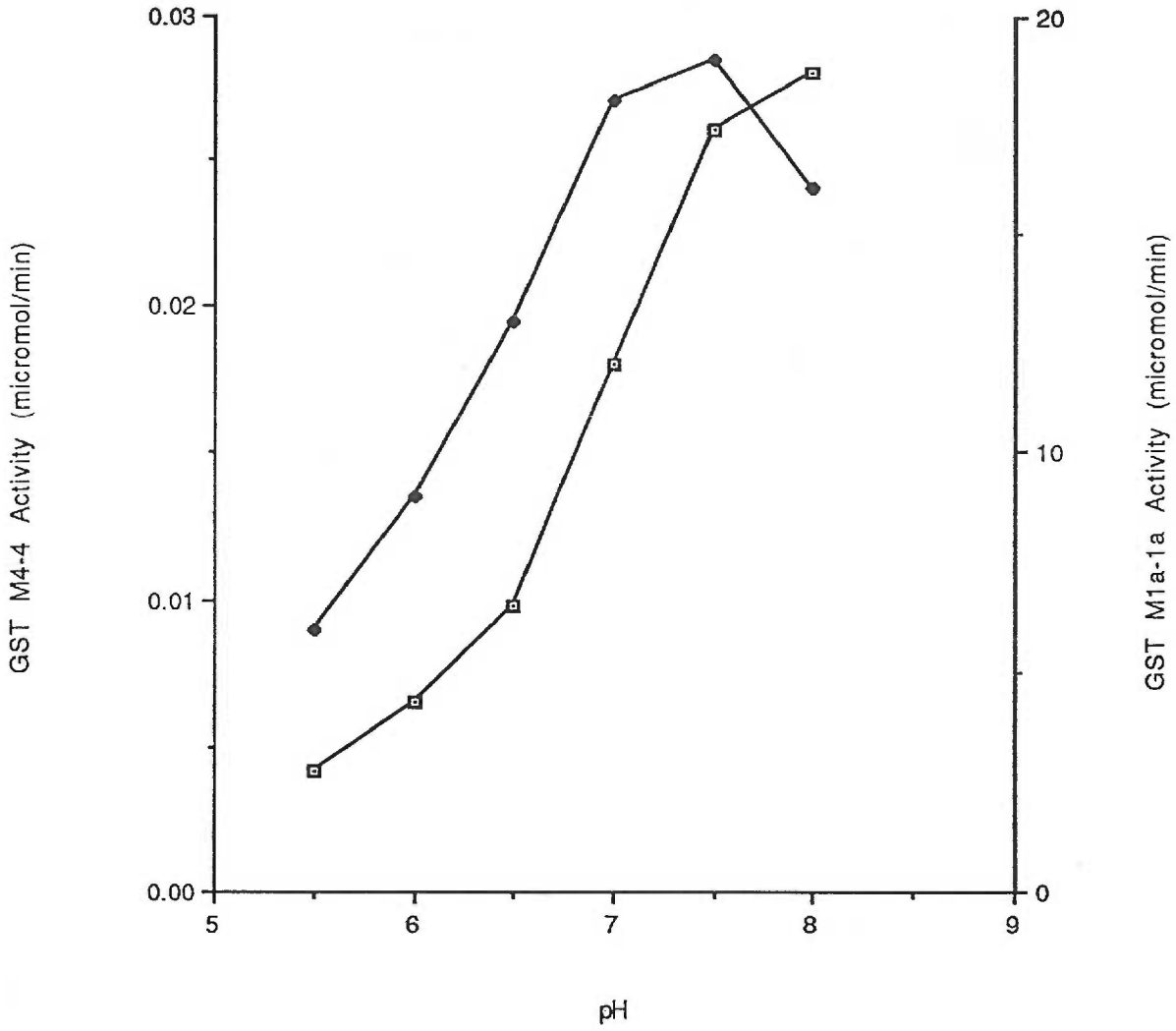

0.03
0.02
0.01
0.00
20
10
0
5
6
7
8
9
pH
GST M4-4 Activity (micromol/min)
GST M1a-1a Activity (micromol/min)

Table 3: Assays for enzyme activity toward various substrate.

Substrate	Substrate conc. (mM)	GSH conc. (mM)	Buffer	Wave-lenght (nm)
1-chloro-2-4-dinitrobenzene	1.0	1.0	0.1 M sodium phosphate (pH 6.5)	340
trans-4-phenylbut-3-en-2-one	0.05	0.25	0.1 M sodium phosphate (pH 6.5)	290
1,2-epoxy-3-(p-nitrophenoxy)-propane	0.5	5.0	0.1 M sodium phosphate (pH 6.5)	360
p-nitrobenzyl chloride	0.25	5.0	0.1 M sodium phosphate (pH 6.5)	310
ethacrynic acid	0.2	0.25	0.1 M sodium phosphate (pH 6.5)	270
(2S,3S)-3-(4-nitrophenyl)-glycidol	1.0	2.0	0.1 M sodium phosphate (pH 6.5)	340
(2R,3R)-3-(4-nitrophenyl)-glycidol	1.0	2.0	0.1 M sodium phosphate (pH 6.5)	340
1,2-dichloro-4-nitrobenzene	1.0	5.0	0.1 M sodium phosphate (pH 8.0)	344
bromosulfothalein	0.03	5.0	0.1 M sodium phosphate (pH 7.5)	330
hydrogen peroxide or cumene hydroperoxide	1.5	1.0	*0.1 M sodium phosphate (pH 7.0)	340
Δ^5-androstene-3,17-dione	0.068	0.1	25 mM Tris(HCl), 12.5 mM K_2HPO_4 (pH 8.3)	248
2-cyano-1,3-dimethyl-1-nitrosoguanidine	1.0	1.0	0.1 M sodium phosphate (pH 7.5)	340

* Also included in this assay are 0.3 U/ml glutathione reductase and 0.1mM NADPH

Table 4: Physicochemical properties of recombinant Mu class enzymes

Enzyme	Apparent subunit M_r	[a]Apparent native M_r	Isoelectric point
GST M1a-1a	26,700[b]	45,000[b]	6.2
GST M1b-1b	26,600[b]	45,000[b]	5.6
GST M2-2	26,500	45,000	5.25
GST M3-3	26,800	[c]	[c]
GST M4-4	26,700	45,000	5.20

[a] The values given have an estimated S. D. of ± 3000.
[b] Values reported in Widersten, M. et al., Biochem. J. **276**, 519-524.
[c] Unable to determine these values due to unusual properties of protein.

Table 5: Specific Activities (μmol min^{-1} mg^{-1}) of recombinant Mu class enzymes with various substrates.

Substrate	GST M1a-1a	GST M2-2	GST M3-3	GST M4-4
CDNB	60.71 ± 6.67	220 ± 4.7	7.26 ± 0.7	1.39 ± 0.21
DCNB	0.022 ± 0.002	2.1 ± 0.10	0.041 ± 0.01	<0.006
Ethacrynic acid	0.057 ± 0.006	0.166 ± 0.018	0.226 ± 0.03	0.117 ± 0.03
para-Nitrobenzyl chloride	0.190 ± 0.006	0.174 ± 0.002	0.031 ± 0.005	0.04 ± 0.03
trans- PBO	0.057 ± 0.043	<0.007	<0.030	<0.004
NPEP	0.065 ± 0.004	<1.9	<0.3	<0.2
(2*S*,3*S*)-NPG	<0.110	<0.130	0.342 ± 0.023	<0.031
(2*R*,3*R*)-NPG	0.157 ± 0.028	0.900 ± 0.089	0.120 ± 0.012	0.050 ± 0.006
Androstene dione	0.035 ± 0.007	<0.07	<0.016	<0.003
Cyano DMNG	<0.21	117.31 ± 7.94	<0.34	<0.03
BSP	<0.067	<0.231	<0.05	<0.03
trans-Stilbene oxide	5.2[a]	[b]	[b]	0.00014
Hydrogen peroxide	<0.113	[b]	0.5 ± 0.1	<0.107
Cumene hydro peroxide	0.620 ± 0.007	[b]	0.3 ± 0.2	<0.05

Abbreviations: CDNB,1-chloro-2,4-dinitrobenzene; DCNB, 1,2-dichloro-4-nitrobenzene; *trans*-PBO, *trans*-4-phenylbut-3-en-2-one; NPEP, 1,2-epoxy-3-(*p*-nitrophenoxy)propane; Cyano DMNG, 2-cyano-1,3-dimethyl-1-nitroguanidine; BSP, bromosulfophthalein; (2*S*,3*S*)-NPG, (2*S*,3*S*)-3-(4-nitrophenyl)-glycidol; (2R,3R)-NPG, (2*R*,3*R*)-3-(4-nitrophenyl)-glycidol; androstene dione, Δ^5-androsten-3,17-dione.

[a] reported in Mannervik, B. & Danielson U.H., (1988) CRC Crit. Rev Biochem Mol. Biol. 23, 283-337.
[b] value not determined for this enzyme.

Table 6: Kinetic constants for 1-chloro-2,4-dinitrobenzene

GST...	k_{cat} (sec^{-1})	$K_m^{[GSH]}$ (mM)	$K_m^{[CDNB]}$ (mM)
M1a-1a*	193.3 ± 25	0.16 ± 0.05	0.65 ± 0.11
M4-4	28.9 ± 12.5	1.84 ± 1.80	8.98 ± 4.25

* These values were reported in Warholm et al. (1983) Biochemistry 22, 3610-3617.

Table 7: Inhibition characteristics of recombinant GST Mu class enzymes

The values given are based upon five replicates of each inhibitor concentration used to generate the standard curve.

	I_{50} (μM)			
Inhibitor	**GSTM1a-1a**	**GSTM2-2**	**GSTM3-3**	**GSTM4-4**
Bromosulfophthalein	2.75	1.93	14.5	12
Cibacron Blue	0.391	0.0843	0.383	4.18
Rose Bengal	0.786	1.57	1.34	1.77
Hematin	0.265	0.357	0.0808	10.1
Indomethacin	385	112	>400*	257**
Tributyltin acetate	0.549	5.94	0.869	1.07
Triethyltin bromide	1.75	4.34	3.00	0.341
Triphenyltin chloride	0.173	15.2	0.317	20
S-Hexylglutathione	6.98	61.2	59.1	13.5
S-Benzylglutathione	7.02	45.9	53.1	4.64
S-(*p*-Bromobenzyl) glutathione	0.525	12.2	31	2.7
S-(2,4-Dinitrophenyl) glutathione	137	253	104	165
S-Ethacrynic acid glutathione	1.38	7.56	2.72	4.09
Ethacrynic acid	2.51	8.18	5.16	4.63

I_{50} is the concentration giving 50% inhibition in the standard assay with 1 mM 1-chloro-2,4-dinitrobenzene and 1 mM GSH in 0.1 M sodium phosphate buffer, pH 6.5, at 30°C. The average coefficient of variation is 6 %.

* Not an inhibitor at 400 μM.

** I_{20} value

III. SUMMARY AND CONCLUSIONS

The GSTs are a major contributor to protection of mammals against the toxic effects, including cancer, of exposure to environmental chemicals. Study of all classes of GSTs will eventually provide significant insights on the risk involved with exposure to toxic substances, and could influence clinical decisions and public policy concerning safe exposure limits. Also, rodent studies suggest that the levels and therefore the protective effect of these enzymes vary substantially on exposure of animals, and presumably humans, to various chemicals. The ability to modulate the expression of particular GST isozymes could be used to increase protection of individuals who, for occupational or other reasons, are exposed to the toxic substrates of these enzymes. The GSTs are implicated in anticancer drug resistance to a wide variety of commonly used drugs. Using knowledge gained about enzymatic inhibitors and the regulation of gene expression, activity and expression of the GSTs could be modulated during cancer therapy to circumvent drug resistance to anticancer drugs.

The human Mu class GSTs are thought to be involved in detoxification of potential carcinogens, drug resistance to certain anticancer drugs and there is a correlation between the deletion of the Mu class gene GSTM1 and lung and other cancers. The objective of this thesis project was to further characterize the Mu class GSTs and contribute to an understanding of their role in these phenomena.

A PCR-based assay was designed which can be used in epidemiological studies to investigate the correlation between a genetic predisposition to lung cancer and the common GSTM1 gene deletion. Several groups have used this assay to define the correlation between the gene deletion, corresponding lack of enzyme activity and bladder and lung cancer in large numbers of affected individuals (Brockmoller et al., 1992, Bell et al., 1992). This assay was also used by us to determine the relationship between the presence or absence of GSTM1 protein in breast carcinomas and this gene deletion (See Appendix A.). It was found that breast carcinomas can be classified as either GSTM1a/GSTM1b protein positive or negative, and this phenotype appears to be determined by the heredity of the individual who develops carcinoma, rather than being a characteristic of the tumor. This finding could be important if GSTM1 is involved in drug resistance as suspected. A person lacking the GSTM1 gene would not be expected to develop drug resistance due to overexpression of GSTM1, so testing a breast cancer patient's lymphocyte DNA for the GSTM1 deletion could be helpful when determining which anticancer drug would be most effective for treatment.

The genomic DNA and the corresponding cDNA sequences were determined for a new member of the human GST gene superfamily, GSTM4. The genomic sequence consists of eight exons, with splice junctions identical to the rat Mu class gene Yb2, and an upstream region which contains the consensus sequence for an antioxidant responsive element (ARE). The cDNA for GSTM4 was isolated from a library derived from the cervical carcinoma cell line, HeLa. Northern

blots of human RNA, probed with a unique 5' region of the GSTM4 cDNA, demonstrated the presence of GSTM4 mRNA in human heart, placenta, lung, brain, liver, skeletal muscle, pancreas, testis, cerebral cortex, uterus, ovary, a lymphoblastoid cell line, and four carcinoma cell lines. Using PCR amplification of a unique region of the GSTM4 gene from a panel of rodent-human somatic cell hybrid DNAs, the chromosomal location of the GSTM4 gene was determined to be chromosome 1.

GSTM4 cDNA sequences were cloned into an *E. coli* expression vector, foreign gene expression was induced and recombinant GSTM4-4 protein was purified to yield a protein of approximately 27.5 kd which has a relatively low specific activity with the standard GST substrate CDNB (1.39 $\pm$ 0.21 U/mg protein). This may explain why GST M4-4 has not previously been purified from tissues since crude fractions from tissues are usually tested for activity with CDNB to aid in the isolation of new GSTs. A detailed comparison of the activity with various substrates and inhibitors was performed between the recombinant GSTM4-4 and the other recombinant human Mu class GSTs, GSTM1a-1a, GSTM2-2 and GSTM3-3. Although GST M4-4 exhibited relatively low catalytic activity toward CDNB, it showed comparable activity to one or more of the other Mu class isozymes with the substrates ethacrynic acid and *para*-nitrobenzyl chloride. This suggests that GST M4-4 may be very active with a biological substrate, despite being a poor catalyst for the conjugation of GSH to CDNB. A true biological substrate for GSTM4-4 probably exists. It is not yet possible to predict what this substrate is, however the future elucidation of the three-

diminsional stucture of a human Mu class isozyme may provide clues. The inhibitors determined for GSTM4-4 and the other Mu class enzymes may prove useful in isolating particular isozymes from heterogeneous mixtures, predicting the presence of heterodimeric enzymes and in treating anticancer drug resistance. The results of these experiments support the hypothesis that there may be marked differences in the substrate and inhibitor specificities of GSTM4-4 and the other Mu class isozymes despite their high degree of amino acid sequence identity. This in turn suggests that of the differences found in the primary structure, many must occur in the hydrophobic substrate binding region (H-site). We propose that diversification of duplicate genes of GSTs may occur primarily in the substrate binding regions to cope with an increasing number of foreign compounds.

In conclusion, a human Mu class GST gene, GSTM4, and the corresponding cDNA have been cloned and sequenced. Details of the expression patterns of GSTM4 mRNA and the properties of the GSTM4-4 enzyme have been revealed. Obviously further studies are needed to clearly define the biological role of GSTM4-4 and the mechanism of regulation of the GSTM4 gene. The data presented here provides the foundation necessary to begin those studies.

References

Bell, D. A., Taylor, J. A., Paulson, D. F., Robertson, C. N., Mohler, J. L. Miller, C. R., & Lucier, G. W. (1992) printed in abstract form in the Proceedings of the American Association for Cancer Research **33**: 291

Brockmoller, J. , Gross, D., Kerb, R., Drakoulis, N., & Roots, I. (1992) Biochem. Pharmacol. **43**:647-650

IV. APPENDIX

A. The absence of GSTM1a/GSTM1b in 50% of breast carcinomas is due to heredity.

GSTM1a and GSTM1b are both alleles of the human GSTM1 gene locus, one of several human genes that code for human Mu class GSTs (DeJong et al., 1991). However, GSTM1a or GSTM1b are expressed in lymphocytes in only 31-60% of humans (Seidegard et al., 1985). In individuals who lack GSTM1a and GSTM1b, the deficiency appears to be due to homozygous deletion of the GST1 gene that codes for GSTM1a and GSTM1b (Seidegard et al., 1988). The frequency of Mu class GST in human breast carcinomas, when detected by Western blotting of breast carcinoma extracts using anti-Mu class GST antisera, is was similar to that of GSTM1a/GSTM1b in the total human population suggesting that the Mu class GST detected is GSTM1a/GSTM1b (Shea et al., 1990). This in turn suggests that the presence or absence of this enzyme in breast carcinomas might be determined by the presence or absence of the gene for GSTM1 in the particular individual developing the breast carcinoma.

To test this possibility, a PCR-based assay for the presence of the GSTM1 gene in human DNA samples was used (Comstock et al., 1990). In this assay, primers hybridizing to the 5' region of exon 4 and the 3' region of exon 5 amplify a 273 bp product from target DNA derived from lymphocytes of individuals who carry the GSTM1 gene. This PCR product consists of the DNA sequence of exon 4, intron 4, and exon 5 of GSTM1 (Seidegard et al., 1988, Lai et al., 1988) An appropriately sized product was not amplified from DNA of

individuals whose lymphocytes do not contain Mu class GST by Western blotting. When this assay was applied to study human breast carcinomas, DNAs from 7 carcinomas positive for Mu class GST on Western blot using anti-Mu class antisera were all positive on the PCR assay and only one of the 11 DNAs from Western blot-negative tumors produced an appropriately sized DNA fragment ($P < 0.001$, χ^2 test) (Figure 23).

There are several human GSTs of the Mu class which cross react immunologically and are coded for by several Mu class genes. Therefore it is not possible to determine which of the Mu class GSTs is present in the breast carcinoma extracts by Western blotting alone. Only about half of normal individuals have detectable GSTM1a/GSTM1b activity in peripheral mononuclear leukocytes or liver. The presence or absence of this enzyme is inherited as an autosomal dominant marker and the absence of this enzyme has recently been shown to be due to a homozygous gene deletion (Seidegard et al., 1985) The similar frequency for Mu class GST positivity among breast carcinomas and for GSTM1a/GSTM1b positivity in liver suggests that the absence of Mu class GST in breast carcinomas (and possibly other neoplasms) may simply be determined by the absence of a copy of the GSTM1 gene in the germ line of the individual rather than being a characteristic of the carcinoma itself. The results of the studies using the polymerase chain reaction are consistent with this hypothesis. We have found that the target sequence for PCR, a portion of the GSTM1 gene locus that codes for the GSTM1a/GSTM1b alleles is present in all breast carcinoma positive for Mu class protein on Western blotting and

absent in all but one of the carcinomas negative for Mu class protein. The single disparate sample (PCR positive but GSTM1a/GSTM1b negative on Western blot) may simply be due to GSTM1 protein degradation during tissue handling and storage. Alternatively, this result may represent a tumor that contains the GSTM1 gene but is deficient in one or more of the steps required in production of a stable GSTM1a/GSTM1b protein. The conclusion made by this study was that breast carcinomas can be classified as either GSTM1a/GSTM1b positive or negative, and this phenotype appears to be determined by the heredity of the individual who develops carcinoma.

References

Comstock, K.E., Sanderson, B.J.S., Claflin, G., & Henner, W.D. (1990) Nucleic Acids Res. **18**: 3670

DeJong, J. L., Mohandas, T., & Tu, C.-P. D. (1991) Biochem. Biopys. Res. Comm. **180**:15-22

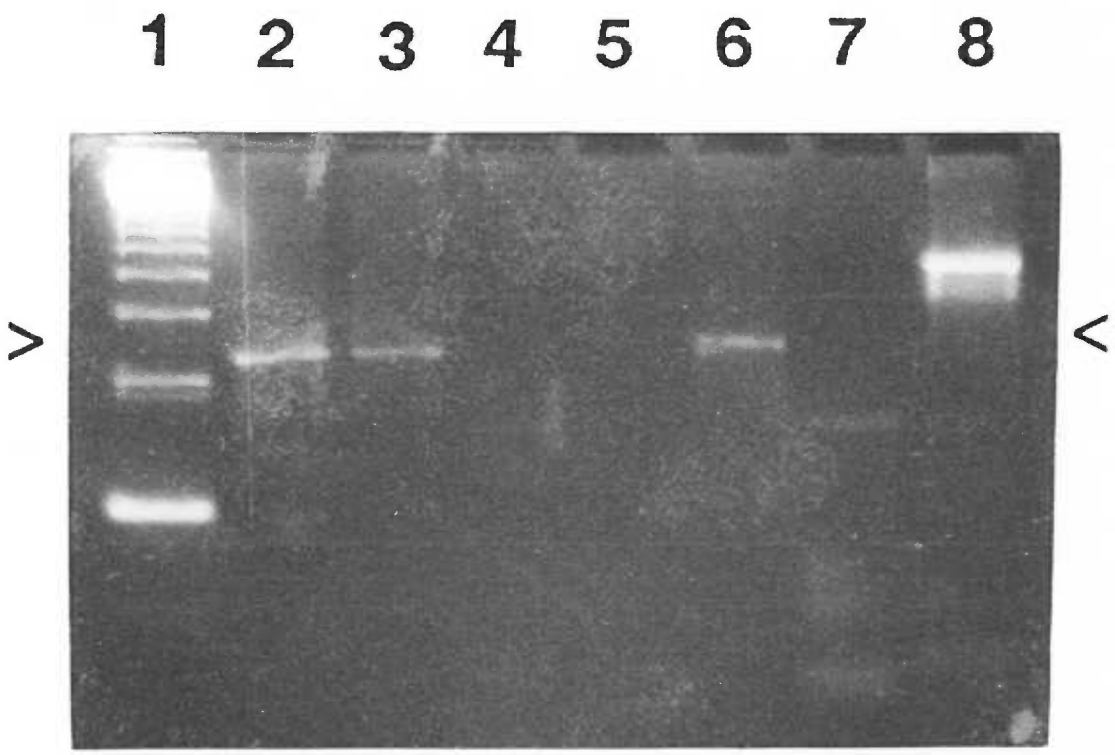
Lai, H.-C. J., Qian, B., Grove, G., &Tu, C.-P. D. (1988) J. Biol. Chem. **263**:11389-11395

Seidegard, J., & Pero, R. W. (1985) Hum. Genet. **69**:66-68

Seidegard, J., Vorachek, W. R., Pero, R. W., & Pearson, W. R. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**:7293-7297

Shea, T. C., Claflin, G., Comstock, K. E., Sanderson, B. J. S., Burstein, N. A., Keenan, E. J., Mannervik, B., & Henner, W. D.(1990) Cancer Res. **50**:6848-6853

Figure 16. GSTM1 PCR products resolved by agarose gel electrophoresis. PCR reactions were done as described previously (Manuscript #1). A 10 µl aliquot of the reaction mixture was resolve by agarose gel electrophoresis. *Lane 1* contains DNA standards (123 bp ladder, Bethesda Research Laboratory, Gaithersburg, MD.); *lanes 2, 3, 6* and *lanes 4 and 5* contain PCR products of target DNA from Mu class GST-postive and -negative carcinomas, respectively; *lane 7* is the product of the PCR reaction in the absence of target DNA; *lane 8* contains a 500 bp product of PCR of a λ bacteriophage DNA with λ primers. *Arrowheads,* the expected position for a 273-base pair DNA product.


1 2 3 4 5 6 7 8
>