Identification and characterization of GSTT3, a third murine Theta class glutathione transferase

Marjorie COGGAN*, Jack U. FLANAGAN*, Michael W. PARKER†, Vanicha VICHAI‡, William R. PEARSON‡ and Philip G. BOARD*[1]

*Molecular Genetics Group, John Curtin School of Medical Research, Australian National University, P.O. Box 334, Canberra, Australian Capital Territory, 2601, Australia, †Biota Structural Biology Laboratory, St. Vincent's Institute of Medical Research, Fitzroy, Victoria, 3065, Australia, and ‡Department of Biochemistry, University of Virginia, Charlottesville, VA 22908, U.S.A.

A novel Theta class glutathione transferase (GST) isoenzyme from mouse termed mGSTT3 has been identified by analysis of the expressed sequence tag database. The gene encoding *mGSTT3* is clustered with the *mGSTT1* and *mGSTT2* genes on chromosome 10 and has an exon/intron structure that is similar to that of the other Theta class genes. mGSTT3 is expressed strongly in the liver and to a decreasing extent in the kidney and testis. Recombinant mGSTT3-3 expressed in *Escherichia coli* had a substrate-specificity profile that differed significantly from that of GSTT1-1 and GSTT2-2 isoenzymes. A molecular model of mGSTT3 suggested that, in comparison with GSTT2, a decrease in volume of the hydrophobic substrate-binding site and the loss of the sulphate-binding pocket prevents its use of the GSTT2 substrate 1-menaphthyl sulphate.

Key words: database analysis, enzyme, glutathione transferase gene.

INTRODUCTION

The glutathione transferases (GSTs) play a significant role in the metabolism of a wide range of electrophilic compounds that include mutagens, carcinogens and some therapeutic agents. In mammals the soluble GSTs can be grouped into at least eight classes, termed Alpha, Mu, Pi, Theta, Kappa, Sigma, Zeta and Omega [1–4]. Although the chloride intracellular channel ('CLIC') proteins are additional members of the soluble GST structural family, their enzymic potential has not yet been investigated [5]. Variable numbers of genetically distinct isoenzymes have been described within each class and, so far, the genes encoding each member of a class have been clustered on a distinct chromosome [6–10]. Although there are many overlaps, different members of a class can have quite distinct substrate preferences. For example, within the Theta class two distinct isoenzymes termed GSTT1-1 and GSTT2-2 have been described in humans, mice and rats [11–16]. Despite their sequence and structural similarities, GSTT1 and GSTT2 isoenzymes have distinct substrate profiles. GSTT1-1 isoenzymes have characteristic substrate preferences for dichloromethane (DCM), 1,2-epoxy-3-(*p*-nitrophenoxy)propane (EPNP) and 4-nitrobenzyl chloride (4-NBC), whereas GSTT2-2 isoenzymes have significant activity with 1-menaphthyl sulphate (MS).

The structure of human GSTT2-2 (hGSTT2-2) was determined by X-ray crystallography [17] and the structure of hGSTT1-1 has been predicted by molecular modelling [18]. Unlike the Alpha, Mu and Pi class GSTs, which contain a characteristic active-site tyrosine residue, the Theta class GSTs have an active-site serine residue that appears to stabilize the glutathione bound in the active site as a thiolate anion. The role of the active-site residues and the reaction mechanism of the Theta class GSTs has been examined in detail by a number of mutagenic studies [15,19–23].

Given the striking differences in substrate specificity between members of the different classes and between members of the same class [1], it is probable that additional GST classes and individual class members have remained undiscovered, because their substrates and catalytic activities are unknown. Analysis of the expressed sequence tag (EST) database has provided a structural approach to the identification of new members of the GST family that is not dependent on prior knowledge of an enzyme's catalytic activity. The discovery of the Zeta and Omega classes has demonstrated the utility of this approach [3,4]. In a previous study of the murine Theta class GSTs [16], we noted several EST clones that appeared to encode Theta class GSTs, but differed significantly from the sequences of mouse GSTT1 (mGSTT1) or GSTT2 (mGSTT2). In the present study, we have identified a novel third Theta class GST gene (*GSTT3*) in the mouse and characterized its product GSTT3-3. Although sharing a high level of sequence identity with GSTT1 and GSTT2, the new isoenzyme does not share their substrate specificities.

EXPERIMENTAL

Cloning and sequencing

EST clones were identified by searching the mouse EST database (http://www.ncbi.nlm.nih.gov/) with the mGSTT1 protein sequence using the TFASTY program [24]. The EST clones (AA871247 and AI385844) were obtained from the I.M.A.G.E. consortium (http://image.llnl.gov/) and sequenced on both strands. Partial genomic sequence was identified in the GenBank® database by use of the BLAST program [25] and the cDNA sequence derived from AI385844. The size and sequence of intron 1 were determined by PCR amplification from Balb/c genomic DNA using the primers mT3exon1F (5′-CCCC-TTCCAGCTGCGTACC) and intron 1R (5′-GGGTTCTTC-TATCCAGACC). All sequencing was carried out using a Thermosequenase cycle sequencing kit (Amersham Biosciences, Castle Hill, New South Wales, Australia).

Abbreviations used: BAC, bacterial artificial chromosome; CDNB, 1-chloro-2,4-dinitrobenzene; CPK Corey–Pauling–Koltun; DCM, dichloromethane; DCNB, 1,2-dichloro-4-nitrobenzene; DDCT, D-dopachrome tautomerase; EA, ethacrynic acid; EPNP, 1,2-epoxy-3-(*p*-nitrophenoxy)propane; EST, expressed sequence tag; GST, glutathione transferase; H-site, hydrophobic substrate-binding site; MS, 1-menaphthyl sulphate; 4-NBC, 4-nitrobenzyl chloride; NBD-Cl, 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole; 4-NPA, 4-nitrophenylacetate; poly(A)$^+$, polyadenylated.

[1] To whom correspondence should be addressed (e-mail Philip.Board@anu.edu.au).

In order to express recombinant mGSTT3, the coding sequence was amplified from the EST clone AI385844 and ligated between the *Sph*I and *Hin*dIII sites in the pQE-30 vector (Qiagen, Clifton Hills, Victoria, Australia) and transformed into *Escherichia coli* M15[pREP4] host cells. The primers used for the amplification were mT3ExF (5′-CACGGATCCGGTCTGGAGCTCTACC-TG) and mT3ExR (5′-CTAGAAGCTTTCAGTGTAACAAA-CACTG) and contained *Sph*I and *Hin*dIII sites respectively. The resulting plasmid was termed pQEmGSTT3.

Southern and Northern blot hybridization

To test for cross hybridization between the murine GST Theta genes, a Southern blot was prepared using mouse genomic DNA. The DNA was digested with the restriction enzymes *Hin*dIII, *Eco*RI, *Bam*HI and *Pst*I (New England Biolabs, Beverly, MA, U.S.A.) and then probed with full-length cDNA for mGSTT1, mGSTT2 or GSTT3. The blots were hybridized in 2 × PE [where 1 × PE is 0.133 M sodium phosphate buffer and 1 mM EDTA (pH 6.9)] containing 7 % (w/v) SDS and 1 % (w/v) BSA at 65 °C overnight. The blots were then washed twice with 2 × SSPE [where 1 × SSPE is 0.01 M sodium phosphate buffer, 0.18 M NaCl and 1 mM EDTA (pH 7.7)] containing 0.1 % SDS at 21 °C for 10 min each, followed by one wash with 1 × SSPE containing 0.1 % SDS at 65 °C for 15 min. The blots were then autoradiographed for between 1 to 3 days.

A multiple-tissue Northern blot of mouse polyadenylated [poly(A)$^+$] RNA was obtained from ClonTech (Palo Alto, CA, U.S.A.). DNA derived from bases 692–1021 of mGSTT3 cDNA was used as a hybridization probe for mGSTT3 mRNA, as this region has low sequence identity with the *mGSTT1* and *mGSTT2* genes.

cDNA probes were radioactively labelled using nick translation (Amersham Biosciences) and were hybridized by the ClonTech ExpressHyb protocol for radioactively labelled cDNA probes. After hybridization, the blot was washed for 40 min at 21 °C in 2 × SSC [where 20 × SSC is 0.3 M sodium citrate and 0.3 M NaCl (pH 7.0)] containing 0.05 % SDS with three changes of solution, followed by two washes at 50 °C for 40 min in 0.1 × SSC containing 0.1 % SDS. The blots were then autoradiographed for between 1 h to overnight.

Expression and purification of recombinant mGSTT3-3

A culture of pQEmGSTT3 in M15[pREP4] cells was grown overnight in the presence of isopropyl thio-β-D-galactoside and processed by the methods described by Whittington et al. [16]. The recombinant enzyme was purified by nickel–nitrilotriacetate-immobilized metal affinity chromatography as previously described [16], with the exception that buffer A [50 mM sodium phosphate buffer and 300 mM NaCl] was adjusted to pH 6.5 and the purified enzyme was dialysed against a solution containing 10 mM Tris/HCl, 1 mM EDTA and 0.5 mM 2-mercaptoethanol (pH 8.25).

Enzyme activity assays

All assays were carried out at 37 °C. Activities toward 1-chloro-2,4-dinitrobenzene (CDNB), 1,2-dichloro-4-nitrobenzene (DCNB), EPNP, 4-NBC, 4-nitrophenyl acetate (4-NPA) and ethacrynic acid (EA) were measured as described previously [1,26]. Activity towards MS was determined by the method of Gillham [27], and glutathione peroxidase activity was measured as described by Lawrence and Burk [28]. The conjugation of GSH to 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole (NBD-Cl) was measured by the method of Ricci et al. [29]. Activity towards DCM was determined by the formation of formaldehyde as described previously [16]

Molecular modelling

Construction of atomic models of the isoforms mGSTT1, mGSTT2 and mGSTT3 was based on the suite of structures solved for the hGSTT2 enzyme (PDB accession numbers 1LJR, 2LJR and 3LJR) [17] using Modeller 4.0 [30] with default settings, and an alignment of mammalian Theta class sequences prepared using Clustal X software [31]. Manual inspection of the Clustal X sequence alignment was carried out to ensure that insertions and deletions occur within the loop regions. The model mGSTT1 and mGSTT3 structures were primarily based on the GSH binary complex solved for hGSTT2. However, the mGSTT1 helix 4 and the C-terminal helix (helix 9) were based on those regions modified for a previous homology model of hGSTT1 [18]. As the amino acid substitutions between mGSTT3 and the other Theta class sequences within helix 4 were more similar to those in the GSTT1 isoforms, the modified helix 4 structure used in the hGSTT1 model was also used as a template for the mGSTT3-3 helix 4. Even though mGSTT3 is three residues shorter than the template structure, helix 9 is proposed to extend one residue beyond the C-terminus of the template. To accommodate an extra residue in helix 9 and remove any unfavourable interactions, helix 5 was constructed such that the N-terminal bends away from the dimer interface. As mGSTT3 is the only Theta class isoform possessing a glycine residue in the middle of helix 5 (Gly-142), this site was used to introduce the bend. Construction of the bent helix 5 was carried out manually based on the bent helix 5 of the hGSTO1 structure [4]. Since it is highly likely that mGSTT2-2 would utilize MS as a substrate, the mGSTT2 model structure was based on the menaphthylglutathione conjugate structure obtained for hGSTT2 [17]. After initial refinement, the quality of stereochemical parameters was assessed with PROCHECK [32]. Manual adjustment of side-chain rotomers was carried out to alleviate identified problem areas, followed by further rounds of model building and refinement, again using default settings. Comparison of the mGSTT1, mGSTT3 and hGSTT2 sequences illustrates a one-residue insertion in the helix 4/5 loop region and a four-residue deletion in the helix 8/9 proline-rich linker in the mGSTT1 and mGSTT3 sequences. These regions were defined as loops, thus are less reliable than the more conserved regions of the isoforms. The stereochemical quality of the final models is good: stereochemical parameters, such as $\varphi - \psi$ angles, side-chain χ angle values, peptide bond planarity, α-carbon tetrahedral distortions and non-bonded interactions, are all better than or within the allowed ranges according to PROCHECK [32]. The correctness of the models is supported by three-dimensional–one-dimensional scores, which do not fall below 0.07 [33].

RESULTS

Identification of mGSTT3

Several EST clones that differed from mGSTT1 and mGSTT2 were identified by database analysis. Two clones (AI385844 and AA871247) were obtained from the I.M.A.G.E. consortium for further study, and preliminary sequencing indicated that clone AI385844 contained the entire coding region and the most extensive 5′ non-coding sequence. This clone was termed pmGSTT3 and was completely sequenced on both strands. The cDNA

```
                                    gatctccgaagccactcctccctcaacctttttgtgagtggagggaaagaaagtg -401
aatttgagggcagttagcacggctaatctgcccctaaattcaccaatagagacgccggttgtgtcagtttttcctttact -321
gggttccgttcctcatttggctgaagcacggaggtacaggggccagtttggctccaaaccaagactcgggctcatagttc -241
aggttcatctgtttcgggggttaacttgcagtccggcactgattggcccaaagttcttgtggtcatccaatgggaaggaa -161
gattgggatcgagctccgatcgaggctaagaggcacttgaggagctgggacccagccgttcctgctcgccgttgcggctg -81
ctccggcttctgtccccagaggcagttcacccgaggctagttcccgaggtctgccagcggctactcccacagcctccgac -1

ATG GGT CTG GAG CTC TAC CTG GAC CTG ATG TCC CAA CCC TGC CGT GCT GTC TAC ATC TTC  +60
M   G   L   E   L   Y   L   D   L   M   S   Q   P   C   R   A   V   Y   I   F   20

GCC AAG AAG AAC GGC ATC CCC TTC CAG CTG CGT ACC ATA GAG CTG CTT AAA GGT CAG CAG  +120
A   K   K   N   G   I   P   F   Q   L   R   T   I   E   L   L   K   G   Q   Q   40

TAC ACA GAT TCC TTC GCC CAG GTG AAC CCT TTG AGG AAG GTG CCA GCT TTG AAG GAT GGG  +180
Y   T   D   S   F   A   Q   V   N   P   L   R   K   V   P   A   L   K   D   G   60

GAC TTC GTC TTG GCA GAG AGT GTG GCC ATC TTG TTG TAT CTG AGT AGA AAG TAC AAG GCC  +240
D   F   V   L   A   E   S   V   A   I   L   L   Y   L   S   R   K   Y   K   A   80

CCC GAC CAC TGG TAC CCT CAG GAT CTG CAG ACT CGA GCT CGT GTA GAC GAG TAC CTG GCA  +300
P   D   H   W   Y   P   Q   D   L   Q   T   R   A   R   V   D   E   Y   L   A   100

TGG CAG CAC ACA GCC CTG CGG AGT TGT TGC ACC AGG GCC ATG TGG CAG AAG ATG ATG TTC  +360
W   Q   H   T   A   L   R   S   C   C   T   R   A   M   W   Q   K   M   M   F   120

CCT GTG TTC CTG GGA CAG CCC GTA CCT CCT GAG ATG CTG GCA TCC ACC TTG GCT GAA CTG  +420
P   V   F   L   G   Q   P   V   P   P   E   M   L   A   S   T   L   A   E   L   140

GAT GGG TGC CTG CAG GTG CTA GAG GAC AAG TTC CTG CGG AAC CAG GCC TTC CTT ACA GGA  +480
D   G   C   L   Q   V   L   E   D   K   F   L   R   N   Q   A   F   L   T   G   160

TCC CAT ATT TCT GTG GCT GAC TTG GTG GCC ATC ACA GAG CTG ATG CAT CCT GTC AGT GCT  +540
S   H   I   S   V   A   D   L   V   A   I   T   E   L   M   H   P   V   S   A   180

GGC TGC AAA ATC TTT GAG AGC CGA CCC AAA CTG GCT GCC TGG CGT CAG AGA GTA GAA GCC  +600
G   C   K   I   F   E   S   R   P   K   L   A   A   W   R   Q   R   V   E   A   200

GAA GTG GGG GAG AGC CTC TTC CAG GAG GCC CAT GAA GTT GTC CTG AAG GCC AAG GAT ATG  +660
E   V   G   E   S   L   F   Q   E   A   H   E   V   V   L   K   A   K   D   M   220

CCT CCC TTG ATG GAC CCT GCC TTG AAA GAG AAG CTG AAG CTC TCT GTG CAG TGT TTG TTA  +720
P   P   L   M   D   P   A   L   K   E   K   L   K   L   S   V   Q   C   L   L   240

CAC TGA gggaatagctcgaactgggagacagacggtcaagggaaactcggggccgtgcaagcgaagagatgaggctgt +798
H   *                                                                           242

ctggctctttggctgtcactaggacatagggtctctggtccgcaatttcctctctgcatcacctgtgttccctctgtgtc +878
cccacctggcttgtgctgtaacatcagttcttatctgacttctggagagagctccaggaaaatacttgggagaaacgtat +958
ctggccatattccttttctttgaaacttacaaaccattgaaccttcccccatcctccaagccatatttgtatatacatg +1038
tatatacacatatacatataattgatatctggtaccacgtgaccctctgtgctcctcaaccctggtgcctcccttcctca +1118
gttttggcatgcaggatctatgatctctcagttgaatacatgcaagagacggggattgactagtgaaagacctttgtcct +1198
ctgcgtctttgcttcctgccataggtggggtgcagggctggttcctttgtgggagtggaagttgcttggtatgtcctgac +1278
tttccaaattcttgtcccaagcattctaaccctgactgctgccttctacctgtctccggctccaaactgtcaagctactt +1358
aatcttgtgccataagacgtcttttatgtgccttacctaataaatacattgttaacaga(n=25)                +1417
```

Figure 1 Nucleotide sequence and deduced amino acid sequence of mGSTT3

The majority of the sequence (−65 to +1417 bp) is derived from the EST clone AI385844, and the remaining 5′ non-coding sequence is derived from the longest EST clone (AI605608). The complete cDNA sequence has GenBank® accession number AF508157. The stop codon is indicated by an asterisk and the poly(A)$^+$ addition signal is underlined. n = 25 represents the number of adenines in the poly(A)$^+$ tail.

in pmGSTT3 is 1858 bp long and consists of 456 bp of 5′ non-coding sequence, 723 bp of coding sequence and 694 bp of 3′ non-coding sequence. The encoded protein consists of 241 amino acid residues and has a deduced molecular mass of 27385 Da (Figure 1). Comparison of the mGSTT3 sequence with the mGSTT1 and mGSTT2 sequences described previously [16] and shown in Figure 2 revealed 69% identity with mGSTT1 and 52% identity with mGSTT2.

Characterization of the *mGSTT3* gene

The BLAST [25] program was used to search the GenBank® database with the mGSTT3 cDNA. This search identified a bacterial artificial chromosome (BAC) sequence (AC009361) that contained the *mGSTT3* gene, with the exception of the 5′ flanking region, exon 1 and part of intron 1. The size and sequence of the putative intron 1 was determined by amplification of DNA from Balb/c genomic DNA. The pooled data indicate that the gene is composed of five exons and four introns and has a similar organization to that of the *mGSTT1* and *mGSTT2* genes (Figures 3 and 4) [16,34]. The BAC containing the partial *mGSTT3* gene maps to chromosome 10 B5-C1, the same region to which we have mapped the *mGSTT1* and *mGSTT2* genes previously [16,34]. However, the *mGSTT1* and *mGSTT2* genes are not included in the 143 kb of BAC sequence downstream of *mGSTT3*. Further database searches have identified two draft sequence contigs from chromosome 10 (GenBank® accession numbers AC068241 and AC007961),

```
                    10        20        30        40        50        60
                     .         .         .         .         .         .
                     ↓
hGSTT2-p    MGLELFLDLVSQPSRAVYIFAKKNGIPLELRTVDLVKGQHKSKEFLQINSLGKLPTLKDG
mGSTT2-p    MGLELYLDLLSQPSRAVYIFAKKNGIPFQTRTVDILKGQHMSEQFSQVNCLNKVPVLKDG
mGSTT1-p    MVLELYLDLLSQPCRAVYIFAKKNNIPFQMHTVELRKGEHLSDAFARQVNMKKVPAMMDG
mGSTT3-p    MGLELYLDLMSQPCRAVYIFAKKNGIPFQLRTIELLKGQQYTDSFAQVNPLRKVPALKDG

hGSTT2-p    DFILTESSAILIYLSCKYQTPDHWYPSDLQARARVHEYLGWHADCIRGTFGIPLWVQVLG
mGSTT2-p    SFVLTESTAILIYLSSKYQVADHWYPADLQARAQVDEYLGWHADNIRGTFGVLLWTKVLG
mGSTT1-p    GFTLCESVAILLYLAHKYKVPDHWYPQDLQARARVDEYLAWQHTGLRRSCLRALWHKVMF
mGSTT3-p    DFVLAESVAILLYLSRKYKAPDHWYPQDLQTRARVDEYLAWQHTALRSCCTRAMWQKMMF

hGSTT2-p    PLIGVQVPE-EKVERNRTAMDQALQWLEDKFLGDRPFLAGQQVTLADLMALEELMQPVAL
mGSTT2-p    PLIGVQVPQ-EKVERNRDRMVLVLQQLEDKFLRDRAFLVGQQVTLADLMSLEELMQPVGL
mGSTT1-p    PVFLGEQIPPETLAATLAELDVNLQVLEDKFLQDKDFLVGPHISLADLVAITELMHPVGG
mGSTT3-p    PVFLGQPVPPEMLASTLAELDGCLQVLEDKFLRNQAFLTGSHISVADLVAITELMHPVSA

hGSTT2-p    GYELFEGRPRLAAWRGRVEAFLGAELCQEAHSIILSILEQAAKKTLPTPSPEAYQAMLLR
mGSTT2-p    GYNLFEGRPQLTAWRERVEAFLGAELYQEAHSTILSILGQAAKKMLPVPPPEVHASMQLR
mGSTT1-p    GCPVFEGHPRLAAWYQRVEAAVGKDLFREAHEVILKVKDCP-----PADLIIKQKLMPRV
mGSTT3-p    GCKIFESRPKLAAWRQRVEAEVGESLFQEAHEVVLKAKDMP-----PLMDPALKEKLKLS

hGSTT2-p    IARIP
mGSTT2-p    IARIP
mGSTT1-p    LTMIQ
mGSTT3-p    VQCLLH
```

Figure 2 Alignment of amino acid sequences from the mouse Theta class GSTs with the hGSTT2 sequence

The alignment is based on that published previously in association with the model of hGSTT1 [18], which is based on the hGSTT2 crystal structure [17]. The alignment of the C-terminus of the mGSTT3 isoform has been defined through comparative modelling. mGSTT3 has 69% identity with GSTT1 and 52% identity with mGSTT2. The conserved active site Ser-11 is indicated (↓).

which confirm the presence of *mGSTT1* immediately upstream of *mGSTT3* and indicates that *mGSTT2* is further upstream of *mGSTT1*. We have used the available sequence in these contigs and BAC AC009361 to construct a map of the Theta genes on mouse chromosome 10 (Figure 4). It is interesting to note that the D-dopachrome tautomerase (*DDCT*) gene that is associated with this gene complex in humans is also present in the mouse Theta gene cluster [9]. It should be noted that this interpretation of the genomic organization of the murine Theta class genes is based in part on draft sequence contigs that may be subject to change during refinement of the genome sequence data.

Tissue expression of mouse Theta class GSTs

To determine the tissues in which Theta class GSTs are expressed, a Northern blot of mRNA was sequentially hybridized with gene-specific cDNA probes. DNA derived from bases 692–1021 of the mGSTT3 cDNA was used to detect transcripts of *mGSTT3*. This region was selected because of its low degree of sequence identity with the *mGSTT1* and *mGSTT2* genes. Similar probes were selected to identify the mGSTT1 and mGSTT2 sequences. Southern blots of mouse genomic DNA hybridized with cDNA probes for mGSTT1, mGSTT2 or mGSTT3 showed hybridization to different genomic fragments, indicating that there is no cross-hybridization between the different genes and that the hybridization to Northern blots is gene specific (results not shown). As shown in Figure 5, there was a relatively high level of message for each isoenzyme in the liver and moderate levels of expression in the kidney. In general, mGSTT3 appears to be expressed more widely than mGSTT1 or mGSTT2, with at least trace levels observed in most tissues. It is also notable that, in contrast with mGSTT1 or mGSTT2, there was a moderate level of mGSTT3 expression in the testis.

Characterization of recombinant mGSTT3-3

The coding region of mGSTT3 was cloned into the plasmid pQE30 that allows expression of recombinant proteins with an N-terminal 6 × His tag, permitting rapid purification of the expressed protein on nickel–nitrilotriacetate–agarose. We have previously used this approach for the expression and purification of mGSTT1-1 and hGSTT2-2 [15,16] and demonstrated that the N-terminal extension does not significantly alter activity. In addition, since the N-terminal is spatially well removed from the active site, it would not be expected to interfere with catalysis. The purified mGSTT3-3 was assayed for activity with a range of compounds that are substrates for other GSTs. The data compiled in Table 1 allow the comparison of the values obtained for mGSTT3-3 with activities reported for GSTT1-1 and GSTT2-2 from several species. Of particular note, mGSTT3-3 had no detectable activity with either DCM or MS, which are highly specific substrates for GSTT1-1 and GSTT2-2 respectively. Although EPNP is a significant substrate for GSTT1 isoenzymes and is not utilized by GSTT2 isoenzymes, mGSTT3-3 exhibits a modest level of activity. It is also notable that mGSTT3-3 utilized 4-NBC at a higher rate than either the GSTT1-1 or GSTT2-2 isoenzymes.

Molecular modelling of mGSTT1, mGSTT2 and mGSTT3

To gain a greater understanding of the structural features that contribute to variations in the substrate selectivity between the

```
gatctccgaagccactcctccctcaacctttttttgtgagtggagggaaagaaagtgaatt      -397
tgagggcagttagcacggctaatctgcccctaaattcaccaatagagacgccggttgtgt      -337
cagtttttcctttactgggttccgttcctcatttggctgaagcacggaggtacaggggcc      -277
agtttggctccaaaccaagactcgggctcatagttcaggttcatctgtttcgggggttaa      -217
cttgcagtccggcactgattggcccaaagttcttgtggtcatccaatgggaaggaagatt      -157
gggatcgagctccgatcgaggctaagaggcacttgaggagctgggacccagccgttcctg      -97
ctcgccgttgcgggctgctccggcttctgtccccagaggcagttcacccgaggctagttcc      -37
cgaggtctgccagcggctactcccacagcctccgacATG GGT CTG GAG CTC TAC       +18
                                  Ex1   M   G   L   E   L   Y       +6
CTG GAC CTG ATG TCC CAA CCC TGC CGT GCT GTC TAC ATC TTC GCC       +63
L   D   L   M   S   Q   P   C   R   A   V   Y   I   F   A       +21
AAG AAG AAC GGC ATC CCC TTC CAG CTG CGT ACC ATA GAG CTG CTT       +108
K   K   N   G   I   P   F   Q   L   R   T   I   E   L   L       +36
AAA G gtgggctcagttggcaggct.....2204bp.....ctctggtctct            +2748
K   G                                                            +38
ttcctcag GT CAG CAG TAC ACA GAT TCC TTC GCC CAG GTG AAC CCT       +2794
     Ex2    Q   Q   Y   T   D   S   F   A   Q   V   N   P       +50
TTG AGG AAG GTG CCA GCT TTG AAG GAT GGG GAC TTC GTC TTG GCA       +2839
L   R   K   V   P   A   L   K   D   G   D   F   V   L   A       +65
GAG AGgt aacatgttccctgcttgc......1166bp......tctctcttcattcc      +4043
E   S                                                            +67
tccag T GTG GCC ATC TTG TTG TAT CTG AGT AGA AAG TAC AAG GCC       +4090
  Ex3   V   A   I   L   L   Y   L   S   R   K   Y   K   A       +80
CCC GAC CAC TGG TAC CCT CAG GAT CTG CAG ACT CGA GCT CGT GTA       +4135
P   D   H   W   Y   P   Q   D   L   Q   T   R   A   R   V       +95
GAC GAG TAC CTG GCA TGG CAG CAC ACA GCC CTG CGG AGT TGT TGC       +4180
D   E   Y   L   A   W   Q   H   T   A   L   R   S   C   C       +110
ACC AGG GCC ATG TGG CAG AAG gtgagttgtggggatgctcc.............     +4221
T   R   A   M   W   Q   K                                        +117
....656bp.....................ttcttgctggtccatcag ATG ATG TTC      +4906
                                          Ex4   M   M   F        +120
CCT GTG TTC CTG GGA CAG CCC GTA CCT CCT GAG ATG CTG GCA TCC       +4951
P   V   F   L   G   Q   P   V   P   P   E   M   L   A   S       +135
ACC TTG GCT GAA CTG GAT GGG TGC CTG CAG GTG CTA GAG GAC AAG       +4996
T   L   A   E   L   D   G   C   L   Q   V   L   E   D   K       +150
TTC CTG CGG AAC CAG GCC TTC CTT ACA GGA TCC CAT ATT TCT GTG       +5041
F   L   R   N   Q   A   F   L   T   G   S   H   I   S   V       +165
GCT GAC TTG GTG GCC ATC ACA GAG CTG ATG CAT gtga gtgtcgtggag      +5089
A   D   L   V   A   I   T   E   L   M   H                        +176
ttgga..........820bp...........tgactcct tctccccacag CCT GTC       +5940
                                            Ex5   P   V          +178
AGT GCT GGC TGC AAA ATC TTT GAG AGC CGA CCC AAA CTG GCT GCC       +5985
S   A   G   C   K   I   F   E   S   R   P   K   L   A   A       +193
TGG CGT CAG AGA GTA GAA GCC GAA GTG GGG GAG AGC CTC TTC CAG       +6030
W   R   Q   R   V   E   A   E   V   G   E   S   L   F   Q       +208
GAG GCC CAT GAA GTT GTC CTG AAG GCC AAG GAT ATG CCT CCC TTG       +2075
E   A   H   E   V   V   L   K   A   K   D   M   P   P   L       +223
ATG GAC CCT GCC TTG AAA GAG AAG CTG AAG CTC TCT GTG CAG TGT       +6120
M   D   P   A   L   K   E   K   L   K   L   S   V   Q   C       +238
TTG TTA CAC TGA gggaatagctcgaactgggagacagacggtcaagggaaactcgg      +6176
L   L   H   *                                                    +241
ggccgtgcaagcgaagagatgaggctgtctggctctttggctgtcactaggacatagggt      +6236
ctctggtccgcaatttcctctctgcatcacctgtgttccctctgtgtccccacctggctt      +6296
gtgctgtaacatcagttcttatctgacttctggagagagctccaggaaaatacttgggag      +6356
aaacgtatctggccatattccttttctttgaaactttacaaaccattgaaccttccccca      +6416
tcctccaagccatatttgtatatacatgtatatacacatatacatataattgatatctgg      +6476
taccacgtgaccctctgtgctcctcaaccctggtgcctcccttcctcagttttggcatgc      +6536
aggatctatgatctctcagttgaatacatgcaagagacggggattgactagtgaaagacc      +6596
tttgtcctctgcgtctttgcttcctgccataggtggggtgcagggctggttcctttgtgg      +6656
gagtggaagttgcttggtatgtcctgactttccaaattcttgtcccaagcattctaaccc      +6716
tgactgctgccttctacctgtctccggctccaaactgtcaagctacttaatcttgtgcca      +6776
taagacgtcttttatgtgccttacctaataaa                                 +6808
```

Figure 3 Nucleotide sequence and intron/exon structure of *mGSTT3*

The nucleotides are numbered from the ATG translation start site to the poly(A)$^+$ signal (underlined). All splice sites obey the GT/AG rule and exons are indicated Ex1–Ex5. The bases in the exons and introns are in upper- and lower-case letters respectively. The stop codon is indicated by an asterisk.

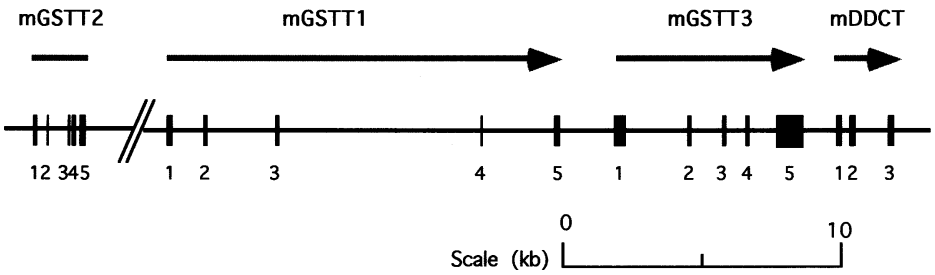

Figure 4 Organization of the mouse Theta-class *GST* and *DDCT* gene cluster

A schematic diagram to show the relative positions and the orientation of the genes. The exons are numbered 1–5. The orientation of *mGSTT2* is unknown and its distance upstream of *mGSTT1* is less than 140 kb.

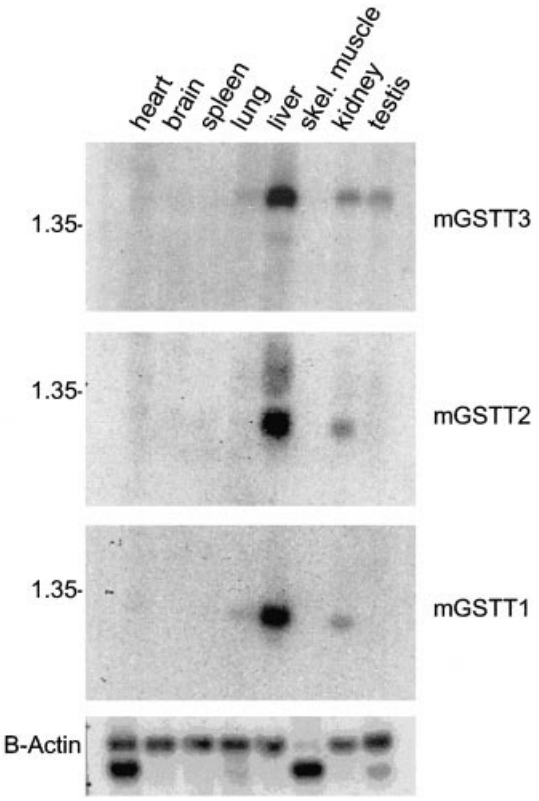

Figure 5 Tissue distribution of mouse Theta class GST mRNA expression

The blotted RNA was sequentially hybridized with a cDNA probes encoding mGSTT1, mGSTT2 and mGSTT3, as described in the text. The blot was subsequently hybridized with a β-actin cDNA to evaluate RNA loading.

different murine Theta class GSTs, we constructed atomic models of mGSTT1, mGSTT2 and mGSTT3 on the basis of the crystal structure of hGSTT2 and a homology model of hGSTT1. Views of the modelled structures highlighting similarities and differences between the three murine enzymes in the active site are shown in Figure 6. Detailed comparisons of the residues co-ordinating GSH (G-site) and residues contributing to the hydrophobic substrate-binding site (H-site) are shown in Tables 2 and 3. As can be expected from the high sequence identity of the four isoforms over the N-terminal domain, many of the residues involved in co-ordinating GSH in the G-site are conserved. Specifically, the Glu-66 OE1 and OE2 carboxylates, the Ser-67 backbone amide and side-chain hydroxyl, as well as the amino side chain of Arg-107, are involved in contacts with the polar and charged groups on the γ-glutamyl moiety of GSH. The backbone carbonyl and amide groups of Val-54 form a β-sheet-like interaction with the backbone of the cysteinyl residue, whereas the thiol hydrogen bonds to the hydroxyl group of the catalytically important Ser-11 [15]. Thus all three mouse isoforms appear to utilize an N-terminal region serine hydroxyl to stabilize the reactive GSH thiolate anion. Notably, in mGSTT2, like the template structure, the free amide group of the γ-glutamyl moiety of GSH is co-ordinated to Asp-104 from the opposite subunit. The substitution of Asp-104 for a threonine residue in the mGSTT1 and mGSTT3 sequences indicates that a similar interaction is not observed in these isoforms. As can be seen from Table 2, the co-ordination of the GSH glycyl moiety differs significantly between the three modelled structures, as well as the hGSTT2 template structure. Significantly, an interaction between position 41 and the glycyl carboxyl group is only found in hGSTT2 and mGSTT3 (lysine and tyrosine residues respectively). However, it is possible for the C-terminal helix 9 residues, including Lys-230 and/or Arg-234 in mGSTT1 and Ser-235 in mGSTT2, to interact with the glycyl carboxyl, thus maintaining a co-ordination point. The mGSTT3 C-terminal helix 9 residue Lys-231, which is equivalent to the mGSTT1 Lys-230, provides another possible co-ordination point for the glycyl carboxyl in this isoform. These differences in co-ordination of the glycyl moiety may facilitate differences in GSH affinity for each of the isoforms, due to alterations in the nature of the contact and possibly slight changes in the position of GSH in the G-site. The replacement of the mGSTT1, mGSTT2 and hGSTT2 His-40 with Gln in mGSTT3 appears to have been necessary in order to avoid unfavourable steric interactions with Tyr-41.

Comparison of the hGSTT2 template and the three modelled mouse H-sites illustrate changes between the isoforms that provide some structural basis for the observed difference in substrate specificities (Table 3). Very little difference occurs between the mGSTT2 and hGSTT2 in the H-sites. The sulphate-binding pocket, consisting of Gln-12, Arg-107, Gly-108, Gly-111, Trp-115 and Arg-239 is conserved, consistent with the potential of mGSTT2 to bind and conjugate MS. However, the domain 1 substitutions of Leu-10 to Val and Leu-36 to Val between mGSTT2 and hGSTT2 decrease the volume of the hydrophobic pocket associated with the co-ordination of the menaphthyl ring by two methylene groups in mGSTT2. The other residues

Table 1 Specific activity of Theta class GSTs with a range of substrates

All results are expressed as μmol/min per mg of protein and are means ± S.D. The data for mGSTT3-3 are the means ± S.D. of at least three determinations. ND, not detected.

Substrate	mGSTT1*	rGSTT1†	rGSTT2‡	hGSTT2§	mGSTT3
CDNB	0.03 ± 0.01	< 0.5	0.7 ± 0.1	ND	0.52 ± 0.01
DCNB	ND		ND	ND	ND
EA	ND		ND	0.29 ± 0.02	0.017 ± 0.005
EPNP	90.6 ± 16.2	180	0.9 ± 0.1	ND	5.48 ± 0.77
NBD-Cl	0.03 ± 0.01			0.057 ± 0.004	0.021 ± 0.001
4-NBC	8.2 ± 1.0	86	3.7 ± 0.2		18.26 ± 1.70
4-NPA	0.09 ± 0.01				ND
MS	0.021 ± 0.004		4.3 ± 0.3	0.317 ± 0.028	ND
Cumene hydroperoxide	2.9 ± 0.1	41	2.4 ± 0.3	6.89 ± 0.48	1.13 ± 0.03
tert-Butyl hydroperoxide	1.7 ± 0.2			0.406 ± 0.042	ND
DCM	4.6 ± 0.2	11		ND	ND
trans-Oct-2-enal	ND			0.068 ± 0.005	ND
trans-Non-2-enal	ND			0.119 ± 0.010	ND
trans,*trans*-Nona-2-enal	ND			0.050 ± 0.003	ND
trans,*trans*-Deca-2-enal	ND			0.113 ± 0.015	ND

* Data from [16]; † data from [11]; ‡ data from [37]; and § data from [38].

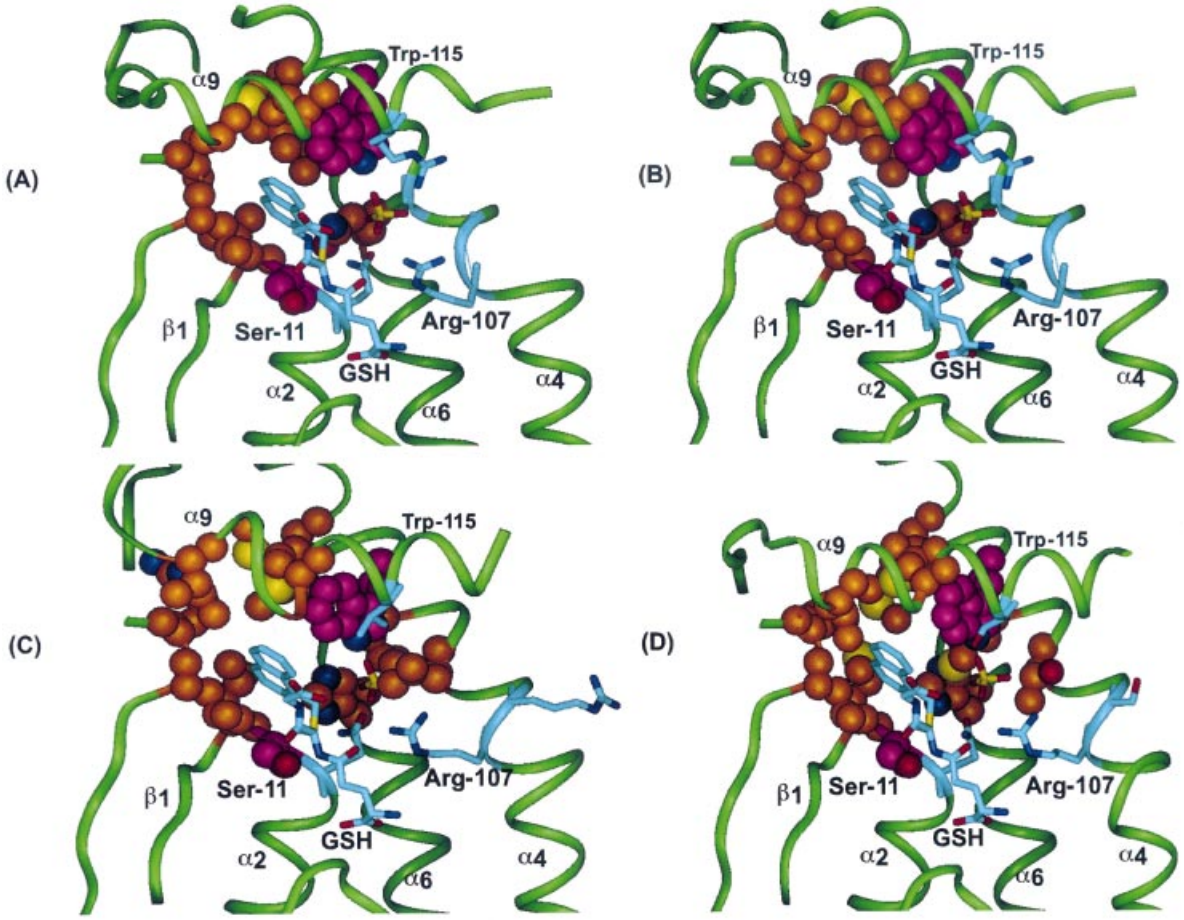

Figure 6 Comparison of the active-site models of mGSTT1, mGSTT2 and mGSTT3 with the crystal structure of hGSTT2

Ribbon diagrams of the active sites of (**A**) the template structure, hGSTT2-2, and the models of mGSTT2-2 (**B**), mGSTT1-1 (**C**) and mGSTT3-3 (**D**). In each diagram, menaphthylglutathione and MS are shown in a stick representation. The residues represented in the brown CPK (Corey–Pauling–Koltun) model include the H-site residues from Table 3 that line the hydrophobic substrate-binding pocket associated with the naphthalene ring of the menaphthylglutathione present in the hGSTT2-2 structure. The size of the H-site decreases in volume in all structures in comparison with the template structure, and a distinct change in H-site topology between the four structures is also apparent. Residues Ser-11 and Trp-115 are represented as a purple CPK model and form a line of division between the hydrophobic pocket and the sulphate-binding pocket of the Theta class H-site. The residues defining the sulphate-binding pocket are represented as cyan sticks. The composition of the sulphate-binding pocket changes dramatically between h/mGSTT2-2 and the mGSTT1-1 and mGSTT3-3 structures. The residue at position 111 in the mGSTT1-1 and mGSTT3-3 structures appears to be more associated with the H-site than the sulphate-binding site, and thus it is represented as a brown CPK model.

Table 2 Residues co-ordinating GSH in molecular models of murine Theta class GSTs and hGSTT2

GSH moiety	mGSTT1	mGSTT2	mGSTT3	hGSTT2
Glycyl	Arg-234	Leu-238	Leu-234	Leu-238
Glycyl	Lys-230	Ser-235	Lys-231	Ala-235
Glycyl	His-40	His-40	Gln-40	His-40
Glycyl	Leu-41	Met-41	Tyr-43	Lys-41
Glycyl	Lys-53	Lys-53	Lys-53	Lys-53
Cysteinyl	Ser-11	Ser-11	Ser-11	Ser-11
Cysteinyl	Val-54	Val-54	Val-54	Val-54
γ-Glutamyl	Glu-66	Glu-66	Glu-66	Glu-66
γ-Glutamyl	Ser-67	Ser-67	Ser-67	Ser-67
γ-Glutamyl	Arg-107	Arg-107	Arg-107	Arg-107
γ-Glutamyl	Thr-104*	Thr-104*	Asp-104*	Thr-104*

* Contact from the second subunit in the dimer.

Table 3 Residues contributing to the H-site in molecular models of murine Theta class GSTs and hGSTT2

mGSTT1	mGSTT2	mGSTT3	hGSTT2
Leu-10	Leu-10	Met-10	Val-10
Gln-12*	Gln-12*	Gln-12*	Gln-12*
Leu-35	Ile-35	Leu-35	Leu-35
Arg-36	Leu-36	Leu-36	Val-36
Arg-107*	Arg-107*	Arg-107*	Arg-107*
–	Gly-108*	–	Gly-108*
Leu-111*	Gly-111*	Thr-111*	Gly-111*
Leu-114	Leu-114	Met-114	Leu-114
Trp-115*	Trp-115*	Trp-115*	Trp-115*
Val-118	Val-118	Met-118	Val-118
Met-119	Leu-119	Met-119	Leu-119
His-176	Gln-176	His-176	Gln-175
Ile-227	Val-232	Leu-228	Ala-232
Leu-231	–	–	–
Met-232	Met-236	Leu-232	Met-236
Val-235*	Arg-239*	Ser-235*	Arg-239*

* Residues defining the sulphate-binding pocket in hGSTT2-2.

that define the walls of the buried active-site cavity remain unchanged.

Significantly, the low activity of mGSTT1 towards MS can be primarily attributed to the disruption of the sulphate-binding pocket facilitated by the hGSTT2 to mGSTT1 substitutions Gly-108 to Arg and Gly-111 to Leu. These substitutions, in particular, are proposed to alter the confirmation of the central turns of helix 4, resulting in the loss of position 108 as a contributor to the H-site, as well as repositioning residue 111 in the sulphate-binding cavity of the GSTT2 isoforms. In addition, Arg-239 in the C-terminal helix 9 of hGSTT2 is replaced by Val-235 in mGSTT1, resulting in the loss of a sulphate anion co-ordination site. Although the hydrophobic pocket co-ordinating the menaphthyl ring in the GSTT2 isoforms can still be defined in mGSTT1, the hGSTT2 to mGSTT1 substitutions in this region decrease its volume, potentially reducing the ability of double-ringed structures to bind mGSTT1. In total, the substitutions between hGSTT2 and mGSTT1, including Val-10 to Leu, Gly-111 to Leu, Gln-175 to His-176, Ala-232 to Ile-227 and the addition of Leu-231, decrease the H-site cavity size by 12 methylene groups. The hydrophobicity of the mGSTT1 active site is also increased by the mGSTT2 to mGSTT1 substitutions Arg-239 to Val-235 and Arg-242 to Met-239 in helix 9 at the C-terminus. In contrast, the Gln-175 to His-176 substitution may provide an extra charge at the base of the mGSTT1 H-site. This residue has been proposed previously [18] to be involved in the dehalogenase and epoxide ring opening activities observed for the hGSTT1 isoform, thus it is possible that the mGSTT1 may utilize the same mechanism for these reactions.

The substrate preferences observed for mGSTT3 are most like those of mGSTT1, implying a relationship between the active sites of the two isoforms. The most obvious similarity between mGSTT1 and mGSTT3 is the decrease in volume of the active-site cavity by ten methylene groups when mGSTT3 is compared with the template structure hGSTT2. The hGSTT2 to mGSTT3 substitutions responsible are Val-10 to Met, Val-36 to Leu, Gly-111 to Thr, Leu-114 to Met, Val-118 to Met, Gln-175 to His-176 and Ala-232 to Leu-228. In addition, the hGSTT2 to mGSTT3 substitutions Gly-108 to Ser and Gly-111 to Thr are likely to facilitate a disruption of the sulphate-binding pocket found in GSTT2 isoforms in a manner similar to Arg-108 and Leu-111 in mGSTT1. The loss of the sulphate-binding pocket occurs in combination with the hGSTT2 to mGSTT3 substitution, Arg-239 to Ser-235, which removes the C-terminal helix sulphate co-ordination point. The mGSTT3 to mGSTT1/mGSTT2/hGSTT2 substitutions Met-10 to Leu/Leu/Val, Leu-36 to Arg/Leu/Val, Met-114 to Leu/Leu/Leu, Met-118 to Val/Val/Val, Met-119 to Met/Leu/Leu and Leu-228 to Ile/Val/Ala indicate that the hydrophobic pocket co-ordinating the MS naphthalene ring in hGSTT2 is at its smallest in mGSTT3. This would seem to provide a mechanism for the discrimination against larger Theta class substrates, like MS, as well as causing a possible shift in the binding site of other aromatic Theta class substrates like EA, EPNP, NBD-Cl, 4-NBC and 4-NPA. The substitutions between mGSTT3 and mGSTT1 of Met-10 to Leu, Thr-111 to Leu and Met-118 to Val alter the active-site topology around His-176, the proposed site of interaction of DCM and EPNP in the hGSTT1 isoform [18]. In particular, the mGSTT3 Met substitutions may prevent the involvement of His-176 in the conjugation reaction. This is supported to an extent by the absence of mGSTT3-mediated conjugating activity toward DCM and a severe reduction toward EPNP.

DISCUSSION

Previous studies in a number of laboratories have identified Theta class GSTs in mice, rats and humans and have shown that they can be grouped into either the GSTT1 or GSTT2 subclasses [11–16]. The GSTT1 and GSTT2 isoenzymes are characterized by distinct substrate preferences. As the Theta class GSTs are not readily purified by glutathione-affinity chromatography, it is possible that there are additional members of this class that have remained undiscovered, because they do not catalyse reactions with known substrates. Our previous examination [16] of mouse EST sequences identified several that appeared to be similar but distinct from those encoding GSTT1 and GSTT2 transcripts, suggesting the existence of an additional Theta class GST gene. The present study has confirmed the existence of an *mGSTT3* gene that maps to the same region of chromosome 10 to which we have previously mapped the *mGSTT1* and *mGSTT2* genes [16,34]. The *mGSTT3* gene appears to have a similar structure to that of the *mGSTT1* and *mGSTT2* genes. The clustering of all members of the mouse Theta class within a distinct region of chromosome 10 is consistent with the class specific distribution of GST genes commonly found in mammals. Although we

scanned the human EST and genome databases and failed to find evidence for an additional Theta class locus in humans, we noted a significant difference in the arrangement of the Theta class genes between humans and mice. The *hGSTT2* and *DDCT* genes are located together but in different orientations [9]. Furthermore, in humans the *GSTT2* and *DDCT* genes are duplicated, giving rise to a *GSTT2* pseudogene. In the mouse, we identified the *mDDCT* gene approx. 750 bp downstream of the *mGSTT3* and not in proximity of the *mGSTT2* gene. Consequently, it is clear that there have been significant rearrangements within the Theta class gene clusters between humans and mice, and this may be indicative of additional variation in other species.

It is evident from the Northern blot that GSTT3 is expressed in a number of tissues, although the predominant expression is clearly in the liver. This is consistent with a role for mGSTT3-3 in detoxification processes. The expression of mGSTT3 in the testis is a notable departure from the expression patterns observed for mGSTT1 and mGSTT2 and implies a metabolic role or an endogenous substrate that differs from the other Theta class GSTs. Recombinant GSTT3-3 expressed in *E. coli* had no activity with the GSTT1-1-specific substrate DCM and no activity with the GSTT2-2-specific substrate MS, demonstrating that it is functionally distinct. Although mGSTT3-3 has relatively high activity with 4-NBC, this activity is not specific, as both GSTT1- and GSTT2-type enzymes have been shown to utilize 4-NBC as a substrate (Table 1). More extensive studies will need to be undertaken to determine the normal substrates for mGSTT3-3. In humans, the *GSTT1* gene is frequently deleted without any immediately obvious consequence [14,35]. Similarly, despite its relatively high expression in mouse testis, it would appear that GSTT3-3 may also be dispensable, as it is not present in humans. Although the Northern blots are not strictly quantitative, the relative levels of Theta class GST gene expression in the liver appears to be similar, suggesting that the different isoenzymes are approximately equally represented.

Homology modelling of the three mouse isoenzymes has allowed the detailed comparison of their active sites. The modelled structures are compatible with our previous observations that the conserved N-terminal region serine residue is responsible for the stabilization of the bound glutathione as a thiolate anion [15,36]. The reduction in volume of the H-sites in mGSTT1-1 and mGSTT3-3, compared with hGSTT2-2, appears to restrict access of bulky substrates. The restricted substrate access and the substitution of essential residues in the sulphate-binding pocket of hGSTT2-2 both contribute to an explanation for the inability of mGSTT1-1 and mGSTT3-3 to utilize MS as a substrate.

REFERENCES

1 Mannervik, B. and Widersten, M. (1995) Human glutathione transferases: classification, tissue distribution, structure and functional properties. In Advances in Drug Metabolism in Man (Pacifici, G. M. and Fracchia, G. N., eds.), pp. 407–460, The European Commission, Brussels

2 Hayes, J. D. and Pulford, D. J. (1995) The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit. Rev. Biochem. Mol. Biol. **30**, 445–600

3 Board, P. G., Baker, R. T., Chelvanayagam, G. and Jermiin, L. S. (1997) Zeta, a novel class of glutathione transferases in a range of species from plants to humans. Biochem. J. **328**, 929–935

4 Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V. et al. (2000) Identification, characterization and crystal structure of the Omega class glutathione transferases. J. Biol. Chem. **275**, 24798–24806

5 Dulhunty, A., Gage, P., Curtis, S., Chelvanayagam, G. and Board, P. (2001) The glutathione transferase structural family includes a nuclear chloride channel and a ryanodine receptor calcium release channel modulator. J. Biol. Chem. **276**, 3319–3323

6 Board, P. G. and Webb, G. C. (1987) Isolation of a cDNA clone and localization of human glutathione S-transferase 2 genes to chromosome band 6p12. Proc. Natl. Acad. Sci. U.S.A. **84**, 2377–2381

7 Xu, S., Wang, Y., Roe, B. and Pearson, W. R. (1998) Characterization of the human class Mu glutathione S-transferase gene cluster and the GSTM1 deletion. J. Biol. Chem. **273**, 3517–3527

8 Board, P. G., Webb, G. C. and Coggan, M. (1989) Isolation of a cDNA clone and localization of the human glutathione S-transferase 3 genes to chromosome bands 11q13 and 12q13–12q14. Ann. Hum. Genet. **53**, 205–213

9 Coggan, M., Whitbread, L., Whittington, A. and Board, P. (1998) Structure and organization of the human Theta-class glutathione S-transferase and D-dopachrome tautomerase gene complex. Biochem. J. **334**, 617–623

10 Blackburn, A. C., Woollatt, E., Sutherland, G. R. and Board, P. G. (1998) Characterization and chromosome location of the gene *GSTZ1* encoding the human Zeta class glutathione transferase and maleylacetoacetate isomerase. Cytogenet. Cell Genet. **83**, 109–114

11 Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M. and Ketterer, B. (1991) Theta, a new class of glutathione transferases purified from rat and man. Biochem. J. **274**, 409–414

12 Hussey, A. J. and Hayes, J. D. (1992) Characterization of a human class-Theta glutathione S-transferase with activity towards 1-menaphthyl sulphate. Biochem. J. **286**, 929–935

13 Hiratsuka, A., Sebata, N., Kawashima, K., Okuda, H., Ogura, K., Watabe, T., Satoh, K., Hatayama, I., Tsuchida, S., Ishikawa, T. et al. (1990) A new class of rat glutathione S-transferase Yrs–Yrs inactivating reactive sulfate esters as metabolites of carcinogenic arylmethanols. J. Biol. Chem. **265**, 11973–11981

14 Pemble, S., Schroeder, K. R., Spencer, S. R., Meyer, D. J., Hallier, E., Bolt, H. M., Ketterer, B. and Taylor, J. B. (1994) Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. Biochem. J. **300**, 271–276

15 Tan, K. L., Chelvanayagam, G., Parker, M. W. and Board, P. G. (1996) Mutagenesis of the active site of the human Theta-class glutathione transferase GSTT2–GSTT2: catalysis with different substrates involves different residues. Biochem. J. **319**, 315–321

16 Whittington, A., Vichai, V., Webb, G., Baker, R., Pearson, W. and Board, P. (1999) Gene structure, expression and chromosomal localization of murine Theta class glutathione transferase mGSTT1-1. Biochem. J. **337**, 141–151

17 Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelvanayagam, G., Tan, K. L., Board, P. G. and Parker, M. W. (1998) Human theta class glutathione transferase: the crystal structure reveals a sulfate-binding pocket within a buried active site. Structure (London) **6**, 309–322

18 Flanagan, J. U., Rossjohn, J., Parker, M. W., Board, P. G. and Chelvanayagam, G. (1998) A homology model for the human theta-class glutathione transferase T1-1. Proteins: Struct., Funct., Genet. **33**, 444–454

19 Flanagan, J. U., Rossjohn, J., Parker, M. W., Board, P. G. and Chelvanayagam, G. (1999) Mutagenic analysis of conserved arginine residues in and around the novel sulfate binding pocket of the human Theta class glutathione transferase T2-2. Protein Sci. **8**, 2205–2212

20 Jemth, P. and Mannervik, B. (1999) Fast product formation and slow product release are important features in a hysteretic reaction mechanism of glutathione transferase T2-2. Biochemistry **38**, 9982–9991.

21 Jemth, P. and Mannervik, B. (2000) Active site serine promotes stabilization of the reactive glutathione thiolate in rat glutathione transferase T2-2. Evidence against proposed sulfatase activity of the corresponding human enzyme. J. Biol. Chem. **275**, 8618–8624

22 Caccuri, A. M., Antonini, G., Board, P. G., Flanagan, J., Parker, M. W., Paolesse, R., Turella, P., Federici, G., Lo Bello, M. and Ricci, G. (2001) Human glutathione transferase T2-2 discloses some evolutionary strategies for optimization of substrate binding to the active site of glutathione transferases. J. Biol. Chem. **276**, 5427–5431

23 Caccuri, A. M., Antonini, G., Board, P. G., Flanagan, J., Parker, M. W., Paolesse, R., Turella, P., Chelvanayagam, G. and Ricci, G. (2001) Human glutathione transferase T2-2 discloses some evolutionary strategies for optimization of the catalytic activity of glutathione transferases. J. Biol. Chem. **276**, 5432–5437

24 Pearson, W. R. (1994) Using the FASTA program to search protein and DNA sequence databases. Methods Mol. Biol. (Totowa, N.J.) **25**, 365–389

25 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402

26 Habig, W. H., Pabst, M. J. and Jakoby, W. B. (1974) Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. J. Biol. Chem. **249**, 7130–7139

27 Gillham, B. (1973) The mechanism of the reaction between glutathione and 1-menaphthyl sulphate catalysed by a glutathione S-transferase from rat liver. Biochem. J. **135**, 797–804
28 Lawrence, R. A. and Burk, R. F. (1976) Glutathione peroxidase activity in selenium-deficient rat liver. Biochem. Biophys. Res. Commun. **71**, 952–958
29 Ricci, G., Caccuri, A. M., Lo Bello, M., Pastore, A., Piemonte, F. and Federici, G. (1994) Colorimetric and fluorometric assays of glutathione transferase based on 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole. Anal. Biochem. **218**, 463–465
30 Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. **234**, 779–815
31 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**, 4876–4882
32 Laskowski, R. A., Moss, D. S. and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures. J. Mol. Biol. **231**, 1049–1067
33 Luthy, R., Bowie, J. U. and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. Nature (London) **356**, 83–85
34 Whittington, A. T., Webb, G. C., Baker, R. T. and Board, P. G. (1996) Characterization of a cDNA and gene encoding the mouse theta class glutathione transferase mGSTT2 and its localization to chromosome 10B5-C1. Genomics **33**, 105–111
35 Chenevix-Trench, G., Young, J., Coggan, M. and Board, P. (1995) Glutathione S-transferase M1 and T1 polymorphisms: susceptibility to colon cancer and age of onset. Carcinogenesis **16**, 1655–1657
36 Caccuri, A. M., Antonini, G., Nicotra, M., Battistoni, A., Bello, M. L., Board, P. G., Parker, M. W. and Ricci, G. (1997) Catalytic mechanism and role of hydroxyl residues in the active site of theta class glutathione S-transferases. Investigation of Ser-9 and Tyr-113 in a glutathione S-transferase from the Australian sheep blowfly, *Lucilia cuprina*. J. Biol. Chem. **272**, 29681–29686
37 Jemth, P., Stenberg, G., Chaga, G. and Mannervik, B. (1996) Heterologous expression, purification and characterization of rat class Theta glutathione transferase T2-2. Biochem. J. **316**, 131–136
38 Tan, K. L. and Board, P. G. (1996) Purification and characterization of a recombinant human Theta-class glutathione transferase (GSTT2-2). Biochem. J. **315**, 727–732

Received 21 December 2001/3 May 2002; accepted 31 May 2002
Published as BJ Immediate Publication 31 May 2002, DOI 10.1042/BJ20011878