Rat glutathione S-transferase M4-4: an isoenzyme with unique structural features including a redox-reactive cysteine-115 residue that forms mixed disulphides with glutathione

Hailing CHENG*, Tatyana TCHAIKOVSKAYA†, Yen-Sheng L. TU*, Jason CHAPMAN*, Biao QIAN*, Wei-Mei CHING‡, Ming TIEN*, Jonathan D. ROWE†, Yury V. PATSKOVSKY†, Irving LISTOWSKY† and Chen-Pei D. TU*¹

*Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, U.S.A., †Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461, U.S.A., and ‡Viral and Rickettsial Diseases Department, Infectious Disease Directorate, US Naval Medical Research Center, 8901 Wisconsin Avenue, Bethesda, MD 20889, U.S.A.

Although the existence of the rat glutathione S-transferase (GST) M4 (*rGSTM4*) gene has been known for some time, the corresponding protein has not as yet been purified from tissue. A recombinant rGSTM4-4 was thus expressed in *Escherichia coli* from a chemically synthesized *rGSTM4* gene. The catalytic efficiency (k_{cat}/K_m) of rGSTM4-4 for the 1-chloro-2,4-dinitrobenzene (CDNB) conjugation reaction was 50–180-fold less than that of the well-characterized homologous rGSTM1-1, and the pH optimum for the same reaction was 8.5 for rGSTM4-4 as opposed to 6.5 for rGSTM1-1. Molecular-modelling studies predict that key substitutions in the helix α 4 region of rGSTM4-4 account for this pK_a difference. A notable structural feature of rGSTM4-4 is the Cys-115 residue in place of the Tyr-115 of other Mu-class GSTs. The thiol group of Cys-115 is redox-reactive and readily forms a mixed disulphide even with GSH; the S-

glutathiolated form of the enzyme is catalytically active. A mutated rGSTM4-4 (C115Y) had 6–10-fold greater catalytic efficiency than the wild-type rGSTM4-4. Trp-45, a conserved residue among Mu-class GSTs, is essential in rGSTM4-4 for both enzyme activity and binding to glutathione affinity matrices. Antibodies directed against either the unique C-terminal undecapeptide or tridecapeptide of rGSTM4 reacted with rat and mouse liver GSTs to reveal an orthologous mouse GSTM4-4 present at low basal levels but which is inducible in mouse liver. This subclass of rodent Mu GSTs with redox-active Cys-115 residues could have specialized physiological functions in response to oxidative stress.

Key words: glutathiolation, mass spectrometry, oxidative stress, structural model, synthetic gene.

INTRODUCTION

Glutathione S-transferases (GSTs, EC 2.5.1.18) are important detoxification enzymes that catalyse the conjugation of GSH to a variety of electrophilic substrates; some GSTs also have selenium-independent GSH peroxidase activities [1,2]. Mammalian GSTs are products of gene superfamilies [3–6]. The cytosolic GSTs are either homodimeric or heterodimeric proteins that are found in many species ranging from bacteria to humans [2]. The diversity of the GST genetic reservoir and the inducibility of many GSTs by xenobiotic compounds such as antioxidants, carcinogens, chemotherapeutic agents, barbiturates, plant herbicide safeners and other compounds, render GST systems part of an organism's defence network against chemical insults (for recent reviews see [7,8]). As many as eight classes of cytosolic GSTs, designated Alpha, Mu, Pi, Sigma, Theta, Zeta, Kappa and Omega have been identified in mammals [2,7,9].

The rat (r) Mu-class GST gene family consists of at least six different genes. The testis-specific *rGSTM5* is representative of one subfamily, and the rest (*rGSTM1–M4* and *rGSTM6* subunits) belong to the other subfamily [10,11]. We have previously characterized a genomic clone λ GTR15-2, which potentially encodes a new rat GST subunit [12,13]. The predicted amino acid

sequence of the putative rGSTM4 subunit (originally designated as GST-Y_{b4}, GenBank accession number AF 106661), is most closely related to that of the rGSTM1 subunit, and the deduced 3'-untranslated region (UTR) of the putative rGSTM4 mRNA also showed substantial homology to that of the rGSTM1 mRNA. Because the rGSTM4-4 protein has not yet been purified in an active form, the biochemical properties of this putative GST are unexplored. In this report, a catalytically active recombinant rGSTM4-4 protein has been characterized and shown to exhibit some unique properties. The enzyme closely resembles a predicted mouse GST isoenzyme, and can exist as a mixed disulphide of mono- and di-S-glutathiolated forms.

EXPERIMENTAL

Materials

Radioactive nucleotides were purchased from ICN (Irvine, CA, U.S.A.). Restriction enzymes were products of New England Biolabs (Beverly, MA, U.S.A.) or American Allied Biochemical (Aurora, CO, U.S.A.). T4 polynucleotide kinase and T4 DNA ligase were products of New England Biolabs and Promega (Madison, WI, U.S.A.), respectively. Oligonucleotides longer than 40 bases were purified through polyacrylamide gels by their

Abbreviations used: GST, glutathione S-transferase; r, rat; h, human; m, mouse; CDNB, 1-chloro-2,4-dinitrobenzene; GS-DNB, S-(2,4-dinitrophenyl)glutathione; BHA, 2(3)-t-butyl-4-hydroxyanisole; ESI-MS, electrospray ionization MS; UTR, untranslated region; TFA, trifluoroacetic acid; RMSD, root-mean-square deviation; G-site, GSH-binding domain.

¹ To whom correspondence should be addressed (e-mail unh@psu.edu).

```

1  GGGAAATTCAGGAGAACAGAAAT  ATG  GCT  ATG  ATA  CTG  GGA  TAT  TGG  AAC  GTT  CGC  GGA  CTT  ACT   64
1  M  A  M  I  L  G  Y  W  N  V  R  G  L  T   14

65  CAC  CCC  ATC  CGC  CTG  CTC  CTG  GAA  TAC  ACA  GAT  TCA  AAC  TAT  GAG  GAG  AAG  AGA  TAC  GTC  124
15  H  P  I  R  L  L  L  E  Y  T  D  S  N  Y  E  E  K  R  Y  V   34

125  ATG  GGG  GAT  GCC  CCC  AAC  TTT  GAC  AGA  AGC  CAG  TGG  CTG  AGT  GAG  AAA  TTC  AAT  CTT  GGC  184
35  M  G  D  A  P  N  F  D  R  S  Q  W  L  S  E  K  F  N  L  G   54

185  CTG  GAC  ATT  CCC  AAT  TTG  CCC  TAC  TTA  ATT  GAT  GGG  TCA  CAC  AAG  GTC  ACC  CAG  AGC  AAT  244
55  L  D  I  P  N  L  P  Y  L  I  D  G  S  H  K  V  T  Q  S  N   74

245  GCC  ATC  CTG  CGC  TAC  CTT  GGC  CGG  AAG  CAC  AAC  CTG  TGT  GGG  GAG  ACA  GAG  GAG  GAG  AGG  304
75  A  I  L  R  Y  L  G  R  K  H  N  L  C  G  E  T  E  E  E  R   94

305  ATT  CGT  GTG  GAC  ACT  TTG  GAG  AAC  CAG  GTC  ATG  GAC  ACC  CGC  ATA  CAT  CTC  ATG  ATA  GTT  364
95  I  R  V  D  T  L  E  N  Q  V  M  D  T  R  I  H  L  M  I  V  114

365  TGC  TGC  AGT  CCT  GAC  TTT  GAG  AAG  CAG  AAG  CCA  GAG  TTC  TTG  AAG  TCC  ATC  CCG  GAG  AAG  424
115  C  C  S  P  D  F  E  K  Q  K  P  E  F  L  K  S  I  P  E  K  134

425  ATG  AAA  ATC  TAT  TCA  GAG  TTC  CTG  GGC  AAG  CGA  CCA  TGG  TTT  GCA  GGG  GAC  AAG  GTC  ACC  484
135  M  K  I  Y  S  E  F  L  G  K  R  P  W  F  A  G  D  K  V  T  154

485  TAT  GTG  GAT  TTC  CTT  GCT  TAT  GAC  ATT  CTT  GAC  CAG  TAC  CGT  ATG  TTT  GAG  CCC  GAG  TGC  544
155  Y  V  D  F  L  A  Y  D  I  L  D  Q  Y  R  M  F  E  P  E  C  174

545  CTG  GAC  GCC  TTC  CCA  AAC  CTG  AAG  GAC  TTC  CTG  GCC  CGC  TTC  GAG  GGC  CTG  AAG  AAG  ATC  604
175  L  D  A  F  P  N  L  K  D  F  L  A  R  F  E  G  L  K  K  I  194

605  TCT  GCT  TAC  ATG  AAG  AGT  AGC  AGC  TTT  CTC  CCA  AGA  CCT  GTG  TTT  ACT  AAG  ATA  CCG  CAG  664
195  S  A  Y  M  K  S  S  S  F  L  P  R  P  V  F  T  K  I  P  Q  214

665  TGG  GGC  ACT  GAT  TAG  GCCCTCGCATGCTGGATCCGCG   702
215  W  G  T  D  *

```

Figure 1 Nucleotide sequence of the synthetic rat GST-Y_{b4} (*rGSTM4*) cDNA and its deduced amino acid sequence

The duplex Y_{b4} sequence is divided into 19 synthetic oligonucleotides as follows (numbers with primes refer to the complementary sequence): oligonucleotide 1, 1–80; 2, 81–160; 3, 161–240; 4, 241–320; 5, 321–400; 6, 401–480; 7, 481–560; 8, 561–640; 9, 641–702; 10, 1'–40'; 11, 41'–120'; 12, 121'–200'; 13, 201'–280'; 14, 281'–360'; 15, 361'–440'; 16, 441'–520'; 17, 521'–600'; 18, 601'–680'; 19, 681'–702'. The underlined sequences are the *EcoRI* and *BamHI* sites for cloning.

commercial suppliers. The 19 oligonucleotides, which encompass both strands of the *rGSTM4* coding sequence with a ribosome-binding site for expression in *Escherichia coli*, and flanking *EcoRI* and *BamHI* sites for cloning (see Figure 1), were purchased from Integrated DNA Technologies (Coralville, IA, U.S.A.). The four pairs of oligonucleotides used to restore deleted nucleotides (G323 and T465) and to correct missense mutations (T160 and T240) in the initial Y_{b4} (*rGSTM4*) cDNA clone were products of Life Technologies (Gaithersburg, MD, U.S.A.).

The vector pKK223-4 for *rGSTM4* expression is a derivative of pKK223-3 (Amersham Pharmacia Biotech, Piscataway, NJ, U.S.A.), where a second *BamHI* site outside the multiple-cloning-site region is converted into a *SmaI* site. GSTs from rat liver, rat testis, mouse liver, human liver and other rat tissues were purified through *S*-hexyl-GSH-affinity chromatography as described previously [14–18]. Rat *rGSTM3*-specific antiserum and the rat Mu-class common antiserum were those described previously [18]. The ECL immunoblot-analysis system was purchased from Amersham Pharmacia Biotech.

Synthesis of the rat Y_{b4} (*rGSTM4*) coding sequence for expression in *E. coli*

Kinase reactions for phosphorylating oligonucleotides 2–18 (see Figure 1; 0.8 nmol each) were carried out in 10 µl of 1 × New England Biolabs T4 ligase buffer and 10 units of T4 polynucleotide kinase at 37 °C for 30 min (1.0 µmol of ATP versus 0.8 nmol of 5' ends). The reactions were terminated by heating at 65 °C for 20 min. Oligonucleotides 1 and 19 (0.8 nmol each) in 10 µl of 1 × New England Biolabs T4 ligase buffer were combined with the 17 kinase reactions and adjusted to 50 mM NaCl in a volume of 200 µl. They were divided equally into two tubes

containing 4 µM of each oligonucleotide. One tube of the mixture was boiled at 100 °C for 10 min and quickly cooled in ice water before ethanol precipitation. The precipitated oligonucleotides were dissolved in a final volume of 400 µl of New England Biolabs T4 ligase buffer containing 80 Weiss units of T4 DNA ligase. After incubation at 20 °C for 12 h, the T4 ligase was inactivated at 65 °C for 10 min and 12 µl of 5 M NaCl, 80 units of *EcoRI* and 80 units of *BamHI* were added to the tube for incubation at 37 °C overnight. The reaction products were purified by phenol extraction once, chloroform/isoamyl alcohol extraction twice and precipitated by ethanol. The pellet was washed with 70 % ethanol, dried and resuspended in 100 µl of TE buffer (10 mM Tris/HCl, pH 8.0/1 mM EDTA) to a concentration of 4 µM of annealed *rGSTM4* gene fragments. Approximately equal molar amounts of the *rGSTM4* fragments (≈ 0.25 µg) and vector DNA (≈ 1.5 µg of *EcoRI*- and *BamHI*-digested pKK223-4) were ligated in 20 µl with 0.1 Weiss unit of T4 DNA ligase at 16 °C overnight. Then, 50–100 ng of the ligated DNA were used to transform *E. coli* DH5α. Plasmid DNAs from transformants were analysed by *EcoRI*–*BamHI* double digestion for the presence of a 702 bp insert. One such transformant, designated Y_{b4}-12, was chosen randomly for DNA sequence analysis on both strands. The results revealed two deletions at positions 323 (G) and 465 (T) and two substitutions at 160 (G → T) and 240 (G → T) relative to the desired *rGSTM4* sequence (Figure 1).

Mutagenesis of Y_{b4}-12 to restore the open reading frame: insertions at nucleotides 323 and 465

Insertion of nucleotide 323 was accomplished according to the procedure of Deng and Nickoloff [19]. Two oligonucleotides, Y_{b4}-323 (5'-GGACACTTTGGAGAACCAGGTC-3') and

pKK(*Afl*III → *Bgl*II) (5'-CAGGAAAGAAGATCTGAGCAA-AAG-3'; with the replaced nucleotides shown in bold), for elimination of a restriction site (*Afl*III) were phosphorylated by T4 polynucleotide kinase and ATP. The protocol of Clontech (Palo Alto, CA, U.S.A.) was followed except for the reaction buffers. The template, *Y_{b4}*-12 (0.1 µg), was mixed with 0.1 µg of phosphorylated *Y_{b4}*-323 oligonucleotide and 0.1 µg of phosphorylated pKK(*Afl*III → *Bgl*II) oligonucleotide in 20 µl of annealing buffer (20 mM Tris/HCl, pH 7.5/10 mM MgCl₂/50 mM NaCl). The mixture was boiled at 100 °C for 3 min and chilled in ice water for 5 min. To the annealing reaction (30 µl final volume) were added 3 µl of 10× synthesis buffer (170 mM NaCl/33 mM MgCl₂/20 mM dithiothreitol/10 mM ATP/1 mM of each dNTP/250 µg/ml BSA), 4 units of T4 DNA polymerase and 4 units of T4 DNA ligase. The reaction was incubated at 37 °C for 2 h and heated at 70 °C for 10 min to stop the reaction. To the reaction were added 0.3 µl of 5 M NaCl, 1 µl of 1 M Tris/HCl, pH 8.0, and 5 units of *Afl*III before incubation at 37 °C for 1 h. A fraction of the mixture (2 µl) was transformed into *E. coli* *BMH*71-18 to enrich for the mutants. A fraction of the total plasmid DNA was digested with *Afl*III again and transformed into *E. coli* *DH5α*. Plasmid DNA from each of the 12 transformants was analysed for loss of the *Afl*III site and gain of a *Bgl*II site before DNA sequence analysis. The correct plasmid was designated *Y_{b4}*-323.

Insertion of a T residue at position 465 of *Y_{b4}*-323 was carried out according to the protocol of Quik Change[®] of Stratagene (La Jolla, CA, U.S.A.) with two complementary oligonucleotides, *Y_{b4}*-465-sense (5'-CGACCATGGTTTGCAGGGGAC-3') and *Y_{b4}*-465-antisense (5'-GTCCCCTGCAAACCATGGTCG-3'). PCR was carried out for 28 cycles. After removal of the template DNA by *Dpn*I digestion, an aliquot (≈ 50–100 ng) was used to transform *E. coli* *DH5α*. The transformants were analysed for rGSTM4-sized (≈ 25 kDa) protein expression by SDS/PAGE before DNA sequence analysis. The correct clone was designated *Y_{b4}*-7 (rGSTM4-4-W45C, S72I). The substitutions at positions 160 and 240 were corrected on *Y_{b4}*-7 DNA by the same procedure using the following pairs of primers in the order of: 240-sense (5'-GTCACCCAGAGCAATGCCATCC-3') with 240-antisense (5'-GGATGGCATTGCTCTGGGTGAC-3') (*Y_{b4}*-W45C), and 160-sense (5'-GAAGCCAGTGGCTGAGTGAGAA-3') with 160-antisense (5'-TTCTCACTCAGCCACTGGCTTC-3'). The final expression clone of rGSTM4 was designated pGTR401-KK. The C115Y and other mutants of rGSTM4 were constructed on pGTR401-KK by the same procedure.

Expression construct of rGSTM1 (*Y_{b1}*)

The coding sequence of rGSTM1 was linked by PCR to a ribosome-binding-site sequence nine nucleotides upstream of the ATG initiation codon [20]. The following pairs of oligonucleotides: *Y_{b1}* forward, 5'-GGAATTCAGGAGAACAGAA-ATATGCCTATGATACTGGGATAC-3', and *Y_{b1}* reverse, 5'-CGGGATCCGCCAGTGTAGCAAGGGCC-3' were used for PCR amplification on the plasmid template pGTR200 [21]. The PCR products were digested with *Eco*RI and *Bam*HI and cloned into similarly digested pKK223-4. Expression clones were identified by 1-chloro-2,4-dinitrobenzene (CDNB) sensitivity of the transformants (*Y_{b1}*) on LB agar plates containing ampicillin (100 µg/ml) and CDNB (20 µg/ml), followed by analyses of enzyme activities in the sonicated crude extracts and SDS/PAGE [22]. After sequence analysis, a transition mutation (A → T) was identified, resulting in rGSTM1-1(M34L). This expression clone was designated pGTR102-KK.

Purification of and enzymic assays for recombinant rGSTM1-1 and rGSTM4-4

The procedure described by Lee et al. [22] was carried out with 6 × 800 ml of overnight cultures grown at 30 °C. After dialysis of the sonicated crude extract, the supernatant was cleared by centrifugation and loaded on to an *S*-hexyl-GSH-Sepharose 4B column (4 ml bed volume). GSH was deleted from the standard elution buffers for the purification of rGSTM4-4 and its derivatives. The eluted proteins were dialysed against 1 litre of 25 mM Tris/HCl, pH 8.0, with two changes and stored at -20 °C in aliquots. The GSTs were electrophoretically homogeneous by SDS/PAGE. A glutathione-deficient *E. coli* strain, JTG10 (*gshA*::Tn10), was used to prepare non-glutathiolated rGSTM4-4 [23]. The yield was considerably lower. Enzyme assays for GST activities were those originated in [24] as described in [15,16]. Blank rates were subtracted from all rate measurements under each assay condition. All recombinant GSTs in this investigation were stable in Bis-Tris propane sulphate buffers of pH 6–9.5 for at least 15 min at 4 °C.

Synthesis of rGSTM4- and rGSTM1-specific peptide antigens

The peptide epitopes chosen for rGSTM1 and rGSTM4 subunits were from the divergent C-terminal positions, NH₂-Thr-Pro-Ile-Phe-Ser-Lys-Leu-Ala-Gln-Trp-Ser-Asn-Lys-COOH (tridecapeptide) and NH₂-Val-Phe-Thr-Lys-Ile-Pro-Gln-Trp-Gly-Thr-Asp-COOH (undecapeptide), respectively. Multiple antigen peptides were prepared according to the method developed by Tam [25]. The inert octa-branched matrix core (Aminotech, Ontario, Canada) was linked to *p*-benzyloxybenzyl alcohol resin with a substitution density in the range of 0.6–0.8 mol/g (Norva-biochem, La Jolla, CA, U.S.A.). The rGSTM1 tridecapeptide and the rGSTM4 undecapeptide were synthesized on to the octa-branched matrix core individually, using an automated solid-phase peptide synthesizer (model 430A, Applied Biosystems, Foster City, CA, U.S.A.). After acetylation of the N-terminus, deprotection was carried out in a solution containing 0.75 g of crystalline phenol, 0.25 ml of 1,2-ethanedithiol, 0.5 ml of thioanisole, 0.5 ml of water and 10 ml of trifluoroacetic acid for 2–3 h. The crude peptides were first dissolved in a minimum amount of DMSO diluted with 8 M urea and dialysed against 8 M urea overnight, with one change for 8 h, against 0.1 M acetic acid overnight and then against water before lyophilization. Amino acid compositions for each product were consistent with the predicted sequences. These products were used to raise antibodies in white female New Zealand rabbits at the Central Biological Laboratory of The Pennsylvania State University (University Park, PA, U.S.A.). A second rabbit anti-rGSTM4 antiserum was raised against the C-terminal tridecapeptide (residues 205–217) coupled to keyhole limpet haemocyanin at the Albert Einstein College of Medicine (Bronx, NY, U.S.A.; see [18] for experimental methods).

MS and HPLC analysis

MS analysis was performed using a quadrupole ion-trap mass spectrometer (LCQ, Finnigan Corp., San Jose, CA, U.S.A.). GST subunits that were resolved initially by HPLC methods (Hewlett Packard HP1090) were injected into the mass spectrometer as a continuous flow of a solution of 50% acetonitrile/water and 0.1% formic acid with a needle voltage of 4.5 kV on the LCQ. Molecular masses for the GST subunits and their glutathiolated adducts were determined with an accuracy of 2–3 amu. A Hewlett Packard HP1090 instrument with a diode array detector was used for HPLC analysis. GST subunits were

resolved using a Vydac C_4 reversed-phase column; the protein in aqueous solution containing 0.1 % trifluoroacetic acid (TFA) was applied to the column (4.6 mm $\times$ 250 mm) and eluted with a linear gradient of acetonitrile/0.1 % TFA at a flow rate of 0.75 ml/min. The gradient was initiated at 20 % acetonitrile and increased to 40 % after 10 min, followed by an increase of 0.5 %/min for 30 min to 55 % acetonitrile at 40 min. Absorbance was recorded at 214 nm and 280 nm to monitor protein peaks.

Kinetic analysis

A stopped-flow apparatus (KinTek Instruments, State College, PA, U.S.A.) was used to determine the initial rates because of the low CDNB-conjugating activity of rGSTM4-4 and its mutants and the short duration of linear product formation. The operation of the stopped-flow apparatus and data collection has been described in Kuan et al. [26]. Kinetic parameters were analysed by KaleidaGraph[™] 3.0 (Synergy Software, Reading, PA, U.S.A.).

Modelling of the rGSTM4-4 tertiary and quaternary structures

On the basis of substantial amino acid sequence identities (over 80 %) between the rGSTM4, rGSTM1 and human (h) GSTM2 subunits, co-ordinate files of the latter two protein structures (PDB accession numbers 6GSY [27] and 2GTU [28]) were used for modelling of the three-dimensional structure of rGSTM4-4. Because the overall folds of hGSTM2-2 and rGSTM1-1 are essentially the same, an initial model of rGSTM4-4 was constructed using co-ordinates of the unliganded hGSTM2-2 dimer (PDB entry 2GTU) and refined using standard X-PLOR (version 3.851) protocols with a structure file taken from this PDB entry. For convenience in the model, Arg-31(A) and Arg-167(B) were replaced with Lys and Gln residues, respectively (A and B refer to the different subunits of a dimer), because of disallowed interactions along the crystallographic symmetry axis. A total of 12087 reflections with sigma cutoff 3 and a resolution range of 2.8–8.0 were used. The refinement was terminated when the R free value was lowered to 0.31 (R factor was 0.24) and no further improvement was observed. The atomic co-ordinates of GSH were taken from the PDB entry 6GSY and those of S -(2,4-dinitrophenyl)glutathione (GS-DNB) from the PDB entry [28,29]. Ligands were placed in the rGSTM4-4 active site after alignment of three-dimensional protein structures of ligand-free rGSTM4-4 and rGSTM1-1 with ligand, using INSIGHTII software (Biosym, San Diego, CA, U.S.A.). Since Cys-115 was suspected to form a mixed disulphide with GSH, allowed rotamers of the Cys-115 side chain were used to build a model of the S -glutathiolated rGSTM4-4. At the final step, all of the models with bound ligands were subjected to refinement using X-PLOR [30], but ligands were refined as rigid bodies even though protein atoms were allowed to move freely. Non-crystallographic symmetry operators were applied at the final energy-minimization step and atomic co-ordinates of ligands were also included in this refinement. The stereochemistry of the final model was verified by PROCHECK software. Standard deviations were: distances of 0.01 Å, angles of 1.44°, dihedral angles of 22.08° and improper angles of 1.29°. A Ramachandran plot statistic showed that 88 % of residues are in the most favoured regions and the rest are in allowed regions. It is noteworthy that the two protein monomers with bound ligands of the model rGSTM4-4 dimer are practically identical, with an overall root-mean-square deviation (RMSD) value between backbone atoms of 0.25 Å.

RESULTS

Nature of the rGSTM4 gene

The nucleotide sequence of the *rGSTM4* gene [12,13] spans at least 6.2 kb on λ GTR15-2. The exons were identified based on homologies with the *rGSTM2* genomic sequence, and cDNAs of *rGSTM1* and *rGSTM3* [12,17,31]. The *rGSTM4* structural gene contains no apparent defects, with an open reading frame of 218 codons, which differs from the *rGSTM1* sequence by a predicted 40 amino acid substitutions (see Figure 1). The sequence differences between rGSTM4 and rGSTM2 and between rGSTM4 and rGSTM3 subunits are greater (53 substitutions for both). The close evolutionary relationship between the *rGSTM1* and *rGSTM4* genes is further supported by the extensive homology in their 3'-UTRs. The last 33 nucleotides in the *rGSTM1* cDNA differ from the corresponding region of *rGSTM4* by only a single nucleotide (C versus T at position 1040 of the pGTR200 cDNA [13,21]). A notable distinction in the predicted rGSTM4

Table 1 Substrate specificity of recombinant rGSTM1-1 and rGSTM4-4

Specific activities (μ mol/min per mg) were determined at pH 6.5 with 1 mM GSH, except for 1,2-dichloro-4-nitrobenzene (DCNB; pH 7.5) and the peroxides (pH 7.0), ND, not detectable. The errors are averages of the deviations from the mean; they are not included when less than 0.005.

Substrate	Specific activity (μ mol/min per mg)	
	rGSTM1-1(M34L)	rGSTM4-4
CDNB	121.2 $\pm$ 2.1	13.8 $\pm$ 0.06
DCNB	1.83 $\pm$ 0.08	0.11 $\pm$ 0.03
<i>p</i> -Nitrobenzyl chloride	1.75 $\pm$ 0.04	ND
1,2-Epoxy-3-propane	0.18 $\pm$ 0.02	0.09
Ethacrynic acid	0.09	ND
Δ^5 -Androstene-3,17-dione	0.02	0.01
Bromosulphophthalein	0.05	0.02
Cumene hydroperoxide	0.89 $\pm$ 0.03	0.24 $\pm$ 0.01
<i>t</i> -Butyl hydroperoxide	0.06	ND
H ₂ O ₂	ND	ND

Figure 2 pH profiles for CDNB-conjugation activity of recombinant rGSTM1-1 and rGSTM4-4

Enzyme activities were measured by a KinTec stopped-flow apparatus in Bis-Tris propane sulphate buffer. The pH was adjusted by addition of sulphuric acid. The ionic strength of each buffer was adjusted with saturated sodium sulphate to the same value measured by a conductivity meter (15000 μ mho). The concentrations of substrates were 10 mM GSH and 1 mM CDNB. The GST concentrations were 35.7 μ M for rGSTM1-1(M34L) and 108 μ M for rGSTM4-4.

Table 2 Kinetic parameters for recombinant rGSTM4-4 and rGSTM1-1(M34L)

The molecular mass of rGSTM4-4 is calculated to be 51 024 Da and that of rGSTM1-1(M34L) is 51 496 Da.

	K_m (mM)		k_{cat} (s^{-1})	$(k_{cat}/K_m) \times 10^{-3}$ ($M^{-1} \cdot s^{-1}$)	
	GSH (1 mM CDNB)	CDNB (10 mM GSH)	CDNB (10 mM GSH)	GSH (1 mM CDNB)	CDNB (10 mM GSH)
rGSTM4-4	2.9 ± 0.46	1.06 ± 0.36	7.6 ± 1.4	2.6	7.1
rGSTM4-4(C115Y)	4.4 ± 0.6	0.25 ± 0.11	18.7 ± 0.6	4.2	74.8
rGSTM1-1(M34L)	0.28 ± 0.03	0.03 ± 0.01	39.2 ± 2.8	140	1307

Figure 3 HPLC and ESI-MS analysis of rGSTM4-4; evidence for glutathiolation

A preparation of rGSTM4-4 (50 μ g) in aqueous 0.1% TFA was applied on a Vydac C_4 reversed-phase HPLC column (4.6×250 mm). Elution was with a linear gradient of acetonitrile/0.1% TFA, at a flow rate of 0.75 ml/min. The solid line represents the pattern obtained for rGSTM4-4 and the dotted line the same preparation preincubated with 1.0 mM GSH for 1 h. The insets are transformed ESI-MS for the two indicated peaks after treatment with GSH. Molecular masses of each major component are shown, and the arrow on the left-hand inset indicates the expected mass of a non-glutathiolated subunit (25512 Da) not present in the S-glutathiolated peak.

mRNA sequence is a stretch of 45 nucleotides in the 3'-UTR, which is duplicated as an imperfect direct repeat containing six substitutions. Alternatively, the 3'-UTR of the rGSTM1 mRNA has a 46-nucleotide deletion relative to the rGSTM4 mRNA, including 22 nucleotides in the 3' half of the first repeat and 24 nucleotides in the 5' half of the second repeat in the rGSTM4 sequence. This homology in the 3'-UTR does not extend to any other pairwise comparison of the rat Mu-class genes. Such a close evolutionary relationship between the *rGSTM1* and *rGSTM4* genes perhaps suggests a functional importance for the 3'-UTR sequences. The results outlined below establish that the rat *rGSTM4* gene encodes a functional protein that is, however, considerably different from that of the closely homologous *rGSTM1*.

Properties of recombinant rGSTM4-4

The substrate-specificity patterns of rGSTM4-4 are qualitatively similar to those of rGSTM1-1(M34L) except that rGSTM4-4 is

generally less active; rGSTM4-4 exhibits catalytic activity with CDNB among traditional GST substrates (Table 1), and has detectable GSH peroxidase activities with cumene hydroperoxide but not t-butyl hydroperoxide or hydrogen peroxide. The *rGSTM4* gene thus encodes an enzymically active GST. The pH optimum for the CDNB conjugation activity is 8.5 for rGSTM4-4 relative to 6.5 for rGSTM1-1(M34L). At pH 7, rGSTM4-4 has only about 35% of the maximal activity observed at pH 8.5 (Figure 2). The kinetic data shown in Table 2 indicate that despite their close sequence homologies, the K_m^{GSH} value for rGSTM4-4 is 10-fold higher than that of rGSTM1-1, and rGSTM1-1, in turn, is about 184 times more efficient (k_{cat}/K_m) for CDNB conjugation than is rGSTM4-4.

The recombinant rGSTM4-4 isolated from *E. coli* DH5 α and purified using S-hexyl-GSH-affinity matrices by elution with S-hexylGSH yielded a single HPLC component (Figure 3), with only trace amounts of an incompletely processed subunit with N-formylmethionine at its N-terminus. The molecular mass determined by electrospray ionization MS (ESI-MS) of 25512 Da

Table 3 Effect of dithiothreitol on the CDNB-conjugation activity of rGSTM4-4 and its mutant

Specific activity ($\mu\text{mol}/\text{min}$ per mg) at pH 6.5 with 10 mM GSH and 1 mM CDNB is shown. Fold increases in activity are shown in parentheses.

Dithiothreitol (mM) ...	Specific activity ($\mu\text{mol}/\text{min}$ per mg)			
	0	5	50	100
rGSTM4-4	2.94 $\pm$ 0.02 (1)	5.62 $\pm$ 0.02 (1.9)	6.12 $\pm$ 0.03 (2.0)	5.5 $\pm$ 0.01 (1.9)
rGSTM4-4(C115Y)	8.12 $\pm$ 0.02 (1)	8.77 $\pm$ 0.01 (1.1)	8.87 $\pm$ 0.00 (1.1)	7.90 $\pm$ 0.00 (1.0)

is precisely the mass deduced from the gene sequence for the rGSTM4 subunit (accession number AF106661). After pre-incubation of the enzyme with GSH for short periods of time, however, two peaks were resolved by HPLC. One component had identical retention time and molecular mass to the natural unmodified rGSTM4 subunit (Figure 3). The second component, which appeared only after incubation with GSH, had a subunit molecular mass of 25817 Da and small amounts with a mass of 26124 Da (Figure 3); these masses are consistent with the rGSTM4 subunit containing one and two covalently linked GSH molecules, respectively. The rGSTM4-4 enzyme expressed and purified from a glutathione-deficient strain of *E. coli* (see the Experimental section) also yielded a single component with a molecular mass of 25512 Da. Incubation of that particular GSH-free preparation with 10 mM GSH for 1 h at 25 °C resulted in the formation of mono- and di-S-glutathiolated rGSTM4 subunits, as determined by ESI-MS analysis of the resultant HPLC fractions (results not shown).

A rGSTM4-4(C115Y) mutant, with a molecular mass of 25572 Da, had 10-fold greater catalytic efficiencies for CDNB conjugation relative to the wild-type rGSTM4-4 (Table 2). After incubation with GSH, only a small amount of an S-glutathiolated product (molecular mass 25876 Da) was obtained for the C115Y mutant (results not shown). This result indicates that a major site of glutathiolation of wild-type GSTM4-4 shown in Figure 3 is likely to be Cys-115. Cys-86, which is near the surface of the protein, probably represents the minor glutathiolation site. The pH profile of rGSTM4-4 with a Tyr-115 residue, however, was very similar to that of the wild-type enzyme, and the K_m^{GSH} values for the mutant remained high. The extensive S-glutathiolation is specific for rGSTM4-4 since rGSTM1-1(M34L) prepared under the same conditions was not glutathiolated after incubation with GSH, nor were rGSTM1-1 and rGSTM2-2 or rGSTM3-3 purified from rat tissue [10]. Nevertheless, S-glutathiolation does not ablate the CDNB-conjugation activity of rGSTM4-4. Treatment of glutathiolated rGSTM4-4 with dithiothreitol (5, 50, 100 mM for 1 h at room temperature) resulted in a nearly 2-fold increase in CDNB-conjugation activity (Table 3). There was no evidence for sulphenic acid or other partially oxidized intermediates of Cys-115 by ESI-MS analysis. Because S-glutathiolation occurred rapidly after incubation of the enzyme with GSH, it is likely to occur during the course of the assay conditions.

The penultimate intermediate in the synthesis and expression of rGSTM4-4 was an rGSTM4-4(W45C) mutant. *E. coli* expressing rGSTM4-4(W45C) did not exhibit CDNB-conjugating activities, and the recombinant mutant rGSTM4-4(W45C) protein had an electrophoretic mobility in SDS gels similar to that of the wild-type rGSTM4 subunit (on the basis of its reaction with antiserum against the rGSTM4 undecapeptide), but did not adhere to S-hexyl-GSH affinity matrices (results not shown). This result indicates that Trp-45 is essential for GSH binding, or that Trp-45 serves important function(s) in the proper folding and/or

Figure 4 Immunoblot analysis of various GSTs with peptide-sequence-specific antisera, against rGSTM1-1 (Y_{b1}) or rGSTM4-4 (Y_{b4}) subunits

(A) Specificity of rGSTM1-1 antibody on a Western blot. Lanes 1–5 are recombinant rGSTM1-1, 0.05, 0.1, 0.25, 0.5 and 1 μg , respectively; lanes 6–10 are recombinant rGSTM4-4, 0.5, 1, 2, 4 and 8 μg , respectively. (B) Specificity of rGSTM4-4 antibody on an immunoblot. Lanes 1–5 are recombinant rGSTM1-1, 0.5, 1, 2, 4 and 8 μg , respectively; lanes 6–10 are recombinant rGSTM4-4, 0.05, 0.1, 0.25, 0.5 and 1 μg , respectively. (C) Reactions of rGSTM1-1 antibody with GSTs from various rat tissues. Lane 1, brain GSTs (Br), $\approx$ 2 μg ; lane 2, recombinant rGSTM4-4 (Y_{b4}), 1 μg ; lane 3, control liver (C-Li), 12.7 μg ; lane 4, phenobarbital-treated liver (PB-Li), $\approx$ 10 μg ; lane 5, lung (Lu), $\approx$ 2 μg ; lane 6, kidney (Ki), $\approx$ 2 μg ; lane 7, rat testis (Tes), $\approx$ 10 μg ; lane 8, recombinant rGSTM1-1 (Y_{b1}), 1 μg . (D) Reactions of rGSTM1-1 antibody with GSTs. Lane 1, control liver, 12.7 μg ; lane 2, phenobarbital-treated liver, $\approx$ 10 μg ; lane 3, mouse liver (M-Li), 8 μg ; lane 4, human liver (Hu-Li), 9.2 μg ; lane 5, rabbit liver (Rab-Li), $\approx$ 6 μg ; lane 6, rat liver (R-Li; Sigma), $\approx$ 2 μg ; lane 7, rat testis (R-Tes), $\approx$ 17 μg . (E) Reactions of rGSTM4-4 antibody with GSTs. Lane 1, brain GSTs, $\approx$ 2 μg ; lane 2, recombinant rGSTM4-4, 1 μg ; lane 3, control liver, 12.7 μg ; lane 4, phenobarbital-treated liver, $\approx$ 10 μg ; lane 5, lung, $\approx$ 2 μg ; lane 6, mouse liver, 8 μg ; lane 7, rat testis, $\approx$ 17 μg ; lane 8, human liver, 9.2 μg ; lane 9, recombinant rGSTM1-1, 1 μg ; lane 10, rabbit liver, $\approx$ 6 μg .

catalysis of rGSTM4-4. Although the carboxylate of the glycyl moiety of bound GSH is less than 3 Å from the indole nitrogen of Trp-45 (see Figure 5, below), until now no function has been assigned to this residue. Whether Trp-45 is essential for activity of other Mu-class GSTs remains to be determined experimentally.

Expression of rGSTM4-4 in rat tissue

The expression patterns of rGSTM4-4 among the major rat tissues/organs were examined by immunoblotting using antisera

Figure 5 Comparison of active sites of rGSTM1 and rGSTM4 subunits

Alignment of active-site residues of the rGSTM4 subunit (lighter-shade residues that are identified by label) and rGSTM1 subunit (darker residues, structure from obtained PDB accession number 6GSY) with a bound GS-DNB molecule. The orientation of Met-111 of rGSTM4 is compared with that of Ile-111 of rGSTM1 in their juxtaposition with respect to the catalytically important Arg-107 residue. A ribbon representation of the helical $\alpha 4$ segment containing Met-104, Ile-108, Met-111, Ile-112 and Cys-115 residues is shown. All structures were drawn using the Insight II program.

raised against the divergent C-terminal undecapeptide and a second antiserum against the C-terminal tridecapeptide of the rGSTM4 subunit (see Figure 6, later). Only four of the 11 amino acids are identical to those of the C-terminal sequence of rGSTM1, but since rGSTM1 subunits are expressed abundantly in many rat tissues, the possibility of immunological cross-reactivity had to be ruled out. The antisera against the C-terminal tridecapeptide of rGSTM1 and an rGSTM3 (Y_{b3})-specific antiserum [18] were therefore used as controls for the abundant rGSTM1-1 and other major Mu-class GSTs in rat tissues. Furthermore, since recombinant rGSTM4-4 can be retained by *S*-hexyl-GSH-affinity matrices, rat-tissue GSTs purified by *S*-hexyl-GSH- or GSH-affinity chromatography were also used in the immunoblotting analysis for rGSTM4-4. Under the experimental conditions employed, the rGSTM1 antiserum had an at least 80-fold selectivity for rGSTM1 over rGSTM4 subunits (i.e. no cross-reactions with 8 μ g of recombinant rGSTM4 when 0.1 μ g of the rGSTM1 subunit shows a positive signal; Figure 4A). The selectivity of rGSTM4 peptide antiserum was also at least 80-fold in favour of rGSTM4 (Figure 4B).

Results shown in Figure 4(C) confirmed that rGSTM1 subunits are expressed in rat brain, liver, lung, kidney and testis; these data are consistent with the Northern hybridization results for rat-tissue mRNAs [21]. Cross-reactions with the rGSTM1 antisera were also observed in mouse liver, rabbit liver and, to a lesser extent, human liver GSTs purified by *S*-hexyl-GSH-affinity chromatography (Figure 4D). The close sequence homology

between a mouse Mu-class GST, GT8.7 (pGT875 [32]; pmGT10 [33]) and rGSTM1-1 in general (93% identity), and only two substitutions in the C-terminal tridecapeptide in particular, could explain the strong cross-reaction. A major Mu-class human liver GST, hGSTM1a-1a (or hGSTM1b-1b), which shares less sequence identity with rGSTM1-1 in the C-terminal tridecapeptide region, also cross-reacted to a slight extent with the anti-rGSTM1 antiserum. The positive signal with rabbit liver GSTs supports the notion that a Mu-class GST, perhaps with considerable sequence homology to the rGSTM1 C-terminal tridecapeptide, exists in rabbit.

Only rat liver and mouse liver GSTs (Figure 4E), however, gave positive signals with the rGSTM4-peptide-specific antiserum. The lack of rGSTM4 signals in immunoblots of other rat tissues (e.g. kidney, lung, brain, heart, spleen and testis) could be due to very low levels of expression or a lack of expression of rGSTM4 in those tissues. The strong signal with mouse liver GSTs is consistent with the presence of the mouse GST GT9.3 (pGT55, see [34]), which shares 94% sequence homology with the rGSTM4 subunit and has only a single substitution [Ala $\rightarrow$ Pro (rGSTM4) at position 212] in the C-terminal undecapeptide (see the Discussion).

Predicted active-site structure of rGSTM4-4

The GSH-binding domain (G-site) of the rGSTM4-4 enzyme consists of highly conserved residues found in the N-terminal α -

β domain of other Mu-class GSTs [35–41]. The hydroxy group of Tyr-6 is within hydrogen-bonding distance of the sulphur atom of bound GSH (3.6 Å). Additional hydrogen-bond interactions are predicted between GSH and the side chains of Trp-7, Arg-42, Trp-45, Lys-49, Asn-58, Gln-71 and Ser-72 (Figure 5). Crystallographic refinement of the rGSTM4 model in the presence of GS-DNB produced a very similar conformation of the G-site compared with that of the ligand-free enzyme, with an overall RMSD value of about 0.3 Å. In addition, RMSD values between G-site residues of rGSTM4 and hGSTM2 or rGSTM1 subunits do not exceed 0.45 Å. Some other amino acid substitutions found in rGSTM4 subunits, however, could have significant effects on its properties (see the Discussion).

The binding sites for the electrophilic co-substrates (H-sites) are far more variable among the Mu-class GSTs, even though the overall protein folds are essentially the same. In the model of the rGSTM4-4 with bound GS-DNB product (Figure 5), the H-site for the enzyme is lined with the non-polar side chains of residues Leu-12, Met-104, Ile-108, Met-111, Ile-112, Cys-115, Phe-208 and Ile-211. It is noteworthy that only Leu-12, which is also a part of the G-site, and Met-104 and Phe-208 are conserved among the Mu-class GSTs, but that the other amino acid residues are variable. In addition, Pro-212 of rGSTM4 (most of the other enzymes have Ala-212) and Trp-214 (which is conserved) may stabilize the conformation of the C-terminal loop (residues 207–214), and thus increase the rigidity of the H-site.

The dimer interface of rGSTM4-4

Although there are several amino acid substitutions, the three-dimensional structure of the dimer interface for rGSTM4-4 is similar to that of all other Mu-class enzymes. It is noteworthy, however, that the key Phe-56 residue characteristic of the Mu GSTs is replaced with Ile-56 in rGSTM4-4. Additional substitutions at the dimer interface for rGSTM4-4 are Ile-69 by Val and Ile-98 by Thr. These substitutions should not interfere with protein folding but may change the relative positions of one monomer with respect to the other at the dimer interface. The hydrophobic surfaces involved in interactions at the interface are thus shorter compared with those of hGSTM2-2 and rGSTM1-1. A Phe-56 side chain of the other GSTs has about a 1.5 times larger surface area, providing tighter van-der-Waals interactions within the key-lock interface than does Ile-56 of rGSTM4-4. Also, the replacement of Ile-69 (rGSTM1, hGSTM2) with Val-69 (rGSTM4) and of Ile-98 (rGSTM1, hGSTM2) with Thr-98 (rGSTM4) greatly decreases the compactness of the dimer interface, because these residues participate in direct contacts between two monomers. For instance, Val-69(A) is within 4–4.3 Å of Ile-94(B), and Thr-98(B) is within the same distance of Ala-74(A). Residues 94 and 98 of the same protein subunit also maintain hydrophobic interactions. Because of these substitutions, there are fewer polar interactions and the packing of rGSTM4-4 should be looser within the interface, as compared with the other Mu-class enzymes. The model predicts that the two protein monomers of rGSTM4-4 can oscillate perpendicular to the 2-fold non-crystallographic symmetry axis by up to 3.5° with respect to each other. Therefore, rGSTM4-4 has the potential to change conformation during catalysis and/or ligand binding by packing more tightly.

DISCUSSION

The presence of a Cys-115 residue in place of the conserved active-site Tyr-115 of other mammalian Mu-class GSTs is one of

several distinctive structural features of the rGSTM4 subunit. In fact, the only other known GST subunit with a cysteine residue at that position is a mouse GST (GT9.3) described previously [32–34]. GT9.3 is probably the orthologue of rGSTM4 and we therefore suggest that the mouse counterpart be designated as the mouse (m) GSTM4 subunit. Neither of these GSTs has ever been purified from rodent tissue, and no known human Mu-class GST counterpart with a Cys-115 residue has been identified. Because mammalian Mu-class GST genes have undergone gene conversion [12], it is often difficult to assign orthologous relationships between the GSTs of different species. However, recent studies indicate that the *hGSTM3*, *mGSTM5* and *rGSTM5* represent an example of orthologous Mu-class genes on the basis of their extensive sequence homologies, tissue-specific patterns of expression and tail-to-tail orientation of their genes relative to the other members of the Mu-class GST gene cluster [10,11]. The rGSTM4 subunit studied here and the mGSTM4 subunit [32,34] have 94% amino acid and 91% mRNA sequence identities, including the characteristic Cys-115 residue; in all likelihood they are authentic GST orthologues (Figure 6).

A probable reason that the rGSTM4-4 and mGSTM4-4 proteins have not as yet been isolated from tissue extracts is their low abundance. Indeed, the retention times of recombinant rGSTM4 and natural rGSTM1 subunits are very similar and they are difficult to resolve by HPLC methods under various conditions. The minor rGSTM4 could thus be obscured (J. D. Rowe and I. Listowsky, unpublished work). It is noteworthy, therefore, that a 30-fold induction of the mGSTM4 mRNA (GT9.3) was observed in livers of mice treated with 2(3)-t-butyl-4-hydroxyanisole (BHA) [32]. In this context, the ease with which rGSTM4-4 forms mixed disulphides with GSH suggests that this GST subclass can be sensitive to cellular redox status and participate in a cell's oxidative response. The activation of rat microsomal GST, by S-glutathiolation by GSSG [42] in the oxidative environment of the membrane system, probably occurs by a common thiol/disulphide-exchange reaction. Low levels of S-glutathiolation of GSTs have also been detected previously among several human Mu-class GSTs [10]. A GST from *Proteus mirabilis* (Pm GSTB1-1) and an Omega-class human GST (GSTO1-1) also have a covalently bound GSH at their active-site Cys-10 and Cys-32 residues, respectively [9,43]. For bacterial GST the covalent GSH is at the G-site and impinges on conjugating activity. By contrast, rGSTM4-4 has a highly reactive Cys-115 residue near its active site that reacts with reduced GSH under relatively mild oxidizing conditions (aerobic conditions for 1 h, Figure 3). This enzyme also allows for a second non-covalent GSH bound at the G-site to function catalytically. Glutathiolation of rGSTM4-4 could conceivably optimize the enzyme for certain physiological functions, as yet to be defined. Alternatively, since the mGSTM4-4 mRNA is inducible by BHA [32], it may be presumed that rGSTM4-4 can also be induced by certain types of redox-active compound and that rGSTM4-4 can function in the oxidative response of rat tissue(s) [44]. This rat and mouse subclass of Mu GSTs with Cys-115 residues could have unique physiological functions that are yet to be determined.

The pH optimum of pH 8.5 did not change for the CDNB-conjugation reaction by introducing the C115Y mutation. It is likely therefore that the amino acid differences in the $\alpha 4$ helical region, and their predicted impact on the function of Arg-107, accounts for the higher pK_a of catalysis of rGSTM4-4 (Figure 2). In this context it is noteworthy that the K_m^{GSH} values for the rGSTM4-4 enzyme are also unusually high (Table 2). A structural explanation for these observations can be based on the fact that the side chain of Arg-107 of rGSTM4 is actually buried behind the Met-111 and Ile-108 side chains (Figure 5). Based on

		30		60
rGSTM4	AMILGYWNVRLTHPIRLLLEYTDSNYEEKRYVMGDAPNFDRSQWLSEKFNGLDIPNLP			
rGSTM1	PMILGYWNVRLTHPIRLLLEYTDSYEEKRYAMGDAPDYDRSQWLNEKFKLGLDFPNLP			
mGSTM4	PMTLGYWNVRLTHPIRLLLEYTDSYEEKRYVMGDAPNFDRSQWLSEKFNGLDIPNLP			
		90		120
rGSTM4	YLIDGSHKVTQSNAILRYLGRKHNLCGETEEERIRVDTLENQVMDTRIHLMI VCCSPDFE			
rGSTM1	YLIDGSRKITQSNAIMRYLARKHHLCGETEEERIRADIVENQVMDNRMLIMLCYNPDFE			
mGSTM4	YLIDGSHKVTQSNAILRYLGRKHNLCGETEEERIRVDTLENQVMDTRIQLMIVCCSPDFE			
		150		180
rGSTM4	KQKPEFLKSIPEKMKIYSEFLGKRPWFAGDKVTYVDFLAYDILDQYRMFEPECLDAFPNL			
rGSTM1	KQKPEFLKTIPEKMKIYSEFLGKRPWFAGDKVTYVDFLAYDILDQYHIFEPKCLDAFPNL			
mGSTM4	KQKPEFLKAIPEKMKIYSEFLGKRPWFAGDKVTYVDFLAYDILDQYRMFEPKCLDAFPNL			
		210		
rGSTM4	KDFLARFEGLKKISAYMKSSSFLPRPVFTKIPQWGTD			
rGSTM1	KDFLARFEGLKKISAYMKSSRYLSTPIFSKLAQWSNK			
mGSTM4	RDFLARFEGLKKISAYMKSSRFLPRPVFTKIAQWGTD			

Figure 6 Sequence alignment of related Mu-class GSTs

The amino acid sequence of the rGSTM4 subunit reported here is compared with those of rGSTM1 [21] and mGSTM4 [32–34] subunits. The shaded residues indicate amino acids that differ from those of rGSTM4.

Figure 7 Topological representation of GSH covalently linked to Cys-115 of the rGSTM4 subunit

A surface model displaying the catalytic cleft of the rGSTM4 subunit is shown for two structures of the covalently bound GSH. Structures (A) and (B) represent two energetically favourable configurations for GSH covalently linked to Cys-115 (see text). In structure (A), a non-covalently bound GSH at the active site is shown below the covalently bound GSH; in this structure binding of CDNB is sterically restricted by the orientation of the GSH. Structure (B) allows for binding of CDNB.

Figure 8 Representation of the dimer interface of rGSTM4-4

Each subunit containing GSH covalently linked to Cys-115 via a mixed disulphide (darker space-filling model) and non-covalently bound GS-DNB molecules (lighter spheres) are viewed near their interface. Note the close spatial proximity of the covalently linked GSH molecules of the two subunits.

the assumption that non-polar side chains within the active site are not allowed to move freely with respect to each other, no conformationally allowed rotamers of Arg-107 can be constructed in which the guanidino group is situated closer than 5 Å from the sulphur atom of enzyme-bound GSH. Modelling attempts to modify relative positions of the Met-111 and Ile-108 side chains to allow Arg-107 to be shifted closer to the SH group of bound GSH greatly increased the solvent-accessible area within the active site, which is thermodynamically unfavourable. For instance, the Met-111 side-chain atoms are within van-der-Waals distances of the Ile-108 and Phe-208 side chains; movement of Met-111 is likely to distort the overall protein fold. The presence of a Met-111 residue is unique for the rGSTM4 subunit, and can account for the inability of Arg-107 to function as a counter ion to promote GSH thiolate anion formation as it does for other GSTs [28].

The hydrophobic active-site pocket of rGSTM4-4 is relatively wide open and can accommodate bulky ligands. One wall of the H-site consists of Met-104, Ile-108, Met-111 and Ile-112 side chains, and the other side is formed by Val-9, Thr-209, Ile-211 and probably Met-34. Although the Cys-115 side chain is relatively solvent-accessible, in the ligand-free enzyme it is predicted that the SH group of Cys-115 is hydrogen-bonded to a backbone oxygen atom of Ile-112. When GS-DNB is incorporated in the active site of the enzyme, the dinitrophenyl moiety is relatively distant (more than 3.5 Å) from surrounding amino acid residues, although the Met-111 side chain may potentially stabilize the dinitrobenzene ring via van-der-Waals interactions. In general, the contribution of this moiety to the overall stabilization energy can be significant, because binding of the ligand produces a decrease in water accessibility within the very hydrophobic H-site surface of rGSTM4-4 ($\approx 115 \text{ Å}^2$). In

the presence of GSH alone about $70\text{--}75 \text{ Å}^2$ of the protein surface area is hidden from solvent.

Cys-115 can non-enzymically form a mixed disulphide with GSH (Figure 3), and models for the S-glutathiolated rGSTM4-4 predict the existence of at least two different conformational states for the covalently linked ligand. One model (Figure 7A) was refined with a GSH–Cys-115 mixed disulphide actually situated within the active-site pocket; although the covalently linked GSH should not interfere with the non-covalent GSH binding, this structure can potentially shut off the H-site for some substrates, including CDNB (see Figure 7A). In this model, the covalently linked GSH moiety blocks access to Met-111 and Phe-208. Another allowed model (Figure 7B) places the covalently linked ligand facing away from the active site, filling in the dimer interface. It is noteworthy that the overall protein conformation in these two models is essentially the same, with RMSD values (excluding the covalently linked GSH moieties) between all atomic co-ordinates of 0.24 Å (which is actually less than estimated co-ordinate error values). Additional alternative models can be constructed, but these require that a majority of the active-site side chains be allowed to drastically change conformations, conditions that are unlikely to yield a catalytically active enzyme.

In the first model (Figure 7A), the covalently linked GSH moiety does not interact with neighbouring residues, except for electrostatic interactions with the guanidino group of Arg-42. Since the active site is largely hydrophobic, additional electrostatic interactions are not expected. In the second model (Figure 7B), the covalently linked GSH molecules of the dimer do not interact electrostatically with side-chains of the protein, but may maintain hydrogen bonds between each other near the dimer interface (Figure 8). For instance, these interactions involve

the carboxy groups of the γ -Glu moieties of the GSH at one end, and carboxy groups of the Gly moieties of the GSH at the other. Both S-glutathiolated protein monomers appear to be identical. In this model, the covalently linked GSH moiety is located closer to the H-site residues Met-104, Ile-108 and Met-111, but does not interact directly with any of them (see Figure 7B).

Modification of Cys-115 in either way should stabilize an overall dimeric structure of rGSTM4-4 by decreasing the hydrophobic protein surface area exposed to solvent by 150 Å², thus providing a stabilization energy value of at least 1–2 kcal/mol or more (Figure 8). Assuming that the covalently bound GSH is not subject to rapid oxidation–reduction exchange, the modification of Cys-115 (Figure 7A) should greatly interfere with catalysis, reducing the apparent affinity constant for the second substrate and turnover numbers. Model A therefore does not allow the enzyme to accommodate GSH derivatives larger than S-methylglutathione at the active site. By contrast, the model in Figure 7(B) actually does not interfere with GSH and GS-DNB binding, thereby theoretically allowing catalysis to occur. Model B thus is the more likely structure for an active glutathiolated rGSTM4-4.

We thank Edward M. Nieves of the Laboratory for Macromolecular Analysis and Proteomics of Albert Einstein College of Medicine (Bronx, NY, U.S.A.) for the ESI-MS, C. Channa Reddy of The Pennsylvania State University (University Park, PA, U.S.A.) for several GST samples, Donna Carey of The Pennsylvania State University for raising antibodies against GST-Y_{b1} and GST-Y_{b4} peptides, and Huangen Ding for the glutathione-deficient *E. coli* strain JTG10. The secretarial assistance of Ms Eileen McConnell is also acknowledged. This work was supported in part by a grant (ES 02678) from the National Institute of Environmental Health Sciences, USPHS (to C.-P. D. T.) and a grant (CA 42448) from the National Cancer Institute (to I. L.). J. C. was supported by the Center of Molecular Toxicology, Environmental Research Resources Institute of The Pennsylvania State University in the spring of 1998.

REFERENCES

- Jakoby, W. B. and Ziegler, D. M. (1990) The enzymes of detoxication. *J. Biol. Chem.* **265**, 20715–20718
- Hayes, J. D. and Pulford, D. (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
- Lai, H.-C. J. and Tu, C.-P. D. (1986) Rat glutathione S-transferases supergene family: characterization of an anionic Y_b subunit cDNA clone. *J. Biol. Chem.* **261**, 13793–13799
- Daniel, V. (1993) Glutathione S-transferases: gene structure and regulation of expression. *Crit. Rev. Biochem. Mol. Biol.* **28**, 173–207
- Listowsky, I. (1993) Glutathione S-transferases: intracellular binding, detoxification, and adaptive responses. In *Hepatic Anion Transport and Bile Secretion: Physiology and Pathophysiology* (Tavolini, N. N and Berk, P. D., eds.), pp. 397–405, Raven Press, New York
- Rushmore, T. H. and Pickett, C. B. (1993) Glutathione S-transferases, structure, regulation, and therapeutic implications. *J. Biol. Chem.* **268**, 11475–11478
- Hayes, J. D. and McLellan, L. I. (1999) Glutathione and glutathione-dependent enzymes represent a coordinately regulated defence against oxidative stress. *Free Radical Res.* **31**, 273–300
- Armstrong, R. N. (1997) Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem. Res. Toxicol.* **10**, 2–18
- Board, P. G., Coggan, M., Chelvanayagam, G., Eastel, S., Jermin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffior, 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
- Rowe, J. D., Patskovsky, Y. V., Patskovsky, L. N., Novikova, E. and Listowsky, I. (1998) Rationale for reclassification of a distinctive subdivision of mammalian class Mu glutathione S-transferases that are primarily expressed in testis. *J. Biol. Chem.* **273**, 9593–9601
- Patskovsky, Y., Huang, Y. Q., Takayama, T., Listowsky, I. and Pearson, W. R. (1999) Distinctive structure of the human GSTM3 gene-inverted orientation relative to the mu class glutathione transferase gene cluster. *Arch. Biochem. Biophys.* **361**, 85–93
- Lai, H.-C. J., Qian, B., Grove, G. and Tu, C.-P. D. (1988) Gene expression of rat glutathione S-transferases: evidence for gene conversion in the evolution of the Y_b multigene family. *J. Biol. Chem.* **263**, 11389–11395
- Qian, B. and Tu, C.-P. D. (2000) An analysis of the rat glutathione S-transferase Y_{b4} gene sequence. *Clin. Chem. Enzymol. Commun.* **8**, 335–347
- Tu, C.-P. D., Weiss, M. J., Li, N. and Reddy, C. C. (1983) Tissue specific expression of rat glutathione S-transferases. *J. Biol. Chem.* **258**, 4659–4662
- Tu, C.-P. D. and Reddy, C. C. (1985) On the multiplicity of rat liver glutathione S-transferases. *J. Biol. Chem.* **260**, 9961–9964
- Reddy, C. C., Burgess, J. R., Gong, Z.-Z., Massaro, E. J. and Tu, C.-P. D. (1983) Purification and characterization of individual glutathione S-transferases from sheep liver. *Arch. Biochem. Biophys.* **224**, 87–101
- Abramovitz, M. and Listowsky, I. (1987) Selective expression of a unique glutathione S-transferase Yb3 gene in rat brain. *J. Biol. Chem.* **262**, 7770–7773
- Abramovitz, M., Ishigaki, S., Felix, A. M. and Listowsky, I. (1988) Expression of an enzymatically active Yb3 glutathione S-transferase in *Escherichia coli* and identification of its natural form in rat brain. *J. Biol. Chem.* **263**, 17627–17631
- Deng, W. P. and Nickoloff, J. A. (1992) Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal. Biochem.* **200**, 81–88
- Tang, A. H. and Tu, C.-P. D. (1994) Biochemical characterization of *Drosophila* glutathione S-transferases D1 and D21. *J. Biol. Chem.* **269**, 27876–27884
- Lai, H.-C. J., Grove, G. and Tu, C.-P. D. (1986) Cloning and sequence analysis of a cDNA for a rat liver glutathione S-transferase Y_b subunit. *Nucleic Acids Res.* **14**, 6101–6114
- Lee, H.-C., Toung, Y.-P. S., Tu, Y.-S. L. and Tu, C.-P. D. (1995) A molecular genetic approach for the identification of essential residues in human glutathione S-transferase function in *Escherichia coli*. *J. Biol. Chem.* **270**, 99–109
- Greenberg, J. T. and Demple, B. (1986) Glutathione in *Escherichia coli* is dispensable for resistance to H₂O₂ and gamma radiation. *J. Bacteriol.* **168**, 1026–1029
- Habig, W. H. and Jakoby, W. B. (1981) Assays for differentiation of glutathione S-transferases. *Methods Enzymol.* **77**, 398–405
- Tam, J. P. (1988) Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5409–5413
- Kuan, I. C., Johnson, K. A. and Tien, M. (1993) Kinetic analysis of manganese peroxidase. The reaction with manganese complexes. *J. Biol. Chem.* **268**, 20064–20070
- Ji, X., Zhang, P., Armstrong, R. N. and Gilliland, G. L. (1992) The three-dimensional structure of a glutathione S-transferase from the mu gene class: structural analysis of the binary complex of isoenzyme 3–3 and glutathione at 2.2-Å resolution. *Biochemistry* **31**, 10169–10184
- Patskovsky, Y. V., Patskovska, L. N. and Listowsky, I. (2000) The enhanced affinity for thiolate anion and activation of enzyme-bound glutathione is governed by an arginine residue of human Mu class glutathione S-transferases. *J. Biol. Chem.* **275**, 3296–3304
- McCallum, S. A., Hitchens, T. K., Torborg, C. and Rule, G. S. (2000) Ligand-induced changes in the structure and dynamics of a human class Mu glutathione S-transferase. *Biochemistry* **39**, 7343–7356
- Bruger, A. T. (1992) XPLOR: a system for x-ray crystallography and NMR, Version 3.1, Yale University Press, New Haven, CT
- Li, N., Toung, Y.-D., Zarlingo, R., Weiss, M. J., Grove, G. W., Reddy, C. C. and Tu, C.-P. D. (1987) The rat glutathione S-transferase gene superfamily: molecular basis of gene multiplicity. *Fed. Proc.* **46**, 2186
- Pearson, W. R., Reinhart, J., Sisk, S. C., Anderson, K. S. and Adler, P. N. (1988) Tissue-specific induction of murine glutathione transferase mRNAs by butylated hydroxyanisole. *J. Biol. Chem.* **263**, 13324–13332
- Townsend, A. J., Goldsmith, M. E., Pickett, C. B. and Cowan, K. H. (1989) Isolation, characterization, and expression in *Escherichia coli* of two murine Mu class glutathione S-transferase cDNAs homologous to the rat subunits 3 (Y_{b1}) and 4 (Y_{b2}). *J. Biol. Chem.* **264**, 21582–21590
- Reinhart, J. and Pearson, W. R. (1993) The structure of two murine class-mu glutathione transferase genes coordinately induced by butylated hydroxyanisole. *Arch. Biochem. Biophys.* **303**, 383–393
- Johnson, W. W., Liu, S., Ji, X., Gilliland, G. L. and Armstrong, R. N. (1993) Tyrosine 115 participates both in chemical and physical steps of the catalytic mechanism of a glutathione S-transferase. *J. Biol. Chem.* **268**, 11508–11511
- Fu, J.-H., Rose, J. P., Tam, M. F. and Wang, B.-C. (1994) New crystal forms of a mu class glutathione S-transferase from rat liver. *Acta Crystallogr.* **D2**, 219–224
- Raghunathan, S., Chandross, R. J., Kretsinger, R. H., Allison, T. J., Penington, C. J. and Rule, G. S. (1994) Crystal structure of human class mu glutathione transferase GSTM2-GSTM2. Effects of lattice packing on conformational heterogeneity. *J. Mol. Biol.* **238**, 815–832
- Ji, X., Johnson, W. W., Sesay, M. A., Dickert, L., Prasad, S. M., Armon, H. L., Armstrong, R. N. and Gilliland, G. L. (1994) Structure and function of the xenobiotic substrate binding site of a glutathione S-transferase as revealed by X-ray crystallographic analysis of product complexes with the diastereomers of 9-(S-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene. *Biochemistry* **33**, 1043–1052

-
- 39 Wilce, M. C. J. and Parker, M. W. (1994) Structure and function of glutathione *S*-transferases. *Biochim. Biophys. Acta* **1205**, 1–18
- 40 Patskovsky, Y. V., Patskovska, L. N. and Listowsky, I. (1999) Functions of His107 in the catalytic mechanism of human glutathione *S*-transferase hGSTM1a-1a. *Biochemistry* **38**, 1193–1202
- 41 Liu, S., Zhang, P., Ji, X., Johnson, W. W., Gilliland, G. L. and Armstrong, R. N. (1992) Contribution of tyrosine 6 to the catalytic mechanism of isoenzyme 3–3 of glutathione *S*-transferase. *J. Biol. Chem.* **267**, 4296–4299
- 42 Dafre, A. L., Sies, H. and Akerboom, T. (1996) Protein *S*-thiolation and regulation of microsomal glutathione transferase activity by the glutathione redox couple. *Arch. Biochem. Biophys.* **332**, 288–294
- 43 Rossjohn, J., Polekhina, G., Feil, S. C., Allocati, N., Masulli, M., Di Ilio, C. and Parker, M. W. (1998) A mixed disulfide bond in bacterial glutathione transferase: functional and evolutionary implications. *Structure* **6**, 721–734
- 44 Thomas, J. A., Poland, B. and Honzatko, R. (1995) Protein sulfhydryls and their role in the antioxidant function of protein *S*-thiolation. *Arch. Biochem. Biophys.* **319**, 1–9
-

Received 15 December 2000/1 February 2001; accepted 12 March 2001