Identification of cDNAs encoding two human Alpha class glutathione transferases (GSTA3 and GSTA4) and the heterologous expression of GSTA4-4

Philip G. BOARD

John Curtin School of Medical Research, Australian National University, GPO Box 334, Canberra ACT 2601, Australia

The Expressed Sequence Tag database has been searched for examples of previously undescribed human Alpha class glutathione transferases. An incomplete transcript of the previously described *GSTA3* gene was identified in a cDNA library derived from 8–9 week placenta. This indicates that the *GSTA3* gene is functional and is possibly under specific developmental regulation. A second cDNA, termed GSTA4, was identified in a brain cDNA library. The encoded GSTA4-4 enzyme was expressed in *Escherichia coli* and was found to be immunologically distinct from GSTA1-1 and to have high activity with alk-2-enals. Although GSTA4-4 appears to be functionally similar to the mouse GST5.7 and rat GST8-8 Alpha class enzymes, sequence comparisons and phylogenetic analysis suggest that GSTA4-4 may be a member of a distinct Alpha class subgroup.

INTRODUCTION

Over the past 10 years cDNAs encoding twelve distinct soluble glutathione transferases (GSTs) have been described in humans. These are members of the Alpha, Mu, Pi, Theta, Kappa and Zeta evolutionary classes [1–3]. Members of each class tend to have high sequence identity (>60%) which falls to less than 25% between classes. Studies so far suggest that the Mu class is the most complex with five genes being identified in a cluster on chromosome 1 [4,5]. Other classes, such as Pi and Zeta, are clearly much less complex with only single genes being identified in humans [3,6,7].

Although only two functional Alpha class GST genes have been positively identified in humans by the cloning of their corresponding cDNAs [8–10], there is considerable evidence suggesting that this class may be more complex. Southern blots probed with a GSTA1 cDNA have generated an extensive pattern of hybridizing fragments, suggesting the presence of multiple Alpha class genes and/or pseudogenes [9]. Genomic cloning studies have identified a number of clones that contain exons that do not correspond to the GSTA1 and GSTA2 sequences [8,11,12]. These studies also identified a complete gene termed *GSTA3* that could potentially be expressed but for which no corresponding cDNA or protein product has ever been identified [11]. Studies in rats and mice have identified Alpha class GSTs that have high activity with 4-hydroxynonenal [13,14], and a human enzyme with these properties has been identified using antibodies to murine GST5.7 [15]. Given the genomic and immunological evidence it is possible that several additional Alpha class GST genes may be expressed in a tissue-specific and/or developmentally regulated manner that has prevented detection of their cDNAs.

The recent expansion of the Expressed Sequence Tag (EST) database has made it possible to search numerous human tissues from different stages of development for evidence of expression of any gene for which some related protein coding DNA sequence is available. Because of the possibility of there being additional undescribed human Alpha class GSTs, the EST database was searched with human and rodent Alpha class GST cDNA sequences and two new human cDNAs were identified. One was an incomplete transcript of the previously described *GSTA3* gene and the second showed similarity with the mouse GST5.7 and rat 8 sequences [11,13,14]. This study reports the cDNA sequence of each clone and the heterologous expression of the human GST with similarities to the mouse GST5.7 and rat 8-8 enzymes.

MATERIALS AND METHODS

cDNA clone identification and sequencing

Two human cDNAs with sequences related to previously identified Alpha class GSTs were identified by a BLAST search of the EST database through the facility provided by the National Centre for Biotechnology Information located at the World Wide Web site http://www.ncbi.nlm.nih.gov/BLAST/. The EST (yw63d06.r1) corresponding to the *GSTA3* gene was isolated from an 8 to 9 week placental cDNA library constructed by Bento Soares and M. Fatima Bonaldo (Merck and Co., Whitehouse Station, NJ, U.S.A.). This cDNA clone subsequently referred to as pGSTA3. The EST (yn64c05.r1) with similarities to the mouse GST5.7 and rat 8 GSTs was isolated from an adult human brain cDNA library constructed by Bento Soares and M. Fatima Bonaldo. This clone is subsequently referred to as pGSTA4. Both cDNA clones were made available by the IMAGE consortium [16].

The cDNAs were sequenced on both strands by the use of specific oligonucleotide primers and the deletion of fragments from either end by specific restriction enzyme digests. All sequencing was carried out using a Thermo Sequenase cycle sequencing kit (Amersham Life Sciences).

Heterologous expression of GSTA4-4

The protein encoded by the pGSTA4 cDNA was expressed in *E. coli* as a ubiquitin fusion in the plasmid pRB269 [16]. The

Abbreviations used: GST, glutathione transferase; EST, expressed sequence tag.

ubiquitin moiety was co-translationally removed by a co-expressed ubiquitin-specific protease Ubp1 as previously described in detail [16,17]. Briefly a pair of oligonucleotide primers, A4 EXA 5′-GAATCCGCGGTGGTATGGCAGCAAGGCCC-3′ and A4.2 5′-CTACAGGACATGAGTAGCC-3′, were used to amplify a 951 bp fragment from the pGSTA4 cDNA and to add a *Sac*II restriction enzyme site and sequence encoding the two C-terminal glycine residues of ubiquitin at the 5′ end immediately up stream of the GSTA4 ATG translation start codon. The amplified sequence was digested with *Sac*II and *Bam*HI and cloned into pBluescript SK- and the 181 bp region from *Sac*II to an *Nco*I site was sequenced to ensure no errors had been introduced during the PCR amplification procedure. Subsequently a 571 bp *NcoI–Bam*H1 fragment was removed and replaced by the equivalent unamplified fragment from the original pGSTA4 cDNA clone. The 752 bp *Sac*II–*Bam*H1) fragment was then excised from pBluescript SK- and ligated between the *Sac*II and *Bam*H1 sites of pRB269. The resulting clone pGSTA4 EX expressed a ubiquitin–GSTA4 fusion protein when grown in the presence of IPTG. When the ubiquitin-specific protease Ubp1 was co-expressed in these cells the ubiquitin moiety was rapidly removed from the N-terminus of the GST.

E. coli expressing recombinant GSTA4-4 were collected by centrifugation and after resuspension in 20 mM $NaPO_4$, pH 7.2 the cells were lysed by passage through a Sorval Ribi cell disrupter. The cellular debris was removed by centrifugation and the recombinant enzyme was purified from the supernatant by glutathione agarose affinity chromatography [18]. The pure enzyme was eluted from the glutathione agarose in 5 mM GSH, 50 mM Tris, pH 9.6, and was dialysed against 20 mM Na_3PO_4, pH 7.2. Recombinant GSTA1 was purified in a similar manner after expression in *E. coli* from a previously described clone [19].

Enzyme characterization

GST assays using 1-chloro-2,4-dinitrobenzene, 1,2-dichloro-nitrobenzene, ethacrynic acid, 1,2-epoxy-3-(4-nitrophenoxy)propane, 4-phenylbut-3-en-2-one and β-napthylacetate as substrates were previously described in detail by Mannervik and Widersten [20]. Activity with 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole was determined spectrophotometrically at 419 nm in sodium acetate buffer, pH 5.0, as described by Ricci and co-workers [21]. GST assays using *trans*,*trans*-alka-2,4-dienals and *trans*-alk-2-enals have been described by Brophy and co-workers [22]. The determination of GST activity with 1-menaphthyl sulphate was essentially as described by Gillham [23], and glutathione peroxidase activity with *t*-butyl hydroperoxide and cumene hydroperoxide was determined by the procedure described by Beutler [24].

After separation of the recombinant protein by SDS polyacrylamide-gel electrophoresis [25] the gels were either stained with Coomassie Blue or were electroblotted on to nitrocellulose filters [26]. The blots were developed with antiserum raised in rabbits against purified recombinant GSTA1-1 [19] and goat anti-(rabbit IgG) alkaline phosphatase conjugate (Sigma). The isoelectric point of GSTA4-4 was determined by isoelectric focusing on Phast gels (Pharmacia) using standard proteins of known pI and previously purified and characterized GSTs. Amino terminal sequencing was carried out by the Australian National University Biomolecular Resource Facility.

Phylogenetic analysis

The alignment of the amino acid sequences was carried out manually in the Genetic Data Environment 2.2 [27] and between sequence distances were computed using Kimura's method [28]. The phylogeny of the sequences was estimated using a neighbour-joining method [29].

RESULTS

cDNA sequences

The cDNA sequences of the GSTA3 and GSTA4 cDNAs are shown in Figures 1 and 2. The GSTA3 cDNA is 743 bp in length and appears to be incomplete. The first 26 nucleotides do not match any known GST sequence but after nucleotide 27 the translated sequence matches the start of exon 3 of the *GSTA3* gene [11]. The deduced amino acid sequence contains five amino acid substitutions compared with the sequence deduced from the *GSTA3* gene. As this is the only cDNA derived from the *GSTA3* gene so far identified it is not yet clear if the differences are owing to genetic polymorphism or sequencing errors. However, since the *GSTA3* gene sequence was only obtained by single pass automated sequencing, it is possible that there may be errors in the original sequence. Comparison of the available sequence with GSTA1 revealed 93% identity, confirming that *GSTA3* is very closely related to GSTA1 and GSTA2.

The GSTA4 cDNA is 1255 bp in length and encodes a protein of 222 residues with a deduced molecular mass of 25.687 Da. Comparison of this protein with GSTA1 reveals only 52% identity. However, comparison of the GSTA4 amino acid sequence with that of mouse GST5.7, rat 8 and chicken CL-3 [30] reveals 58, 56 and 57% respectively (Figure 3). A neighbour-joining tree constructed from the alignment in Figure 3 is shown in Figure 4 and indicates that while GSTA3 is closely related to GSTA1 and GSTA2, GSTA4 appears to be equidistant from the

```
                                   <
GCAGAGACCGGCTAGACTTTACTCAATTTGAAGAGAAATTTATAGGATCTGCAGAAGATT     60
                               F  E  E  K  F  I  G  S  A  E  D  L    41
                        ><
TGGGAAAGTTAAGAAATGATGGGAGTTTGATGTTCCAGCAAGTACCAATGGTTGAGATTG    120
  G  K  L  R  N  D  G  S  L  M  F  Q  Q  V  P  M  V  E  I  D     61

ATGGGATGAAGTTGGTACAGACCAGAGCCCTTCTCAACTACATTGCCAGCAAATACAACC    180
  G  M  K  L  V  Q  T  R  A  L  L  N  Y  I  A  S  K  Y  N  L     81
                                  ><
TCTACGGGAAAGACATAAAGGAGAGAGCCCTAATTGATATGTATACAGAAGGTATGGCAG    240
  Y  G  K  D  I  K  E  R  A  L  I  D  M  Y  T  E  G  M  A  D    101

ATTTGAATGAAATGATCCTTCTTCTGCCCTTATGTCGACCTGAGGAAAAAGATGCCAAGA    300
  L  N  E  M  I  L  L  L  P  L  C  R  P  E  E  K  D  A  K  I    121
                                                        ><
TTGCCTTGATCAAAGAGAAAACAAAAAGTCGCTATTTCCCTGCCTTCGAAAAAGTGTTAC    360
  A  L  I  K  E  K  T  K  S  R  Y  F  P  A  F  E  K  V  L  Q    141

AGAGCCATGGACAAGACTACCTTGTTGGCAACAAGCTGAGCCGGGCTGACATTAGCCTGG    420
  S  H  G  Q  D  Y  L  V  G  N  K  L  S  R  A  D  I  S  L  V    161

TGGAACTTCTCTACTATGTGGAAGAGCTTGACTCCAGCCTTATCTCCAACTTCCCTCTGC    480
  E  L  L  Y  Y  V  E  E  L  D  S  S  L  I  S  N  F  P  L  L    181
     ><
TGAAGGCCCTGAAAACCAGAATCAGCAACCTGCCCACGGTGAAGAAGTTTCTACAGCCTG    540
  K  A  L  K  T  R  I  S  N  L  P  T  V  K  K  F  L  Q  P  G    201

GCAGCCCAAGGAAGCCTCCCGCAGATGCAAAAGCTTTAGAAGAAGCCAGAAAGATTTTCA    600
  S  P  R  K  P  P  A  D  A  K  A  L  E  E  A  R  K  I  F  R    221

GGTTTTAATAAAGCAGCCATGGAGGCTAAGAACATGCAAGACCAATATTCTAAAGTTTTG    660
  F  *                                                           222

CAACAATGAAGTGCTTTACTTAAGTGTTGATTGTGCCTGTTGTAAAGCTAATGAACCCTT    720
TCCAATTATATGCTAATTAAATAATAAAACTCCTATTTGCT                        762
```

Figure 1 Nucleotide sequence of a GSTA3 cDNA and its deduced amino acid sequence

The first 26 nucleotides do not code for any recognizable GST sequence. The first recognizable amino acid sequence coincides exactly with the start of exon 3 of the *GSTA3* gene [11]. Inverted arrows above the nucleotide sequence show the positions of splice junctions. The stop codon is indicated by an asterisk. The poly(A) addition signal is underlined. The sequence has accession number AF020919.

```
GGGACCGCTGACCTGGCGCTTTGTGCGGCTCCAGGCCTCCGAGTGACTCCAGAAAGCCTG    60
AAAAGCTATCATGGCAGCAAGGCCCAAGCTCCACTATCCCAACGGAAGAGGCCGGATGGA   120
          M  A  A  R  P  K  L  H  Y  P  N  G  R  G  R  M  E     17

GTCCGTGAGATGGGTTTTAGCTGCCGCCGGAGTCGAGTTTGATGAAGAATTTCTGGAAAC   180
 S  V  R  W  V  L  A  A  A  G  V  E  F  D  E  E  F  L  E  T     37

AAAAGAACAGTTGTACAAGTTGCAGGATGGTAACCACCTGCTGTTCCAACAAGTGCCCAT   240
 K  E  Q  L  Y  K  L  Q  D  G  N  H  L  L  F  Q  Q  V  P  M     57

GGTTGAAATTGACGGGATGAAGTTGGTACAGACCCGAAGCATTCTCCACTACATAGCAGA   300
 V  E  I  D  G  M  K  L  V  Q  T  R  S  I  L  H  Y  I  A  D     77

CAAGCACAATCTCTTTGGCAAGAACCTCAAGGAGAGAACCCTGATTGACATGTACGTGGA   360
 K  H  N  L  F  G  K  N  L  K  E  R  T  L  I  D  M  Y  V  E     97

GGGGACACTGGATCTGCTGGAACTGCTTATCATGCATCCTTTCTTAAAACCAGATGATCA   420
 G  T  L  D  L  L  E  L  L  I  M  H  P  F  L  K  P  D  D  Q    117

GCAAAAGGAAGTGGTTAACATGGCCCAGAAGGCTATAATTAGATACTTTCCTGTGTTTGA   480
 Q  K  E  V  V  N  M  A  Q  K  A  I  I  R  Y  F  P  V  F  E    137

AAAGATTTTAAGGGGTCACGGACAAAGCTTTCTTGTTGGTAATCAGCTGAGCCTTGCAGA   540
 K  I  L  R  G  H  G  Q  S  F  L  V  G  N  Q  L  S  L  A  D    157

TGTGATTTTACTCCAAACCATTTTAGCTCTAGAAGAGAAAATTCCTAATATCCTGTCTGC   600
 V  I  L  L  Q  T  I  L  A  L  E  E  K  I  P  N  I  L  S  A    177

ATTTCCTTTCCTCCAGGAATACACAGTGAAACTAAGTAATATCCCTACAATTAAGAGATT   660
 F  P  F  L  Q  E  Y  T  V  K  L  S  N  I  P  T  I  K  R  F    197

CCTTGAACCTGGCAGCAAGAAGAAGCCTCCCCCTGATGAAATTTATGTGAGAACCGTCTA   720
 L  E  P  G  S  K  K  K  P  P  P  D  E  I  Y  V  R  T  V  Y    217

CAACATCTTTAGGCCATAAAACAACACATCCATGTGTGAGTGACAGTGTGTTCCTAGAGA   780
 N  I  F  R  P  *                                              222

TGGTATTGTCTACAGTCATGTCTTAATGGATCCCAGCTCTGTCATGGTGCTATCTATGTA   840
TTAAGTTGGGTCCTAAGTTGGGTCTTTTGTGTCAAAGAGATCATCTCTTCTAGAAATATC   900
AACCTTTTTTGTCCAGTAAATAATTGTTAGGGGATCTTTATTGGAAAACTTTTTTGGAGA   960
GGCTGGTATTTAAGTTAGATCTGATTGGGCTACTCATGTCCTGTAGCCAGTTCATCCTCA  1020
TAATAAGAATGGGCAGGATCTCTTGTTCTCTCCTGAGTGTCTTTCTACTCTCCTGAGCGT  1080
CTTTCTGCTCTCCTTATCCTGTTCTCTTATCCTTATCCCCTCCAGTCTCTGCCTAATTTT  1140
TAGTGTTTAATAACAACCGAATGTCTAGTAAATGACTCTCCTCTGAGCTGTAATAAATAA  1200
AATGGTAGTAATGAATGCAATCAGTGTTAGCCAAAATAAAGATTTATGAGTCATT       1255
```

Figure 2 Nucleotide sequence of a GSTA4 cDNA and its deduced amino acid sequence

The translation termination codon is shown by an asterisk. The poly(A) addition signal is underlined. The sequence has accession number AF020918.

```
C1     MSGK.V...A.T.........L........E.K...K..D.Q..KSDGS.
C3     MA.K.V.Y.F....K...I..L........E.V....R..YE.LLQSGI.
r8     MEVK...Y.FQ.......I..L..T.....E......R..YELK.KDGC.
m5.7   MA.K...Y.F........I..L........E......R..YE.M.KDG..
4      MAARPKLHYPNGRGRMESVRWVLAAAGVEFDEEFLETKEQLYKLQDGNHL
3      M.GK.....F.......PI..L........E.K.IGSA.D.G..RNDGS.
2      M.EK.....S.I......I..L........E.K.IKSA.D.D..RNDGY.
1      M.EK.....F.A......T..L........E.K.IKSA.D.D..RNDGY.


C1     ..............MV...A..N.I.G.Y..Y..D....AL....VE.LA
C3     M..................A..N.I.G.Y..Y..D....A......G..D
r8     ..G...L......L.T...A..S.L.A.Y..Y..D....VR....AD..Q
m5.7   ..G...L......M.T...A..S.L.A.Y..Y..D....VR....AD..Q
4      LFQQVPMVEIDGMKLVQTRSILHYIADKHNLFGKNLKERTLIDMYVEGTL
3      M..................AL.N...S.Y..Y..DI...A.....T..MA
2      M..................A..N...S.Y..Y..DI..KA.....I..IA
1      M..................A..N...S.Y..Y..DI..RA.....I..IA


C1     ..Y..IM.NVVQPA.KKEEHLA.ALD..AN........V.KD..HD....
C3     ..MGF.LSF...SAE.KV.QCAFVVE..TS....AY..V.KD...D....
r8     ..MMMI.GA..KA.QEKEESLALAVKR.KN..........KD..EA....
m5.7   ..MMMIAVA..KT.KEKEESYDLILSR.KT..........KD..EA....
4      DLLELLIMHPFLKPDDQQKEVVNMAQKAIIRYFPVFEKILRGHGQSFLVG
3      ..N.MILLL.LCR.EEKDAKIALIKE.TKS....A...V.QS...DY...
2      ..G.MILLL..TQ.EEQDAKLALIQE.TKN....A...V.KS...DY...
1      ..G.MILLL.V.P.EEKDAKLALIKE.IKN....A...V.KS...DY...


C1     .K..R...H..E....V..SK.DA.AK..L..SFKART.......K..Q.
C3     .R..W..IH..EA..MV...KSDA..G..L..AFKKRI.S.....K..A.
r8     ....W..IQ..EA..MV..VSAPV..D..L..AFKTRI.......K..Q.
m5.7   ....W..IQ..EA..MV..LSAPV..D..L..AFKTRI.......K..Q.
4      NQLSLADVILLQTILALEEKIPNILSAFPFLQEYTVKLSNIPTIKRFLEP
3      .K..R..IS.VELLYYV..LDSSLI.N..L.KALKTRI..L..V.K..Q.
2      .K..R..IH.VELLYYV..LDSSLI.S..L.KALKTRI..L..V.K..Q.
1      .K..R..IH.VELLYYV..LDSSLI.S..L.KALKTRI..L..V.K..Q.


C1     ..QR..RLEEKDIPRLMAIFH
C3     ...R..IS.DK..ET.RRVL.MYYDVKPH
r8     ..QR.....GH..DV.RTVLK.
m5.7   ..QR.....GP..EV.RIVLKF
4      GSKKKPPPDEIYVRTVYNIFRP
3      ..PR...A.AKALEEARK...F
2      ..PR...M..KSLEESRK...F
1      ..PR...M..KSLEEARK...F
```

Figure 3 An alignment of the amino acid sequences of human GSTA1 [9], GSTA2 [10], GSTA3, GSTA4, mouse GST5.7 [14], rat GST 8 [13], chicken GSTCL-1 [37] and chicken GSTCL-3 [30]

Residues identical with those in GSTA4 are shown as a (●). The first 29 residues of GSTA3 were not encoded in the cDNA and are derived from exon 2 of the corresponding *GSTA3* gene [11].

other human sequences and the mouse GST5.7, rat GST8 and chicken GST CL-3 cluster.

Expression and characterization of recombinant GSTA4-4

GSTA4-4 was expressed in *E. coli* as a ubiquitin fusion protein that is co-translationally cleaved at the C-terminus of the ubiquitin moiety leaving the free GST subunit. The GST was purified by affinity chromatography on glutathione agarose. As shown in Figure 5, the recombinant protein was essentially pure as determined by Coomassie Blue staining after SDS polyacrylamide-gel electrophoresis. Amino terminal sequencing revealed that the ubiquitin had been completely removed, revealing an N-terminal methionine (31 %) or an alanine residue (59 %). Thus there appears to have been some minor processing of the GSTA4 N-terminal after cleavage of the ubiquitin. Examination of Figure 5 shows that the recombinant GSTA4 appears to be slightly smaller than the recombinant GSTA1 run on the same gel. The deduced molecular masses differ by only 73 Da (GSTA1 = 25.614, GSTA4 = 25.687 kDa) and indicate that GSTA1 is the smaller of the two. Thus it seems likely that this difference in migration on the SDS gel may represent minor differences in their interactions with the gel. The possibility of C-terminal processing has not been excluded but the removal of the C-terminal helix has a detrimental effect on activity [31] and this does not appear to be a factor here (see below.)

The pI of GSTA4-4 was determined by isoelectric focusing on polyacrylamide gels and was estimated at pH 8.3, which is not quite as high as that determined for GSTA1-1 (pH 8.9) [32]. Western blots of GSTA1 and GSTA4 developed with antiserum raised against recombinant GSTA1-1 failed to show any cross-reactivity with GSTA4 (Figure 6).

The activity of GSTA4-4 was evaluated with a variety of substrates and is compared with GSTA1-1 in Table 1. There were a number of striking differences in the activity profiles. While GSTA1-1 was notably active with substrates such as 1-chloro-2,4-dinitrobenzene, 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole and cumene hydroperoxide, GSTA4-4 was considerably less active. In contrast GSTA4-4 showed very high activity with the reactive

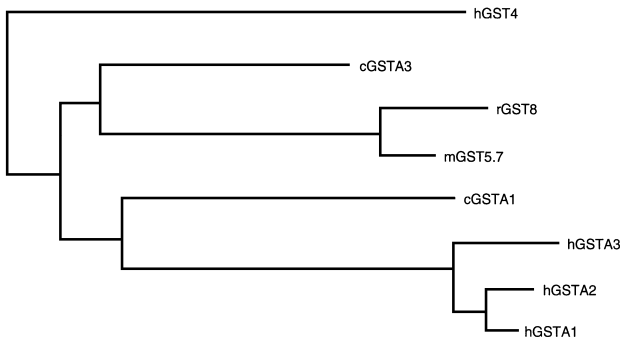

Figure 4 A neighbour-joining tree estimating the phylogenetic relationships between selected Alpha class GSTs

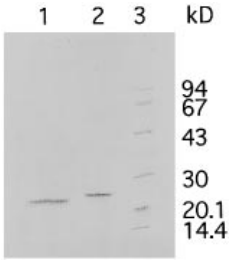

Figure 5 SDS/PAGE of purified recombinant GSTA4 (lane 1) recombinant GSTA1 (lane 2) and molecular size standards (lane 3)

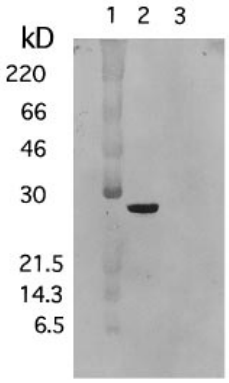

Figure 6 Western blot of recombinant GSTA1 (lane 2) and recombinant GSTA4 (lane 3)

The blot was probed with rabbit antiserum to recombinant GSTA1. Prestained molecular mass markers are shown in lane 1. Equivalent amounts of protein were added to lanes 2 and 3.

Table 1 Activity of recombinant GSTA1-1 and GSTA4-4 with various substrates

Results are means ± S.D. calculated from four determinations. N.d., not detectable.

	Specific activity [μmol · min^{-1} · (mg^{-1} of protein)]	
Substrate	GSTA1-1	GSTA4-4
1-chloro-2,4-dinitrobenzene	94.1 ± 6.66	5.3 ± 1.29
1,2-Dichloronitrobenzene	0.17 ± 0.07	N.d.
7-Chloro-4-nitrobenzo-2-oxa-1,3-diazole	66.8 ± 3.91	17.3 ± 3.54
1,2-Epoxy-3-(4-nitrophenoxy)propane	N.d.	N.d.
4-Phenylbut-3-en-2-one	0.018 ± 0.001	0.056 ± .009
Ethacrynic acid	0.17 ± 0.078	0.50 ± 0.09
p-Nitrobenzyl chloride	0.038 ± 0.09	0.053 ± 0.03
Menaphthyl sulphate	N.d.	N.d.
trans, *trans*-Hexa-2,4-dienal	0.09 ± 0.041	0.46 ± 0.109
trans, *trans*-Hepta-2,4-dienal	0.26 ± 0.030	0.43 ± 0.033
trans, *trans*-Deca-2,4-dienal	0.17 ± 0.042	0.24 ± 0.045
trans-Hex-2-enal	0.18 ± 0.04	4.24 ± 0.633
trans-Oct-2-enal	0.37 ± 0.116	43.3 ± 8.51
trans-Non-2-enal	0.86 ± 0.311	54.0 ± 3.59
t-Butyl hydroperoxide	N.d.	N.d.
Cumene hydroperoxide	19.8 ± 1.81	0.20 ± 0.34

carbonyl compounds such as *trans*-oct-2-enal and *trans*-non-2-enal. Activity increased with carbon chain length but was notably lower with *trans*,*trans*-alkadienals.

DISCUSSION

In this study the EST database was searched for cDNAs encoding additional human Alpha class GSTs. Two previously undescribed cDNAs were identified. The first encodes GSTA3 and, unfortunately, it is incomplete. However, the identification of this cDNA has clearly demonstrated for the first time that the *GSTA3* gene we previously described [11] is transcribed. The identification of a single GSTA3 transcript in an 8–9 week placental library suggests that this gene is probably developmentally regulated. In previous unpublished studies, we failed to find a copy of this cDNA in adult liver, brain, testis and heart libraries. It may be significant that the start of the recognizable coding sequence in the GSTA3 cDNA corresponds with the first residue of exon 3 of the *GSTA3* gene. This suggests that the transcript that generated the cDNA had been incorrectly processed. Further studies are required to obtain a complete cDNA so that the GSTA3-3 enzyme can be heterologously expressed and characterized in greater detail.

The second cDNA isolated encoded a GST we have termed GSTA4-4. It is clear that GSTA4 is considerably diverged from GSTA1, GSTA2 and GSTA3 since it is immunologically distinct and has only about 52% amino acid sequence identity with this group compared with > 90% identity within the GSTA1, GSTA2 and GSTA3 group. It appears that there are similar subdivisions within other GST classes. For example, GSTM3 is a clear outlier compared with the other members of the Mu class [4,33] and the Theta class can be divided into the GSTT1 and GSTT2 subgroups that have distinct sequence differences and catalytic properties [34,35]. In the case of the Theta subclasses there are clearly orthologous enzymes in a number of mammalian species [36]. In the present example the human GSTA4-4 enzyme has functional similarities with the mouse GST5.7 [14] and the rat GST8-8 [13] enzymes. These rodent enzymes have characteristically high activity with 4-hydroxynonenal which is similar to the high activity of GSTA4-4 with *trans*-2-nonenal. However, despite this functional similarity the phylogenetic analysis suggests that GSTA4 is almost equally diverged from the other human Alpha class enzymes and the rodent enzymes. Thus it appears that GSTA4 may be a member of a distinct subgroup within the Alpha class. GSTA4-4 can be compared with GST5.8, an Alpha class GST previously purified from human liver [15]. GST5.8 was identified by cross-reactivity with antiserum to mGST5.7 and although there are many interesting similarities with GSTA4-4, especially the high activity with the reactive carbonyl compounds and relatively low activity with 1,chloro-2,4-dinitrobenzene, there are two notable exceptions that suggest that these may be separate isoenzymes. First the pI of the GSTA4-4 expressed in this study is considerably more basic (pI 8.3) than the isoenzyme purified from human liver (pI 5.8). Secondly, two peptides sequenced from the human liver GST 5.8 protein [15], vary distinctly from the sequence obtained from the GSTA4 cDNA (Figure 7). The available data therefore suggest that there may be more than one human GST that utilizes reactive carbonyl compounds; however, no other cDNAs with these characteristics were identified in this study.

The alkenals are known products of oxidative metabolism [38,39] and enzymes such as GSTA4-4 appear to play a significant role in the protection of intracellular components against oxidative damage. It might therefore be expected that GSTA4-4 would be widely expressed in a number of tissues. The GSTA4 cDNA clone described here was isolated from a brain library, and during our database searches similar clones were identified in libraries derived from 8 week embryo, foetal brain, foetal heart, foetal liver, placenta, infant brain, pregnant uterus, breast,

```
hGST5.8     IAVAPFKTPKEKAELYDLIK
hGSTA4-4    LIMH--LK-DDQQKEVVNMA
mGST5.7     ------------E-S----L

hGST5.8     VEELSAPVLNEFPLLQAFIK
hGSTA4-4    L--KIPNI--A--FL-EYTV
mGST5.7     ---------SD-------KT
```

Figure 7 Comparison of peptide sequences identified in human GST5.8 [15] with the equivalent sequences from mouse GST5.7 [14] and human GSTA4

Sequence identity with GST5.8 is indicated by a dash (–).

testis and senescent fibroblasts. It is of interest that no clones were detected in adult liver libraries. Further studies using specific anti-GSTA4-4 serum are needed to gain a better understanding of the tissue distribution of this enzyme.

I am grateful to Ms M. Coggan and Ms L. Langton for technical assistance in undertaking this study. Dr. S. Easteal kindly provided assistance with the phylogenetic analysis.

REFERENCES

1 Hayes, J. D. and Pulford, D. J. (1995) CRC Crit. Rev. Biochem. **30**, 445–600
2 Pemble, S. E., Wardle, A. F. and Taylor, J. B. (1996) Biochem. J. **319**, 749–754
3 Board, P. G., Baker, R. T., Chelvanayagam, G. and Jermiin, L. (1997) Biochem. J. **328**, 929–935
4 Pearson, W. R., Vorachek, W. R., Xu, S., Berger, R., Hart, I., Vannais, D. and Patterson, D. (1993) Am. J. Hum. Genet. **53**, 220–233
5 Ross, V. L., Board, P. G. and Webb, G. C. (1993) Genomics **18**, 87–91
6 Cowell, I. G., Dixon, K. H., Pemble, S. E., Ketterer, B. and Taylor, J. B. (1988) Biochem. J. **255**, 79–83
7 Board, P. G., Coggan, M. and Woodcock, D. M. (1992) Genomics **14**, 470–473
8 Tu, C.-P. and Qian, B. (1986) Biochem. Biophys. Res. Commun. **141**, 229–237
9 Board, P. G. and Webb, G. C. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 2377–2381
10 Rhoads, D. M., Zarlengo, R. P. and Tu, C.-P. D. (1987) Biochem. Biophys. Res. Commun. **145**, 474–481
11 Suzuki, T., Johnston, P. N. and Board, P. G. (1993) Genomics **18**, 680–686

Received 10 September 1997; accepted 27 October 1997

12 Röhrdanz, E., Nguyen, T. and Pickett, C. B. (1992) Arch. Biochem. Biophys. **298**, 747–752
13 Stenberg, G., Ridderstrom, M., Engstrom, S., Pemble, S. E. and Mannervik, B. (1992) Biochem. J. **284**, 313–319
14 Zimniak, P., Singhal, S. S., Srivastava, S. K., Awasthi, S., Sharma, R., Hayden, J. B. and Awasthi, Y. C. (1994) J. Biol. Chem. **269**, 992–1000
15 Singhal, S. S., Zimniak, P., Sharma, R., Srivastava, S. K. and Awasthi, Y. C. (1994) Biochim. Biophys. Acta **1204**, 279–286
16 Lemnon, G., Auffray, C., Polymeropoulos, M. and Soares, M. B. (1996) Genomics **33**, 151–152
17 Baker, R. T., Smith, S. A., Marano, R., McKee, J. and Board, P. G. (1994) J. Biol. Chem. **269**, 25381–25386
18 Simmons, P. C. and Vander Jagt, D. L. (1977) Anal. Biochem. **82**, 334–341
19 Board, P. G. and Pierce, K. (1987) Biochem. J. **248**, 937–941
20 Mannervik, B. and Widersten, M. (1995) in Advances in Drug Metabolism in Man (Pacific, G. M. and Fracchia, G. N., eds.), pp. 407–459, Commission of the European Communities, Luxemburg
21 Ricci, G., Caccuri, A. M., Lo Bello, M. A. P., Piemonte, F. and Federici, G. (1994) Anal. Biochem. **218**, 463–465
22 Brophy, P. M., Southan, C. and Barrett, J. (1989) Biochem. J. **262**, 939–946
23 Gillham, B. (1971) Biochem. J. **121**, 667–672
24 Beutler, E. (1975) Red Blood Cell: A Manual of Biochemical Methods, 2nd edn., pp. 71–72, Grune and Stratton, New York
25 Laemmli, U.K. (1970) Nature (London) **227**, 680–685
26 Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4350–4354
27 Smith, S. W., Overbeek, R., Woese, C. R., Gilbert, W. and Gillevet, P. M. (1994) Comput. Appl. Biosci. **10**, 671–675
28 Kimura, M. and Ohta, T. (1969) Genetics **61**, 763–771
29 Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. **4**, 406–425
30 Chang, L-H., Fan, J.-Y., Liu, L.-F., Tsai, S.-P. and Tam, M. F. (1992) Biochem. J. **281**, 545–551
31 Board, P. G. and Mannervik, B. (1991) Biochem. J. **275**, 171–174
32 Stockman, P. K., McLellan, L. I. and Hayes, J. D. (1987) Biochem. J. **244**, 55–61
33 Campbell, E., Takahashi, Y., Abramovitz, M., Peretz, M. and Listowsky, I. (1990) J. Biol. Chem. **265**, 9188–9193
34 Pemble, S. E., Schroeder, K. R., Spencer, S. R., Meyer, D. J., Hollier, E., Bolt, H. M., Ketterer, B. and Taylor, J. B. (1994) Biochem. J. **300**, 271–276
35 Tan, K. L. and Board, P. G. (1996) Biochem. J. **315**, 727–732
36 Whittington, A. T., Webb, G. C., Baker, R. T. and Board, P. G. (1996) Genomics **33**, 105–111
37 Liu, L. F., Wu, S. H. and Tam, M. F. (1993) Biochim. Biophys. Acta **1216**, 332–334
38 Benedetti, A., Comparti, M. and Esterbauer, H. (1980) Biochim. Biophys. Acta **620**, 281–296
39 Esterbauer, H., Cheesman, K. H., Dianzani, M. U., Poli, G. and Slater, T. F. (1982) Biochem. J. **208**, 129–140