**Drosophila** glutathione S-transferase 1-1 shares a region of sequence homology with the maize glutathione S-transferase III

YANN-PYNG S. TOUNG*, TAO-SHIH HSIEH†, AND CHEN-PEI D. TU*‡

*Department of Molecular and Cell Biology, Pennsylvania State University, University Park, PA 16802; and ‡Department of Biochemistry, Duke University School of Medicine, Durham, NC 27709

Communicated by James C. Wang, September 22, 1989 (received for review June 25, 1989)

**ABSTRACT** We have characterized a *Drosophila* glutathione S-transferase (RX:glutathione R-transferase, EC 2.5.1.18) cDNA encoding a protein of 209 amino acids. The cDNA was expressed in *Escherichia coli* harboring the expression plasmid construct pGTDml-KK. The active enzyme, designated as *Drosophila* glutathione S-transferase 1-1, had a specific activity toward 1-chloro-2,4-dinitrobenzene comparable to that for the mammalian glutathione S-transferase III. *Drosophila* glutathione S-transferase 1-1 had no obvious homology to any mammalian or parasitic glutathione S-transferases. The gene was found to be a member of a multigene family.

The glutathione S-transferases (RX:glutathione R-transferases EC 2.5.1.18; GSTs) are a family of multifunctional proteins (for recent reviews, see refs. 1 and 2). High multiplicity of GSTs with overlapping substrate specificities may be essential to their multiple roles in xenobiotic metabolism, drug biotransformation, and protection against peroxidative damage. This isozyme is ubiquitous among eukaryotes. The rat and human GSTs are products of their respective gene superfamilies (3-6). In plants, GSTs are involved in the detoxification of certain herbicides. Different specificities of plant GSTs are thought to be the basis of selective actions of some herbicides (7-10). It has also been proposed that resistance to certain selected pesticides in insects may be related to changes in their GST expression (11, 12). The parasitic helminths of the genus Schistosoma have surface antigens that are glutathione S-transferases. Acquired immunity in mice, rats, hamsters, and monkeys against this antigen from *Schistosoma japonicum* or *Schistosoma mansoni* has provided protection against schistosomiasis, a chronic debilitating disease in several parts of the world (13-15). The major squid lens crystallins may be themselves GSTs or evolutionarily related to GSTs (16, 17). In this communication, we report the molecular characterization of a *Drosophila* GST gene§ and its heterospecific expression in *Escherichia coli*.

**MATERIALS AND METHODS**

**Materials.** Chemicals, S-hexylglutathione (GSH)-linked agarose and antibodies were purchased from Sigma and/or Merck. Radioactive nucleotides were products of Amersham or DuPont/NEN. 32P-labeled protein A was purchased from ICN. Restriction endonucleases and T4 DNA ligase were products from New England Biolabs. A nick-translation kit was purchased from BRL. Kc, cells were grown in spinner flasks to a density of $\approx 5 \times 10^6$ cells per ml, according to a published procedure (18). Twelve- to 16-hour-old *Drosophila* embryos (Oregon R) were collected and washed before use. A *Drosophila* head cDNA library in the Agtll vector was provided by P. Salvaterra of the Beckman Research Institute (Duaire, CA) (19). A genomic library in the AEMBL4 vector was constructed from Oregon R genomic DNA as described (20).

**Purification of GSTs from Kc Cells and *Drosophila* Embryos.** One unit of GST activity catalyzes the formation of 1-chloro-2,4-dinitrobenzene (CDNB)-GSH conjugate at a rate of 1 μmol/min. Kc cells ($\approx 1 \times 10^{12}$) were lysed with Nonidet P-40 and cycles of freezing-and-thawing. After removal of nuclei by centrifugation at 5000 × g (6000 rpm, SS34 rotor, Sorvall centrifuge), the supernatant fractions were brought to 70% saturation by adding solid ammonium sulfate. The precipitated proteins, which contained nearly all the GST activities determined by CDNB conjugation, were recovered by centrifugation (8000 × g for 20 min). The precipitate was dissolved in a minimum volume of 25 mM Tris-HCl (pH 8.0) and dialyzed against the same buffer (1 liter) overnight with one change. Traces of denatured proteins were removed from the dialyzed fraction by centrifugation before affinity chromatography on S-hexyl-GSH (3-ml bed volume, Econo-column from Bio-Rad). Approximately 35% of the CDNB-conjugation activities appeared in the flow-through fractions that did not bind to a second column of S-hexyl-GSH-linked agarose; these fractions were not processed further for the results reported here. The affinity column was washed with 25 mM Tris-HCl (pH 8.0) buffer that contained 0.2 M KCl, and GST activities were eluted according to published procedures (21, 22). The eluted fractions contained two major bands on SDS/PAGE: one band at 23.4 kDa and the other band at 28.5 kDa (data not shown). This enzyme sample was dialyzed against 50 mM 2-(N-morpholino)ethanesulfonic acid (Mes) (pH 6.1) and further purified by fast protein liquid chromatography on a Mono Q column. The elution was carried out with a gradient of 0–0.3 M KCl in Mes buffer (pH 6.1). The first activity peak (peak I), which appeared just before the beginning of the gradient, contained a single band (23.4 kDa) on SDS/PAGE. The rest of the GST activities were eluted between 60 mM and 140 mM KCl in two overlapping activity peaks, peak II and peak III. Fractions of peak III contained four bands between 23.4 kDa and 31 kDa on SDS/PAGE. Peak I and peak III GSTs were used separately to raise polyclonal antibodies in rabbits as described.

Abbreviations: GST, glutathione S-transferase; GSH, glutathione; CDNB, 1-chloro-2,4-dinitrobenzene.

§To whom reprint requests should be addressed at: Department of Molecular and Cell Biology, 6 Allhouse Laboratory, Pennsylvania State University, University Park, PA 16802.

The sequence reported in this paper has been deposited in the GenBank data base (accession no. X14233).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
Fig. 1. Characterization of cDNA clone pGTDml. (A) Nucleotide and deduced amino acid sequence. The ATG initiation codon is boxed, and the termination codon is indicated by an asterisk. The putative poly(A) addition signal is underlined. Arg-14, Pro-54, Asp-58, Ala-68, Ile-69, Gly-144, Asp-151, and Leu-160 are the eight residues identical to those conserved in mammalian GSTs (5, 6). (B) Partial restriction map and sequencing strategy. Only restriction sites pertinent to subcloning into M13mp18/19 phage are shown. Arrows and lines represent the direction and extent of the sequence determinations. The protein coding region is shown by the solid black bar.
Biochemistry: Youn et al.

RESULTS

Characterization of pGTDm1 and pGTDm1-1 Clones. DNAs were isolated from AGTDm1 and AGTDm1-1 and analyzed after EcoRI digestion. Each clone contained an insert DNA of ~750 base pairs (bp) in size. These clones were subcloned into M13mp18 and -mp19 vectors for sequence analysis either directly or after further digestion according to the strategy in Fig. 1. The two cDNAs were identical in sequence except that the pGTDm1-1 insert (data not shown) did not have six adenylate residues at the 3’ end (Fig. 1). The open reading frame of 209 amino acids had a calculated Mr of 23,839. The deduced N-terminal sequence Met-Val-Asp-Phe-Tyr-Tyr-Leu-Pro-Gly-Ser-Ser-Pro-Cys-Arg-Ser-Val... was consistent with the protein sequence determined from peak I GST purified from Kc0 cells or Drosophila embryos (Xaa-Xaa-Asp-(Ser)-Tyr-Leu-Pro-Gly-Ser-Xaa-Pro-(Trp)-Xaa-Ser-Xaa-Ile-Met-Thr-Ala-Xaa). The sequence match at the N-terminal region, the size of the open reading frame, and its positive identification by antibody against purified Drosophila GST made pGTDm1 a probable candidate for a Drosophila GST cDNA clone.

Heterospecific Expression of pGTDm1 cDNA in E. coli. The two constructs in E. coli JM105 in opposing orientations, pGTDm1-KK and pGTDm1-KK’, were induced separately by isopropyl β-D-thiogalactoside (32, 33). The GST activity against CDNB could be detected only in sonicated extracts of cultures of pGTDm1-KK at a level of 0.09 unit per mg of protein. After purification of this activity from large-scale induced cultures (4-liter cultures) by affinity chromatography and fast protein liquid chromatography Mono Q column, we recovered the activity (17% yield) in a fraction containing a single band on SDS/PAGE. Its mobility (23.4 kDa) was identical to that of the purified embryo GST (Fig. 2). Thus, we have proven that the pGTDm1 cDNA insert encodes a Drosophila GST subunit, designated as subunit 1. When the fast protein liquid chromatography-purified GST (both E. coli-expressed and Kc0 cell peak I enzyme) was chromatographed on a Sephadex G-100 column, CDNB-conjugation activities appeared at elution volumes of Mf, 48,500 relative to marker proteins (aprotinin, cytochrome c, carbonic anhydrase, bovine serum albumin). Therefore, GST subunit 1

Table 1. Substrate specificities of E. coli-expressed Drosophila GST 1-1

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Specific activity, μmol/min per mg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Chloro-2,4-dinitrobenzene</td>
<td>24.08 ± 2.23*</td>
</tr>
<tr>
<td>Ethacrynic acid</td>
<td>0.24</td>
</tr>
<tr>
<td>1,2-Dichloro-4-nitrobenzene</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>p-Nitrobenzyl chloride</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>1,2-Epoxy-3-(p-nitrophenoxy)propane</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Cumene hydroperoxide</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

*Average of three determinations.
most probably exists as active dimers, designated as Drosophila GST 1-1. The substrate specificity of GST 1-1 is described in Table 1. CDNB was the most active substrate, but this protein (at 19.6–392 nm enzyme concentration) did not have any detectable GSH peroxidase activity against cumene hydroperoxide relative to a similar concentration of human GST 1-1 (33).

**Genomic Complexity Detected by pGTDm1 cDNA.** Genomic Southern blot analysis with pGTDm1 cDNA insert as a probe revealed multiple bands upon various restriction endonuclease digestions (Fig. 3). The 2.0-kilobase (kb) fragment in λGTDm101 was also strongly positive toward the 3' noncoding probe of pGTDm1 (data not shown). In the EcoRI-digested sample, the 2.8-kb band has not been cloned in the two genomic clones λGTDm101 and λGTDm102. Other restriction digestions revealed multiple bands >4 kb in size (Fig. 3A).

**FIG. 3.** Southern analysis of Drosophila genomic DNA and two Drosophila genomic clones, λGTDm101 and λGTDm102. (A) Drosophila genomic DNA (=2.6 μg each) was digested overnight with 200 units each of BamHI (B) (lane 1), EcoRI (E) (lane 2), HindIII (H) (lane 3), and PsiI (P) (lane 4) and separated by gel (0.8% agarose) electrophoresis. After transfer to a nitrocellulose membrane, the DNA was hybridized with the 3P-labeled EcoRI insert of pGTDm1. (B) EcoRI restriction pattern of two genomic clones, AGTDm101 (~1 μg) (lane 1) and AGTDm102 (~1 μg) (lane 2). (C) Autoradiogram of Southern analysis of AGTDm101 (lane 1) and AGTDm102 (lane 2) revealed by the pGTDm1 cDNA probe. Numbers at left of A and B indicate size markers (λ-HindIII) in kb. Numbers at right of C indicate sizes of positive DNA fragments. Lines between B and C indicate corresponding EcoRI fragments. Hybridizations were done in 5x SSC (1x SSC is 0.15 M sodium chloride, 0.015 M sodium citrate, pH 7)/5x Denhardt's solution (1x Denhardt's solution is 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin)/50 mM sodium pyrophosphate/0.1% NaDdSO4/denatured salmon testis DNA at 100 μg/ml, 40% formamide (vol/vol) at 40°C for 40 hr. The cDNA probe had a specific activity of 2.3 x 10^6 dpm/μg of DNA. The filters were washed at room temperature three times in 2x SSC/0.1% NaDdSO4 for 90 min and then washed in 0.1x SSC/0.1% NaDdSO4 at 55°C for 2 hr.

**FIG. 4.** Partial amino acid sequence comparison between maize GST III and Drosophila GST 1-1. The two sequences (in one-letter code) are aligned for maximal homology. Identical amino acids are represented by boldfaced letters. Those amino acid residues that are in the same group of side-chain type (i.e., small polar, S, G, D, and N; large polar, E, Q, K, and R; intermediate polarity, Y, H, and W; large nonpolar, F, M, L, I, and V; and small nonpolar, C, P, A, and T) are labeled with asterisks.

**DISCUSSION**

We have isolated and characterized a cDNA encoding the Drosophila GST subunit 1 of Mr = 23,839, which contains 209 amino acids. The CDNB conjugation activity (24 μmol/min per mg of protein) of Drosophila GST 1-1 is comparable to that for mammalian GSTs and is approximately twice the reactivity of heterodimeric GST of Drosophila larvae and adults, as characterized by Cochrane et al. (34). The Drosophila GST did not have as broad a substrate specificity pattern as most mammalian GSTs relative to the substrates tested (Table 1). The antibody against GST 1-1 GST does not cross-react with other Drosophila GSTs. It is clear that Drosophila GST is an isozyme family composed of at least three classes of subunits of 23.4 kDa, 28.5 kDa, and 35 kDa (ref. 34 and this study).

The size of Drosophila GST subunit 1 (209 amino acids) makes it most analogous to the rat and human GST P or π subunits (210 amino acids). These proteins, however, do not share much amino acid sequence homology. Indeed, Drosophila GST subunit 1 does not share any extended amino acid sequence homology to the mammalian or parasitic GST sequences in the literature, except for 8 (Arg-14, Pro-54, Asp-58, Ala-68, Ile-69, Gly-144, Asp-151, and Leu-160 in Fig. 1) of the 12 amino acid residues conserved in all GST sequences known to date (2–6). Interestingly, subunit 1 does have significant homology to the maize GST III sequence, as shown in Fig. 4 (35). In a region of 44 amino acids, 22 amino acids were identical between the two GSTs at homologous positions (50% identity). In addition, there are 7 amino acids with similar side-chain groups and 6 more with similar polarity for a maximum of 79% sequence homology. This conservation is striking considering the evolutionary difference between maize and Drosophila melanogaster. It is therefore tempting to propose functional significance for this stretch of 44 amino acids.

The major difference in chromatographic behavior between maize GST III and Drosophila GST 1-1 is that maize GST III can bind to a GSH–agarose affinity column but not to a 5-hexyl-GSH–agarose affinity column (ref. 36; K. P. Timmerman and C.P.D.T., unpublished results). Neither enzyme has GSH peroxidase activity. Therefore, these 44 conserved residues may be important in binding xenobiotic substrates. On the other hand, the difference in chromatographic behavior may reflect a subtle difference between the two GSTs in binding the common substrate GSH. The evolutionary significance of this sequence conservation remains to be elucidated.

The pGTDm1 cDNA has apparently detected a multigene family on genomic Southern blots. We have also isolated two different, but potentially overlapping, genomic clones (Fig. 3). Partial sequence analysis of the genomic clone λGTDm102 has revealed sequences homologous to the N-terminal region of pGTDm1 cDNA sequence (C.P.D.T., N. M. Simkovich, Y-P.S.T., and L. Tu, unpublished results). The lack of immunological cross-reactivity between antibody against peak 1 GST and other potential GST subunits suggests that the 28.5-kDa band, if indeed a GST subunit, may not be closely related to GST subunit 1. Therefore, Drosophila GSTs may be encoded by different gene families, which likely
constitute a gene superfamly, as is the case for mammalian systems. GSH S-transferase(s) has been suggested to impart pesticide resistance against dichlorodiphenyltrichloroethane (DDT) and organophosphorus compounds in houseflies (Musca domestica L.). Molecular analysis of Drosophila GSTs should provide further understanding of a family of important xenobiotic metabolizing enzymes and their relationship to pesticide resistance.

We thank Dr. P. Salvaterra for providing theagt1 cDNA library, Loretta Tu for skillful technical assistance, Dr. N. N. Aronson for critical reading of the manuscript, and Eileen McConnell for typing the manuscript. This work was supported by grants from the U.S. Public Health Services [ES02678, ES00140 (RCDA), and USDA 86-CRCR-1-2148 to C-P.D.T. and GM29906 to T-s.H.]. Part of this work was carried out in 1987 when T-s.H. and C-P.D.T. were in residence at the Institute of Molecular Biology, Academia Sinica in Taipei. The support of Academia Sinica is also gratefully acknowledged.