Multiple human liver glutathione S-transferases (GST, EC 2.5.1.18) with overlapping substrate specificities may be essential to their multiple roles in xenobiotic metabolism, drug biotransformation and protection against peroxidative damage. Multiple GSTs have been purified from human liver and several cDNAs (Tu et al., 1986) were resolved by chromatofocusing (Vander Jagt et al., 1985). They are composed of at least two classes of subunits, H\(_4\) (M, 26,000) and H\(_2\) (M, 27,500), according to their relative electrophoretic mobilities to rat GSTs (Vander Jagt et al., 1985; Tu et al., 1986). We have demonstrated earlier a definitive immunological cross-reactivity between the human liver and various rat GSTs and a nucleotide sequence homology between their respective cDNAs (Tu et al., 1986). Human liver GSTs lack a mobility class equivalent to the rat liver GST\(_{4\alpha}\) (M, 28,000) yet Y\(_1\) cDNA (pGTR262) hybridized to the H\(_4\) subunit cDNA (pGTH1) as strongly as did the Y\(_2\) cDNA (pGTR261) (Lai et al., 1984; Tu et al., 1984; 1986). Furthermore, the H\(_4\) subunit 1 cDNA, pGTH1, selected rat Y\(_1\) and Y\(_2\) subunit mRNAs with almost equal efficiency in hybrid-selected translation in vitro (Tu et al., 1986). In this communication, we provide a molecular basis to these observations by determining the nucleotide sequence of the H\(_4\) subunit 1 cDNA, pGTH1.

The cDNA sequence of pGTH1 was determined by a combination of the chemical method (Maxam & Gilbert, 1980) and the chain-termination method after subcloning into the EcoRI site of the M13 mp18 vector (Sanger et al., 1977; Messing, 1983). The cDNA is 810 nucleotides long, containing 66 nucleotides in the 5' non-coding region, a 222 amino acid open reading frame, and a 222 amino acid non-coding region (Fig. 1). The coding region nucleotide sequence is 80% base-for-base identical to the rat liver Y\(_1\) cDNA (pGTB42) (Telakowski-Hopkins et al., 1984) and much shorter than the corresponding region (101bp) of the Y\(_1\) cDNA pGTR261 (Lai et al., 1984). They share short, dispersed sequence homology, however. For example, nucleotides 756-797 of pGTB42 (Fig. 1) are identical to the pGTR261 sequence (Messing, 1983).

Abbreviations used: GST, glutathione S-transferase; bp, base pairs.

![Fig. 1. Nucleotide sequence of the H\(_4\) subunit 1 cDNA (pGTH1) and its deduced amino acid sequence are compared with the rat Y\(_1\) subunit (pGTB261) and the rat Y\(_2\) subunit (pGTR262). The initiation codons (ATG) are boxed and the poly(A) addition signals (AATAAA) are underlined.](image-url)
Fig. 2. Nucleotide sequence match between pGTH1 cDNA and rat Y, Y, and corn GST cDNA sequences

Regions listed here have more than 50% sequence homologies. Nucleotides are numbered according to Fig. 1 (for pCTH1) and the original publications: A, pCTR187; B, pGTR200; C, pGTR5; D, pMON9000. Non-homologous nucleotides in pGTH1 are represented in lower case letters.

with the DNA blot hybridization signals we observed before (Tu et al., 1986). This conserved region encodes amino acid residues 70-95, which are highly conserved throughout the GST supergene family (Fig. 3). This region of pGTH1 cDNA has almost equal extent of homologies (59%) with the rat placental GST subunit Y, cDNA pGP5 (Fig. 2) (Suguoka et al., 1985). More interestingly, the human GST sequence has detectable homology (>50%) to a corn GST cDNA sequence pMON9000 (Shah et al., 1986) in three regions (Fig. 2, 52%, 55% and 64% respectively), possibly reflecting the common recognition of the essential substrate GSH and/or the hydrophobic nature of some part of the xenobiotic substrate.

The rat GSTs are encoded by a supergene family (Lai & Tu, 1986). The Y, subunit cDNAs (e.g. pGTR187 and pGTR200) did not hybridize to the Y, or Y, subunit cDNAs in Southern hybridizations (Ding et al., 1985; Lai & Tu, 1986; Lai et al., 1986). They do have overlapping substrate specificities against some common substrates such as 1-chloro-2,4-dinitrobenzene, however (Tu & Reddy, 1985).

After careful comparison of their coding sequences, significant conservation of the amino acid residues were identified (~29%) (Lai & Tu, 1986). We found that the H1 subunit 1 sequence has the majority of these conserved amino acids in rat GSTs. The tabulated GST sequences in Fig. 3 showed 27 identical amino acid residues and 29 others in the same side chain groupings (Swanson, 1984; Doolittle, 1985) for a total of 25.7% homology between this H1 subunit and various rat GST subunits.

The amino acid positions 70-95 in the GST supergene family (Fig. 3) may be of unusual significance since they are conserved exceedingly well throughout evolution. It is very likely that they are important residues for GSH binding and/or substrate binding and catalysis. They should provide a reasonable target for mutational studies in vitro.

We gratefully acknowledge the financial support of the United States Public Health Service for this work.

Fig. 3. Amino acid sequence conservation between the human GST H subunit 1 and the rat GST supergene family

Identical amino acids are boxed. Amino acids of the same groups (i.e. small polar, S, G, D, N; large polar, E, Q, K, R; intermediate polarity, Y, H, W; large non-polar, F, M, L, I, V; small non-polar, C, P, A, T) are labelled with asterisks (*). The daggers (†) denote amino acids of the same group in three out of the five subunit classes, Y, Y, Y, Y, and H.


Received 8 January 1987