Construction and characterization of a plasmid containing complementary DNA to mRNA encoding the *N*-terminal amino acid sequence of the rat glutathione transferase Ya subunit

John B. TAYLOR,*† Roger K. CRAIG,* Dennis BEALE‡ and Brian KETTERER*

**Courtauld Institute of Biochemistry, The Middlesex Hospital Medical School, Mortimer Street, London W1P 7PN, U.K., and ‡Agricultural Research Council, Institute of Animal Physiology, Babraham, Cambridgeshire CB2 4AT, U.K.*

(Received 31 October 1983/Accepted 15 December 1983)

Free polyribosomal poly(A)-containing RNA isolated from normal rat liver was used to prepare a complementary DNA plasmid library in the *Pst*I site of the plasmid pAT153. A plasmid pGSTR155 complementary to mRNA coding for a glutathione transferase Ya subunit was selected by differential hybridization *in situ* and preliminary characterization was performed by hybrid-selected mRNA translation, immunoprecipitation and polyacrylamide-gel electrophoresis of the product synthesized *in vitro*. The nucleotide sequence of the complementary DNA contained within pGSTR155 was determined and shown to contain a single open reading frame corresponding to the first 129 amino acids of the *N*-terminus of the Ya subunit and a further 63 nucleotides upstream of the initiating methionine codon.

GSH transferase activity is widespread if not universal in the plant and animal world (for review see Jakoby & Habig, 1980). The most thoroughly investigated activity is that of rat liver, which contains at least seven GSH transferases. They are all dimers, and four subunits designated Ya, Yb¹, Yb² and Yc have been characterized according to their rate of migration on SDS/polyacrylamide gels, substrate specificities and tryptic peptide maps. This has shown that the subunits fall into two related families, namely Ya and Yc on the one hand and Yb¹ and Yb² on the other (Bass *et al.*, 1977; Beale *et al.*, 1982, 1983; Ketterer *et al.*, 1982; Mannervik & Jensson, 1982). As their name implies, the function common to all these rat liver enzymes is the catalysis of the GSH conjugation of electrophiles, most of which arise from the oxidative metabolism of xenobiotics. The multiplicity of enzymes is assumed to result from the need to conjugate numerous types of substances differing in the nature of their electrophilic centres and their molecular structures (see Ketterer *et al.*, 1983b).

Abbreviations used: GSH, reduced glutathione; SDS, sodium dodecyl sulphate; DBM, diazobenzyloxymethyl; cDNA, complementary DNA.

† To whom correspondence and reprint requests should be addressed.

The Ya and Yc subunits possess an additional enzyme activity, namely GSH peroxidase activity towards organic peroxides. Ya also has a high-affinity binding site for the endogenous ligands haem and bilirubin (Ketterer *et al.*, 1983b). Ya is not only the most versatile of subunits, but also the most abundant in the normal liver and together with Yb¹ is inducible by agents such as phenobarbitone which induce the other xenobiotic-metabolizing enzymes. Both the abundance of Ya and its unique combination of functions recommend it for structural studies.

Partial nucleotide sequences of two cloned mRNA species coding for different subunits of GSH transferases of rat liver have been reported. Kalinyak & Taylor (1982) have described a cDNA sequence complementary to a portion of mRNA encoding either a GSH transferase Ya or Yc subunit. The other cloned cDNA sequence (Tu *et al.*, 1982) was also complementary to a mRNA encoding either a Ya or Yc subunit, and contained sequence encoding the C-terminal part of the protein and some of the 3'-untranslated region of the mRNA. Comparison of the two clones showed a common region of sequence of 352 base-pairs in length differing by only nine bases distributed throughout the overlap.

Here we describe the isolation of a cDNA clone

containing novel sequence coding for the first 129 amino acids of the *N*-terminal portion of a GSH transferase Ya subunit. A further sequence of 63 nucleotides upstream of the first in-phase methionine codon has also been determined. A run of 149 nucleotides at the 3'-end of the cloned cDNA sequence is identical with that described by Tu *et al.* (1982).

Materials and methods

Materials

Restriction enzyme *Nco*I was obtained from New England Biolabs through CP Laboratories, Bishop's Stortford, Herts., U.K. Avian myoblastosis virus reverse transcriptase (batch no. G-91180) was the gift of Dr. J. W. Beard, Life Sciences Inc., St. Petersburg, FL 33707, U.S.A. Oligo(dG)₁₂₋₁₈ was purchased from Boehringer, U.K. 'Biodyne A' transfer membranes (1.2 µm) were obtained from Pall Ultrafine Filtration Corp., Glen Cove, NY 11542, U.S.A. A nick-translation kit was obtained from Amersham International, Amersham, Bucks., U.K. All other enzymes, chemicals and solvents were obtained from sources described previously (see Craig *et al.*, 1981; Allison *et al.*, 1981).

Buffers and medium

Buffers used during the isolation of polyribosomes from rat liver are essentially as described by Sala-Trepat *et al.* (1978) and were as follows: buffer SSB(A) which was 25 mM-Tris/HCl, pH 7.5 at 4°C, containing 100 mM-KCl, 40 mM-NaCl, and 7.5 mM-MgCl₂; buffer SSB(B) which was sterile SSB(A) containing 50 mM-NH₄Cl, 0.5 mg of heparin (lithium salt)/ml, 1.5 mg of yeast tRNA/ml and 6 mM-2-mercaptoethanol; and buffer SSB(C) which was 10 mM-Tris/HCl, pH 7.6 at 20°C, containing 100 mM-NaCl, 1 mM-EDTA, 0.5% SDS and 0.5% sarcosyl. SSB(A) and SSB(C) also contained diethylpyrocarbonate (0.05%) and were sterilized by autoclaving at 15 lbf/in² (104 kPa) for 15 min.

Other buffers and media were as described previously by Craig *et al.* (1981). All restriction enzyme incubations were performed using the reaction conditions recommended by the manufacturers.

Isolation of poly(A)-containing RNA from rat liver

Livers were obtained from male Wistar rats weighing 180–200 g inbred at the Courtauld Institute of Biochemistry. Free polyribosomes were isolated using a modification of the centrifugation conditions of Ramsey & Steele (1977). Briefly, the liver [1 g wet wt. of tissue/3 ml of SSB(A)] was minced with scissors on ice, washed twice with ice-cold SSB(A), resuspended in an equal volume of

SSB(B) and homogenized with ten strokes in a Potter homogenizer (0.38 mm clearance). The resulting homogenate was centrifuged at 16000 *g*_{av} for 10 min at 4°C in a MSE 21 rotor. The supernatant was collected, made 3 mg/ml with respect to heparin and centrifuged at 100000 *g*_{av} for 12 min at 5°C in a Beckman SW27 rotor. A superficial layer of fat was removed by adsorption to sterile glass wool, and the remaining supernatant, which contained the free polyribosomes, was decanted and layered over a discontinuous gradient composed of 3 ml of 2 M-sucrose and 2 ml of 1.38 M-sucrose in SSB(B) and centrifuged at 130000 *g*_{av} for 17 h at 4°C in a Beckman 60Ti rotor. The pellet was collected and resuspended in SSB(C) (2.5 ml/10 g wet wt. of tissue), made 150 µg/ml with respect to proteinase K and 2% (w/v) with respect to SDS and then incubated first at 20°C for 2 h, then at 45°C for 40 min. The fractions were extracted with an equal volume of phenol/chloroform (1:1, v/v) and the RNA was recovered by ethanol precipitation of the aqueous layer. Isolation of poly(A)-containing RNA from free polyribosomal RNA was then performed as previously described in Craig *et al.* (1976).

Cell-free protein synthesis and product analysis

Cell-free protein synthesis was carried out in a nuclease-treated reticulocyte lysate as described by Pelham & Jackson (1976) and SDS/polyacrylamide-gel electrophoretic analysis of [³⁵S]methionine-labelled peptides was performed by using slab gels as described previously (Craig *et al.*, 1979). Fluorography followed the procedure of Bonner & Laskey (1974), except that dimethyl sulphoxide was replaced by glacial acetic acid throughout (Burckhardt *et al.*, 1979). Antibody precipitation procedures were as described by Craig *et al.* (1976).

Construction of a cDNA library from poly(A)-containing RNA isolated from normal rat liver

Low specific radioactivity ³H-labelled DNA (720 c.p.m./ng) complementary to poly(A)-containing RNA from free polyribosomes isolated from normal rat liver was synthesized as described previously (Craig *et al.*, 1981). The resulting cDNA was then extended at the 3'-OH termini with dC residues (15–18) essentially as described by Land *et al.* (1981). Single-stranded cDNA (2.1 pmol; average length 500 nucleotides) was dissolved at a concentration of 3.3 µg/ml in 140 mM-cacodylic acid/60 mM-Tris base adjusted with KOH to a final pH of 7.6, containing 2.5 mM-MnCl₂ and 0.1 mM-dithiothreitol, in a final volume of 100 µl, and the mixture was added to 1 µmol of freeze-dried [³H]-dCTP (20 Ci/mmol). The reaction mixture was pre-incubated for 5 min on ice, for 5 min at 30°C, and the reaction was then started by the addition of 7.5

units of terminal deoxynucleotidyl transferase. The reaction was allowed to continue at 30°C for 8 min, then stopped by addition of 100 µl of 10 mM-Tris/HCl and 10 µl of 200 mM-EDTA. This mixture was extracted with an equal volume of phenol/chloroform (1:1, v/v), the aqueous phase retained and the cDNA recovered by ethanol precipitation as described previously (Craig *et al.*, 1981).

The recovered dC-tailed single-stranded cDNA was dissolved in 50 µl of water. Double-stranded cDNA was synthesized as described by Land *et al.* (1981) using oligo(dG)₁₂₋₁₈ hybridized to the 3'-dCMP-tailed cDNA to prime the synthesis of the second strand. Hybridization of oligo(dG)₁₂₋₁₈ was carried out in 74.5 µl of 50 mM-Tris/HCl, pH 7.6, containing 8 mM-MgCl₂, 25 mM-KCl and 40 µg of oligo(dG)/ml over a decreasing temperature gradient. Thus incubations were at 68°C and 55°C for 5 min each, at 50°C for 10 min and 43°C for 15 min, then the mixture was chilled on ice. The reaction volume was then increased to 150 µl by addition of dithiothreitol and all four deoxynucleoside triphosphates to final concentrations of 10 mM and 1 mM respectively. Reverse transcriptase (28 units) was then added, and mixture was incubated for 15 min at 37°C, and then 45 min at 43°C. The incubation was terminated by the addition of 50 µl of 10 mM-Tris/HCl, pH 7.6, containing 1 mM-EDTA, 10 µl of 20% (w/v) SDS, 20 µl of 200 mM-EDTA and 20 µg of *Escherichia coli* tRNA. The mixture was then extracted with phenol/chloroform and the nucleic acid recovered by ethanol precipitation as described above.

The double-stranded cDNA was made flush-ended by using S₁ nuclease and a size-selected population (400–3000 base pairs) was prepared by agarose-gel electrophoresis as previously described (Craig *et al.*, 1981). This was then extended at the 3'-OH termini with dC residues (12–15) using the reaction conditions described by Allison *et al.* (1981).

Recombinant DNA plasmid construction, transformation and selection

Supercoiled pAT153 was restricted with the restriction endonuclease *Pst*I and oligo(dG)₁₂₋₁₄ homopolymer tails were added to the protruding 3'-OH termini by using calf thymus terminal transferase as described by Allison *et al.* (1981). Annealing of poly(dC)-tailed cDNA to poly(dG)-tailed pAT153 and the transformation of *E. coli* RR1 with the chimaeric plasmids were as described elsewhere (Allison *et al.*, 1981).

Identification of plasmids containing GSH transferase cDNA sequences

Primary selection. In total 4000 transformants,

obtained from 10 ng of double-stranded cDNA, were stored on nitrocellulose filters (Hanahan & Meselson, 1980) and duplicate filters were used for colony hybridization according to a modified procedure of Grunstein & Hogness (1975) as described by Craig *et al.* (1981). Colonies were selected which hybridized strongly to ³²P-labelled cDNA synthesized from free polyribosomal poly(A)-containing mRNA isolated from phenobarbitone-treated rat liver, but less strongly to a ³²P-labelled cDNA synthesized from free polyribosomal poly(A)-containing RNA isolated from normal liver.

Secondary selection. Selected colonies (72) were grown in liquid culture (5 ml) and plasmid DNA was prepared (5–10 µg) according to the method of Birnboim & Doly (1979). The plasmid DNA from each selected colony was partially restricted with the restriction endonuclease *Hae*III, heat denatured, then bound to DBM paper discs as described elsewhere (Alwine *et al.*, 1977; Wahl *et al.*, 1979). Filter-bound DNA preparations were then incubated with free polyribosomal poly(A)-containing RNA isolated from phenobarbitone-treated rats and the hybridized poly(A)-containing RNA species complementary to the immobilized cloned DNA were eluted and then recovered by ethanol precipitation, all as described by Craig *et al.* (1981). The hybrid-selected poly(A)-containing RNA species were used to direct protein synthesis in the nuclease-treated rabbit reticulocyte system and the products were analysed by SDS/polyacrylamide-gel electrophoresis and fluorography as described above.

Recombinant plasmid DNA judged to contain cDNA sequence complementary to GSH transferase mRNA by the criteria of hybrid-selected cell-free translation were retransformed into *E. coli* HB101, then grown on a large scale (2 litres), and the plasmid DNA isolated as described elsewhere (Craig *et al.*, 1981). Hybrid-selected cell-free translation was repeated, except that the resulting [³⁵S]methionine-labelled translation products were immunoprecipitated with antibodies raised against total rat liver GSH transferases, prior to electrophoretic analysis of the products.

Northern blotting

Poly(A)-containing RNA was analysed under denaturing conditions by electrophoresis through 1.1% (w/v) agarose gels containing 2.2 M-formaldehyde essentially as described by Maniatis *et al.* (1982). The RNA samples were heated at 60°C for 2 min, then cooled to room temperature before loading onto the gel. Samples were electrophoresed at 150 V for approx. 2 h.

The RNA was directly transferred from the agarose gel to 'Biodyne A' transfer membrane by using the conditions described by Thomas (1980).

After transfer overnight, the membrane was baked at 80°C for 1 h, then prehybridized at 42°C for 4 h in 10 ml of hybridization buffer composed of 50 mM-Tris/HCl, pH 7.6, containing 1 M-NaCl, 0.1% Na₂P₂O₇, 10% (w/v) dextran sulphate, 100 µg of single-stranded salmon sperm DNA/ml (200–300 base-pair length), ×10 Denhardt's reagent (Denhardt, 1966) and 50% (v/v) formamide. The membrane was hybridized overnight at 42°C in 10 ml of hybridization buffer containing 20 ng of [³²P]dCTP-labelled nick-translated total plasmid DNA/ml (specific radioactivity 1–2 × 10⁸ c.p.m./µg) prepared by using an Amersham reagent kit and then was washed at 52°C for 90 min in 500 ml of 5 mM-Tris/HCl, pH 7.5, containing 25 mM-NaCl, 0.1% SDS and 1 mM-EDTA with three changes of buffer. Finally the membrane was blotted dry on Whatman 3MM paper and autoradiographed at –70°C with Kodak 'Lanex' intensifier screens.

Nucleotide sequence analysis

The nucleotide sequence of cDNA inserted in pGSTr155 was determined in its entirety on both strands by using the chemical DNA sequencing procedure of Maxam & Gilbert (1980). Fragments were labelled at 3'-OH termini by using T4 DNA polymerase and at 5'-phosphate termini using T4 polynucleotide kinase as described by Hall *et al.* (1982).

Tryptic peptide analysis

Tryptic peptides from 3 mg of GSH transferase B homodimer (YaYa) were prepared, separated on Whatman 3MM paper, and identified, according to methods described by Beale & Smith (1976). Spots were cut out and the peptides were eluted with 6 M-HCl, then hydrolysed *in vacuo* at 105°C for 18 h. The amino acid composition of the peptides was determined with a Locarte amino acid analyser (Locarte Co. Instrument Makers, London W12, U.K.).

Results and discussion

Characterization of a poly(A)-containing RNA population from rat liver and establishment of the presence of mRNA encoding GSH transferase subunits

Free polyribosomal poly(A)-containing RNA isolated from livers taken from either normal or phenobarbitone-treated rats was translated in the rabbit reticulocyte lysate system, and the resulting [³⁵S]methionine-labelled proteins were analysed by SDS/polyacrylamide-gel electrophoresis and fluorography. This demonstrated that both RNA preparations directed the synthesis of a large

spectrum of proteins over and above the background (Fig. 1a, track 1), but that significant quantitative differences were present when the cell-free translation products directed by free polyribosomal liver mRNA isolated from a normal animal were compared with a similar preparation from a phenobarbitone-treated animal (compare Fig. 1a, tracks 2 and 3). In particular, the induced preparation directed the synthesis of two predominant proteins of estimated *M_r* 25000 and 26500. The synthesis of these proteins was also directed by mRNA isolated from normal rat liver, but their relative proportion when compared with other proteins synthesized by the same preparation appeared somewhat lower. Both these proteins, and an additional minor component of *M_r* 28000, were selectively precipitated by antibodies raised against a total GSH transferase preparation (Fig. 1a, track 4). Further experiments using an antiserum raised specifically against the Yb¹ and Yb² subunits but not the Ya or Yc subunits resulted in the selective immunoprecipitation of the *M_r* 26500 component (Fig. 1b, track 1), thereby identifying this as a GSH transferase Yb subunit synthesized *in vitro*. Since the immunoprecipitation experiments described here have been performed using cell-free products directed by free polyribosomal poly(A)-containing RNA isolated from phenobarbitone-treated animals, and Ya and Yb¹ are the only GSH transferase subunits induced by phenobarbitone feeding (Ketterer *et al.*, 1983a), the polypeptides of *M_r* 25000 and 26500 must represent the subunits of Ya and Yb¹ respectively, and phenobarbitone feeding gives rise to the induction of their respective mRNAs. The minor immunoprecipitated polypeptide of *M_r* 28000 would appear to be representative of the Yc subunit (Fig. 1a, track 4; Fig. 1b, track 2) since this is known to be immunoprecipitated by the antiserum against the total GSH transferases, and moreover to be present at reduced level in phenobarbitone-induced liver when compared with Ya and Yb subunits (Bass *et al.*, 1977).

Having established the presence of mRNA directing the synthesis of the Ya and Yb transferase subunits in rat liver mRNA preparations, we have constructed a cDNA library using free polyribosomal poly(A)-containing RNA isolated from normal rat liver, with the intention to isolate and characterize recombinant plasmids containing cDNA representative of each GSH transferase subunit mRNA.

Construction of recombinant cDNA plasmids using free polyribosomal poly(A)-containing RNA and identification of a plasmid containing DNA complementary to a GSH transferase subunit mRNA

Double-stranded cDNA was synthesized as des-

Fig. 1. Identification of a recombinant plasmid containing GSH transferase Ya subunit cDNA sequence
Poly(A)-containing RNA preparations were added to a nuclease-treated reticulocyte lysate, and the [35 S]methionine-labelled cell-free translation products were separated by SDS/polyacrylamide-gel electrophoresis, then visualized by fluorography. Arrows denote the Ya, Yb and Yc subunits synthesized *in vitro*, and also shown are the relative mobilities of the following [14 C]methylated standard marker proteins of known M_r , obtained from Amersham International: lysozyme (M_r , 14 300), carbonic anhydrase (M_r , 30 000), ovalbumin (M_r , 46 000), bovine serum albumin (M_r , 69 000). (a) Lane 1, reticulocyte lysate background with no added mRNA. Lanes 2 and 3, [35 S]methionine-labelled peptides directed by free polyribosomal poly(A)-containing RNA from normal and phenobarbitone-induced rat liver respectively. Lane 4, the antibody precipitate of products visualized in lane 3. The antibody was raised against total rat liver GSH transferases. (b) Antibody precipitates of total [35 S]methionine-labelled peptides (see lane 3, in a). The antibodies were raised against GSH transferase A (recognizing Yb¹ and Yb²) (lane 1) and total rat liver GSH transferases (lane 2). (c) Lane 1, [35 S]methionine-labelled peptides directed by free polyribosomal poly(A)-containing RNA isolated from phenobarbitone-induced rat liver. Lanes 2 and 3, [35 S]methionine-labelled peptides directed by sequence-specific free polyribosomal poly(A)-containing RNA selected by hybridization to partially restricted denatured pAT153 DNA and pGStr155 DNA respectively, immobilized on DBM-paper. Lanes 4, 5 and 6, the antibody precipitates of cell-free products visualized in lanes 1, 2 and 3 respectively. The antibody was raised against total rat liver GSH transferases.

cribed in the Materials and methods section and used for the construction, using homopolymer tailing, of chimaeric plasmid DNA in the *Pst*I site of pAT153. This DNA was then used to transform *E. coli* RR1. Clones containing DNA complementary to liver mRNA enriched by phenobarbitone induction were selected by differential hybridization using 32 P-labelled cDNA synthesized from free polyribosomal poly(A)-containing RNA isolated from normal and induced rat livers.

From this primary screening 72 colonies which appeared to hybridize 32 P-labelled cDNA synthesized from the induced poly(A)-containing RNA more strongly were selected for small-scale rapid plasmid preparations. These preparations were then used for secondary screening by hybrid-selected translation. From this secondary screening, 12 plasmids were chosen which selected RNA yielding cell-free translation products of mobility corresponding to M_r 25 000–30 000 (J. B. Taylor, unpublished results). These 12 plasmids were re-

transformed into *E. coli* HB101 because yields of plasmid DNA from large-scale plasmid preparations were greater than when the plasmids were in *E. coli* RR1.

Plasmid DNA from large scale preparations was then used to repeat the hybrid-selected translation experiments, and, in this instance, the identity of the cell-free translation product was determined by antibody precipitation and SDS/polyacrylamide-gel electrophoresis. The results (Fig. 1c) show that one of the twelve recombinants examined (pGStr155) hybridized an mRNA species which directed the synthesis of a protein of the expected electrophoretic mobility of a Ya subunit (compare tracks 2 and 3). Moreover, this peptide was immunoprecipitated by antisera raised against total GSH transferases (see Fig. 1a, track 4). Thus, whereas this antiserum selectively precipitates three GSH transferase subunits whose synthesis was directed by liver mRNA isolated from the phenobarbitone-treated animal (Fig. 1c, tracks 1

Fig. 2. Observed restriction nuclease sites and the sequencing strategy of cDNA contained within pGSTr155, and also the nucleotide sequence of pGSTr155 cDNA insert and the corresponding amino acid sequence

The open rectangular box represents the cDNA insert, the solid line represents pAT153 sequences. ●, 3'-labelled; ○, 5'-labelled; □, G-C tails. Either total plasmid or a *Hha*I fragment (820 base pairs) were used as sequencing substrates. The deduced amino acid sequence was assigned as described in the text. Underlined is an in-phase stop codon upstream from the first subsequent in-phase methionine codon. The latter has been called 1 in the amino acid sequence and corresponds to nucleotides at position 64–66 in the cDNA sequence.

and 4), immobilized pGSTr155 selectively hybridizes mRNA which directs the synthesis of the smaller (M_r 25000) or Ya subunit (Fig. 1c, tracks 3 and 6). Controls using immobilized pAT153 DNA (Fig. 1c, tracks 2 and 5) resulted in no specific hybridization of mRNA species which directed antibody precipitable material.

Sequence analysis of a recombinant plasmid containing DNA which is complementary to the 5'-portion of a mRNA encoding the Ya subunit of GSH transferase

On the basis of hybrid-selected cell-free translation and selective use of well-characterized antisera (Ketterer *et al.*, 1976), we have, in the absence

of amino acid sequence data for any of the GSH transferase subunits, identified the cDNA cloned within pGSTr155 to represent part of the mRNA encoding the Ya subunit. In order to establish part at least of the primary amino acid sequence of the Ya subunit we have determined the complete nucleotide sequence of the cDNA contained in pGSTr155. This was determined by the method of Maxam & Gilbert (1980), using the strategy outlined in Fig. 2. Sequence was determined on both strands, and all restriction sites used were confirmed by overlapping sequence ladders labelled at separate sites. The final sequence (see Fig. 2) shows that we have cloned 452 base pairs (exclud-

Fig. 3. Size estimation of rat liver GSH transferase B mRNA

Free polyribosomal poly(A)-containing RNA (1 μ g) isolated from normal rat liver was analysed under denaturing conditions in 1.1% (w/v) agarose gels. The RNA was transferred to 'Biodyne A' transfer membrane, hybridized with 32 P-labelled pGSTr155 and autoradiographed. The 32 P-labelled DNA hybridized to two diffuse species of RNA of approx. 980 and 1150 nucleotides. The thickness of the bands suggests some heterogeneity within the population. DNA fragments of known size were the products of a *Hinf*I digest of pAT153 (1631, 517, 396 and 298 base pairs), treated in the same way as the RNA (see the Materials and methods section).

ing G-C tails), representing approx. 50% of the mature mRNA as judged by Northern blotting analysis (see Fig. 3). The latter revealed that the cloned cDNA sequence hybridized strongly to a diffuse and therefore probably heterogeneous mRNA species of 980 ± 75 nucleotides present in free polyribosomal poly(A)-containing RNA from normal liver, whilst additional weak and diffuse hybridization was observed to other larger poly(A)-containing RNA species of 1150 ± 100 nucleotides in length. This larger species is possibly the Yc mRNA, since significant homology at the amino acid level between Ya and Yc subunits has already been described (Beale *et al.*, 1982).

Examination of the nucleotide sequence of the

cDNA cloned into pGSTr155 revealed the presence of a single open reading frame encoding 129 amino acids starting 64 nucleotides in from one end of the cloned cDNA with a methionine (AUG), followed by a second in-phase AUG triplet 45 nucleotides further downstream. It is interesting to note that the presence of a 10-nucleotide inverted repeat sequence (residues 44–53, 79–88; Fig. 2) at the 5'-end of the deduced mRNA sequence. The significance of this hairpin structure, which has a free energy (ΔG) of -65.5 kJ when calculated according to the values of Tinoco *et al.* (1973), is unknown. If such a structure is formed *in vivo* then the first in-phase AUG codon lies in the non-base-paired part of the hairpin loop and the second lies downstream to the stem of the loop, suggesting a possible role in the initiation of protein synthesis.

The 149 nucleotides at the 3'-end of the cDNA contained in pGSTr155 (303–452) correspond exactly with those at the 5'-end of the GSH transferase cDNA sequence described by Tu *et al.* (1982), whilst 251 nucleotides at the 3'-end of pGSTr155 (202–452) correspond to the GSH transferase sequence described by Kalinyak & Taylor (1982), with the exception of the seven differences also observed and discussed in detail by Tu *et al.* (1982). Taking all the above observations into account, and bearing in mind the cloning strategy used, which enriches for 5'-regions of mRNA species (see Land *et al.*, 1981), we would conclude that we have cloned the 5'-portion of the mRNA encoding the Ya subunit of GSH transferase B.

Establishment of the N-terminal amino acid sequence of the Ya subunit of GSH transferase

It is not possible to assign definitively the initiating methionine codon from nucleotide sequence data alone, since examples of initiation of translation at the second as opposed to the first in-phase AUG proximal to the 5'-end of mammalian mRNA species have been reported (see Hall *et al.* 1982). Past structural studies of GSH transferase B (YaYa) have not identified the N-terminal amino acid because it is blocked in an unknown manner (see Beale *et al.*, 1982). Consequently, in order to determine whether the predicted N-terminal amino acid sequence occurs in the native subunit, we have examined the amino acid composition of tryptic peptides of the GSH transferase B homodimer (YaYa) which were characterized as described previously (Ketterer *et al.*, 1976, and see the Materials and methods section). One such peptide, which gave colour reactions for tyrosine, histidine and arginine but only a very weak reaction with ninhydrin, was considered to be a candidate for the N-terminal peptide on the supposition that the first in-phase methionine codon is used *in vivo* (see Fig.

2). Further analysis of this peptide showed the amino acid composition to be (Gly, Lys, Pro, Val, Leu, His, Tyr, Phe, Asn, Ala, Arg) (see the Materials and methods section). This would correspond to the amino acid composition of the predicted sequence 1–13 with the exception that methionine and serine were absent (note: a Lys-Pro peptide bond is resistant to trypsin digestion). This result indicates that initiation *in vivo* occurs at the first in-phase AUG, because a peptide of this amino acid composition does not occur elsewhere in the deduced sequences (the present work; Tu *et al.*, 1982; Kalinyak & Taylor, 1982). This conclusion assumes that the initiating methionine is cleaved by aminopeptidase (Jackson & Hunter, 1970) and that the serine residue represents the blocked *N*-terminal amino acid, which we presume is not recovered after acid hydrolysis. In addition two further peptides were identified, namely (Gly, Arg; representing residues 14 and 15) and (Met, Glu, Cys, Ile, Arg; representing residues 16–20). These conform with an *N*-terminal sequence which includes the methionine coded for by the second in-phase AUG codon and so point to the use of the first in-phase AUG codon to initiate translation *in vivo*.

In conclusion we have described the construction of a rat liver cDNA library and the selection and characterization of a recombinant plasmid containing a cDNA sequence encoding the amino acid sequence of the *N*-terminal portion of a GSH transferase Ya subunit. The availability of this cloned sequence should enable us to screen libraries for recombinants containing full-length cDNA representative of Ya subunits, which may then be examined by nucleotide sequence analysis in order to investigate reported heterogeneity of the Ya subunit (Beale *et al.*, 1982) and the apparent heterogeneity at the mRNA level (see Fig. 3). It is likely that such studies will also result in the selection of recombinants containing Yc cDNA sequences, since sequence similarities have been reported at the amino acid level (Beale *et al.*, 1982), whilst homology may exist at the nucleotide level (see Tu *et al.*, 1982). The determination of the primary amino acid sequence of the Ya and Yc subunits, and also the Yb subunits of rat GSH transferases, should form the basis of subsequent enzyme structure and function analysis with respect to tissue and substrate specificity (see Mannervik *et al.*, 1983; Meyer *et al.*, 1983; Tu *et al.*, 1983). The availability of cDNA hybridization probes will provide insight into the organization of the genes, and into the mechanisms involved in the induction of this gene family.

This work was supported by the Cancer Research Campaign. We also thank Dr. L. Hall and Dr. M. S.

Davies for their advice and assistance during these experiments, and Professor P. N. Campbell for his enthusiastic support and encouragement.

References

- Allison, J., Hall, L., MacIntyre, I. & Craig, R. K. (1981) *Biochem. J.* **199**, 725–731
- Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5350–5354
- Bass, N. M., Kirsch, R. E., Tuff, S. A., Marks, I. & Saunders, S. J. (1977) *Biochim. Biophys. Acta* **492**, 163–175
- Beale, D. & Smith, I. (1976) in *Chromatographic and Electrophoretic Techniques* 2, 4th edn. (Smith, I., ed.), pp. 31–45, Heinemann, London
- Beale, D., Ketterer, B., Carne, T., Meyer, D. & Taylor, J. B. (1982) *Eur. J. Biochem.* **126**, 459–463
- Beale, D., Meyer, D. J., Taylor, J. B. & Ketterer, B. (1983) *Eur. J. Biochem.* **137**, 125–129
- Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523
- Bonner, W. M. & Laskey, R. A. (1974) *Eur. J. Biochem.* **46**, 83–88
- Burckhardt, J., Telford, J. & Birnstiel, M. L. (1979) *Nucleic Acids Res.* **6**, 2963–2971
- Craig, R. K., Brown, P. A., Harrison, O. S., McIlreavy, D. & Campbell, P. N. (1976) *Biochem. J.* **160**, 57–74
- Craig, R. K., Boulton, A. P., Harrison, O. S., Parker, D. & Campbell, P. N. (1979) *Biochem. J.* **181**, 737–756
- Craig, R. K., Hall, L., Parker, D. & Campbell, P. N. (1981) *Biochem. J.* **194**, 989–998
- Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun.* **23**, 641–646
- Grünstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961–3965
- Hall, L., Craig, R. K., Edbrooke, M. R. & Campbell, P. N. (1982) *Nucleic Acids Res.* **10**, 3503–3515
- Hanahan, D. & Meselson, M. (1980) *Gene* **10**, 63–67
- Jakoby, W. B. & Habig, W. H. (1980) *Enzymatic Basis of Detoxification*, Academic Press, New York
- Jackson, R. & Hunter, T. (1970) *Nature (London)* **227**, 672–676
- Kalinyak, J. E. & Taylor, J. M. (1982) *J. Biol. Chem.* **257**, 523–530
- Ketterer, B., Tipping, E., Beale, D. & Meuwissen, J. A. T. P. (1976) in *Glutathione: Metabolism and Function* (Arias, I. M. & Jakoby, W. B., eds.), pp. 243–257, Raven Press, New York
- Ketterer, B., Beale, D. & Meyer, D. (1982) *Biochem. Soc. Trans.* **10**, 82–84
- Ketterer, B., Beale, D., Taylor, J. B. & Meyer, D. J. (1983a) *Biochem. Soc. Trans.* **11**, 466–467
- Ketterer, B., Coles, B. & Meyer, D. J. (1983b) *Environ. Health Perspect.* **49**, 59–69
- Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schütz, G. (1981) *Nucleic Acids Res.* **9**, 2251–2266
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning—A Laboratory Manual*, Cold Spring Harbor Laboratory, New York
- Mannervik, B. & Jensson, H. (1982) *J. Biol. Chem.* **257**, 9909–9912

- Mannervik, B., Guthenberg, C., Jensson, H., Warholm, M. & Alin, A. (1983) in *Extrahepatic Drug Metabolism and Chemical Carcinogenesis* (Rydström, J., Montelius, J. & Bengtsson, M., eds.), pp. 153–162, Elsevier Science Publishers, Amsterdam
- Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560
- Meyer, D. J., Christodoulides, L. G., Nyan, O., Schuster-Bruce, R. & Ketterer, B. (1983) in *Extrahepatic Drug Metabolism and Chemical Carcinogenesis* (Rydström, J., Montelius, J. & Bengtsson, M., eds.), pp. 189–190, Elsevier Science Publishers, Amsterdam
- Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67**, 247–256
- Ramsey, J. C. & Steele, W. J. (1977) *Biochem. J.* **168**, 1–8
- Sala-Trepat, J. M., Savage, M. J. & Bonner, J. (1978) *Biochim. Biophys. Acta* **519**, 173–193
- Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5201–5205
- Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41
- Tu, C.-P. D., Weiss, M. J., Karakawa, W. & Reddy, C. C. (1982) *Nucleic Acids Res.* **10**, 5407–5419
- Tu, C.-P. D., Weiss, M. J., Li, N.-Q. & Reddy, C. C. (1983) *J. Biol. Chem.* **258**, 4659–4662
- Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3683–3687