Characterization of the rat glutathione S-transferase Yc_2 subunit gene, *GSTA5*: identification of a putative antioxidant-responsive element in the 5′-flanking region of rat *GSTA5* that may mediate chemoprotection against aflatoxin B_1

David J. PULFORD and John D. HAYES

Biomedical Research Centre, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, Scotland, U.K.

We have isolated and characterized genomic DNA encoding the rat glutathione S-transferase Yc_2 subunit. This protein is now referred to as rGSTA5 and is noteworthy because of its high activity towards aflatoxin B_1-8,9-epoxide, its marked inducibility by chemoprotectors, its sex-specific regulation, and its over-expression in hepatoma and preneoplastic nodules. The *rGSTA5* gene, which was isolated on two overlapping bacteriophage λ clones, is approx. 12 kb in length and, unlike other class Alpha genes described to date, it comprises six exons. The transcription start site has been identified 228 bp upstream from the ATG translational initiation codon, and is situated 51 bp downstream from a consensus TATA-box. Deletion analysis, using luciferase reporter constructs, has shown that the region between -177 bp and $+65$ bp from the transcriptional start site contains a functional promoter. Computer-assisted analysis of the upstream sequence has indicated the presence of an antioxidant-responsive element (ARE), and several elements thought to be required for tissue-specific expression of the enzyme. In addition, several putative oestrogen-responsive half sites were observed in both upstream and intronic sequences.

INTRODUCTION

The glutathione S-transferase (GST) isoenzymes are a supergene family of detoxication proteins which are universally distributed in nature (for a review, see [1]). GST enzymes possess a broad specificity and catalyse the conjugation of GSH with a wide spectrum of lipophilic compounds which contain an electrophilic centre. Their substrates include carcinogens, pesticides, therapeutic drugs and products of oxidative stress. A remarkably large number of transferases have been described and at least five classes of soluble GST have been identified; these are designated Alpha, Mu, Pi, Sigma and Theta [1,2]. In most mammalian species, each GST class is typically represented by between three and eight genes [1,3].

Evidence suggests that the level of GST expression can represent an important factor in determining the sensitivity of organisms to noxious chemicals [1]. The laboratory of Wattenberg [4], as well as that of Talalay [5], has demonstrated that many of the chemicals which protect against chemical carcinogenesis induce GST activity. Such chemoprotectors include coumarins, dithiolethiones, indoles, isothiocyanates, organosulphides and phenolic antioxidants [6]. As these chemicals inhibit the initiation stage of chemical carcinogenesis they have been referred to as blocking agents [7]. While many studies have reported a correlation between GST induction and failure of carcinogens to produce tumours in various target tissues, few attempts have been made to identify the GST genes responsible for the resistant phenotype. Most of the studies into the mechanisms of chemoprotection have employed 1-chloro-2,4-dinitrobenzene (CDNB) as a model substrate to monitor GST levels and, consequently, relatively little is known about the regulation of specific GST subunits by anti-carcinogenic agents.

In an attempt to identify which GSTs are important in chemoprotection against the potent hepatocarcinogen aflatoxin B_1 (AFB_1), the isoenzymes which metabolize AFB_1-8,9-epoxide have been purified from rats fed on diets containing the antioxidant chemoprotector ethoxyquin [8]. This study resulted in the isolation of two inducible heterodimeric class Alpha GSTs which are essentially absent from the livers of normal adult male rats fed on control diets. Although these two GSTs contribute little to the total hepatic GST activity towards CDNB they were found to have at least 50-fold greater transferase activity towards AFB_1-8,9-epoxide than the constitutively expressed enzymes. Both of the rat enzymes were found to contain the Yc_2 subunit (now called rGSTA5 [1]) and it was therefore postulated that this polypeptide is responsible for the high activity of GST Ya_1Yc_2 and Yc_1Yc_2 (now called rGSTA1–5 and rGSTA3–5, respectively) towards epoxidated AFB_1. Molecular cloning and expression of recombinant rGSTA5-5 confirmed that the A5 subunit has substantial activity towards AFB_1-8,9-epoxide [9]. Indeed, it has been estimated that the rGSTA5 subunit exhibits about 180-fold greater activity towards AFB_1-8,9-epoxide than rGSTA3 and 1000-fold greater activity than rGSTA1 and rGSTA2 (previously called Ya_1 and Ya_2, respectively).

Since the identification of GSTA5 as the major AFB_1-metabolizing transferase subunit in ethoxyquin-treated rat liver [8], interest has focused on the regulation of this polypeptide. It has been found that GSTA5 is inducible by many chemoprotector blocking agents including butylated hydroxyanisole (BHA), butylated hydroxytoluene, ethoxyquin, indole-3-carbinol, 1,2-dithiole-3-thione, oltipraz and phenobarbital [9–13]. In addition to induction by chemoprotectors, GSTA5 is developmentally regulated and is subject to a sexual dimorphism [9,10]. Furthermore, GSTA5 is expressed in preneoplastic nodules and hepatomas caused by exposure to AFB_1 [8,14].

As a first step towards gaining an understanding of the molecular mechanisms involved in regulating the GSTA5 subunit, genomic clones have been isolated that encompass *GSTA5*.

Abbreviations used: AFB_1, aflatoxin B_1; ARE, antioxidant-responsive element; BHA, butylated hydroxyanisole; CDNB, 1-chloro-2,4-dinitrobenzene; DTT, dithiothreitol; ERE, oestrogen-responsive 'half site' element; ethoxyquin, 6-ethoxy-1,2-dihydro-2,2,4-trimethylquinoline; GST, glutathione S-transferase; TPA, tetradecanoylphorbol-13-acetate ('PMA'); TRE, TPA-responsive element; XRE, xenobiotic-responsive element.

One of the clones which has been obtained contains about 5 kb of 5′-flanking sequence, and within this region *cis*-acting enhancers which are likely to be responsible for the induction of GSTA5 by chemoprotectors have been identified.

MATERIALS AND METHODS

Chemicals

All chemicals were obtained from Sigma Chemical Co. Ltd. (Poole, Dorset, U.K.). Radioactive isotopes; [α-^{32}P]dCTP (specific radioactivity, 3000 Ci/mMol), [γ-^{32}P]ATP (specific radioactivity, 3000 Ci/mmol) and [γ-^{33}P]ATP (specific radioactivity, 1000–3000 Ci/mmol) were obtained from Amersham International (Amersham, Bucks., U.K.). The 20 × SSC stock solution comprises 3 M NaCl, 0.3 M sodium citrate.

Plasmids, bacterial strains, and general molecular biology techniques

Escherichia coli strain XL1-Blue MRA [Δ(*mcrA*)*183*, Δ(*mcrCB-hsdSMR-mrr*)*173*, *endA1*, *supE44*, *thi-1*, *gyrA96*, *relA1*, *lac*] was used during the screening of a rat genomic λ DASH library. Restriction endonucleases were obtained from Gibco BRL (Paisley, Scotland, U. K.). Recombinant plasmids constructed during the present study are described in Table 1. Unless otherwise stated, plasmids were transformed into *E. coli* strain NM522 {*supE thi-1* Δ(*lac-proAB*) Δ (*mcrB-hsdSM*) *5* (r_k^- m_k^-) [*F′ proAB lac* I^q*Z*Δ*M*15]} supplied by Stratagene (Cambridge, U.K.). All bacterial transformations and alkaline lysis mini-preparations of plasmid DNA were carried out using standard techniques [15]. Qiagen column procedures for large-scale preparation of double-stranded template DNA were followed according to the manufacturer's instructions (Qiagen Ltd., Dorking, U.K.). Single-stranded DNA template was prepared according to the protocol of Stratagene for isolation of bacteriophage DNA.

For Southern analysis, DNA fragments, obtained by restriction enzyme digestion, were separated on 1 % agarose gels in TAE buffer (40 mM Tris/CH_3OOH, 1 mM EDTA) and transferred to Hybond-N nylon membrane (Amersham International).

Screening of a rat genomic library

A Fisher rat genomic library in the λ DASH vector was obtained from Stratagene. Library screening was carried out using an overnight culture of XL1-Blue MRA grown at 30 °C in 50 ml of LB-liquid containing 0.2 % maltose and 10 mM $MgSO_4$. Cells were harvested by centrifugation at 2600 ***g*** and resuspended at a D_{600} of 0.5 in ice-cold 10 mM $MgSO_4$. Portions (600 μl) of resuspended cells were each infected with 50000 plaque-forming units (pfu) of the λ library and incubated at 37 °C for 15 min prior to the addition of 6.5 ml of top-agarose (heated to 48 °C) containing 10 mM $MgSO_4$. These mixtures of infected *E. coli* and top-agarose were plated immediately on to prewarmed 150-mm-diam. LB–agarose plates and incubated at 37 °C overnight. Following a 2 h incubation at 4 °C, the plaque-containing plates were overlayed with a Hybond-N filter for 1 min. Duplicate filters were prepared in the same manner but were left in contact with the plates for 2 min. Both sets of filters were treated for 8 min at 25 °C in a solution of 1.5 M NaCl/0.5 M NaOH to denature bacterial and phage proteins, and were then placed in a solution of 1.5 M NaCl, 0.5 M Tris/HCl, pH 8.0, for 5 min to neutralize alkaline conditions. After a brief wash in a solution of 0.2 M Tris/HCl, pH 7.5, 2 × SSC (i.e. 0.3 M NaCl, 30 mM sodium citrate), filters were dried and the adsorbed DNA was UV cross-linked in a UV Stratalinker 2400 (Stratagene). Plates were stored at 4 °C.

Filters were prehybridized at 42 °C in a solution of 6 × SSC, 5 × Denhardt's, 0.1 % SDS containing 500 μg of denatured salmon sperm DNA for at least 2 h. The *Eco*RI fragment from λJH36 (EMBL Data Bank accession number X78847) that contains the entire coding sequence for rat GSTA5 was isolated, labelled with [α-^{32}P]dCTP, using the random-primed DNA labelling kit (Boehringer Mannheim, Lewes, East Sussex, U.K.) and used to screen filters. Hybridization conditions were the same as those used for prehybridization except that filters were left overnight at 42 °C. Filters were washed in a solution of 0.1 × SSC/0.1 % SDS at between 55 °C and 65 °C before being exposed to Kodak X-OMAT film. Following autoradiography, putative positive clones were identified, isolated from the original plates and subjected to further rounds of screening until plaque purity was achieved. Bacteriophage λ DNA was prepared by the method of Sambrook et al. [15].

Table 1 Plasmids used and generated in this study

Plasmid	Description	Source
pBluescript	General purpose vector for isolation of double- or single-strand DNA templates.	Stratagene.
pYc36	*GSTA5* cDNA isolated from λJH36 and inserted into *Eco*RI site of pBluescript.	[9]
pDP1	4 kb fragment containing exons 5 and 6 from λDP131 inserted into the *Xba*I site of pBluescript. Used to generate DNA sequence data and for restriction mapping.	This work.
pDP9	2 kb fragment containing exon 6 from pDP1 inserted into the *Eco*RI-*Bam*HI site of pBluescript. Used for the isolation of sense strand template DNA.	This work.
pDP10	2 kb fragment containing exon 6 from pDP1 inserted into the *Eco*RI-*Bam*HI site of pBluescript. Used for the isolation of antisense strand template DNA.	This work.
pDP6	9 kb *Eco*RI fragment containing exons 1, 2 and 3 and 2 kb of 5′ untranslated sequence from λDP644. Subcloned into pBluescript and used to generate DNA sequence data and for restriction mapping.	This work.
pDP812	2 kb 5′ untranslated upstream *Kpn*I fragment from pDP6 subcloned into pBluescript. Used for the isolation of sense strand template DNA.	This work.
pDP84	2 kb 5′ untranslated upstream *Kpn*I fragment from pDP6 subcloned into pBluescript. Used for the isolation of antisense strand template DNA.	This work.
pGL3-Basic	Vector lacking eukaryotic promoter and enhancer sequences but containing luciferase structural gene. Thus expression of luciferase activity depends on insertion of functional promoter.	Promega.
pSV-β-gal	Vector containing the β-galactosidase structural gene under control of the constitutively active SV 40 promoter.	Promega.
pDP-1902A5*Luc*	*Kpn*I fragment from pDP812 inserted into pGL3-Basic. Used to assess promoter activity in luciferase assays.	This work.
pDPΔTATA	Construct containing the region −177 to −1902 of *rGSTA5* inserted into pGL3-Basic.	This work.

Table 2 Oligonucleotide primers used for sequencing 5′ untranslated region of the rat *GSTA5* gene and for primer extension experiments

Primer	Oligonucleotide sequence
DP1	GTTCTAACAGTCTGTCCTAC
DP3	GGAGGATGGTAACGTGTGGG
DP7	CATGGCCACATGAACTGCCTA
DP8	TAGGCAGTTCATGTGGCCATG
DP11	CAAACCCAGAGGACACGGCTG
DP12	GGAGCCCATCCGGTGGCTCCTG
DP20	CAAGGGAAAGAAGTGTTGCTG
DP21	CAGCAACACTTCTTTCCCTTG
DP22	GCCATGCCGGGGAAGCCAGTC
DP23	GACATTTATCAGAAGAGAGTAGG
DP24	GTGAATTTACCTACACACTTTGC
DP25	AACCATGCTGGCCTCCCCTC
DP29	GCATGTTAGCAGCCTAGC
DP30	GCTAGGCTGCTAACATGC
DP31	CAGCCGTGTCCTCTGGGTTTG
DP32	CCTGCTGCCGTGGCCATCGC
DP33	GGTGGCTCCTGATAGGTACC
DP34	GGCCTCTCCTAACAGAACC
DP35	CACAGTCGGGACAAGTAGC
DP37	CAGAATGCAGTAGCCTGAAC
DP38	GACATAAGAGAGGATACTGG
DP39	CGCCCACGCTCTTGTTATC
DP40	GCCTGCTCTGACTCTGACAATCC
DP41	CCTACTCTCTTCTGATAAATG
DP524	CCATCACCAGATAAATTGAGG
DP45PE	GGTACCTATCAGGAGCCACCCTCAGGTTGC

Table 3 Oligonucleotide primers used for sequencing exons and for PCR

Primer	Oligonucleotide sequence
T3	ATTAACCCTCACTAAAG
T7	AATACGACTCACTATAG
JH11	CATGGTGGAGATTGACGGGA
JH13	TTTGGATACTCCAGCGTCTA
JH19	ATGAGGAATTGCCTGGGATGC
DP12	GGAGCCCATCCGGTGGCTCCTG
DP15	TCCCGTCAATCTCCACCATG
DP16	GCATGTAGGGGTAATAGAGAAC
DP22	GCCATGCCGGGGAAGCCAGTC
DPex4	GGAACCGTTACTTCCCTGCC
DPex5	CGAGATAATCTTGTCCGTGGC
KS	CGAGGTCGACGGTATCG

DNA sequence determination

This information was obtained using either double-stranded or single-stranded DNA template; both the Sequenase version 2.0 method of Amersham International and the *fmol* DNA sequencing system of Promega (Southampton, U.K.) were used during the project. The oligonucleotide primers used for sequencing DNA and those employed in PCR and Southern hybridization experiments are shown in Tables 2 and 3.

Computer-aided DNA sequence analysis

In order to identify putative enhancer elements in the rat *GSTA5* gene and its flanking sequences, the Apple MacIntosh MacPattern software programme using the 'Tfd: the transcription factors database' [16] was employed.

Determination of transcription start site

Isolation of total cellular RNA

The determination of the transcription start site, by S1 nuclease protection studies and by primer extension experiments, was carried out using total cellular RNA which was isolated from rat Morris Hepatoma 7800 C1 cells [17] using TRIzol Reagent supplied by Gibco BRL.

Analysis of transcription initiation site by S1 nuclease mapping

The radioactive S1 nuclease protection probe was generated by mixing 10 μg of single-stranded template DNA obtained from pDP812 with 5 pmol of DP45PE in 10 mM Tris/HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$. Primer and template DNA were denatured through incubation at 65 °C for 5 min, and were allowed to anneal by cooling to 35 °C over 30 min. Reaction volumes were increased to 40 μl to allow random hexamer priming to be carried out in 10 mM Tris/HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, containing 0.625 mM each of dATP, dGTP and dTTP, and 1.25 mM dithiothreitol (DTT) (buffer A) and 50 pmol of [α-^{32}P]dCTP. Reactions were started by addition of 2 units (1 μl) of Klenow (Boehringer Mannheim) and were allowed to proceed at 37 °C for 15 min before being terminated by heating to 65 °C for 15 min. The reaction products (in a final volume of 50 μl of buffer A) were digested with *Stu*I (2 h, 37 °C) and purified by electrophoresis on a 6 % polyacrylamide/8 M urea sequencing gel; the S1 nuclease protection probe was located by autoradiography and cut out of the gel.

The S1 nuclease protection probe ($\sim$ 12000 c.p.m.) was allowed to hybridize with a 30 μg aliquot of total cellular RNA from Morris Hepatoma cells in 10 mM Tris/HCl buffer, pH 7.5, containing 300 mM NaCl and 1 mM EDTA, by being placed at 70 °C overnight. Reactions were chilled on ice for 5 min, and 9 vol. of S1 nuclease assay buffer (5 % glycerol, 1 mM $ZnSO_4$, 30 mM sodium acetate, pH 4.5, 50 mM NaCl) containing 50 units of S1 nuclease was added. Reactions were incubated at either 25 °C, 30 °C or 37 °C for 30 min and stopped by addition of 40 μl of 5 M ammonium acetate and 5 μl of 0.5 M EDTA. Following addition of SDS to a final concentration of 0.3 % (v/v), double-stranded DNA–RNA hybrids were precipitated by addition of ethanol. The pellets, which were recovered by centrifugation, were washed in 70 % ethanol and resuspended in 4 μl of formamide dye loading buffer. S1 nuclease-protected products were analysed on 6 % polyacrylamide/8 M urea sequencing gels. DNA sequencing ladders prepared using DP45PE were run alongside primer extension reaction products on the sequencing gel.

Analysis of the transcription initiation site by primer extension

Primer extension experiments were carried out with 10 pmol of end-labelled DP45PE (Table 2) annealed to approx. 20 μg of total cellular RNA. Annealing reactions (10 μl total volume) were preincubated at 65 °C for 5 min prior to addition of 20 μl of reverse transcription reaction 'master mix' that achieved a final reaction mixture (30 μl) comprising 62.5 mM Tris/HCl, pH 8.3, 11.25 mM DTT, 11.25 mM $MgCl_2$, 40 units of RNasin and 5 units of avian myeloblastosis virus (AMV) reverse transcriptase. Reverse transcription reactions were incubated at 52 °C for 30 min before addition of 1 μl of 0.5 M EDTA to terminate the reaction. RNA was hydrolysed by incubating reactions at 65 °C for 30 min after the addition of 2 μl of 6 M NaOH. The RNA-depleted samples were neutralized by addition of 2 μl of 6 M acetic acid. The products of primer extension reactions were purified using the GENO-BIND protocol adapted from the 5′-

AmpliFINDER RACE Kit (Clontech). The primer extension DNA products were ethanol-precipitated, washed in 80 % (v/v) ethanol and resuspended in 4 μl of sterile distilled water. After addition of 3 μl of formamide gel loading dye, primer-extended products were separated on a 7 % polyacrylamide denaturing sequencing gel. DNA sequencing ladders, prepared using the same primers as those used for the primer extension reactions, were run alongside primer extension reaction products on the sequencing gel.

Generation of *in vitro* transcribed RNA

In vitro transcribed RNA was prepared with 1 μg of pDP812 as template in a buffer containing 40 mM Tris/HCl, pH 8.0, 25 mM NaCl, 8 mM $MgCl_2$, 2 mM spermidine, 1 mM NTP, 5 mM DTT and 40 units of RNasin. Reactions were started by addition of 80 units of T3 RNA polymerase and were allowed to proceed at 37 °C for 90 min prior to phenol/chloroform extraction and ethanol precipitation. RNA was washed with 70 % (v/v) ethanol, resuspended in 60 μl of diethylpyrocarbonate-treated water, and stored at −70 °C until required.

Determination of intron sizes using PCR

PCR methods were carried out using oligonucleotide primers described in Table 3 on a Hybaid Omnigene PCR machine. Annealing temperatures (usually T_m −2 to −4 °C) were adjusted for the individual oligonucleotides used. PCRs were usually continued for 30 cycles following an initial denaturation step at 94 °C for 2 min.

Functional analysis of the 5′-flanking region of *rGSTA5*

Generation of pLuc constructs

The *Kpn*I fragment from pDP812, that contains 1902 bp of genomic sequence upstream from the transcription start site and 65 bp of downstream sequence, was subcloned into pGL3-Basic (Promega), using standard methods [15] to generate the plasmid pDP-1902A5*Luc* (Table 1). Orientation of the inserted fragment in pDP-1902A5*Luc* was confirmed by restriction digestion with *Bam*HI. The pDPΔTATA plasmid (Table 1) was constructed by digesting pDP-1902A5*Luc* with *Stu*I and *Sma*I. Following gel purification, the 6.6 kb fragment was re-ligated, to generate a plasmid containing the region 177 bp to 1902 bp upstream from the transcription start site of *GSTA5*.

Cell culture, transient transfection, and reporter gene assays

The human hepatoma cell line, HepG2, was maintained in continuous culture in Dulbecco's modified Eagle's medium with Glutamax-I and 1000 mg/l glucose (Gibco BRL) supplemented with 20 % (v/v) heat-inactivated fetal bovine serum, 100 units/ml penicillin, and 10 μg/ml streptomycin at 37 °C in a humidified atmosphere with 5 % CO_2. Cells were passaged every 5 days with a 1:3 dilution.

All reporter gene plasmids were purified prior to transfection by centrifugation through a caesium chloride gradient containing ethidium bromide. HepG2 cells were seeded at 3×10^5 cells per 5-cm-diam. dish 24 h prior to transfection. One hour before transfection, the medium was replaced with fresh growth medium. Cells were transfected using the calcium phosphate method. Briefly, DNA-$CaCl_2$ solutions were prepared by mixing 25 μl of 2.5 M $CaCl_2$, 22 μl of TE (10 mM Tris/HCl/1 mM EDTA) buffer, pH 8.0, with 4 μg of either pDP-1902A5*Luc*, pDPΔTATA (Table 1) or luciferase control plasmid (pGL3-Basic, Promega), plus 4 μg of pSV-β-gal plasmid (Promega, control for transfection efficiency) in a final volume of 250 μl. DNA was precipitated by mixing with an equal volume of 2 × Hepes-buffered saline (50 mM Hepes, pH 7.1, 280 mM NaCl, 1.5 mM Na_2HPO_4) with constant agitation. The precipitants were incubated at room temperature for 30–40 min before addition to the cell culture medium. After 24 h, cell cultures were washed with 3 × growth medium and left for a further 24 h in fresh growth medium prior to harvesting using lysis buffer (Promega) for luciferase and β-galactosidase enzyme assays. Reporter gene assays for luciferase or β-galactosidase activity were carried out using the luciferase and β-galactosidase assay kits (Promega) according to the manufacturer's instructions.

RESULTS

Isolation of the rat Yc_2 subunit gene, *GSTA5*

Six λ phage genomic clones that contain sequences encoding the rat GSTA5 (Yc_2) subunit have been isolated by screening a total of approx. 6×10^6 plaques. These clones were designated λDP131, λDP231, λDP232, λDP233, λDP234, and λDP644. Initial characterization by Southern blot analysis indicated that λDP131, λDP231, λDP232, λDP233 and λDP234 were identical and contained approx. 12 kb of genomic sequence. However, λDP644 was shown to be distinct from the others and was found to contain an insert of 14.5 kb. Rather than characterize all six clones, it was decided that λDP131 and λDP644 should be studied in greater detail. Southern blotting demonstrated that the clones contain overlapping sequences and together they encompass approx. 26 kb of genomic DNA. When Southern blots were probed with DP1 (oligonucleotide representing nucleotides −46 to −66 of the cDNA, Table 2) or with JH19 (oligonucleotide representing nucleotides +815 to +835 of the cDNA, Table 3), results indicated that λDP644 contains the 5′ region of *GSTA5* while λDP131 contains the 3′ region (results not shown).

Restriction analyses were carried out on λDP131 and λDP644 and the data obtained were confirmed using subcloned restriction fragments (Figure 1). The λDP644 clone was found to contain approx. 5 kb of the 5′-untranslated region. PCR and sequence analysis showed that while λDP131 contains 12 kb of genomic DNA, only about 2.5 kb of the 5′ region represents sequence encoding the *GSTA5* gene. Thus, although the combined extent of genomic DNA contained within both λDP131 and λDP644 is approx. 26 kb, a significant amount of this DNA represents untranslated 5′- and 3′-flanking sequence.

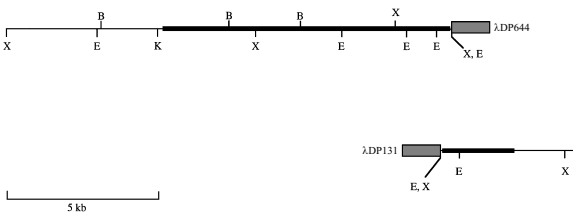

Figure 1 Restriction maps of λDP131 and λDP644

Restriction maps with *Bam*HI (B), *Eco*RI (E), *Kpn*I (K), and *Xba*I (X) for the two λ clones containing the rat *GSTA5* gene. Shaded boxes indicate λ DNA while solid lines indicate the extent of genomic DNA that contains coding sequence for the *GSTA5* subunit. As only the 4 kb *Xba*I fragment has been subcloned from λDP131, the restriction map is incomplete 3′ of this restriction site.

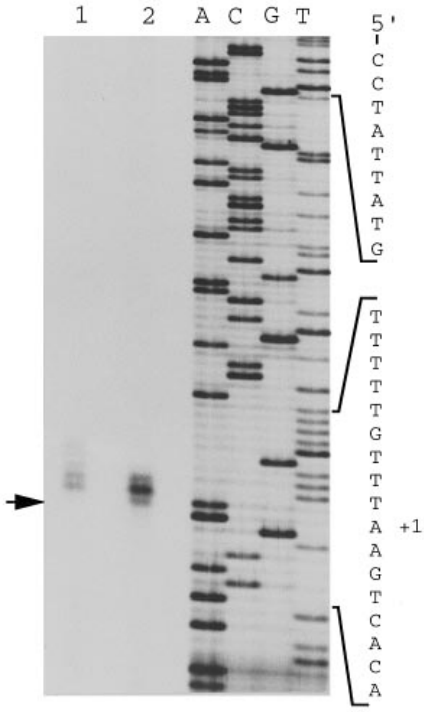

Figure 2 Determination of transcription start site of rat *GSTA5* gene

S1 nuclease protection assays were carried out as described in the Materials and methods section. The lanes were loaded as follows: lane 1, the product obtained by carrying out S1 nuclease digestion at 25 °C; lane 2, the product obtained by carrying out S1 nuclease digestion at 30 °C; the lanes designated A, C, G and T contain sequence ladders obtained by sequencing pDP812 using primer DP45PE. Partial 5′-flanking sequence corresponding to the sequencing ladders of the *GSTA5* gene is shown. The arrow in the left-hand margin indicates the transcription start site, which is labelled as +1 on the corresponding sequencing ladder.

Determination of the transcription start site

Using total RNA isolated from the rat Morris Hepatoma cell line, that has been previously shown to express GSTA5 (R. McLeod and J. D. Hayes, unpublished work), the transcription start site of the *GSTA5* gene was identified by a combination of S1 nuclease protection and primer extension studies. In repeated S1 nuclease protection assays carried out at 25 °C, 30 °C and 37 °C, the transcription start site was found to occur 228 bp upstream from the ATG translation initiation codon (Figure 2). This nucleotide is centred within an A-T-rich region of the DNA and, as the A-T-enriched double-stranded DNA in this region may melt to form S1 nuclease-sensitive single strands, additional experiments were undertaken to establish the authenticity of the S1 nuclease protection data. Confirmation that the region centred 228 bp upstream from the ATG initiation codon is not itself sensitive to S1 nuclease digestion was obtained by repeating the experiment using *in vitro* transcribed RNA, produced from pDP812, that extended upstream from the putative transcription start site by 2 kb. In these experiments the full-length S1 probe was protected from nuclease digestion. Results obtained from primer extension studies using the oligonucleotide DP45PE (Table 2) were in agreement with those obtained for S1 nuclease protection studies. Thus, by primer extension and S1 nuclease protection studies, the transcription start site was identified as being located 228 bp upstream from the ATG initiation codon and is indicated as +1 in Figure 2.

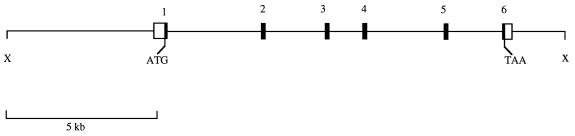

Figure 3 Structural organization of rat *GSTA5* subunit gene

Open and closed boxes represent non-coding and protein coding regions of *GSTA5* mRNA, respectively. Introns separating each of the six exons are indicated.

Rat *GSTA5* gene structure

The structure of the *GSTA5* gene has been determined by restriction mapping, PCR and sequence analysis. A diagrammatic representation of the gene structure is presented in Figure 3. The sizes of the exons are shown in Table 4. Southern blot analysis, using radiolabelled exon-specific oligonucleotides, has also been used to confirm the relative positions of the exons. Figure 4 shows the DNA sequence of each of the six exons of *GSTA5*, starting at the transcription start site and extending beyond the polyadenylation sequence.

Identification of potential minimal promoter in the 5′-flanking region of *GSTA5*

The expression of the GSTA5 subunit has been shown to be induced by exposure of animals to phenobarbital, phenolic antioxidants, indole-3-carbinol and oltipraz [8–13]. In order to establish whether regulation of *GSTA5* expression by xenobiotic compounds is likely to involve similar *cis*-acting elements to those described previously [1], the genomic DNA adjacent to the transcription start site was characterized. Restriction digestion of λDP644 DNA with *Eco*RI yielded a 9 kb fragment that was subcloned into pBluescript to generate pDP6 (Table 1). Using the primers described in Table 2, DNA sequence data were obtained for the 2 kb immediately 5′ of the ATG translation initiation codon (Figure 5). Two putative TATA-box sequences are present 215 bp and 279 bp upstream from the ATG initiation codon. However, as primer extension and S1 nuclease protection experiments indicated that transcription starts 228 bp upstream from the ATG codon (Figure 2), it seems likely that the TATA-box centred 279 bp upstream from the ATG initiation codon is functional. Although transcription is initiated 51 bp downstream from the TATA-box, sequence analysis demonstrated that no CAAT-box is present in this region. A consensus GC-box has been identified in the 5′-flanking region of *GSTA5*, but is located 1661 bp upstream of the transcription start site. It is noteworthy that a 116 bp polypurine sequence is present between −984 bp and −1100 bp with respect to the transcription start site (Figure 5). Within this 116 bp sequence, which contains only three pyrimidine nucleotides, over 60% of the nucleotides are guanine with the remainder being almost exclusively adenine. Guanine bases occur in groups of between two and five that are separated by one or two adenine nucleotides. Also observed within this region are a number of $(dA\text{-}dG)_n$ repeat sequences.

To examine the function of the 5′-flanking region of *GSTA5*, the region spanning −1902 to +65 bp with respect to the transcription start site was isolated and subcloned into pGL3-Basic thus generating pDP-1902A5*Luc* (Table 1). When this reporter plasmid was transfected into HepG2 cells, luciferase activity was 24- to 30-fold higher than observed with pGL3-Basic

Table 4 A comparison of exon and intron sizes from different class Alpha GSTs

Data for *rGSTA2*, *mGstA1*, *hGSTA1* and *hGSTA2* are taken from refs. [27–30] respectively.

Rat GSTA5 (Yc_2) subunit gene		Other mammalian class Alpha GST genes				
Intron	*rGSTA5* (kb)	Intron	*rGSTA2* (kb)	*mGstA1* (kb)	*hGSTA2* (kb)	*hGSTA1* (kb)
		1	2.35	2.00	5.60	4.50
1	4.00	2	3.50	2.50	1.60	1.60
2	2.50	3	0.65	0.80	1.20	1.30
3	1.10	4	2.10	3.70	2.00	2.00
4	2.50	5	1.50	1.90	1.20	1.10
5	2.50	6	0.80	0.90	0.85	1.00
Exon	*rGSTA5* (bp)	Exon	*rGSTA2* (bp)	*mGstA1* (bp)	*hGSTA2* (bp)	*hGSTA1* (bp)
		1	43	42	36	36
1	315	2	109	109	117	117
2	51	3	52	52	52	52
3	134	4	133	133	133	133
4	142	5	142	142	142	143
5	132	6	132	132	132	132
6	500	7	234	234	317	123

(Table 5). Furthermore, when HepG2 cells were transfected with pDPΔTATA (Table 1), a construct lacking the region spanning −177 to +65 bp that includes the consensus TATA-box, luciferase activity was similar to that seen in cell extracts prepared from HepG2 cells transfected with pGL3-Basic (Table 5). Thus, the region spanning −1902 bp to +65 bp of *GSTA5* can support transcription while the sequence between −177 bp and +65 bp contains the elements necessary for basal expression.

Identification of putative enhancers in the 5′-flanking region of *GSTA5*

Computer-aided sequence examination enabled the identification of five further putative enhancer elements within the promoter region in the 5′-flanking region of the *GSTA5* gene (Figure 5). Two of these, located at −392 bp and −497 bp from the transcription start site, are homologous to the consensus binding sequence for liver-specific HNF-5, 5′-TGTTTGC-3′, and may account for the tissue-specific expression of GSTA5 in fetal and adult female rat livers. A third putative enhancer element of particular interest in view of the involvement of phenolic antioxidants in the regulation of GSTA5 expression, is a consensus antioxidant-responsive element (ARE) located at −421 bp (Figure 5); the ARE is present in reverse orientation and is identical to the previously described 5′-TGACNNNGC-3′ consensus that has been shown to be functional in the regulation of transcription of the rat *GSTA2* gene [18,19]. A fourth potential element, 5′-TCACAAAGGAGAAGG-3′, with similarity to the 3′ region of the ACTAAAAGCTGGAGG Barbie box [20], can be identified in *GSTA5* between −1773 bp and −1759 bp, that might be responsible for the induction of this gene by phenobarbital. A fifth element that may be involved in the sexual dimorphic regulation of GSTA5 expression in adult rats is the oestrogen-responsive half site element (ERE), 5′-GGTCA-3′, located between the ARE and the hepatic nuclear factor 5 (HNF-5) element between −485 bp and −481 bp.

The rat GSTA5 subunit has been shown to be induced by structurally diverse xenobiotics, some of which (e.g. 3-methylcholanthrene) might be expected to work through the xenobiotic-responsive element (XRE) [21]. Indeed, an XRE consensus site, 5′-CACGC-3′ [22,23], has been identified in the rat *GSTA2* gene [24]. Sequence analysis and computer-aided examination of the *GSTA5* gene have indicated the presence of the sequence 5′-CACGCACGCACACGCACGCA-3′ within the first intron that contains multiple XRE motifs (Figure 6). This XRE-containing structure is flanked on either side by two ERE motifs and a sequence, 5′-TGAGTCA-3′, which represents a 12-*O*-tetradecanoylphorbol-13-acetate (TPA)-responsive element (TRE) consensus, 5′-T(G/T)A(G/C)TCA-3′ [25,26] (Figure 6).

DISCUSSION

In the present study the entire gene which encodes the rat GSTA5 subunit has been isolated. Although class Alpha GST genes encoding Ya-type subunits (i.e. GSTA1 and GSTA2) have been characterized, the present paper represents the first description of a class Alpha GST gene encoding a Yc-type subunit (i.e. GSTA3 and GSTA5). The rat *GSTA5* gene is approximately 12 kb in size and is thus of similar size to other class Alpha GSTs described to date [24,27–32]. However, unlike other class Alpha GST genes described previously, all of which comprise seven exons, rat *GSTA5* is composed of only six exons. In the *GSTA5* gene the transcription start site and translation initiation codon are both contained within the first exon, while in other class Alpha GST genes the transcription start site and translation initiation codon are represented in exon one and exon two, respectively. By comparison with other class Alpha GST genes which have been characterized hitherto, rat *GSTA5* would appear to lack the intron equivalent to that found in the 5′-flanking non-coding region of other class Alpha GST genes; in rat *GSTA2* and mouse *GstA1* intron 1 is found 22 bp upstream from the ATG initiation codon [27,28] while in human *GSTA1* and *GSTA2* intron 1 is found 30 bp upstream from the initiation codon [29–32]. It is possible that the presence or absence of an intron in the 5′ non-coding region of class Alpha GST genes represents a characteristic difference between those encoding Ya-type (GSTA1 and A2) and Yc-type (GSTA3 and A5) subunits. This difference in the structure of the non-coding region of different

EXON 1

+1 AAGTCACATATTAACCGATGGATACACTAAACTGGTTTCCTGCAACCTGAGGGTGGCTCCTGATAGGTACCAATTTGGA
CCATGGAACAGAGTCCAGGAATGTTTCCGACCCTGCCCTAAAGAAGGCAGACACTTCTTTAGCAGCCGTTGTCCAGACC
CCCTCGTAGGACAGACTGTTAGAACAGGCTGTGCTTCATCTCTGTTTAGAGAACTCAAGCAATTGCTGCC ATG CCG
Met Pro

GGG AAG CCA GTC CTT CAC TAC TTC GAT GGC AGG GGG AGA ATG GAG CCC ATC CGG TGG CTC
Gly Lys Pro Val Leu His Tyr Phe Asp Gly Arg Gly Arg Met Glu Pro Ile Arg Trp Leu

CTG GCT GCA GCT GGA GTA GAG 315
Leu Ala Ala Ala Gly Val Glu 29

EXON 2

316 TTT GAA GAA AAT TTT CTG AAA ACT CGG GAT GAC CTG GCC AGG TTA AGA AGT 366
Phe Glu Glu Asn Phe Leu Lys Thr Arg Asp Asp Leu Ala Arg Leu Arg Ser 46

EXON 3

367 GAT GGG AGT TTG ATG TTT GAA CAA GTG CCC ATG GTG GAG ATT GAC GGG ATG AAG CTG GTG
Asp Gly Ser Leu Met Phe Glu Gln Val Pro Met Val Glu Ile Asp Gly Met Lys Leu Val

CAG ACC AAA GCC ATT CTC AAC TAC ATT GCC ACC AAA TAC AAC CTC TAT GGG AAG GAC ATG
Gln Thr Lys Ala Ile Leu Asn Tyr Ile Ala Thr Lys Tyr Asn Leu Tyr Gly Lys Asp Met

AAG GAG AGA GCC CT 500
Lys Glu Arg Ala Le 91

EXON 4

501 C ATC GAC ATG TAT GCA GAA GGT GTG GCC GAT CTG GAG TTG ATG GTT CTC TAT TAC CCC TAC
u Ile Asp Met Tyr Ala Glu Gly Val Ala Asp Leu Glu Leu Met Val Leu Tyr Tyr Pro Tyr

ATG CCC CCT GGG GAG AAA GAG GCG AGT CTT GCC AAG ATC AAG GAC AAA GCA AGG AAC CGT
Met Pro Pro Gly Glu Lys Glu Ala Ser Leu Ala Lys Ile Lys Asp Lys Ala Arg Asn Arg

TAC TTC CCT GCC TAT GAG AAG 642
Tyr Phe Pro Ala Tyr Glu Lys 138

EXON 5

643 GTG TTG AAG AGC CAC GGA CAA GAT TAT CTC GTT GGC AAC AAG CTG AGC AGG GCT GAT GTT
Val Leu Lys Ser His Gly Gln Asp Tyr Leu Val Gly Asn Lys Leu Ser Arg Ala Asp Val

TCC CTG GTT GAA CTT CTC TAC CAT GTG GAA GAG ATG GAC CCA GGC ATT GTG GAC AAC TTC
Ser Leu Val Glu Leu Leu Tyr His Val Glu Glu Met Asp Pro Gly Ile Val Asp Asn Phe

CCT CTG CTA AAG 774
Pro Leu Leu Lys 182

EXON 6

775 GCC CTG AGA ACC AGA GTC AGC AAC CTC CCC ACA GTG AAG AAG TTT CTT CAG CCT GGC AGC
Ala Leu Arg Thr Arg Val Ser Asn Leu Pro Thr Val Lys Lys Phe Leu Gln Pro Gly Ser

CAG AGG AAG CCT TTT GAT GAT GAG AAA TGT GTA GAA TCA GCG AAG AAG ATC TTC AGT TAA
*Gln Arg Lys Pro Phe Asp Asp Glu Lys Cys Val Glu Ser Ala Lys Lys Ile Phe Ser *** 221*

TTCAGTCAGCTATGGATACACTGTACCCACAAAGCCAGCCTCAGAAAGCTCTGCAACAATGAAGTATTTTGACTAAATG
TTGACCGTACTTATTGGGAGGGTAACATGTTTTCTAAGGCTTCTGTGTTAATTCATATAGACATGACTCATGAGGAATT
GCTGGGATGCCATCTAGTTGAGTTAAAACCTCAATCTCGATCACTTCCTCGGATATTTTCTTAATGTTCAATAAAACAA
AACAAGCTTCTTAGACGCTGGAGTATCCAAACATTGTCATGAAATAGCTGTCATATCCTTGTCAAACAGCGTCACGTAG
AAACCCTCGTGTCAAACTCTCTTACGCAAAAGTAATCTTTCCTTATGGAGAGTGTCCTTTCTCT 1274

Figure 4 Exon sequences of the rat *GSTA5* gene

DNA sequence information representing each of the six exons and the deduced amino acid sequence of the *GSTA5* gene is presented. Position +1 is the transcriptional start nucleotide while other numbers indicate the relative position within the mRNA of each exon splice site. Italicized numbers indicate the relative position of amino acids inclusive of the initiation methionine; this residue is not included in the mature protein.

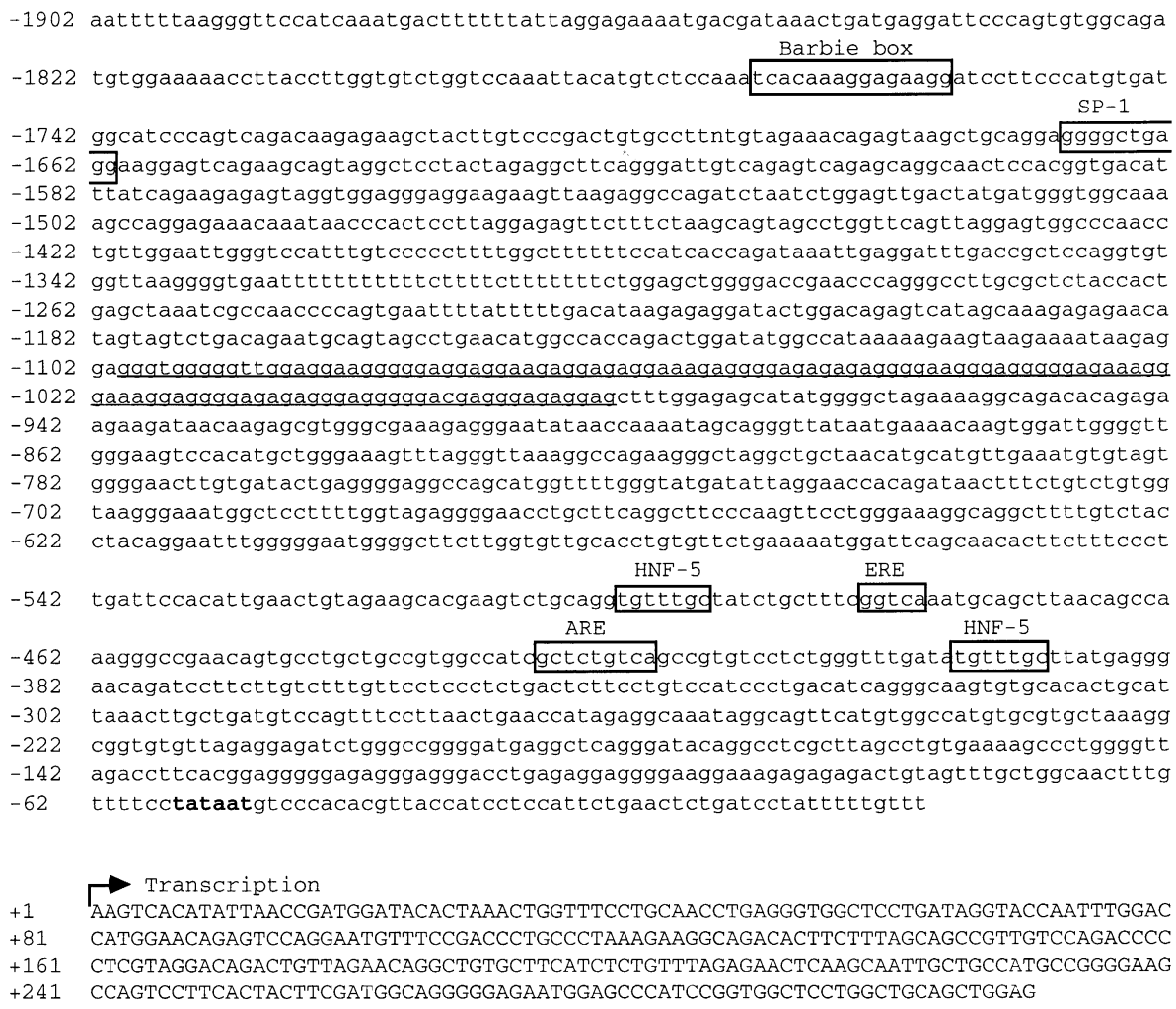

Figure 5 The 5′ upstream sequence of the rat *GSTA5*, indicating transcription start site and putative enhancers

The nucleotide sequence of the 2 kb immediately upstream of the transcription start site of the *GSTA5* gene was determined by the dideoxynucleotide chain termination method. A putative Barbie box and consensus binding sequences for HNF-5, ERE and SP-1 are shown. In addition, a consensus ARE sequence that is in reverse orientation is also indicated. The polypurine G-A tract is represented by the underlined sequence. Some nucleotide sequence 3′ of the transcription start site is also presented.

Table 5 Relative luciferase activity in HepG2 cells transiently transfected with p*Luc* constructs

HepG2 cells were transfected with the indicated plasmids as described in the Materials and methods section. After 24 h cells were harvested and protein extracts were prepared. Cell protein extracts were assayed for luciferase activity, and β-galactosidase activity was measured to normalize for transfection efficiency. Results shown represent the mean of four separate experiments $\pm$ S.E.M.

Construct	Relative luciferase activity
pGL3-Basic	1.75 $\pm$ 0.14
pDP-1902A5*Luc*	27.75 $\pm$ 3.08
pDPΔTATA	0.90 $\pm$ 0.11

class Alpha GST genes is not completely surprising as the rat Ya subunit (*GSTA1* and *A2*) genes and the rat Yc subunit (*GSTA3* and *A5*) genes are located on different chromosomes [33].

The lack of an intron in the non-coding region of *GSTA5* results in exon 1 of this gene being significantly larger than the corresponding exon from other class Alpha GST genes (Table 4). Exons 2, 3, 4 and 5 of *GSTA5*, which form highly homologous regions of the core protein (residues 30–182), have almost identical intron/exon boundaries to exons 3, 4, 5 and 6 from rat *GSTA2*. Conservation of sequence between exon 6 from *GSTA5* and exon 7 from *GSTA2* is high but diverges significantly in the 3′ untranslated region. Exon 6 of rat *GSTA5* is significantly longer than the equivalent exon of other class Alpha genes. Thus, the core coding areas of *GSTA5* and *GSTA2* are significantly conserved at the cDNA and genomic levels, but their 5′- and 3′-flanking regions demonstrate marked divergence. Interestingly, this pattern is also observed when rat *GSTA5* is compared with class Alpha GST from both mouse and human (Table 4). As with other class Alpha *GST*, there appears to be little conservation in the size of introns between the different genes.

Transfection into HepG2 cells of a reporter construct containing the region -1902 bp to $+65$ bp of the *rGSTA5* gene has shown that it can support transcription. Deletion analysis has demonstrated that the elements necessary for basal expression are contained within the region -177 bp to $+65$ bp, and includes the TATA-box located between -56 bp and -51 bp.

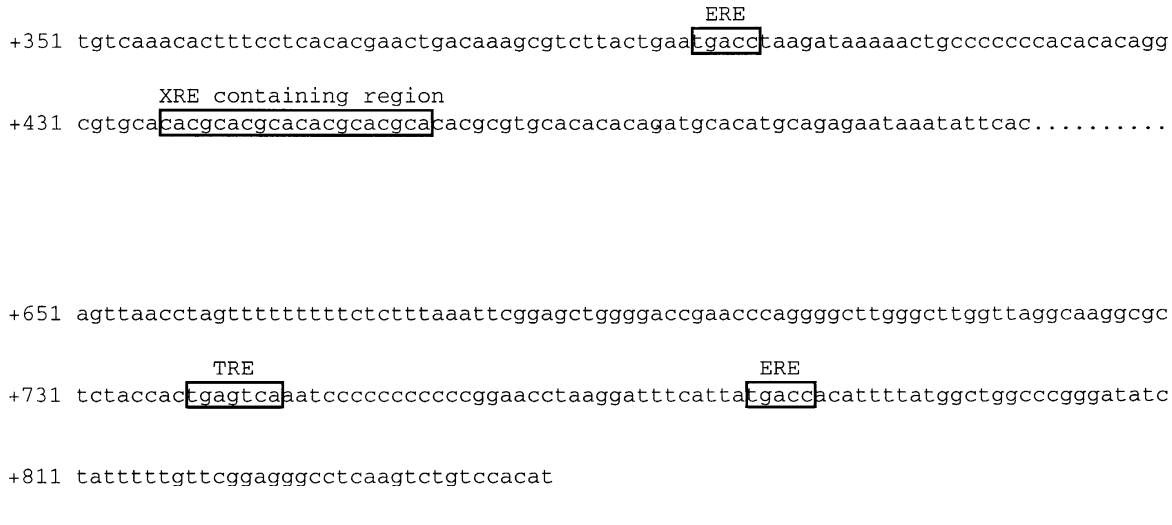

Figure 6 Putative enhancer elements within intron 1 of the rat *GSTA5* gene

The nucleotide sequences representing segments of intron 1 from the rat *GSTA5* gene are shown. Numbers indicate the position of each region relative to the transcription start site. Boxed nucleotides represent sequences corresponding to the consensus motifs for an ERE (oestrogen-responsive half site), a TRE (TPA-responsive element) and an XRE (xenobiotic-responsive element).

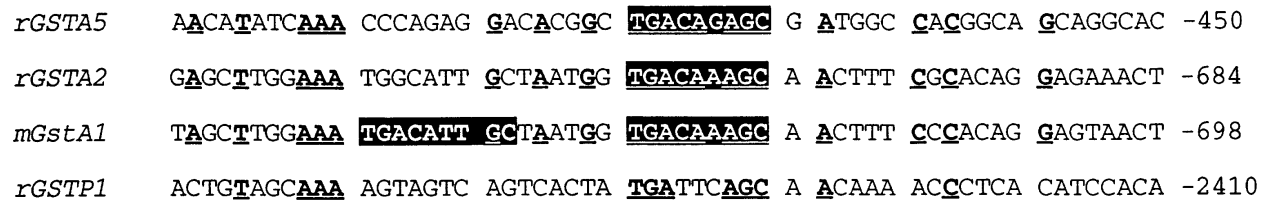

Figure 7 Alignment of *cis*-acting elements from four different antioxidant-responsive *GST* genes

Nucleotides that form the conserved ARE motif, 5′-TGACNNNGC-3′, are indicated in bold inverted type. Numbers indicate the relative position of the 3′-end of each 56-bp element with respect to the transcription start site. Nucleotides in *rGSTA5* present in three or more genes are indicated in bold underlined type. In *rGSTA2* the ARE and some of the conserved flanking nucleotides have been shown to be important in binding nucleoprotein complexes [37]. Sequence information is taken from [18,34–36].

Despite the absence of an intron in the non-coding region of *GSTA5*, analysis of the 5′ untranslated sequence revealed some similarity in the type of enhancer elements within the promoter region of this gene and other class Alpha GST genes. Most significantly the *GSTA5* gene contains a consensus ARE, in reverse orientation, that is located between nucleotide −421 bp and −429 bp. Although it remains to be established whether this ARE is functional, the fact that GSTA5 is inducible by BHA is consistent with the existence of such an element in the *GSTA5* gene. Comparison of the 5′-flanking region of *GSTA5* containing this putative ARE with similar regions in other GST genes [18,34,35] reveals the existence of a significant degree of sequence similarity, suggesting that the consensus ARE along with their flanking sequences may have evolved from a common ancestor. It should be noted that a guanine nucleotide, rather than an adenine nucleotide, is found immediately 3′ of the ARE consensus sequence in the 5′-flanking region of rat *GSTA5* (Figure 7) whereas, as noted by Jaiswal [38], the 3′-flanking nucleotide immediately adjacent to other ARE sequences is occupied by an adenine nucleotide.

Inspection of the flanking region of mouse *GstA1* (Figure 7) reveals the existence of two ARE consensus sequences; the two ARE motifs are tandemly arranged and separated by 6 bp. The demonstration that the mouse *GstA1* gene is positively regulated through this enhancer element by electrophilic compounds and antioxidants alike, led Friling et al. [35] to refer to this enhancer as an electrophile-responsive element (EpRE). Comparison of the EpRE in mouse *GstA1* with the flanking nucleotides of the ARE in rat *GSTA5* (Figure 7) indicates that the promoter region of rat *GSTA5* contains only one ARE consensus sequence. The possession of a single putative ARE in *GSTA5* is similar to the situation that exists in *GSTA2*.

Rat GSTA5 is inducible by phenobarbital [11]. The work of Fulco and his colleagues [20] has shown that the 5′-flanking regions of many barbiturate-inducible genes contain a Barbie box. This box comprises an invariant AAAG motif within an imperfectly conserved 5′-ATCAAAAGCTGGAGG-3′ sequence; typically between 8 and 13 nucleotides, out of the 15 bp that the Barbie boxes comprise, are conserved between different phenobarbital-inducible mammalian genes. A potential Barbie box, 5′-tcacAAAGgaGaAGG-3′, can be identified flanking *GSTA5* between −1773 and −1759 bp (Figure 5).

Elevated levels of rat GSTA5 are found in hepatomas and in neoplastic nodule-bearing livers [8,14]. The expression of this subunit during the early stages of hepatocarcinogenesis may be attributed to the TRE consensus sequence, 5′-TGAGTCA-3′,

found in intron 1 (Figure 6) but further work is required to clarify this point.

A remarkable feature of the 5′ untranslated sequence is the presence of a 116-nucleotide polypurine run between −984 bp and −1000 bp. Polypurine sequences have been described previously [38–40] and have been implicated in the formation of triple-stranded and tetramer-stranded DNA structures. Several laboratories have suggested that repeating units of four guanine residues form a square planar unit leading to the generation of tetramer-stranded DNA [39,40]. In this model, each guanine residue serves as both a donor and acceptor in adjacent G-G Hoogsteen base pairing. The repeating units of four guanine residues can be separated by up to four nucleotides [39,40]. Such guanine-rich sequences occur in immunoglobulin switch regions [41], in gene promoters [42,43] and in telomeric DNA, where they have been implicated in homologous pairing of sister chromatids during the early phases of meiosis [40]. Htun and Dahlberg [38] have suggested that repeat co-polymer sequences of $(dA\text{-}dG)_n$ can assume a triple-stranded hinged DNA (H-DNA) structure that may be involved in relaxing negatively supercoiled DNA. Such relaxation of negative supercoils may enable increased transcriptional activity of downstream sequence [44]. Although the suggestion that these structures could form in this promoter sequence would be speculative, the difficulties from compressions that we encountered in obtaining DNA sequence data may be a result of similar secondary structure forming within the template DNA.

The S1 nuclease protection assay demonstrated that in Morris Hepatoma cells transcription of *GSTA5* is initiated 51 bp downstream of a consensus TATA-box. This location differs from the conventional spacing of 25–30 nucleotides between the transcriptional start site and the TATA-box. Whether the potential ability of regions of the *GSTA5* promoter to form complex secondary structures influences the spacing between the TATA-box and the initiation of transcription is a point that requires further investigation.

In conclusion, this paper describes the isolation and characterization of the structure of the rat *GSTA5* gene encoding the AFB_1-metabolizing rat A5 (Yc_2) subunit. The most notable feature of *GSTA5* is that it comprises only six exons while previously described class Alpha GST genes contain seven exons. However, computer-aided analysis demonstrated that, in common with the rat *GSTA2* gene [15], *GSTA5* contains an ARE and two putative regulatory enhancer elements believed to be responsible for tissue-specific regulation. Like *GSTA2*, the *GSTA5* gene contains sequence with homology to the XRE. However, in *GSTA5*, the four contiguously aligned XRE motifs are found within the first intron and not in the 5′ sequence upstream of the transcription initiation site. In addition, *GSTA5* also contains ERE consensus sequences in both promoter and intronic regions that may have some role in ontogenic and sex-specific regulation of expression of the protein subunit.

We are especially indebted to Dr. Brian McStay for invaluable expert advice in carrying out S1 nuclease protection assays and to Dr. Ronnie McLeod for RNA that he prepared from Morris Hepatoma cells. We also wish to thank Dr. Bo Lungdren for helpful discussions and Dr. Elizabeth Ellis for advice in preparing this manuscript. This work was supported by an MRC project grant (G9310812PA).

REFERENCES

1 Hayes, J. D. and Pulford, D. J. (1995) Crit. Rev. Biochem. Mol. Biol. **30**, 445–600
2 Meyer, D. J. and Thomas, M. (1995) Biochem. J. **311**, 739–742
3 Pickett, C. B. and Lu, A. Y. H. (1989) Annu. Rev. Biochem. **58**, 734–764
4 Sparnins, V. L., Venegas, P. L. and Wattenberg, L. W. (1982) J. Natl. Cancer Inst. **68**, 493–496
5 Benson, A. M., Batzinger, R. P., Ou, S.-Y. L., Bueding, E., Cha, Y.-N. and Talalay, P. (1978) Cancer Res. **38**, 4486–4495
6 Prestera, T., Holtzclaw, W. D., Zhang, Y. and Talalay, P. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 2965–2969
7 Wattenberg, L. W. (1993) Cancer Res. **53**, 5890–5896
8 Hayes, J. D., Judah, D. J., McLellan, L. I., Kerr, L. A., Peacock, S. D. and Neal, G. E. (1991) Biochem. J. **279**, 385–398
9 Hayes, J. D., Nguyen, T., Judah, D. J., Petersson, D. G. and Neal, G. E. (1994) J. Biol. Chem. **269**, 20707–20717
10 Meyer, D. J., Harris, J. M., Gilmore, K. S., Coles, B., Kensler, T. W. and Ketterer, B. (1993) Carcinogenesis **14**, 567–572
11 McLellan, L. I., Judah, D. J., Neal, G. E. and Hayes, J. D. (1994) Biochem. J. **300**, 117–124
12 Stresser, D. M., Williams, D. E., McLellan, L. I., Harris, T. M. and Bailey, G. S. (1994) Drug Met. Dispos. **22**, 392–398
13 Buetler, T. M., Gallagher, E. P., Stahl, D. L., Wang, C., Hayes, J. D. and Eaton, D. L. (1995) Tox. Appl. Pharmacol. **135**, 45–57
14 Hayes, J. D., Judah, D. J. and Neal, G. E. (1993) Cancer Res. **53**, 3887–3894
15 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
16 Ghosh, D. (1992) Nucleic Acids Res. **20**, 2091–2093
17 Sohlenius, A.-K., Andersson, K., Bergstrand, A., Spydevold, Ø. and DePierre, J. W. (1994) Biochim. Biophys. Acta **1213** 63–74
18 Rushmore, T. H., Morton, M. R. and Pickett, C. B. (1991) J. Biol. Chem. **266**, 11632–11639
19 Nguyen, T., Rushmore, T. H. and Pickett, C. B. (1994) J. Biol. Chem. **269**, 13656–13662
20 Liang, Q., He, J.-S. and Fulco, A. J. (1995) J. Biol. Chem. **270**, 4438–4450
21 Hayes J. D., McLeod, R., Pulford, D. J., Judah, D. J., Nguyen, T., Arthur, J. R. and Neal, G. E. (1996) in Glutathione S-Transferases: Structure, Function and Clinical Implications (Vermeulen, N. P. E., Mulder G. J., Nieuwen-Luyse, M., Peters, W. H. M. and Van Bladderen, P. J., eds.), pp. 73–87, Taylor and Francis, London
22 Fujisawa-Sehara, A., Sogawa, K., Yamane, M. and Fujii-Kuryama, Y. (1987) Nucleic Acids Res. **15**, 4179–4191
23 Denison, M. S., Fisher, J. M. and Whitlock, Jr., J. P. (1988) J. Biol. Chem. **263**, 17221–17224
24 Paulson, K. E., Darnell, Jr., J. E., Rushmore, T. H. and Pickett, C. B. (1990) Mol. Cell. Biol. **10**, 1841–1852
25 Angel, P., Imagaw, M., Chiu, R., Stei, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P. and Karin, M. (1987) Cell **49**, 729–739
26 Lee, W., Mitchell, P. and Tjian, R. (1987) Cell **49**, 741–752
27 Telakowski-Hopkins, C. A., Rothkopf, G. S. and Pickett, C. B. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 9393–9397
28 Daniel, V., Sharon, R., Tichauer, Y. and Sarid, S. (1987) DNA **6**, 317–324
29 Suzuki, T., Johnston, P. N. and Board, P. G. (1993) Genomics **18**, 680–686
30 Klone, A., Hussnatter, R. and Sies, H. (1992) Biochem. J. **285**, 925–928
31 Rohardanz, E., Nguyen, T. and Pickett, C. B. (1992) Arch. Biochem. Biophys. **298**, 747–752
32 Rozen, F., Nguyen, T. and Pickett, C. B. (1992) Arch. Biochem. Biophys. **292**, 589–593
33 Yamada, T., Muramatsu, Y., Yasue, M., Agui, T., Yamada, J. and Matsumoto, K. (1992) Cytogenet. Cell Genet. **61**, 125–127
34 Okuda, A., Imagawa, M., Maeda, Y., Sakai, M. and Muramatsu, M. (1989) J. Biol. Chem. **264**, 16919–16926
35 Friling, R. S., Bensimon, A., Tichauer, Y. and Daniel, V. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6258–6262
35a Daniel, V., Tichauer, Y. and Sharon, R. (1988) Nucleic Acids Res. **16**, 351
36 Nguyen, T. and Pickett, C. B. (1992) J. Biol. Chem. **267**, 13535–13539
37 Jaiswal, A. K. (1994) Biochem. Pharmacol. **48**, 439–444
38 Htun, H. and Dahlberg, J. E. (1989) Science **243**, 1571–1576
39 Williamson, J. R., Raghuraman, M. K. and Cech, T. R. (1989) Cell **59**, 871–880
40 Sen, D. and Gilbert, W. (1988) Nature (London) **334**, 364–366
41 Shimizu, A. and Honjo, T. (1984) Cell **36**, 801–803
42 Evans, T., Schon, E., Gora-Maslak, G., Patterson, J. and Efstratiadis, A. (1984) Nucleic Acids Res. **12**, 8043–8058
43 Kilpatrick, M. W., Torri, A., Kang, D. S., Engler, J. A. and Wells, R. D. (1986) J. Biol. Chem. **261**, 11350–11354
44 Borowiec, J. A. and Gralla, J. D. (1985) J. Mol. Biol. **184**, 587–598

Received 22 January 1996/29 March 1996; accepted 22 April 1996