Structure and expression of a cluster of glutathione S-transferase genes from a marine fish, the plaice (*Pleuronectes platessa*)

Michael J. LEAVER*, Joy WRIGHT and Stephen G. GEORGE

NERC Unit of Aquatic Biochemistry, University of Stirling, Stirling FK9 4LA, Scotland, U.K.

Glutathione S-transferases are involved in the detoxification of reactive electrophilic compounds, including intracellular metabolites, drugs, pollutants and pesticides. A cluster of three glutathione S-transferase genes, designated *GSTA*, *GSTA1* and *GSTA2*, was isolated from the marine flatfish, plaice (*Pleuronectes platessa*). *GSTA* and *GSTA1* code for protein products with 76% amino acid identity. *GSTA2* appears to contain a single nucleotide deletion which would render any product non-functional. All of these genes consist of six exons of similar sizes and

greater than 70% nucleotide identity, and are interrupted by five introns of differing sizes. *GSTA* and *GSTA1* mRNAs were present in a range of tissues, while *GSTA2* mRNA was not detected. Expression of *GSTA* mRNA was increased in plaice intestine and spleen by pretreatment with β -naphthoflavone, and expression of both *GSTA* and *GSTA1* mRNAs was increased in plaice liver and gill by pretreatment with the peroxisome-proliferating agent perfluoro-octanoic acid.

INTRODUCTION

Cytosolic glutathione S-transferases (GSTs; EC 2.5.1.18) catalyse the conjugation of reduced glutathione to a variety of compounds containing an electrophilic centre, and are probably present in all eukaryotic cell types, having been found in all organisms examined to date [1]. The GSTs are a multigene family of enzymes; in the rat, one of the most widely studied organisms with regard to GSTs, they may be encoded by at least 20 genes. The products of these genes have been divided into five classes (Alpha, Mu, Pi, Theta and Sigma) based largely on structural and immunological similarities [1,2]. The native enzymes are homodimers or heterodimers of polypeptides from the same class, and individual subunits exhibit different substrate specificities. This diversity of structure within the family of GST isoenzymes provides the capability to conjugate a very broad range of compounds. These include not only reactive products produced endogenously during cellular oxidative processes, such as fatty acid hydroperoxides, hydroxyalkenals and nucleic acid base hydroperoxides and alkenals [3,4], but also a very large range of xenobiotics, including plant metabolites and man-made chemicals such as pesticides, drugs and pollutants [5]. Additional functions of certain mammalian GST isoforms include binding of intracellular metabolites such as haem, bilirubin and steroid hormones, and they may therefore act as intracellular storage and carrier proteins for small lipophilic molecules.

A characteristic feature of several mammalian GST isoforms is their induction following exposure to electrophilic xenobiotics, many of which are often substrates for GST. The discovery of upstream enhancer elements in the rat Alpha-class *GSTA2* and mouse *gstA1* genes that confer inducibility after treatment with a very broad range of electrophilic xenobiotics has been described [6–8]. These sequences have been termed antioxidant response elements (AREs) [6] or electrophile responsive elements [7], and have been shown to interact with a number of transcription factors [5,6] and other unidentified DNA-binding proteins [9].

Less is known about GSTs from other organisms. To date

only Theta-class GST isoforms have been reported from plants and insects [10], while isoforms related to Alpha-, Pi- and Theta-class GSTs have been identified in several fish species [11,12]. The marine pleuronectid flatfish, plaice, is a member of a group of fish which are widespread in the oceans of the world and, in the Northern hemisphere, have been used extensively as indicator species for environmental quality assessments [13,14]. Many of these studies have targeted the xenobiotic-mediated response of an enzyme, cytochrome *P*-4501A, as a measure of exposure to environmental pollution in these fish [15]. However, several other enzymes, including GSTs, are involved in xenobiotic metabolism in vertebrates [16] and may also be regulated by pollutants. Previous studies have shown that plaice liver contains two major GST isoforms, *GSTA* and *GSTB*. *GSTA* does not cross-react with antibodies against representatives of any of the mammalian GST classes, whereas *GSTB* is immunologically related to mammalian Alpha-class isoforms [11]. Subsequent cloning of a cDNA for *GSTA* showed it to be most closely related to GSTs from insects and plants and Theta-class isoforms from mammals [17]. In order to study further the possible roles of *GSTA* in pollutant metabolism and resistance, we have undertaken gene cloning studies. Here we describe the structure and tissue-specific expression of a cluster of three GST genes from plaice, and provide evidence for a novel mechanism of GST regulation by peroxisome proliferators, a class of compounds that includes endogenous polyunsaturated fatty acids and their metabolites, hypolipidaemic drugs and widespread environmental pollutants such as industrial plasticizers.

MATERIALS AND METHODS

Materials

Restriction enzymes, polynucleotide kinase and Moloney murine leukaemia virus reverse transcriptase were purchased from Gibco-BRL. Radioisotopes [α -³⁵S]dATP, [α -³²P]dCTP and [γ -³²P]ATP were purchased from ICN-Flow. Klenow fragment,

Abbreviations used: BNF, β -naphthoflavone; GST, glutathione S-transferase; ARE, antioxidant response element; PFOA, perfluoro-octanoic acid; PPAR, peroxisome-proliferator-activated receptor; PPRe, peroxisome proliferator response element; RT-PCR, reverse transcription-PCR.

* To whom correspondence should be addressed.

The sequences described have been submitted to the Genbank/EMBL/DBJ Nucleotide Sequence Databases; the full sequence of the plaice genomic DNA fragment containing *GSTA*, *GSTA1*, *GSTA2* and *PPTN* and flanking and intervening regions has been assigned accession no. X95199, and the sequence of *GSTA1* cDNA has been assigned accession no. X95200.

sequencing reagents, modified T7 DNA polymerase and S1 nuclease were obtained from Pharmacia Biotech. Bacteriophage λ FIXII and λ Gold *in vitro* packaging extracts were purchased from Stratagene. Oligonucleotides were synthesized by core facilities at the University of Stirling or were purchased from commercial sources. Nylon (Hybond-N) membrane was purchased from Amersham International.

Animals and experimental treatments

Juvenile and adult plaice were obtained by trawl off the Scottish east coast and kept in filtered recirculating seawater at a constant 10 °C for 2 weeks before use. Animals were fed *ad libitum* during this period. β -Naphthoflavone (BNF) and perfluoro-octanoic acid (PFOA) were both administered to male fish only by a single intraperitoneal injection at 50 mg per kg 72 h prior to analysis. The fish were not fed during this period. This BNF dose is typical of that used to induce cytochrome *P*-4501A in fish species [18], while the PFOA dose used is one-quarter of the 30-day LD₅₀ in the rat for a single intraperitoneal injection [19].

Library construction and screening

Molecular biological procedures were carried essentially as described in Sambrook et al. [20]. Plaice genomic DNA was digested with a range of concentrations of *Mbo*I and analysed by agarose gel electrophoresis. Pools containing DNA in the 16–25 kbp size range were then treated with Klenow fragment, dATP and dGTP, resulting in partially filled-in ends. λ FIXII was digested with *Xho*I and partially filled in with Klenow fragment, dCTP and dTTP before ligating to various concentrations of partially digested and filled-in plaice DNA. An aliquot of each of these ligation products was then tested by *in vitro* packaging and plating on *Escherichia coli* P2392. The library, consisting of 8×10^5 recombinants, was then screened with a full-length plaice GSTA cDNA [17] and positive plaques were identified and purified by three successive rounds of plating and hybridization. Conditions for hybridization were as described for blot hybridizations (see below).

Subcloning and sequencing

Restriction fragments of cloned genomic DNA were subcloned into pBluescript KS(+), propagated in *E. coli* XL1Blue and isolated by standard miniprep techniques. These subclones were used to generate overlapping templates for sequencing reactions by nested deletion subcloning using a Pharmacia Nested Deletion Kit. Sequencing reactions were carried out with a Pharmacia Sequencing Kit. Both DNA strands of all exons, exon/intron boundaries and putative promoter regions were sequenced. In some cases only a single strand of the intronic regions was sequenced. Sequence analysis and homology searches were performed with the DNAsis/PROSis software package (Hitachi) and FASTA [21].

Nucleic acid preparation and blot hybridization

Genomic DNA was prepared by lysing whole-blood samples in 100 mM Tris/HCl, pH 8.5, 200 mM NaCl, 50 mM EDTA and 0.2 % SDS, followed by digestion with RNase A for 1 h at 25 °C, then extraction with phenol and phenol/chloroform. Extracted DNA was then either dialysed extensively in 10 mM Tris/HCl and 1 mM EDTA before library construction, or ethanol-precipitated and resuspended in 10 mM Tris/HCl and 1 mM EDTA prior to Southern analysis. For Southern blotting, purified DNA was digested with the appropriate restriction enzymes, fractionated by 0.8 % agarose gel electrophoresis, denatured in

0.4 M NaOH/1.5 M NaCl for 30 min, neutralized in 0.5 M Tris/HCl, pH 7.4, and 1.5 M NaCl and transferred to a nylon membrane (Hybond-N) by capillary blotting in $10 \times$ SSC ($1 \times$ SSC is 0.15 M NaCl and 0.015 M sodium citrate).

RNA was isolated by the acid guanidinium thiocyanate/phenol method [22], and poly(A)⁺ RNA was obtained by adsorption and elution from poly(U) paper (Amersham International). For Northern hybridization, samples were denatured in 1.5 M glyoxal, 40 % DMSO, 40 mM Mops/NaOH, pH 7.0, 10 mM sodium acetate and 1 mM EDTA for 1 h at 50 °C before fractionation in 1.2 % agarose in 40 mM Mops/NaOH, pH 7.0, 10 mM sodium acetate, 1 mM EDTA and 0.5 μ g/ml ethidium bromide. RNA was visualized by UV transillumination to check integrity and loading consistency before vacuum blotting directly to Hybond-N in $10 \times$ SSC.

For both Northern and Southern blots, hybridizations were carried out with [α -³²P]dCTP random-primer-labelled DNA fragments in 50 mM sodium phosphate, pH 7.2, 875 mM NaCl, 5 mM EDTA, $5 \times$ Denhardt's solution, 0.2 % SDS and 100 μ g/ml sonicated denatured herring sperm DNA at 65 °C for 16 h. Filters were then washed in $2 \times$ SSC/0.1 % SDS at room temperature three times for 10 min each and then twice in $0.1 \times$ SSC/0.1 % SDS at 65 °C for 10 min each. Hybridizing DNA fragments were visualized by autoradiography.

Primer extension

Oligonucleotides (5'-GGTTTCAGAGAGAAGCTG for GSTA mRNA or 5'-GGTTTCGAGATGTCGAGC for GSTA1 mRNA) were end-labelled with [γ -³²P]ATP and T4 polynucleotide kinase in 50 mM Tris/HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂ and 1 mM dithiothreitol. Samples of 1 μ g or 5 μ g of RNA were heated to 90 °C with 1 pmol of labelled oligonucleotides and the mixture was allowed to reach 37 °C before adding dNTPs to 1 mM plus 200 units of reverse transcriptase in a final reaction volume of 25 μ l. After 1 h at 37 °C, the reaction was stopped with 1 μ l of 0.5 M EDTA, and 10 μ g of RNase A was added. After incubation at room temperature for 30 min, the reaction mixture was extracted with phenol/chloroform, ethanol-precipitated and resuspended in formamide dye buffer. Samples were run on 6 % polyacrylamide DNA sequencing gels, alongside sequence ladder size markers.

Cloning of cDNA for GSTA1

A 1 μ g sample of adult male plaice liver total RNA was incubated for 5 min at 70 °C in a volume of 15 μ l with 2.5 pmol of a primer (5'-CACAGAAGCAGCTTTC) complementary to bases 746–762 of the putative GSTA1 mRNA predicted from the gene sequence (see Figure 2). After cooling on ice, reverse transcription was carried out in 20 mM Tris/HCl, pH 8.4, 50 mM KCl, 3 mM MgCl₂, 400 mM dNTPs and 200 units of reverse transcriptase in a final volume of 25 μ l at 42 °C for 30 min. The reaction was then heated to 65 °C for 15 min, and treated with 2 units of RNase H for 10 min at 55 °C. cDNA was then purified by binding and elution from a preparatory silica matrix (GlassMAX; Gibco-BRL). One-tenth of the purified cDNA was then subjected to PCR using cloned Pfu polymerase (Stratagene), a primer to bases 33–50 immediately upstream of the ATG initiation codon of GSTA1 (5'-TTCCGGCTCGACATCTCG) and a primer nested inside the site of the cDNA priming site and abutting the GSTA1 termination codon (5'-TTCAAAGTGAGTAAGTT) (see Figure 2). PCR was carried out using one cycle at 95 °C for 5 min, followed by 30 cycles at 95 °C for 1 min, 45 °C for 1 min and 72 °C for 2 min, and a final single cycle at 72 °C for 8 min.

Products were resolved by agarose gel electrophoresis and subcloned to pBluescript KS(+) for sequence analysis.

RESULTS

Three repeated *GSTA* genes are present in the plaice genome

Three distinct clones were isolated from a plaice genomic DNA library after screening by hybridization to the previously characterized *GSTA* cDNA [17], and were designated λ gGSTA3, λ gGSTA15 and λ gGSTA29. After restriction mapping and sequencing of the portions of these clones that hybridized to the *GSTA* cDNA on Southern blots, it was determined that the gene for *GSTA* spanned the overlapping region of the λ gGSTA3 and λ gGSTA29 clones. On further sequencing of λ gGSTA3, other areas with remarkable identity to the *GSTA* cDNA were identified, and complete sequencing of a 13 kb region encompassed by these two clones revealed the existence of two further genes for related GSTs. The three genes were designated *GSTA*, *GSTA1* and *GSTA2* (Figure 1). Only a short region of the λ gGSTA15 clone hybridized to the *GSTA* cDNA. Upon sequencing it was found that this area was 94% identical to the final 160 nucleotides of the *GSTA* cDNA and 200 nucleotides of its downstream genomic flanking region (results not shown), and it appeared to be a truncated, unprocessed copy of part of the last exon and the 3' flanking region of *GSTA* (Figure 1). The three full-length *GSTA* genes possess six exons of similar sizes but have quite different intron sizes. The nucleotide and predicted amino acid sequences of *GSTA* are 96% and 97% identical respectively to the *GSTA* cDNA. These differences probably arise from the fact that different individuals from a wild outbred population of plaice were used to generate the cDNA and genomic libraries (see the Discussion section). The nucleotide and predicted amino acid sequences of *GSTA1* are 77% and 76% identical to the *GSTA* cDNA (Figure 2). When aligned with *GSTA* and *GSTA1*, *GSTA2* appears to have a single nucleotide deletion in exon 4 which would presumably render any predicted product non-functional (Figure 2).

In view of the variation between the plaice GST genes and their cloned cDNAs, the distribution of these genes among other individuals was analysed. Restriction fragments comprising parts of the final, least similar, exons of these three genes were used to generate gene-specific probes (Figure 2), and were hybridized to Southern blots of *Nco*I-digested plaice DNA. The initiation codons of all three genes contain an *Nco*I site and, from consideration of the λ clone maps, each gene can be excised as a discrete *Nco*I restriction fragment of 1430 bp for *GSTA2*, 4225 bp for *GSTA1* and approx. 6500 bp for *GSTA*. On performing this experiment with gene-specific probes for *GSTA* and *GSTA1* on DNA from three individuals (Figure 3), three different restriction

patterns were evident. However, each of the probes specifically hybridized to different sets of DNA fragments. A restriction fragment from the 5' half of the *GSTA* cDNA, encompassing the first four (least divergent) exons of these three genes, when hybridized at low stringency, hybridized to a set of DNA fragments of the same sizes as those observed individually after hybridization with each of the gene-specific probes. Interestingly, the probe derived from the last exon of *GSTA* hybridized to an additional fragment not seen in the other blots, which presumably corresponds to the truncated sequence present in the λ gGSTA15 clone. The gene-specific probe for *GSTA1* hybridized to a single DNA fragment of the same size in all of the individuals tested, which corresponded to the fragment predicted by the λ clone sequences. The gene-specific probe for *GSTA* hybridized to DNA fragments of the predicted size and, in addition, to another one or two fragments, suggesting the presence of another as yet undetected sequence with a high degree of identity to *GSTA*.

A transposon-like sequence is present between *GSTA* and *GSTA1*

A search of the Genbank Nucleotide Sequence Database against *GSTA*, *GSTA1* and *GSTA2* and their intervening and flanking regions revealed the presence of a transposon-like sequence immediately upstream of the *GSTA* gene (Figure 4), designated *PPTN*. This consists of an open reading frame coding for a sequence of 313 amino acid residues with similarity to putative transposases from a number of organisms, all of which display significant identity to the *Tc1* transposon from the nematode *Caenorhabditis*, and with 40% amino acid identity with the predicted amino acid sequence from the *Tc1b* transposase gene of *Caenorhabditis briggsae* [23,24]. However, the plaice transposase coding region was interrupted by a single termination codon which would presumably render the transposon inactive. This transposase-like sequence was flanked by two inverted repeat sequences of 40 bp and 28 bp, with the repeats separated by 143 bp.

GSTA and *GSTA1* are expressed in a wide range of plaice tissues, and are induced by PFOA in gill and liver and by BNF in intestine and spleen

To test for the expression of *GSTA*, *GSTA1* and *GSTA2*, the gene-specific probes were used to probe Northern blots of RNA from various plaice tissues, from either untreated fish or fish treated with BNF, an inducer of mammalian GST isoforms or, in view of the identification of peroxisome proliferator response elements (PPREs) in the *GSTA1* upstream region, with the peroxisome proliferator PFOA. In untreated fish, hybridizing RNAs were present for both *GSTA* and *GSTA1* (Figure 5), and the patterns of hybridization in the different tissues appeared to

Figure 1 Structure of λ clones containing GST genes and a transposon-like sequence

Regions of two overlapping clones (λ gGSTA3 and λ gGSTA29) containing the *GSTA*, *GSTA1*, *GSTA2* and *PPTN* genes are shown, approximately to scale. The solid line indicates the sequenced portion of the clones. Exons are shown as black boxes and are numbered, and the *PPTN* transposon-like sequence is indicated with an open box. A third λ clone (λ gGSTA15), containing a truncated portion of the *GSTA* final exon, is also indicated. Selected restriction sites are also indicated: B, *Bam*HI; N, *Nco*I; H, *Hind*III.

Figure 2 Comparison of GST gene exon sequences

GSTA, GSTA1 and GSTA2 exon sequences are shown in upper case, regions of identity are boxed and flanking sequences are shown in lower case. ▲ symbols mark the positions of intron/exon boundaries; # marks a deletion in exon 4 of GSTA2. Initiation codons, termination codons and putative polyadenylation signals are underlined. Regions used to generate primers for mRNA extension analysis and for RT-PCR of GSTA1 cDNA are shown, with bold arrows indicating the direction of primer extension. The complete sequences of regions used to generate gene-specific probes are enclosed in a bold box. The deduced amino acid sequence for GSTA is shown above the gene sequence, and differences in the GSTA1 deduced amino acid sequence from that of GSTA are shown below the gene sequence.

Figure 3 Southern blot of plaice genomic DNA

Genomic DNA from three individual fish (1, 2 and 3) was restriction-digested with *Nco*I, separated by agarose gel electrophoresis, transferred to a nylon membrane and hybridized to: (A) the 5' portion of *GSTA* cDNA, (B) the gene-specific probe for *GSTA* and (C) the gene-specific probe for *GSTA1*. The positions of DNA size markers are shown, and the approximate positions predicted from λ clone maps for DNA fragments hybridizing to each of the GST probes are indicated.

be similar, although not identical, for these two probes. Interestingly, very little hybridization was evident with either probe in adult female liver. In contrast, the gene-specific probe for *GSTA2* did not hybridize to RNA from any of the tissues tested.

Comparison of the *GSTA* blots with the same blot stripped and reprobed with a rainbow trout tubulin cDNA showed that male plaice treated with BNF showed no differences in *GSTA1* mRNA expression in any of the tissues when compared with untreated fish, but *GSTA* appeared to be induced in intestine and spleen. Fish treated with PFOA showed clear increases in hybridization to *GSTA* and *GSTA1* mRNAs in liver and gill compared with untreated fish. Faint signals arising from high-molecular-mass RNA on PFOA treatment, particularly in brain, were probably the result of non-specific hybridization, since these signals were present after using each of the probes. Increases in tubulin mRNA in liver and intestine of PFOA-treated animals were also evident, and it is possible that these may be the result of peroxisome proliferation in these tissues resulting in an increased requirement for microtubule formation, although no evidence for this is available at present.

Reverse transcription-PCR (RT-PCR) cloning of a *GSTA1* cDNA and primer extension

Further evidence for the expression of *GSTA1* was obtained from RT-PCR analysis of adult male plaice liver RNA. Primers specific for *GSTA1* (Figure 2) produced a single band on gel electrophoresis which, after subsequent subcloning and sequencing, proved to be 99% identical to the transcript predicted for *GSTA1* (results not shown). In addition, transcriptional start sites for *GSTA* and *GSTA1* were determined by primer extension of adult male plaice liver RNA using oligonucleotides comp-

```
CTGCAGCAACATATTAAAGAGTTTTACTGCATTTTGATATTTTCGATAGGTTTTCAGCCATAAAGTGTATATAA
AGTAGGTCATCCACTGACCTGGAGGAGTGGGATTTATGACCTGTACTGCCAACCCAGGAGGGTATCGTTT
GGGCATATTCATGTGTATACAACTATATGTTTGTAGCCTTTGTGTGAAAGTGATCACCTCCATCAAGAAACA
CGCTAATATACAAATAGACCTGATGTATAAGAGGCAACAGATGTTGCTGAAACTCATATATTTCTATTCAGGCC
TTTGTGCGACGCACATAAGATAAAGGTAGTCGGAGTCTATTCATTCATATATGTCACCGGACGTCCTGAACAC
TTTCTTACCTGAAATGACATGTGCCATCATGACTGGGCATCAGTGTATAGAGTTCACTACTGTAATATATAGCTG
AGGTTTCAGSCCTGTCAAAGTGGACTTTTACAGTTCTCAACAGAGGGCAGCACTGGCATCTTGGCTTTTGAATCC
GACCTGTTGAAAGTTTGGATCCGTCTCTTTTATTTTTCACCTTTGTAACCTGCTGCGAAGTGCCAGTAA
TAAGGATATCATATTTGTTTATCCATATGTAGTAACCTATAAATACAACTATTGGTGTGAAATGCATTTTTTTC
CAGTTGATAAAAAAAGTGTATTTTAAATGTTGCCCTGAGTAACCTAAGGAAAGGTTAACTACTTGACAGGGT
GTGAACATATCTTGAGTCAATATCTTTGGTGTACTCACTCAAAATGGGAAGGAGCTATTTCACCAATCCAGACGT
GCAGGAGAGCGTGTGCTATATACAGTGCCTTGCATAAGTATTCACCCCTTTGGACTTTTCTACATTTTGTGCAT
GGTATACACAGATTAATAATTTTATTCATCGTGAGTTTATGTAATGACCAACACAAATATAGTCATCATTTGG
AAGTGGGGGGAATATACATGGATTTTCACTATTTTACAAATAAAAATCTGAAAGTGTGAGTCATATGTAT
TCACCCCTTTTACTGTGAAACCCCTTAAATCCATCATAGAAATGGAAGATATGGCACAACCCGAAAC
CTACCAAGAGGAGGCGCTCCACCAAACTGAAGAGTCGGACAGGAGAAATTAATCAGAGAAGCACCAGGAGG
CCCATGGTTACTCTGGAGGAGTGCAGAGATCCACAGCTGAGGTGGAGAAATCTGTCCACAGGACAACTATAGT
CGCTACTCCACAAATCTGGCTTTATGGAAGAGTGGCAAGAGAAAGCCGTTGTTGAAAGGGATCCATAAAAA
TCCGCTTGGAGTTTGCAGAGCCATGTTGGGAGACACAGCAACATCTGGAAGAAGGTGCTCTGTGTCAGATGAG
ACCAAAATTGAATTTTGGCTCAATGCACAAACCTATGTGTGCGGAAACCAACCTGCCATCAGCTTGAG
CACACATCCCAACAGTGAAACATGGTGGTGTAGCATCATGCTGTGGGGATGCTTCTCTTCAGCAGATACAGG
AAACTGGTCAGATAGAGGAAAGATGGATGGAGCCAAATACAGGAAATCCCTTAAGABAATCTGATGCAGTCT
GCAGAGAGCTTGAGACTGGGCGGAGGTTTCATCTTCCAGCAGGAACATGACCTTAACATACAGCAGAGCTACA
AAGGAATGCTTGGATTAAAGATGTTAATGCTTAAATGGCCAGTCAAGCCAGACCTCAATCCAATAGAG
AATCTATGGCAGACTTGAAGATTGCGGTTCACAGACGGTCTCCATCCAACTGACAGAGCTCATCTTTTGTGC
CAGAGAAGATGGCAAAACCTTTCCATCTCTAGATGTGCAAGCTGTGAGAGACATACCCCAAAAGACTTGCAGCT
ATAATTGCAGCAAGGGGGTCTACCAAGTATTGACACAGGGGGTGAATATAGGCACCCACAGATGTCAB
CTTTTTTGTCTCATTTATGCTGTGTGTCACAAATAAATTTATTTTGACCTCCAAAGTACTATGCATGTTTGT
GATCAACCGGAAAGTTTATTTAAGTCTATTTGAATTCAGCTTAGTAACAATACATACTGGGAAAGTCCAA
GGGGGGTGAATCTATGCAAGGCACTGTAAACAGAGTCTGCCATTCGCGCCCCCTTCCTCACCCTCTCCAC
CACCCTCTCTCTCTCTGAACTCAGAGTACTGCAGCTTCTCTGAAACCATG
```

Figure 4 Sequence of the region upstream of *GSTA*

Putative AREs and CCAAT boxes are indicated, and the transposon sequence is underlined. The first exon of *GSTA* up to and including the protein initiation codon is italicized.

plementary to the sequences immediately upstream of the initiation codons for these genes (Figure 2). The primer extension reactions identified a single transcriptional initiation site for each of these two genes, 72 bp upstream of the initiation codon for *GSTA* and 55 bp upstream of the initiation codon for *GSTA1* (Figure 6).

DISCUSSION

We have isolated from plaice and characterized a 13 kb genomic DNA fragment from plaice containing two intact GST genes, designated *GSTA* and *GSTA1*, and a third, apparently non-expressed, defective gene, *GSTA2*. Each of these genes has five introns and six exons, and the sizes and sequences of the corresponding exons of each gene are very similar, although the sizes and sequences of the corresponding introns are highly divergent. These genes are most similar in sequence to Theta-class GST genes from mammals and Theta-class-related genes from insects and plants. However, the structure of the rat Theta-class *GSTYrs* gene [25] cannot be compared with those of the plaice genes, as the *Yrs* gene contains five exons and four introns and, in so far as the intact proteins can be aligned, there is no particular conservation of the positions of exon/intron boundaries. Theta-class-related GSTs from non-vertebrates have in some cases intronless genes (in insects) [26], or vary in plants from three to ten exons [27,28], and thus their structures are not closely related to these fish genes either.

The predicted transcript for *GSTA* differs from the previously described *GSTA* cDNA [17] by 30 nucleotides, 14 of which occur in the 3' untranslated region (results not shown). These differences result in six differences in the predicted amino acid sequences.

Figure 5 Tissue distribution of GSTA and GSTA1 mRNAs

Samples of 8 μ g of total RNA were loaded per lane: L, liver; K, kidney; I, intestine; S, spleen; G, gill; B, brain; T, testis (or ovarian wall). The sex of the fish is indicated. BNF and PFOA were each administered by a single intraperitoneal injection at 50 mg/kg 72 h prior to analysis

Figure 6 Primer extension of plaice RNA

Radiolabelled oligonucleotide primers complementary to GSTA and GSTA1 mRNAs were extended using reverse transcriptase and electrophoresed on 6% denaturing acrylamide. Lane 1, 1 μ g of plaice RNA-GST primer; lane 2, 0.1 μ g of plaice poly(A)⁺ RNA-GST primer; lane 3, 5 μ g of plaice RNA-GST primer; lane 4, 0.5 μ g of plaice poly(A)⁺ RNA-GST primer; lane 5, 5 μ g of yeast tRNA-GST primer; lanes 6–9, sequencing reaction sizing ladder; lane 10, 5 μ g of yeast tRNA-GST primer; lane 11, 0.5 μ g of plaice poly(A)⁺ RNA-GST primer; lane 12, 5 μ g of plaice RNA-GST primer; lane 13, 0.1 μ g of plaice poly(A)⁺ RNA-GST primer; lane 14, 1 μ g of plaice RNA-GST primer. Sizes of extended products are indicated.

Differences were also observed in the nucleotide sequences of the GSTA1 gene and cDNA (results not shown) which resulted in two changes in the deduced amino acid sequences. Unfortunately,

these results are complicated by the fact that different individuals from a wild, outbred, population were used to generate the cDNA and genomic libraries and as starting material for mRNA analysis. Nevertheless, allowing for some variation between individuals, we conclude that the GSTA and GSTA1 genes and the cDNAs described here and elsewhere [17] are equivalent. Southern blot analysis of restriction-digested plaice DNA suggests that there may be a relatively large inter-individual variation in these genes, since three individuals from the same population of fish gave three different hybridization patterns. We have also observed a high degree of variation when probing plaice DNA with cDNA probes to other xenobiotic-metabolizing genes (M. J. Leaver, J. Wright and S. G. George, unpublished work); furthermore, other workers have reported a high degree of allozyme heterozygosity and chromosomal polymorphism (Robertsonian) in Pleuronectid flatfish and plaice in particular [29–31] when compared with other fish species. Nevertheless the use of gene-specific probes shows that we may be able to account for all but one of the GSTA-related sequences in this species. Interestingly, the truncated portion of the last exon of GSTA, which was present in the λ gGST15 clone and which we originally suspected of being a cloning artifact, was detected on Southern blots of restriction-digested plaice DNA hybridized to the gene-specific GSTA probe that contained the 3' end of GSTA cDNA, but not on an identical blot probed with the less specific 5' half of this cDNA.

It is unlikely that GSTA and GSTA1 are pseudogenes, because they contain typical eukaryotic intron/exon splice junctions and code for full-length GSTs, and both transcripts can be detected on Northern blots of RNA from various plaice tissues hybridized to the gene-specific probes. Very similar, but not identical, patterns of expression were observed for GSTA and GSTA1, each of these genes producing a similar-sized mRNA of around 1000 nucleotides [17] in all of the tissues tested. In untreated fish the relative amounts of GSTA mRNA were higher in liver and gill than those of GSTA1, while GSTA1 mRNA was more evident in brain than was GSTA. There would also appear to be

large sex-related differences in the patterns of hybridization of both *GSTA* and *GSTA1* mRNAs, particularly in liver: expression in the female liver was very low for both *GSTA* and *GSTA1*. This may be largely due to the fact that the fish used for this experiment were adults, collected just prior to spawning; up to this point the female liver has been devoted to production of egg proteins and lipids, and may give completely different results at other times of the year. Similar low levels of other xenobiotic-metabolizing enzymes and total GST activity have been noted in spawning female plaice on other occasions; however, non-spawning adult females and juveniles of both sexes at other times of the year have similar levels of GST activity as males [32]. Further evidence for the expression of *GSTA* and *GSTA1* can be seen from primer extension analysis of RNA from adult male liver, which clearly shows the presence of a single transcriptional start site for each of *GSTA* and *GSTA1*, indicating that both of these mRNAs are present as distinct molecules. In addition, RT-PCR of adult male liver RNA using primers specific to *GSTA1* gave the predicted product, which on sequencing proved to be 99% identical to the sequence predicted from the *GSTA1* gene. The use of a polyclonal antibody raised to *GSTA* gives a single band on Western blots of plaice tissues; however, *GSTA* and *GSTA1* have predicted molecular masses of 25 723 and 25 746 Da respectively, and are sufficiently similar in structure that it is probably not possible to use this particular antibody to discriminate between these two isoforms. In contrast, no *GSTA2*-hybridizing material was evident on Northern blots of the various tissues. Taken together with the nucleotide deletion within exon 4 of this gene, this strongly suggests that *GSTA2* is not a functional gene.

The presence of the transposon-like element, *PPTN*, in the plaice GST gene cluster is intriguing and suggests that a transpositional event has occurred at this gene locus, although the presence of a termination codon within the putative coding region for the plaice transposase indicates that the transposon is no longer functional. Similar transposon-like elements, consisting of a transposase gene or pseudogene flanked by inverted repeats, appear to be very widespread, occurring in nematodes, fruit flies and fish species [33]. The spacing between the plaice transposable element and the transcriptional start site of *GSTA* is less than 30 bp, which suggests that the former might have some role in the regulation of this gene. Interestingly, two possible AREs, with the consensus sequence TGACNNNGC [7], reside within the transposon-like sequence, and a further two potential AREs are situated upstream of the transposon sequence (Figure 4). Mouse *gstA1* and quinone reductase gene AREs (or electrophile responsive elements) differ from those in the plaice sequence in that they are composed of two repeated ARE units separated by 6 nucleotides, whereas the plaice elements are separated by 80–90 nucleotides. Nevertheless, treatment of adult male plaice with BNF resulted in a distinct increase in *GSTA* mRNA in intestine and gill, suggesting that this region may be active as a promoter in certain cell types. The presence of CCAAT boxes adjacent to two of the putative *GSTA* AREs may be significant in determining the tissue-specific inducibility of this gene, as these particular enhancer core sequences are frequently found in the promoters of vertebrate genes that have been shown to bind and be transactivated by a wide range of CCAAT binding proteins that are expressed in a highly tissue-specific manner [34].

The 600 bp DNA sequence between *GSTA2* and *GSTA1* contains an imperfect 40 bp repeat sequence, with each repeat containing a number of recognizable promoter elements, most notably several potential AGGTCA half-sites. These half-sites are arranged either as direct repeats separated by a single nucleotide, suggesting the involvement of peroxisome-

```

GCCGGAGGTTTATGTTTTCAGGTTGTGTGTCATTCTTGGAAAATTTCTCAAGAACGCCTAGAAGGAATTTCTCC
AAATTTGGTACAAATGTTAACTTGGATTACGGATTATCTGATAAGAATGTATGTTCAAGGGTCAAGGTCAAGGT
CACCTGTGACCTCACAATTTTTTTTTTGTCTTTGTAATGCAACAAATGTAACGCCTTAAGAGAACACTTTAAAAATG
TTAATTGAGACTCAGGGAATTAATAATTAGGTTTGTAGTGGTCAAGGATCACAGGTCAAGGTCCGGTGGTTTACACA
AAACAATGCTCAACAAGATGCTCAAACTGAAATTTCTCATGTATTGTACCAATGTCTACTTGGACTCAGATATA
ACTGATTATATTTCAGAGGTTAAAGGCCACGGTGACCTCATATGAGTCTGGAAAGATTATGTGCGACATTCAACCAA
TTTGTACAGTTATTCAGTTCAAGTTGTAGTTTCTTCTCTCCGACAGGAGGTGGCAGTGTGCGCTTCAACCAACCCG
TTGGAATCTCATGAATTGAGTGAAGAACACAGCAGATCTCTTCTTCTTGTGCGTACTCGTGCCTCTCTCT
GGAGCAGTGACAGCGAAGCTCCATTTTTCCGGCTGCAGATCTCGAACCATG

```

Figure 7 Sequence of the region upstream of *GSTA1*

The 40 bp imperfect repeats are bold-underlined, putative PPREs are underlined and oestrogen response elements (ERs) are boxed. The first exon of *GSTA1* up to and including the protein initiation codon is italicized.

proliferator-activated receptors (PPARs) and 9-*cis*-retinoic acid receptors [35–37], or as inverted repeats separated by three nucleotides, the structure of the classical oestrogen response element [38] (Figure 7). A number of mammalian genes involved in mitochondrial, peroxisomal and microsomal lipid metabolism are known to be regulated by the PPAR family of transcription factors via PPREs in the promoter regions of these genes. Three PPARs have been identified in *Xenopus*, human and mouse [37,39,40]. Mouse PPAR α and PPAR γ are both activated by peroxisome proliferators, including fibrates drugs and fatty acids, particularly arachidonate analogues and metabolites [36,41,42]. Recently, another arachidonate-derived compound, a prostaglandin J2 metabolite, has been shown to be a specific endogenous activating ligand for human PPAR γ [43,44]. Although there are no reports of the existence of PPARs in fish species, the presence of PPRE-like elements upstream of the plaice GST genes further suggests a role for regulation by polyunsaturated fatty acids. To date there is scant evidence that either fatty acids or hypolipidaemic agents are able to regulate the expression of any mammalian GST isoforms; indeed, it has been noted that administration of peroxisome-proliferating agents decreases the 1-chloro-2,4-dinitrobenzene-conjugating activity of rat liver [45], although these studies do not rule out regulation of GSTs with low activity against 1-chloro-2,4-dinitrobenzene, such as mammalian Theta-class isoforms, by such agents [46]. In spite of this, the fact that a potent peroxisome-proliferating agent, PFOA [47,48], results in induction of both *GSTA* and *GSTA1* mRNAs in plaice liver and gill, and the presence of multiple PPREs upstream of these genes, suggests that they are indeed regulated in a similar manner to that described for mammalian genes containing PPREs. Interestingly, PFOA treatment induced *GSTA1* and *GSTA* mRNA expression, and both of these genes lie downstream of the PPREs, whereas only the *GSTA* gene lies downstream of any putative AREs and, as noted above, only *GSTA* mRNA was induced by BNF. This could indicate that the expression of these genes may be strictly dependent on the spatial distribution of enhancer sequences in their upstream regions.

The presence of oestrogen response elements in the *GSTA1* upstream region is of great interest, as it may indicate a possible role for this enzyme in sexually maturing female plaice during oocyte development. It is possible that the enormous mobilization of lipid during lipovitellin synthesis in females presents a specific requirement for a GST-mediated detoxification mechanism to deal with lipid peroxidation, since up to 50% of total fish egg lipid is composed of polyunsaturated fatty acids originating from the mother [49]. Interestingly, a recent report has shown that oestrogen response elements are capable of binding PPARs and can promote transactivation of gene expression by peroxisome proliferators in transfection assays using artificial gene constructs

[50]. No obvious identifiable regulatory elements were evident in the region upstream of the *GSTA2* gene, and there was no evidence for its expression on Northern blots of RNA from various tissues.

The fact that the *GSTA* and *GSTA1* genes are constitutively expressed in a very wide range of plaice tissues indicates that they may have a 'housekeeping' function associated with the detoxification of endogenously derived toxins. Various mammalian GST isoforms have been shown to be highly active in conjugating toxic lipid alkenals and purine and pyrimidine alkenals, all of which are believed to arise in cells as a consequence of oxidative processes [3,4]. Interestingly, Theta-class GSTs, to which the plaice isoforms are most closely related, from organisms as diverse as mammals and plants have been shown to be efficient in the metabolism of fatty acid hydroperoxides, alk-2-enals and 4-hydroxyalkenals, which predominantly arise as a consequence of polyunsaturated fatty acid peroxidation [51–53]. Therefore, given the fact that marine fish are particularly rich in polyunsaturated fatty acids [54], a possible role for the plaice *GSTA* enzymes might be in the detoxification of potentially deleterious fatty acid metabolites. It will now be possible to test this hypothesis by assaying heterologously expressed *GSTA* and *GSTA1* for activity with a range of substrates.

In summary, we have isolated a cluster of Theta-class-related GST genes from the marine fish, plaice. One of these genes, *GSTA*, is preceded by a sequence containing multiple ARE-like sequences, and *GSTA* mRNA is increased in intestine and spleen after BNF treatment. mRNAs for *GSTA* and *GSTA1* are induced by treatment with PFOA, a peroxisome-proliferating agent, and a region upstream of both of these genes contains multiple PPRE-like sequences. Given that GSTs from a diverse range of organisms are known to conjugate potentially toxic fatty acid metabolites, the results described here are highly suggestive of a link between the regulation of lipid metabolism and the GST-mediated detoxification of oxidized fatty acid metabolites.

REFERENCES

- Mannervik, B. and Danielson, U. H. (1988) *CRC Crit. Rev. Biochem.* **23**, 283–337
- Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M. and Ketterer, B. (1991) *Biochem. J.* **274**, 409–414
- Esterbauer, H., Cheeseman, K. H., Dianzani, M. U., Poli, G. and Slater, T. F. (1982) *Biochem. J.* **208**, 129–140
- Berhane, K., Widersten, W., Engström, Å., Kozarich, J. W. and Mannervik, B. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1480–1484
- Hayes, J. D. and Pulford, D. J. (1995) *Crit. Rev. Biochem. Mol. Biol.* **30**, 445–600
- Rushmore, T. H., Morton, M. R. and Pickett, C. B. (1991) *J. Biol. Chem.* **266**, 11632–11639
- Friling, R. S., Bergelson, S. and Daniel, V. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 668–672
- Prester, T., Holtzclaw, W. D., Zhang, Y. and Talalay, P. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2965–2969
- Wang, B. and Williamson, G. (1994) *Biochim. Biophys. Acta* **1219**, 645–652
- Buetler, T. M. and Eaton, D. L. (1992) *Environ. Carcinog. Ecotoxicol. Rev.* **C10**, 181–203
- George, S. G. (1994) in *Aquatic Toxicology: Molecular, Biochemical and Cellular Perspectives* (Malins, D. C. and Ostrander, G. K., eds.), pp. 37–85, Lewis Publishers, Boca Raton, Ann Arbor, London, Tokyo
- Dominey, R. J., Nimmo, I. A., Cronshaw, A. D. and Hayes, J. D. (1991) *Comp. Biochem. Physiol.* **100B**, 93–98
- Collier, T. K., Connor, S. D., Eberhart, B. L., Anulacion, B. F., Goksøyr, A. and Varanasi, U. (1992) *Mar. Environ. Res.* **34**, 195–199
- Eggens, M., Bergman, A., Vethaak, D., van der Weiden, M., Celander, M. and Boon, J. P. (1995) *Aquat. Toxicol.* **32**, 211–225
- Bucheli, T. D. and Fent, K. (1995) *Crit. Rev. Environ. Sci. Technol.* **25**, 201–268
- Parke, D. V. (1968) *The Biochemistry of Foreign Compounds*, Pergamon Press, Oxford
- Leaver, M. J., Scott, K. and George, S. G. (1993) *Biochem. J.* **292**, 189–195
- Goksøyr, A., Andersson, T., Buhler, D. R., Stegeman, J. J., Williams, D. E. and Förlin, L. (1991) *Fish Physiol. Biochem.* **9**, 1–13
- Olson, C. T. and Anderson, M. E. (1983) *Toxicol. Appl. Pharmacol.* **70**, 362–372
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Pearson, W. R. and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444–2448
- Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159
- Moerman, D. G. and Waterston, R. H. (1989) in *Mobile DNA* (Berg, D. E. and Howe, M. M., eds.), pp. 537–556, American Society for Microbiology, Washington, DC
- Harris, L. J., Prasad, S. S. and Rose, A. M. (1990) *J. Mol. Evol.* **30**, 359–369
- Ogura, K., Nishiyama, T., Hiratsuka, A., Watabe, T. and Watabe, T. (1994) *Biochem. Biophys. Res. Commun.* **205**, 1250–1256
- Yann-Pyng, S. T., Hsieh, T. and Tu, C. P. (1993) *J. Biol. Chem.* **268**, 9737–9746
- Dudler, R., Hertig, C., Rebmann, G., Bull, J. and Mauch, F. (1991) *Mol. Plant-Microbe Interact.* **4**, 14–18
- Itzhaki, H. and Woodson, W. R. (1993) *Plant Mol. Biol.* **22**, 45–58
- Smith, P. J. and Fujio, Y. (1982) *Mar. Biol.* **69**, 7–20
- Brulé, T. (1989) *J. Fish Biol.* **35**, 607–620
- Fan, Z. and Fox, D. P. (1991) *J. Fish Biol.* **38**, 635–640
- George, S. G., Young, P., Leaver, M. J. and Clarke, D. J. (1990) *Comp. Biochem. Physiol.* **96C**, 185–192
- Henikoff, S. (1992) *New Biol.* **4**, 382–388
- Johnson, B. F. and McKnight, S. L. (1989) *Annu. Rev. Biochem.* **58**, 799–840
- Issemann, I. and Green, S. (1990) *Nature (London)* **347**, 645–650
- Keller, H., Dreyer, C., Medin, J., Mahfoudi, A., Ozato, K. and Wahli, W. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2160–2164
- Kliwer, S. A., Umeson, K., Noonan, D. J., Heyman, R. A. and Evans, R. M. (1992) *Nature (London)* **358**, 771–774
- Evans, R. M. (1988) *Science* **240**, 889–895
- Dreyer, C., Krey, G., Keler, H., Givel, F., Helftenbein, G. and Wahli, W. (1992) *Cell* **68**, 879–887
- Wahli, W., Braissant, O. and Desvergne, B. (1995) *Curr. Biol.* **2**, 261–266
- Yu, K., Bayona, W., Kallen, C. B., Harding, H. P., Ravera, C. P., McMahon, G., Brown, M. and Lazar, M. A. (1995) *J. Biol. Chem.* **270**, 23975–23983
- Amri, E.-Z., Bonino, F., Ailhaud, G., Abumrad, N. A. and Grimaldi, P. A. (1995) *J. Biol. Chem.* **270**, 2367–2371
- Kliwer, S. A., Lenhard, J. M., Willson, T. M., Patel, I., Morris, D. C. and Lehmann, J. M. (1995) *Cell* **83**, 813–819
- Forman, B. M., Tontonoz, P., Chen, J., Brun, R. P., Spiegelman, B. M. and Evans, R. M. (1995) *Cell* **83**, 803–812
- Foliot, A., Touchard, D. and Mallet, L. (1985) *Biochem. Pharmacol.* **35**, 1685–1690
- Hiratsuka, A., Sebata, N., Kawashima, K., Okuda, H., Ogura, K., Watabe, T., Satoh, K., Hatayama, I., Tsuchida, S., Ishikawa, T. and Sato, K. (1990) *J. Biol. Chem.* **265**, 11973–11981
- Ikeda, T., Aiba, K., Fukuda, K. and Tanaka, M. (1985) *J. Biochem. (Tokyo)* **98**, 475–482
- Sohlenius, A. K., Ludgren, B. and Depierre, J. W. (1992) *J. Biochem. Toxicol.* **7**, 205–212
- Silversand, C. and Haux, C. (1995) *J. Comp. Physiol.* **164**, 593–599
- Keller, H., Givel, F., Perroud, M. and Wahli, W. (1995) *Mol. Endocrinol.* **9**, 794–804
- Hiratsuka, A., Kanazawa, M., Nishiyama, T., Okuda, H., Ogura, K. and Watabe, T. (1995) *Biochem. Biophys. Res. Commun.* **212**, 743–750
- Bartling, D., Radzio, R., Steiner, U. and Weiler, E. W. (1993) *Eur. J. Biochem.* **216**, 579–586
- Tan, K. L. and Board, P. G. (1996) *Biochem. J.* **315**, 727–732
- Bell, M. V., Henderson, R. J. and Sargent, J. R. (1986) *Comp. Biochem. Physiol.* **83B**, 711–719