Genomic cloning and characterization of the rat glutathione S-transferase-A3-subunit gene

Nasser FOTOUHI-ARDAKANI*† and Gerald BATIST*†[1]

*McGill Center for Translational Research in Cancer, Lady Davis Institute for Medical Research, Sir Mortimer B. Davis–Jewish General Hospital, 3755 Cote-Ste-Catherine, Montreal, Quebec, Canada H3T IE2, and †Division of Experimental Medicine, McGill University, 3655 Drummond St. West, Montreal, Quebec, Canada H3G IY6

The rat glutathione S-transferase-A3-subunit (*GSTA3*) gene is a member of the class Alpha GSTs, which we have previously reported to be overexpressed in anti-cancer-drug-resistant cells. In this study, we report the isolation and characterization of the entire rat *GSTA3* (rGST Yc_1) subunit gene. The rat *GSTA3* subunit gene is approximately 15 kb in length and consists of seven exons interrupted by introns of different lengths. Exon 1, with a length of 219 bp, contains only the 5′-untranslated region of the gene. Each exon–intron splicing junction exhibited the consensus sequence for a mammalian splice site. The transcription start site and exon 1 of rat *GSTA3* were characterized by a combination of primer extension and rapid amplification of the cDNA ends. Position +1 was identified 219 bp upstream of the first exon–intron splicing junction. The proximal promoter region of the rat *GSTA3* subunit gene does not contain typical TATA or CAAT boxes. A computer-based search for potential transcription-factor binding sites revealed the existence of a number of motifs such as anti-oxidant-responsive element, *ras*-response element, activator protein-1, nuclear factor-κB, cAMP-response-element-binding protein, Barbie box and E box. The functional activity of the regulatory region of the rat *GSTA3* subunit gene was shown by its ability to drive the expression of a chloramphenicol acetyltransferase reporter gene in rat mammary carcinoma cells, and its activity was greater in melphalan-resistant cells known to have transcriptional activation of this gene by previous studies. The structure of the gene, with a large intron upstream of the translation-initiation site, may explain why the isolation of this promoter has been so elusive. This information will provide the opportunity to examine the involvement of the rat *GSTA3* subunit gene in drug resistance and carcinogenesis.

Key words: drug resistance, expression, rGSTA3 isolation, structural organization.

INTRODUCTION

The glutathione S-transferases (GSTs) are a family of enzymes that catalyse the conjugation of glutathione (GSH) with xenobiotics as part of detoxification and drug-resistance pathways [1–3]. The GSTs are present in almost all eukaryotic species and, thus far, six classes of this enzyme have been described, Alpha, Mu, Pi, Sigma, Theta and Zeta [3,4]. In rats, each class consists of several subunits with various nomenclatures in the literature [3]. The rat GSTA3 (rGST Yc_1) subunit belongs to the Alpha class, and is located on chromosome 9 [5]. Several reviews have recently described the GSTs in great detail [1–3].

The development of anti-cancer-drug resistance is a major problem in the treatment of cancer by chemotherapy. A number of pathways have been reported to be involved in drug resistance; among them are alterations in drug transport through P-glycoprotein and multidrug-resistance protein [6,7], increased DNA repair [8,9], changes in DNA topoisomerases [10], resistance to apoptosis [11,12], and alteration of GSH and GSH-associated enzymes [2,13,14]. Several studies have demonstrated the involvement of GST isoforms [1,2,15] in chemotherapy resistance to various drugs, particularly to alkylating agents [16–18].

To study the mechanisms of anti-cancer-drug resistance, we have previously reported the isolation of a drug-resistant rat mammary carcinoma cell line (MatB) by exposing cells to increasing concentrations of the alkylating agent melphalan [melphalan-resistant (MLNr) MatB]. *In vitro* and *in vivo* studies in our laboratory demonstrated increased GST activity, especially elevated GSTA3 subunit levels, in MLNr cells [19,20]. The introduction of the rat liver GSTA3 cDNA (clone pGTB42) into wild-type (WT) MatB cells was reported to confer several-fold resistance to melphalan [19]. Nuclear run-on experiments showed that the induced GSTA3 mRNA in MLNr cells is due to transcriptional activation [21]. We have recently isolated the GSTA3 cDNA overexpressed in these cells and confirmed that in fact it is virtually identical to the previously described GSTA3 subunit of GST mRNA [22] which has been designated GST Yc_1 in the past.

Although at least three different laboratories [22–24] have isolated rat (r) GSTA3 cDNAs, the role of this gene in chemotherapy resistance has not been amenable to study since isolating its promoter has been extremely difficult. We here report the first isolation and characterization of the *rGSTA3* subunit gene, including its regulatory regions. This information will permit detailed studies of the mechanism of overexpression of this gene in drug-resistant tumours.

MATERIALS AND METHODS

Isolation of genomic clone

We have isolated three overlapping clones from the λDASH rat liver genomic library (Stratagene), named λBF7, λBF11 and

Abbreviations used: GST, glutathione S-transferase; r, rat; ARE, anti-oxidant-responsive element; Arnt, aryl hydrocarbon-receptor nuclear transporter; WT, wild-type; MLNr, melphalan-resistant; CAT, chloramphenicol acetyltransferase; UTR, untranslated region; RRE, *ras*-response element; AP-1, activator protein-1; NF-κB, nuclear factor-κB; CREB, cAMP-response-element-binding protein; 5′-RACE, rapid amplification of cDNA 5′ ends.

[1] To whom correspondence should be addressed (e-mail gbatist@onc.jgh.mcgill.ca).

The nucleotide sequences reported in this paper have been submitted to the DDBJ/EMBL/GenBank Nucleotide Sequence Databases with accession numbers AF067442 and AF111160.

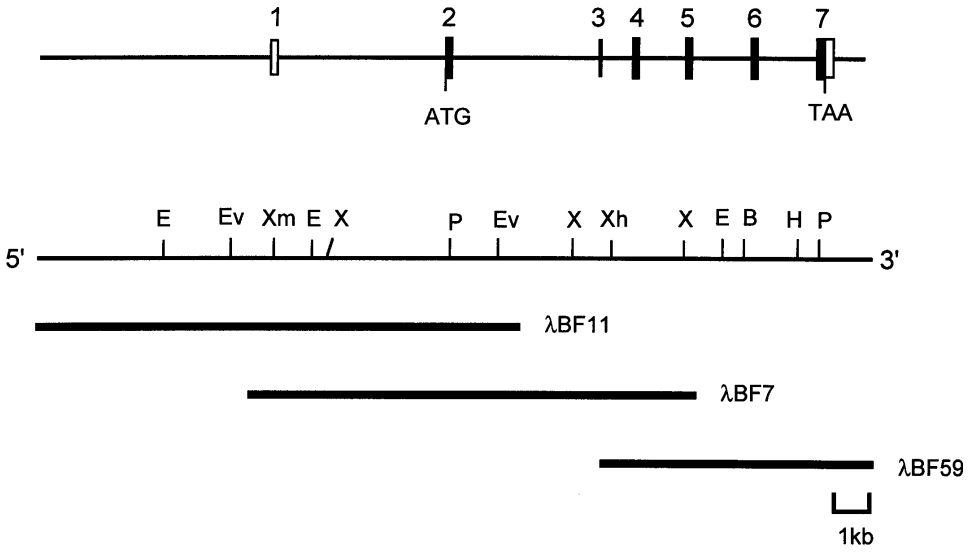

Figure 1 Structure of *rGSTA3* (GST Yc_1) subunit gene and alignment of isolated genomic clones

Three phage clones, λBF7, λBF11 and λBF59, are represented as horizontal lines at the bottom. The seven exons (1–7), six introns and 5′-flanking region of the *rGSTA3* subunit gene are shown to scale at the top. The closed boxes represent coding exons and the open boxes 5′- and 3′- untranslated regions (UTRs). The initiation (ATG) and stop (TAA) codons are indicated. The positions of restriction sites are shown: B, *Bam*HI; E, *Eco*RI; Ev, *Eco*RV; X, *Xba*I; Xh, *Xho*I; P, *Pst*I; H, *Hin*dIII; Xm, *Xmn*I.

Table 1 The oligonucleotide primers used in this study

Ex–Int, primers for exon–intron splicing junction; PE, primer extension; Seq., sequencing of rGSTA3 subclones; PR, primers used to generate cDNA probe for the genomic-library screening; 5UT and 3UT, primers involved in the sequencing of the 5′- and 3′-UTRs; RC; rapid amplification of cDNA 5′ ends (5′-RACE) experiment.

Primer	Sequence	Strand	Assay
Y1–1	5′-GCTGCTCCCACAGTGTCCAGTGGC-3′	Antisense	5UT
Y1–2	5′-CCCCTCCAATGCTGCCAGCAGAAAGC-3′	Antisense	5UT
Y1–3	5′-GGGAGGTGCTATAAAGAGCATACC-3′	Antisense	5UT
Y1–4	5′-GGCTTCGGATGTGCTCTGTGTTCAG-3′	Antisense	PE
Y1–9	5′-GCAGCGGGGACCTTATTGGAC-3′	Sense	PR
Y1–10	5′-GTTAAAAAGCTGCTCCCTCTAAG-3′	Antisense	PR, PE, RC
Y1–91	5′-GCAGCTTTTACCTTATTGGAC-3′	Sense	PR
Y1–11	5′-CTGAGGAAGTGATCATGATTTCAA-3′	Antisense	3UT
Y1–12	5′-TCGAAAGCTTTGCAACAATCGCA-3′	Sense	3UT
Y1–17	5′-CGGTTCCTTGCTTTGTCCTTGATT-3′	Antisense	Ex–Int 4
Y1–23	5′-CCTCTATGGGAAGGACATGAAGG-3′	Sense	Ex–Int 4
Y1–15	5′-CATGGTAGAGAACTTGAACTAGGT-3′	Antisense	Ex–Int 5
Y1–16	5′-CAAGGAACCGTTACTTTCCTGCCTT-3′	Sense	Ex–Int 5
Y1–18	5′-GAAGGAGTGGCGGATCTGGAT-3′	Sense	Ex–Int 5
Y1–19	5′-CCTTCATGTCCTTCCCATAGAGG-3′	Antisense	Ex–Int 3
Y1–13	5′-TGCGATTGTTGCAAAGCTTTCG-3′	Antisense	Ex–Int 6
Y1–14	5′-CTGGACCCCAGCGCTTTGGC-3′	Sense	Ex–Int 6
Y1–22	5′-CATGCCGGGGAAGCCAGTCC-3′	Sense	Ex–Int 2
Y1–24	5′-GCCTGGCCAGGTCATCCCGAG-3′	Antisense	Ex–Int 2
Y1–25	5′-CTCTACTCCAGCTGCAGCC-3′	Antisense	Ex–Int 1
Y1–26	5′-GCAGCTTTTACCTTATTGGAC-3′	Sense	Ex–Int 1
Y1–27	5′-GGCTTCCCCGGCATGGCAGCA-3′	Antisense	Ex–Int 1
T3	5′-ATTAACCCTCACTAAAG-3′		Seq.
T7	5′-TTAATACGACTCACTAT-3′		Seq.

λBF59. The probe for screening the library was initially a PCR-generated fragment of 125 bp located at the 5′ end of rGSTA3 cDNA corresponding to nucleotides −149 to −25 (with the translation start site as +1) [24]. Using this probe, we obtained clones λBF7 and λBF11, which contain the entire 5′-flanking region as well as the first four introns and exons. The rat genomic library was screened following the standard protocols and conditions suggested by the manufacturer (Stratagene). To isolate the remaining 3′ regions of the *rGSTA3* subunit gene, we further screened the library with a probe made from a fragment located at the 3′ end of the λBF7 clone. This resulted in the isolation of the third clone, λBF59 (Figure 1). The clones were subjected to

restriction digestion with *Eco*RI and *Xba*I enzymes and the fragments were subcloned into pBluescriptIIKS plasmid (Stratagene) for analysis. A 4.2 kb *Eco*RI fragment of clone λBF11, which hybridized with the 5′ non-coding cDNA probe, was also subcloned into plasmid pBluescriptIIKS for sequencing (named prA3/4.2E1). This construct comprised about 3 kb of the promoter plus exon 1, as well as 1.2 kb of the first intron. The first intron is located 22 bp upstream of the translation start site.

Southern-blot hybridization of λ phage DNA

Small-scale phage DNAs were extracted from the rGSTA3-positive clones using a Wizard lambda preps kit (Promega). The λDNAs were digested with *Eco*RI and *Xba*I restriction enzymes and Southern-blot hybridization was performed for analysis. In brief, the fractionated λDNA from each clone was separated on a 0.8 % agarose gel, and transferred on to a Hybond-N nylon membrane (Amersham) in 0.4 M NaOH. Hybridization of the membrane with the [α-^{32}P]dCTP-labelled probe was carried out at 42 °C overnight in a solution containing 5 × SSPE (0.9 M NaCl/5 mM NaH_2PO_4/1 mM EDTA), 10 % dextran sulphate, 1 % SDS and 500 μg/ml salmon sperm DNA. The membrane was washed sequentially in 2 × SSC/0.1 % SDS (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), 1 × SSC/0.1 % SDS and 0.5 × SSC/0.1 % SDS at 42 °C for 15 min in each washing solution. The *Eco*RI and *Xba*I fragments were subcloned into the pBluescriptIIKS plasmid for sequencing and mapping of the exon–intron splicing junctions.

DNA sequence analysis

The sequencing of the isolated genomic DNAs of the rat *GSTA3* subunit gene was performed using Sequenase (Amersham) and synthetic oligonucleotide primers (Life Technologies) corresponding to internal sequences (Table 1). Both manual and automated sequencing methods were used to determine the sequence of the 5′-flanking region in both orientations. The BLAST program of the GenBank database [25] was employed to search for sequence identity. For identification of the putative consensus elements and potential transcription-factor-binding sites in the 5′-regulatory region of *rGSTA3*, the TRANSFAC program from the World Wide Web database (http://transfac.gbf.de/TRANSFAC) was used [26].

PCR amplification

For generation of the cDNA probe corresponding to nucleotides –149 to –25 (ATG reference as +1) [24], PCR amplification was performed as described previously [27] using the Y1–9 and Y1–10 primers (Table 1). For determination of the sizes of introns 1 and 2, PCR amplification was performed using the Expand™ Long Template PCR system (Boehringer Mannheim), the λBF7 clone as template and 50 pM of each primer (Life Tchnologies; Table 1). Reaction mixtures were amplified using a programmable thermal cycle (MJ Research, Watertown, MA, U.S.A.). The amplification parameters were: denaturation at 94 °C for 3 min followed by 30 cycles of 94 °C for 30 s, 55 °C for 1 min and 68 °C for 3 min. The amplified products were subjected to agarose-gel electrophoresis and sequencing for verification.

Mapping of the transcription start site

To identify the transcription start site of the *rGSTA3* subunit gene, primer extension was performed following the standard methods [28]. Briefly, primer Y1–10 (Table 1) complementary to the rGSTA3 mRNA was end-labelled with [γ-^{32}P]ATP and then it was annealed to 5 μg of total RNA from MLNr MatB cells at 60 °C. Yeast tRNA was used as control for extension specificity. To prevent premature termination due to the secondary structure of mRNA, reverse transcription was performed using murine leukaemia virus reverse transcriptase (Gibco BRL) at 52 °C for 30 min. The extended product was run on a 6 % sequencing gel along with sequencing reaction of prA3/4.2E1, using Y1–10 as the primer (Figure 2A). This experiment was repeated several times to ensure the reproducibility of the results.

To confirm the transcription start site of the *rGSTA3* subunit gene, the rapid amplification of cDNA 5′ ends (5′-RACE) method was performed following the protocols described by the manufacturer (Clontech). In brief, 10 μg of total RNA from MatB cells was annealed with a gene-specific oligonucleotide Y1–10 corresponding to the 3′ end of the first exon (Figure 2B) and the 5′-end of the cDNA was generated. The synthesized cDNA was ligated to an oligonucleotide adaptor provided. A fraction of this reaction product was used as a template for PCR amplification using the same gene-specific primer (Y1–10) and an anchored primer (primer A). The PCR amplification was carried out as described in the previous section. As a positive control, two gene-specific primers, Y1–91 and Y1–10, were included in PCR amplification. A single oligonucleotide (primer A or Y1–10) was used in PCR reactions as a negative control. The PCR products were separated on a 1 % agarose gel for analysis.

Construction of plasmids

A 4.2 kb *Eco*RI (−2797 to +1413) fragment from clone λBF11 that contained the 5′-regulatory region of *rGSTA3* was subcloned into plasmid pBluescript IIKS for sequencing. We designated this plasmid prA3/4.2E1. The genomic fragment was excised from plasmid prA3/4.2E1 using *Kpn*I and *Xba*I restriction sites and inserted into the pCAT-Basic reporter plasmid (Promega) for analysis of its transcriptional activity in both WT and MLNr MatB cells. The above construct was designated prA3/4.2-CAT (where CAT is chloramphenicol acetyltransferase). A 5′-deletion mutant of the *rGSTA3* regulatory region was constructed by deletion of the 1.7 kb *Eco*RV fragment from plasmid prA3/4.2-CAT. The resulting plasmid was named prA3/2.5-CAT. Similarly, a 3′-deletion mutant of the *rGSTA3* subunit gene was obtained by deletion of the 1.4 kb *Xmn*I/*Sma*I fragment from plasmid prA3/4.2-CAT. This resulted in plasmid prA3/2.7-CAT.

Cell culture and transient transfection

Rat mammary carcinoma cells (WT MatB) and the MLNr subline (MLNr MatB) were grown in minimal essential medium supplemented with 10 % fetal bovine serum (Gibco-BRL) at 37 °C with 5 % CO_2. Transfection of cells with recombinant plasmids was performed using lipofectamine (Gibco-BRL). The plasmids used for transfection were purified using a Qiagen maxi-prep kit (Qiagen). WT and MLNr MatB cells were seeded into 6-well plates at a density of 2.5×10^5 cells per 35-mm well and maintained overnight. The next day the cells were rinsed twice and complete medium was replaced with serum-free minimal essential medium. DNA (3 μg) was incubated with 6 μl of lipofectamine at room temperature for 30 min prior to being added to each respective well. The cells were transferred to a 37 °C incubator for 5 h, and then the complete medium was substituted for 24 h prior to harvest. For control of transfection efficiency, the cells were co-transfected with 1 μg of plasmid containing the β-galactosidase gene (pSV-βgal). The cells were

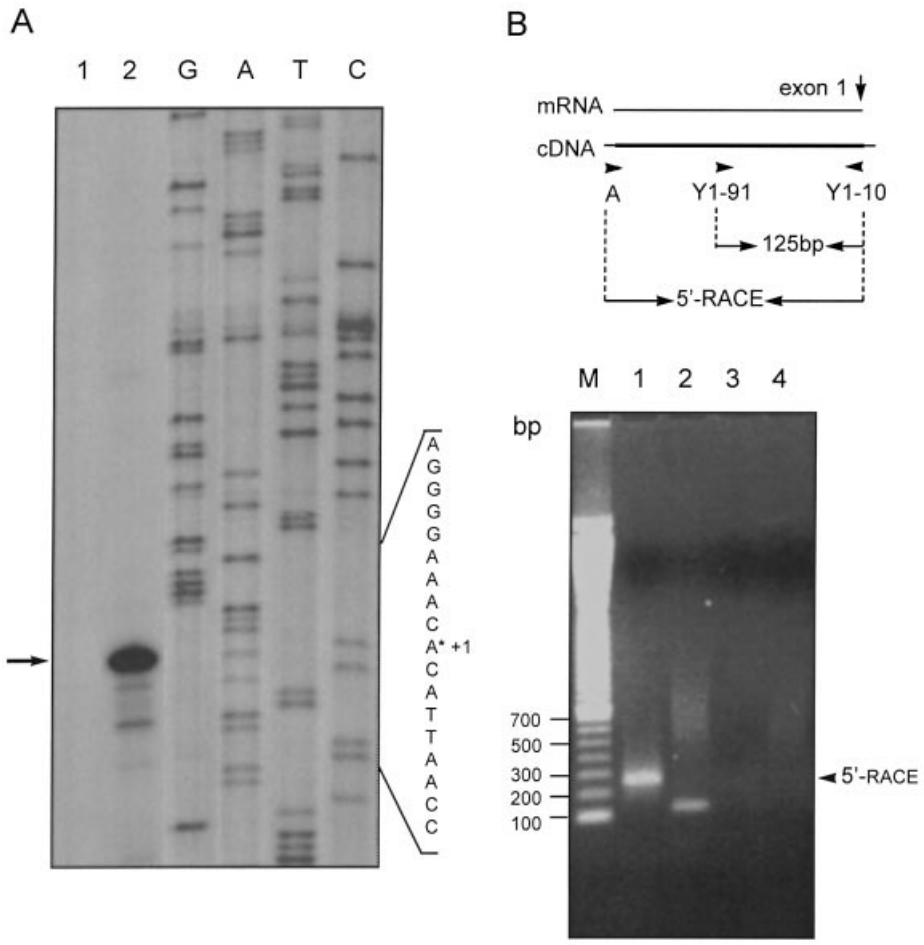

Figure 2 Mapping of exon 1 and determination of transcription-initiation site

(**A**) The transcription start site was mapped by using primer-extension analysis. The primer Y1–10 was end-labelled, annealed with 5 μg of total RNA from MLNr cells, and reverse transcribed. The extended product was run on a 6% sequencing gel. The sequencing reaction of the *rGSTA3* genomic template with Y1–10 primer was used as marker. Lane 1, primer extension with yeast tRNA as control; lane 2, with total RNA from MLNr MatB cells. The position of the transcription start site (+1) is indicated by the arrow. (**B**) Determination of *rGSTA3* transcription-initiation site by the 5′-RACE method. The positions of the three primers and the PCR fragments obtained in this experiment are shown. Primer Y1–10 was used in reverse-transcription reaction and all three primers were included in PCR amplification. M, molecular-size markers in bp; lane 1, 5′-RACE product; lane 2, positive control using two gene-specific primers, Y1–91 and Y1–10; lanes 3 and 4, negative controls using only primer A or Y1–10, respectively.

harvested and prepared for protein assays (Bradford method; [29]), β-galactosidase activity assays and CAT assays [30].

CAT assays

CAT assays were performed as described by Rushmore et al. [30]. In brief, 50 μg of protein was incubated with acetyl-CoA and [^{14}C]chloramphenicol for 4 h at 37 °C. They were then extracted with 1 ml of ethyl acetate and dried by speed vacuum. The chloramphenicol-acetylated products were separated on TLC plates, visualized by exposure to X-ray films, and then quantified by PhosphorImager analysis.

RESULTS AND DISCUSSION

Isolation of *rGSTA3* (rGST-Yc$_1$) genomic clones

A rat liver genomic library (Stratagene) was screened with a PCR fragment located in the 5′ end of rGSTA3 cDNA (previously designated GST Yc$_1$), which resulted in the isolation of two independent genomic clones. Restriction-enzyme mapping and sequencing of these clones revealed that they overlapped, and contained a long 5′-flanking region and the first four introns. To obtain the remaining 3′ end of the gene, further screening of the same library was performed using the most 3′ fragment of clone λBF7 as a probe. Screening approximately 6×10^5 plaques resulted in the isolation of clone λBF59. These three phage clones together contained the entire *rGSTA3* subunit gene (Figure 1).

Structural organization of the *rGSTA3* subunit gene

The structural map of the *rGSTA3* subunit gene was determined by a combination of restriction-enzyme digestion, Southern-blot hybridization, PCR analysis and DNA sequencing. The *rGSTA3* subunit gene is approximately 15 kb in length and consists of seven exons interrupted by six introns of different lengths. The first exon, with the length of 219 bp, contains only 5′-untranslated region (UTR) of the mRNA. The ATG start codon is found 22 bp downstream from the intron 1–exon 2 boundary (Figures 1 and 3); therefore, exons 2–7 provide the coding information for this gene. Interestingly, in rat *GSTA2* and mouse *GSTA2*, intron

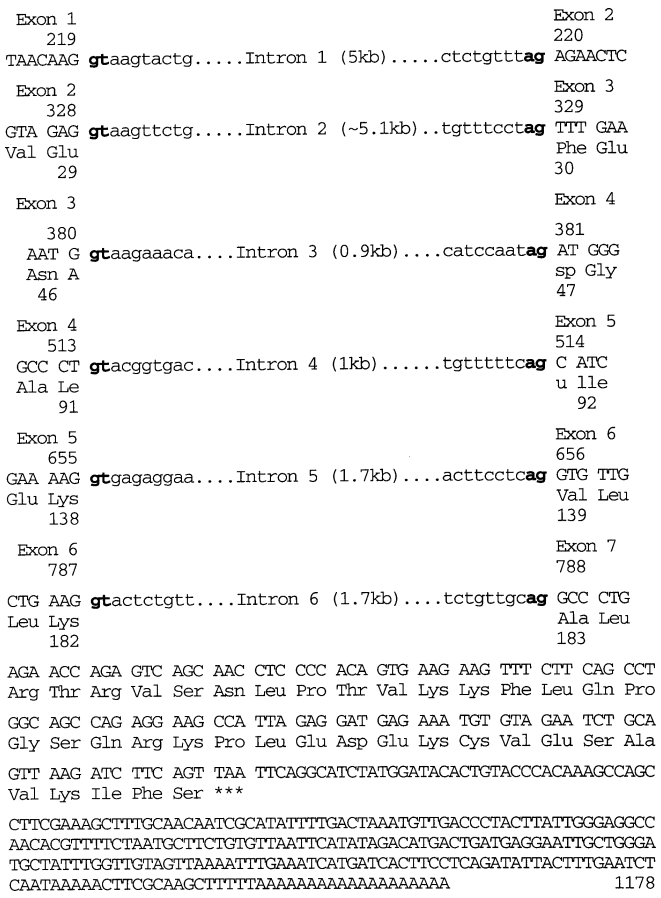

Figure 3 Analysis of exon–intron splice junctions of the *rGSTA3* subunit gene, including the 3′-UTR

Exon sequences, including 3′-UTR (shown in capital letters), and intron sequences (in lower-case letters) were determined by sequence analysis. The deduced amino acid sequence of the rGSTA3 cDNA is displayed below the nucleotide sequence. The numbers above indicate the nucleotide positions of the junctions in reference to the transcription start site, whereas the positions of amino acids (starting at the initiation codon methionine) are shown below. Intron 1 interrupts the 5′-UTR of the gene 22 bp upstream of the ATG start codon. Introns 3 and 4 interrupt codon triplets of the *rGSTA3* subunit gene. Intron sizes were determined by a combination of sequencing (introns 3 and 4), PCR amplification of genomic clones using cDNA primers, and restriction-digest analysis of genomic clones. Full amino acid and nucleotide sequences of exon 7 are shown below. The complete coding sequences of rGSTA3 cDNA have been published previously [23,24].

1 is also located 22 bp upstream from the ATG start codon [31,32]. The class-Alpha *GST* genes characterized thus far, with the exception of *rGSTA5*, have an intron in the 5′-non-coding region [31–34]. The lengths of the coding exons 3–6 of *rGSTA3* are identical to those of other class-Alpha *GSTs* (Table 2). This could have arisen as a result of the ancestral duplication of the GST Alpha gene.

The sequencing analysis of our phage genomic clones (λBF7 and λBF11) revealed that, in both, codon 9 was identical to the rGSTA3 cDNA clone previously reported by Telakowski-Hopkins et al. [23] and was TAC. This contrasts with another rGSTA3 cDNA clone (λJH24) reported by Hayes et al. [24] that contains TAT at codon 9. Moreover, λBF7 and λBF11 differ from λJH24 in four nucleotides located between 120 and 123 bp upstream of the first exon–intron splicing junction (Figure 4). Both of our phage clones contained TTTT, whereas λJH24 has GGGG. These differences might be due to the presence of polymorphisms in different rats.

Introns 1 and 2 of *rGSTA3* were the longest of the six introns present, with estimated lengths of approximately 5 and 5.1 kb, respectively (Figure 1). The sequences of the exon–intron boundaries of this gene were consistent with the consensus splice–site sequences of mammalian genes, containing gt–ag (Figure 3). In general, the organization of the *rGSTA3* subunit gene is similar to that of other mammalian Alpha-class GST genes, with the exception of *rGSTA5*, in terms of numbers of exons and introns (Table 2). However, *rGSTA3*, with a size of about 15 kb, appears to be larger than other members of the class-Alpha GSTs.

Mapping of the transcription start site

The transcription-initiation site of the *rGSTA3* subunit gene was determined by both primer extension and 5′-RACE studies using total RNA from MLNr MatB cells, in which the gene was highly expressed [22]. Primer extension analysis was performed with antisense primer Y1–10 (Table 1). The elongated products were analysed on a 6% sequencing gel (Figure 2A). The end points of the primer extension were determined with the help of a sequencing ladder derived from prA3/4.2E1 as the template and Y1–10 as the primer. The end point of the primer extension, which is the start site for the transcription of the *rGSTA3* subunit gene, was determined to be an adenine (+1), corresponding to 219 nucleotides upstream of the first intron–exon splicing junction (Figure 2A). We obtained the same results when we used the total RNA from rat liver tissue.

The results from primer extension were confirmed by 5′-RACE studies using the gene-specific primer Y1–10 and the anchor primer A (Figure 2B). Thus the combination of results obtained from 5′-RACE and primer-extension analysis were in agreement with each other, and the transcription start site was identified to be located 219 bp upstream of the first exon–intron boundary, and is shown as +1 in Figure 2(A). In this experiment, two gene-specific primers Y1–91 and Y1–10 were used as positive controls and the expected 125 bp fragment was obtained by PCR amplification of the generated cDNA 5′ ends (Figure 2B).

Analysis of the 5′-flanking region of *rGSTA3*

A 4.2 kb *Eco*RI fragment of *rGSTA3* genomic DNA, which spanned sequence −2797 to +1413, was subcloned into pBluescriptIIKS plasmid and sequenced (Figure 4). Analysis of the sequence immediately upstream of the potential transcription start site revealed no canonical TATA or CAAT boxes, characteristic of genes that are widely expressed in different cell types [35]. The computer-based analysis of the 5′-flanking region of the *rGSTA3* subunit gene identified several potential transcription-factor-binding sites. Such motifs included binding sites for the so-called anti-oxidant-responsive elements (ARE), Barbie box, Sp-1 (GC box) and a half-site element for the oestrogen response element. These sequences have been previously found in the promoter of *rGSTA5*, the other member of the same subfamily of the GST Alpha class [33]. In addition, the 5′-upstream region of *rGSTA3* contains consensus sequences for activator protein-1 (AP-1), nuclear factor-κB (NF-κB), E box, cAMP-response-element-binding protein (CREB) and *ras*-response element (RRE).

Recently, it has been demonstrated that in extracts of livers from male rats, the GSTA3 subunit is induced by a number of anti-oxidant and metabolizable xenobiotic agents, such as

Table 2 Comparative analysis of the *rGSTA3* subunit gene with other class-Alpha GST genes

Data for other GST Alpha genes were obtained from previous studies [31–34]. h, Human; m, mouse; NA, not applicable.

		Exon length (bp)						
Exon no.	Mammalian GST Alpha genes...	*rGSTA3*	*rGSTA5*	*rGSTA2*	*mGSTA1*	*hGSTA1*	*hGSTA2*	*hGSTA3*
1		219	315	43	42	36	36	Not known
2		109	51	109	109	117	117	117
3		52	134	52	52	52	52	52
4		133	142	133	133	133	133	133
5		142	132	142	142	143	142	143
6		132	500	132	132	132	132	132
7		391	NA	234	234	123	317	123
		Intron length (kb)						
Intron no.	Mammalian GST Alpha genes...	*rGSTA3*	*rGSTA5*	*rGSTA2*	*mGSTA1*	*hGSTA1*	*hGSTA2*	*hGSTA3*
1		5.00	4.00	2.35	2.00	4.50	5.60	Not known
2		5.10	2.50	3.50	2.50	1.60	1.60	2.50
3		0.90	1.10	0.65	0.80	1.30	1.20	1.30
4		1.00	2.50	2.10	3.70	2.00	2.00	2.40
5		1.70	2.50	1.50	1.90	1.10	1.20	2.10
6		1.70	NA	0.80	0.90	1.00	0.85	1.20

coumarin, ethoxyquin, β-naphthoflavone, *trans*-stilbene oxide, indole-3-carbinol, oltipraz, phenobarbital, benzyl isothiocyanate, butylated hydroxyanisole and diethylmaleate [36]. The putative enhancer elements that are identified in the 5′-flanking region of the *rGSTA3* subunit gene (Figure 4) may be involved in its inducibility by the above agents.

The sequence analysis of the *rGSTA3* promoter revealed the existence of three potential AREs; RTGASNNNGCR [37,38], of which one was located in the 5′-UTR, 30 bp downstream of the transcription start site, and the other two in the distal portion of the 5′-flanking region at −1941 and −2220. Whether one or all of these AREs are involved in the regulation of the *rGSTA3* by the anti-oxidants that induce this gene [36] is a subject of further analysis. The promoter of the human γ-glutamylcysteine synthetase (γ-GCS) heavy-subunit gene contains four AREs, of which only a distal ARE sequence was found to be involved in the constitutive and β-naphthoflavone-induced expression of this gene [38]. The second element was the Barbie box (at −606), reported to be present in the regulatory region of most GSTs with the core sequence of AAAG common in all of them. This element might be responsible for the induction of the GST genes by phenobarbital [33,39]. Seven putative AP-1 consensus sites were found throughout the promoter region of the *rGSTA3* subunit gene, at nucleotides +19, −450, −768, −931, −1924, −1998 and −2644 (Figure 4). The Fos–Jun family of transcription factors and related proteins like Nrl (neural retina leucine zipper) and Maf (musculoaponeurotic fibrosarcoma) have been demonstrated to bind to the AP-1 sites of many promoters, resulting in their activation [40–42]. Another regulatory element in the promoter region was indicated by the presence of two sites for an NF-κB-like element at −714 and −1104. NF-κB activation, in contrast to AP-1 activation [43], has been shown to be inhibited by anti-oxidants [44]. The possibility of interactions between the transcription factors that bind to both AP-1 and NF-κB sites has been previously demonstrated in the human *GSTP1-1* gene [42] and will be explored in the rat *GSTA3* subunit gene. The next potential element was an E-box motif, CACGTG (aryl hydrocarbon-receptor nuclear transporter; Arnt), which is located at nucleotide −1327. It has been shown that this motif is recognized by a number of basic helix–loop–helix transcription factors, such as Max (dimerization partner of Myc onco-protein), USF (upstream stimulatory factor) and Arnt, leading to transcriptional activation of the target genes [45–47]. Two sites for CREB, which *trans*-activates cAMP-responsive genes [48], were also seen in the promoter of the *rGSTA3* subunit gene at −591 and −816. Another element found in the 5′-upstream region of this gene was RRE, which was situated at nucleotide −337 and may function as a silencer element influencing the basal expression of this gene. This element was previously hypothesized to function as a repressor in a Mu-class GST, hGSTYBX [49]. The nucleotide-sequence analysis of the 5′-upstream region of *rGSTA3* also showed the presence of eleven copies of the tetranucleotide GATA within the distal region of the promoter between nucleotides −2226 and −2769. The function of this motif with regard to transcriptional expression of the *rGSTA3* subunit gene is unknown at present. The consensus sequence for hepatocyte nuclear factor-5 (TGTTTGC), which is present in the promoter of the *rGSTA5* and might be a reason for its tissue-specific expression in the rat liver, was interestingly not seen in the 5′-flanking region of the *rGSTA3* we studied here [33].

An interesting motif found within the first intron (between nucleotides +254 and +308) was a 56 bp CA pyrimidine-purine microsatellite repeat (Figure 4). This tandem repeat could contain polymorphism information that would provide an excellent genetic marker for the *rGSTA3* subunit gene, to be used in linkage analysis of physiologically important traits, perhaps relating to detoxification of relevant contaminants.

Functional activity of the *rGSTA3* regulatory region

A 4.2 kb *Eco*RI fragment from the *rGSTA3* subunit gene that included 2.8 kb of the promoter region and 1.2 kb of intron 1

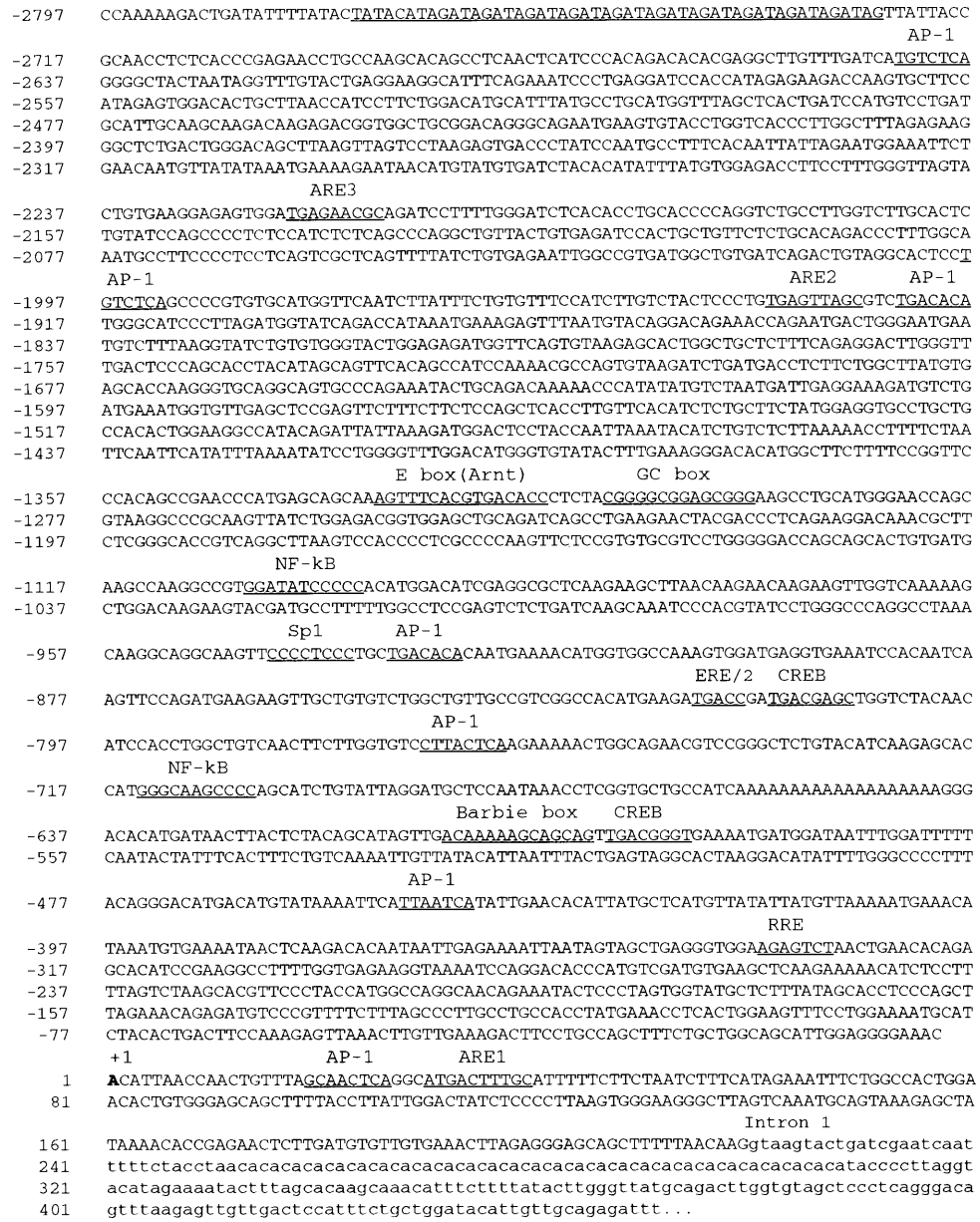

Figure 4 Sequence analysis of the proximal promoter region, the first exon (5′-UTR), and a portion of the first intron of the *rGSTA3* subunit gene

The transcription start site is shown and denoted +1. The putative *cis*-acting elements like anti-oxidant-responsive element (ARE), activator protein-1 (AP-1), Sp-1, nuclear factor-κB (NF-κB), Barbie box, oestrogen response element (ERE; half-site), E box, cAMP-response-element-binding protein (CREB) and *ras*-response element (RRE) are indicated. The nucleotide sequence of the first exon (5′-UTR) and some nucleotide sequence of the first intron are presented. The GATA tetranucleotide repeats located in the distal portion of the promoter is underlined.

was inserted 5′ of the CAT reporter gene (prA3/4.2-CAT). This construct, along with its two deletion mutants, prA3/2.5-CAT and prA3/2.7-CAT (Figure 5A), was transiently transfected into either WT or MLNr MatB cells, which are known to have significantly higher expression of the rGSTA3 mRNA [22]. The transfection efficiency was normalized by co-transfection of a β-galactosidase plasmid driven by the simian virus 40 promoter. The *rGSTA3* construct (prA3/4.2-CAT) exhibited higher promoter activity in MLNr cells compared with WT cells (Figure 5B). These results demonstrate that this region of the *rGSTA3* subunit gene could act as a functional regulatory region that contains the necessary *cis*-acting elements for driving the expression of this gene in MatB cells. At the same time, these results clearly indicate that this region contains enhancer element(s) that appear to be important for its elevated expression in MLNr cells. The identification of *cis*-acting element(s) that are involved in over-expression of the *rGSTA3* subunit gene in MLNr cells is currently under way.

The 5′-deletion mutant of the *rGSTA3* subunit gene (construct prA3/2.5-CAT) revealed significantly higher promoter activity in MLNr cells compared with plasmid prA3/4.2-CAT (Figure 5B). This difference was not significant in WT cells. Thus these data indicate the existence of the regulatory element(s) within the distal region of the *rGSTA3* promoter that affect the expression

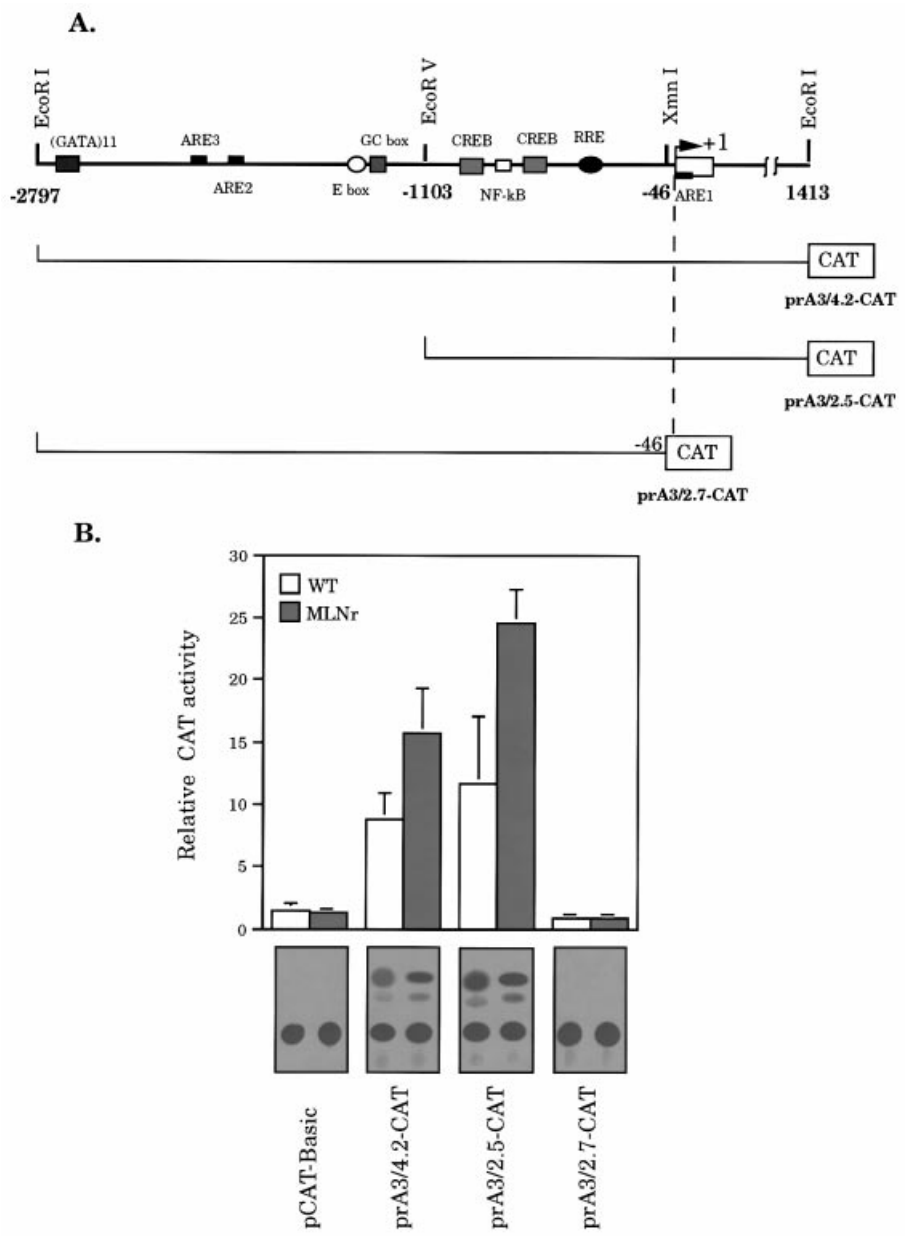

Figure 5 Functional activity of the *rGSTA3* regulatory region

(**A**) Schematic representation of the *rGSTA3* 5′-flanking region/CAT constructs. The restriction enzymes used for generation of the mutant constructs are shown; +1, transcription start site; large open box, mRNA UTR. The putative response elements are indicated: ARE, anti-oxidant-responsive element; RRE, *ras*-response element; NF-κB, nuclear factor-κB; CREB, cAMP-response-element-binding protein. (**B**) Relative CAT activity of the *rGSTA3* upstream sequence. The CAT expression plasmids indicated in (**A**) were used to transfect WT and MLNr MatB cells. The CAT activity was measured as described in Materials and methods. The co-transfected β-galactosidase activity was used to normalize the transfection efficiency. These results are the means of at least three experiments.

of this gene in MLNr cells. Further analysis is needed for identification of the sequences in this region of the promoter that might influence the expression of the *rGSTA3* subunit gene in MLNr cells.

Deletion of the region between nucleotides −46 and +1413 (construct prA3/2.7-CAT), which removes the important region surrounding the transcription start site, resulted in virtually complete elimination of the basal CAT activity (Figure 5B). This is consistent with previous studies demonstrating that a promoter that lacks a TATA box requires an initiator element downstream of the transcription-initiation site for binding of general transcription factors such as TFIID (transcription-initiation factor IID) complex [50].

Conclusions

In this study, we reported the molecular cloning and characterization of the entire *rGSTA3* subunit gene, including its 5′-flanking sequences. We also presented evidence for its enhanced promoter activity in drug-resistant cancer cells. Several potential

binding sites for transcription factors, such as ARE, AP-1, NF-κB, Barbie box, CREB, RRE and E box, were found in the 5′-flanking region of the *rGSTA3* subunit gene. Their possible functions remain to be tested. The availability of the *rGSTA3* genomic clones and the complete sequences of the gene's regulatory regions now make it feasible to initiate detailed investigation of the molecular mechanisms of anti-cancer drug resistance in the GSH- and GST-dependent pathway. We are presently analysing the involvement of *cis*- and *trans*-acting element(s) in transcriptional regulation of the *rGSTA3* subunit gene in anti-cancer-drug-resistant tumour cells. This study also provides new insights into the structural and functional relationships among human and rat GST Alpha-class members.

We thank Dr. Marco Di Fruscro and Dr. Robyn Schecter for critical reading of this manuscript. This work is supported by a grant from the National Cancer Institute of Canada.

REFERENCES

1 Tew, K. D. (1994) Cancer Res. **54**, 4313–4320
2 Shen, H., Kauvar, L. and Tew, K. D. (1997) Oncol. Res. **9**, 295–302
3 Hayes, J. D. and Pulphord, D. J. (1995) Crit. Rev. Biochem. Mol. Biol. **30**, 445–600
4 Board, P. G., Baker, R. T., Chelvanayagam, G. and Jermiin, L. S. (1997) Biochem. J. **328**, 929–935
5 Yamada, T., Muramatsu, Y., Yasue, M., Agui, T., Yamada, J. and Matsumoto, K. (1992) Cytogenet. Cell. Genet. **61**, 125–127
6 Ames, G. F.-L. (1986) Cell **47**, 323–324
7 Fardel, O., Lecureur, V. and Guillouzo, A. (1996) Gen. Pharmacol. **27**, 1283–1291
8 Brown, R., Hirst, G. L., Gallager, W. M., McIlwarth, A. J., Margison, G. P., Van der Zee, A. G. and Anthoney, D. A. (1997) Oncogene **15**, 45–52
9 Barret, J. M. and Hill, B. T. (1998) Anti-cancer Drugs **9**, 105–123
10 Withoff, S., De Jong, S., De Vries, E. G. and Mulder, N. H. (1996) Anti-cancer Res. **16**, 1867–1880
11 Hickman, J. A. (1992) Cancer Metastasis Rev. **11**, 121–139
12 Anthoney, D. A., McIlwarth, A. J., Gallagher, W. M., Edlin, A. R. and Brown, R. (1996) Cancer Res. **56**, 1374–1381
13 Batist, G., Tulpule, A., Sinha, B., Katki, A. G., Meyers, C. E. and Cowan, K. H. (1986) J. Biol. Chem. **261**, 15544–15549
14 Moscow, J. A., Townsend, A. J. and Cowan, K. H. (1989) Mol. Pharmacol. **36**, 22–28
15 Black, S. M. and Wolf, C. R. (1991) Pharmacol. Ther. **51**, 139–154
16 Puchalski, R. B. and Fahl, W. E. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 2443–2447
17 Ban, N., Takahashi, Y., Takayama, T., Katahira, T., Sakamaki, S. and Niitsu, Y. (1996) Cancer Res. **56**, 3577–3582
18 Kuga, T., Sakamaki, S., Matsunaga, T., Hirayama, Y., Kuroda, H., Takashi, Y., Kusakabe, T., Kato, I. and Niitsu, Y. (1997) Hum. Gene Ther. **8**, 1901–1910
19 Schecter, R. L., Woo, A., Duong, M. and Batist, G. (1991) Cancer Res. **51**, 1434–1442
20 Greenbaum, M., Letourneau, S., Assar, H., Schecter, R. L., Batist, G. and Cournoyer, D. (1994) Cancer Res. **54**, 4442–4447
21 Schecter, R.L, Alaoui-Jamali, M. A., Woo, A., Fahl, W. E. and Batist, G. (1993) Cancer Res. **53**, 4900–4906
22 Fotouhi-Ardakani, N., Woo, A., Lewandowska, M., Schester, R. and Batist, G. (1998) J. Biochem. Mol. Toxicol. **12**, 11–17
23 Telakowski-Hopkins, C. A., Rodkey, J. A., Bennett, C. D., Lu, A. Y. and Pickett, C. B. (1985) J. Biol. Chem. **260**, 5820–5825
24 Hayes, J. D., Nguyen, T., Judah, D. J., Petersson, D. G. and Neal, G. E. (1994) J. Biol. Chem. **269**, 20707–20717
25 Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Nucleic Acids Res. **25**, 3389–3402
26 Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L. and Kolchanov, N. A. (1998) Nucleic Acids Res. **26**, 362–367
27 Kirby, G. M., Batist, G., Fotouhi-Ardakani, N., Nakazawa, H., Yamasaki, H., Kew, M., Cameron, R. G. and Alaoui-Jamali, M. A. (1996) Int. J. Cancer **68**, 21–25
28 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: a laboratory Manual, 2nd edn., Cold Spring Harbor Press, Cold Spring Harbor, NY
29 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
30 Rushmore, T. H., Morton, M. R. and Pickett, C. B. (1991) J. Biol. Chem. **266**, 11632–11639
31 Telakowski-Hopkins, C. A., Rothkopf, G. S. and Pickett, C. B. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 9393–9397
32 Daniel, V., Sharon, R., Tichauer, Y. and Sarid, S. (1987) DNA **6**, 317–324
33 Pulphord, D. J. and Hayes, J. D. (1996) Biochem. J. **318**, 75–84
34 Suzuki, T., Johnson, P. N. and Board, P. G. (1993) Genomics **18**, 680–686
35 Melton, D. W., McEwan, C., McKie, A. B. and Reid, A. M. (1986) Cell **44**, 319–328
36 Sherratt, P. J., Manson, M. M., Thomson, A. M., Hissink, E. A. M., Neal, G. E., Van Bladern, P. J., Green, T. and Hayes, J. D. (1998) Biochem J. **335**, 619–630
37 Nguyen, T., Rushmore, T. H. and Pickett, C. B. (1994) J. Biol. Chem. **269**, 13656–13662
38 Mulcahy, R. T., Wartman, M. A., Bailey, H. H. and Gipp, J. J. (1997) J. Biol. Chem. **272**, 7445–7454
39 Liang, Q., He, J.-S. and Fulco, A. J. (1995) J. Biol. Chem. **270**, 4438–4450
40 Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P. and Kerin, M. (1987) Cell **49**, 729–739
41 Kerppola, T. K. and Curran, T. (1994) Oncogene **9**, 675–684
42 Xia, C. L., Hu, J., Ketterer, B. and Taylor, J. B. (1996) Biochem. J. **313**, 155–161
43 Bergelson, S., Pinkus, R. and Daniel, V. (1994) Oncogene **9**, 565–571
44 Schulze-Osthoff, K., Los, M. and Baeuerle, P. A. (1995) Biochem. Pharmacol. **50**, 735–741
45 Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B. and Burley, S. K. (1993) Nature (London) **36**, 38–45
46 Ferre-D'Amare, A. R., Pognonec, P., Roeder, R. G. and Burley, S. K. (1994) EMBO J. **13**, 180–189
47 Antonsson, C., Arulampalam, V., Whitelaw, M. L., Pettersson, S. and Poellinger, L. (1995) J. Biol. Chem. **270**, 13968–13972
48 Yamamoto, K. K., Gonzalez, G. A., Biggs, III, W. H. and Montminy, M. R. (1988) Nature (London) **334**, 494–498
49 Fan, W., Rosario, T., Cooper, T. and Norris, J. S. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 6104–6108
50 Struhl, K. and Moqtaderi, Z. (1998) Cell **94**, 1–4

Received 29 October 1998/7 January 1999; accepted 10 February 1999