Gilbert, L. M., & Gilbert, G. A. (1978) *Methods Enzymol. 48*, 195–212.
Hagel, L., & Andersson, T. (1984) *J. Chromatogr. 285*, 295–306.
Hummel, J. P., & Dreyer, W. J. (1962) *Biochim. Biophys. Acta 63*, 530–532.
Kegeles, G., & Cann, J. R. (1978) *Methods Enzymol. 48*, 248–270.
Kosaka, M., & Solomon, A. (1980) *Am. J. Med. 69*, 145–154.
Maeda, H., Steffen, E., & Engel, J. (1978) *Biophys. Chem. 9*, 57–64.
Natvig, J. B., Gaarder, P. I., & Turner, M. W. (1972) *Clin. Exp. Immunol. 12*, 177–184.
Nichol, L. W., & Ogston, A. G. (1965) *Proc. R. Soc. London, Ser. B 163*, 343–368.
Nisonoff, A., Hopper, J. E., & Spring, S. B. (1975) *The Antibody Molecule*, Academic Press, New York.
Normansell, D. E. (1970) *Immunochemistry 7*, 787–797.
Plaut, A. G., Gilbert, J. V., & Wistar, R. (1977) *Infect. Immun. 17*, 130–135.
Pope, R. M., Teller, D. C., & Mannik, M. (1975) *J. Immunol. 115*, 365–373.
Rialdi, G., Raffanti, S., & Manca, F. (1984) *Mol. Immunol. 21*, 945–948.
Short, M. T., Westholm, F. A., Schiffer, M., & Stevens, F. J. (1985) *J. Chromatogr. 323*, 418–423.
Stevens, F. J., & Schiffer, M. (1981) *Biochem. J. 195*, 213–219.
Stevens, F. J., & Ainsworth, C. F. (1985) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 44*, 1077.
Stevens, F. J., Westholm, F. A., Solomon, A., & Schiffer, M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1144–1148.
Steward, M. W., Turner, M. W., Natvig, J. B., & Gaarder, P. I. (1973) *Clin. Exp. Immunol. 15*, 145–152.
Stone, M. J., & Metzger, H. (1968) *J. Biol. Chem. 243*, 5977–5984.
Sullivan, B., & Riggs, A. (1967) *Biochim. Biophys. Acta 140*, 274–283.
Van Holde, K. E. (1975) *Proteins (3rd Ed.) 1*, 225–291.
Winzor, D. J., & Scheraga, H. A. (1963) *Biochemistry 2*, 1263–1267.
Wittner, M. K., Bach, M. A., & Kohler, H. (1982) *J. Immunol. 128*, 595–599.
Zimmerman, J. K. (1974) *Biochemistry 13*, 384–389.
Zimmerman, J. K., & Ackers, G. K. (1971) *J. Biol. Chem. 246*, 7289–7292.
Zimmerman, J. K., Cox, D. J., & Ackers, G. K. (1971) *J. Biol. Chem. 246*, 4242–4250.

Multiplicity of Glutathione *S*-Transferase Genes in the Rat and Association with a Type 2 Alu Repetitive Element

Gail S. Rothkopf, Claudia A. Telakowski-Hopkins, Ronald L. Stotish, and Cecil B. Pickett*

Department of Molecular Pharmacology and Biochemistry, Merck Sharp & Dohme Research Laboratories, Rahway, New Jersey 07065

Received August 23, 1985

ABSTRACT: Southern blot analysis of rat genomic DNA using glutathione *S*-transferase Ya and Yc cDNA probes was employed to estimate the size of the Ya/Yc multigene family. A minimum of five to seven Ya/Yc genes were detected; at least two of these are Yc genes. The presence of multiple genes was further supported by the isolation of three nonoverlapping genomic clones from a rat *Eco*RI library that hybridized to a Ya cDNA clone, pGTB38. However, not all *Eco*RI bands seen in genomic blots were represented in the clones, suggesting that not all Ya/Yc genes have been isolated. The organization of a Ya gene in one of these *Eco*RI genomic clones, λGTB38-3, and an overlapping clone, λGTB45-1, isolated from a *Hae*III library, was investigated with 5′ and 3′ probes prepared from Ya and Yc cDNA clones. Restriction endonuclease mapping and hybridization studies revealed that the gene spans over 10 kilobases and contains at least three introns. Sequences upstream from the 5′ untranslated region of the gene, and within an intron in the 5′ coding region, were found to contain sequences homologous to a type 2 Alu repetitive element from the rat growth hormone gene [Page, G. S., Smith, S., & Goodman, H. M. (1981) *Nucleic Acids Res. 9*, 2087–2104]. The repetitive sequences in λGTB38-3 were identified by hybridization to a novel Ya cDNA clone, pGTB45. This cDNA clone was isolated from a cDNA library previously described [Telakowski-Hopkins, C. A., Rodkey, J. A., Bennett, C. D., Lu, A. Y. H., & Pickett, C. B. (1985) *J. Biol. Chem. 260*, 5820–5825] with nick-translated intron sequences as probes. pGTB45 is virtually identical with pGTR261 [Tu, C.-P. D., Lai, H.-C. J., Li, N.-Q., Weiss, M. J., & Reddy, C. C. (1984) *J. Biol. Chem. 259*, 9434–9439], except that the 3′ untranslated region extends 231 base pairs beyond the polyadenylation signal of pGTR261. This elongated 3′ untranslated sequence is unique in that it contains a full-length type 2 Alu repetitive element, which includes two additional, overlapping polyadenylation signals.

The glutathione *S*-transferases catalyze the conjugation of the reduced sulfhydryl group of glutathione with electrophilic centers (Jacoby, 1978). Reactive groups of toxic xenobiotic compounds and potentially toxic endogenous metabolic intermediates (Meyer & Ketterer, 1982) are blocked, thus precluding attack on sensitive macromolecules such as protein

* Address correspondence to this author.

0006-2960/86/0425-0993$01.50/0

and nucleic acids. The transferases also serve as binding protins that sequester and facilitate the excretion of toxic hydrophobic molecules through covalent and noncovalent binding (Litwack et al., 1971; Arias et al., 1976; Jacoby, 1978).

The most extensively studied class of these ubiquitous enzymes is that of the rat liver. At least ten glutathione *S*-transferases exhibiting broad, overlapping substrate specificities have been isolated (Jakoby, 1978; Mannervik & Jensson, 1982; Hayes, 1984; Reddy et al., 1984; Sheehan & Mantle, 1984). The enzymes consist of binary combinations of six major subunits, Yα, Ya, Yb$_1$, Yb$_2$, Yc, and Yn, which can be separated by one-dimensional sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) (Bass et al., 1977; Hayes, 1984; Pickett et al., 1984; Reddy et al., 1984; Tu et al., 1984).

Glutathione *S*-transferase B, a heterodimer comprised of two closely related subunits, Ya and Yc, is the most abundant form in rat liver. Peptide mapping experiments and carboxyl-terminal sequence analysis (Beale et al., 1982; Boyer et al., 1983; Ketterer et al., 1983) have indicated that both Ya and Yc subunits are comprised of a mixture of at least two microheterogeneous polypeptides. More recently, application of recombinant DNA techniques toward resolution of microheterogeneous forms has also indicated the existence of at least two highly homologous Ya subunits (Lai et al., 1984; Pickett et al., 1984; Telakowski-Hopkins et al., 1985; Tu et al., 1984). Unfortunately, functional differences between the microheterogeneous subunits have not been elucidated.

Studies demonstrating differential induction of Ya and Yc subunits (Pickett et al., 1981, 1982a,b, 1983) and cDNA sequence comparisons (Telakowski-Hopkins et al., 1985; Tu et al., 1984) rule out posttranslational processing of the product of a single gene as a method of generating multiple subunit forms. Most likely, they are products of multiple closely related genes that evolved through gene duplication and divergence.

In the present study, we present the first characterization of a complex rat glutathione *S*-transferase Ya/Yc multigene family. We also describe a novel cDNA clone, pGTB45, directly related to the Ya clone pGTR261 characterized by Tu et al. (1984). This new clone harbors a complete type 2 Alu repetitive sequence in its 3′ untranslated region. The type 2 Alu sequence provides an additional polyadenylation signal that generates diversity in the Ya mRNA sequences.

Experimental Procedures

Screening of a Rat Genomic Library. One million plaques from a rat *Eco*RI and a *Hae*III genomic library (gift of T. D. Sargent, R. B. Wallace, and J. Bonner) (Sargent et al., 1979) were screened with a modification of the Benton and Davis (1977) procedure allowing in situ amplification of plaques on nitrocellulose filters (Woo, 1979). Both rat liver libraries were constructed by partial digestion of high molecular weight DNA with *Eco*RI or *Hae*III and insertion into λCharon 4A. Plaque hybridization was performed essentially as described by Maniatis et al. (1982). The hybridization buffer contained 50% formamide, 5× Denhardt's solution, 5× SSPE [20× SSPE = 3.6 M NaCl, 200 mM NaH_2PO_4, and 20 mM EDTA (pH 7.4)], 0.1% SDS, and 100 μg/mL denatured salmon sperm DNA. Radiolabeled probe was added at a concentration of 8.3 ng/mL. Hybridization was carried out at 42 °C overnight. Posthybridization washes included three 5-min washes with 2× SSC and 0.1% SDS at room temperature, one wash with 0.1× SSC and 0.1% SDS for 1 h at 68 °C, and a final wash with 0.1× SSC for 1 h at 68 °C. Filters were air-dried and then exposed to Kodak XAR-2 X-ray film with Du Pont Lightning Plus intensifier screens at −70 °C. Positive plaques were picked and replated until 100% of the plaques were positive.

Preparation of Rat Liver DNA. DNA was prepared from pooled livers of starved adult Sprague-Dawley rats as described by Blin and Stafford (1976).

Hybrid Selection Assay. The hybrid select translation procedure for cDNA clones was described in detail in a previous publication (Pickett et al., 1984). The same method was applied to the assay of genomic clone fragments; however, the amount of DNA per assay varied from 6 to 32 μg. Hybrid-selected mRNAs were translated in the rabbit reticulocyte lysate translation system. Translation products were run directly on 10% SDS–polyacrylamide gels or subjected to immunoprecipitation prior to electrophoresis. Radiolabeled polypeptides were visualized by autoradiography.

Preparation of Phage DNA. Recombinant phage were propagated in 1-L cultures with NZYDT medium essentially as described by Blattner et al. (1977) and Maniatis et al. (1982). *Escherichia coli* strain LE392 served as host. Phage were harvested and purified by cesium chloride density gradient centrifugation as described by Yamamoto et al. (1970). DNA was purified by treatment with proteinase K and phenol–chloroform extractions, followed by extensive dialysis against 10 mM Tris (pH 8.0) and 1 mM ethylenediaminetetraacetic acid (EDTA) (Maniatis et al., 1978, 1982). Minipreps of phage DNA from plate lysates were done according to the method of Maniatis et al. (1982).

Southern Blot Analysis. DNA samples for analysis were digested to completion with restriction endonucleases, and digest products were separated on 1% agarose gels in TEA buffer [40 mM tris(hydroxymethyl)aminomethane) (Tris), 20 mM acetic acid, and 2 mM Na_2EDTA, pH 8.1] with 0.5 μg/mL ethidium bromide. Gels were photographed, then denatured, neutralized, and rinsed in 6× SSC. DNA was transferred to nitrocellulose according to Southern (1975). Hybridization of genomic blots was carried out essentially as described by Wahl et al. (1979). The hybridization buffer contained 50% formamide, 10% sodium dextran sulfate, (Pharmacia), 10× Denhardt's solution, 5× SSC, 50 mM Tris (pH 7.5), 10 mM EDTA, 0.5% SDS, and 1 mg/mL denatured salmon sperm DNA. Dextran sulfate was excluded from the hybridization buffer for cloned DNAs. A total of 10^6–10^7 dpm of probe was added to the buffer (50 μL of buffer/cm^2 of filter), and filters were hybridized overnight at 42 °C. Filters were washed twice for 15 min at room temperature in 250-mL volumes of 2× SSC, 1 mM EDTA, and 0.2% SDS and then twice for 15 min at 52 °C in 250-mL volumes of 0.1× SSC, 1 mM EDTA, and 0.2% SDS. Filters were blot-dried, wrapped in Saran Wrap, and exposed to Kodak XAR-2 film at −70 °C with Du Pont Lightning-Plus intensifying screens.

Radiolabeling of Probes. Genomic fragments and cDNA probes were labeled by nick translation with [α-^{32}P]dCTP (3000 Ci/mmol, New England Nuclear) as described by Rigby et al. (1977). Specific activities of probes ranged from 2×10^8 to 2×10^9 cpm/μg.

Preparation and Screening of cDNA Library. A detailed description of the construction of the cDNA library has been published elsewhere (Telakowski-Hopkins et al., 1985). The library was screened as described by Grunstein and Hogness (1975) using ^{32}P-labeled genomic fragments as probes.

Isolation of Rat Liver RNA. Total rat liver RNA was isolated by the guanidine thiocyanate method of Chirgwin et al. (1979). Poly(A+)-RNA was isolated from the total liver RNA preparation by oligo(dT)–cellulose chromatography as

described previously (Pickett & Lu, 1981).

DNA Sequence Analysis. The chemical method of Maxam and Gilbert (1980) was used for DNA sequence analysis. Appropriate restriction fragments were 5′ end-labeled with polynucleotide kinase, or 3′ end-labeled with terminal transferase, and used for DNA sequence analysis.

Subcloning of Genomic Fragments. *Eco*RI fragments of clone λGTB38-3 were mixed with linearized, dephosphorylated pBR325 in a 3:1 molar ratio at a final DNA concentration of 1 μg/30 μL of reaction volume. T4 ligase was added, and the reaction mix was incubated at 15 °C overnight. *E. coli* strain RR_1 cells were transformed with ligated DNAs by the calcium chloride/rubidium chloride procedure (Villa-Komaroff et al., 1978; Kushner, 1978; Norgard et al., 1978).

Recombinant DNA Containment Procedures. All recombinant DNA procedures were carried out in accordance with NIH guidelines under BL-1 containment.

Results

Multiplicity of Glutathione S-Transferase Genes in the Rat Genome. In order to estimate the complexity of the transferase Ya/Yc gene family, Southern blots of genomic DNA were hybridized with probes prepared from four recombinant clones representing three nearly full-length cDNA sequences. pGTB38 (Pickett et al., 1984) and the full-length sequence represented by the two new overlapping clones pGTB45 and pGTB46 (described in detail in a later section) represent two different Ya clones. pGTB46 includes the 5′ untranslated region and complete protein coding region, whereas pGTB45 overlaps the protein coding region and extends through the 3′ untranslated sequence. The fourth clone pGTB42 (Telakowski-Hopkins et al., 1985) contains the entire coding sequence of the Yc subunit and significant amounts of the 5′ and 3′ untranslated regions. The 5′ and 3′ probes were prepared by *Bgl*II/*Cla*I double digests of these cDNAs cloned in plasmid vector pBR322. *Cla*I cleaves once in the vector sequence and *Bgl*II once in each insert sequence, yielding two fragments per recombinant plasmid. The *Bgl*II site occurs in the central protein coding region for Ya clones. In pGTB42, the Yc clone, the *Bgl*II site is found near the junction of protein coding and 3′ untranslated regions.

A comparison of hybridization patterns shown in Figure 1 for all three 5′ probes employed (pGTB38-5′, pGTB46-5′, and pGTB42-5′) reveals a complex pattern in each case, with up to 12 bands of various signal intensities detectable in a single digest. Patterns with Ya probes are virtually identical, a reflection of the 98% sequence homology over 5′ untranslated and protein coding regions (see discussion of pGTB45/46 sequence data). The hybridization pattern of pGTB42-5′, a Yc probe, was very similar to that of Ya probes, a result of the 78% homology between 5′ protein coding regions of pGTB38 and pGTB42 (Telakowski-Hopkins et al., 1985). Although protein coding regions of Ya and Yc clones are highly homologous, there is striking sequence divergence in 3′ untranslated regions (Tu et al., 1984; Telakowski-Hopkins et al., 1985). Consequently, patterns of hybridization with Ya and Yc 3′ probes (Figure 1, pGTB38-3′, pGTB46-3′, and pGTB42-3′) were much less complex. The 3′ probe derived from pGTB38 hybridizes to the greatest number of fragments since this is the only full-length Ya 3′ probe. The complete 3′ seqence for the gene represented by overlapping clones pGTB45 and pGTB46 is present only in the pGTB45-3′ probe. Unfortunately, this probe is not useful since it encodes a repetitive element occurring approximately 100 000 times in the rat genome (see discussion of pGTB45 sequencing data). The result of hybridizations of the pGTB45-3′ probe with

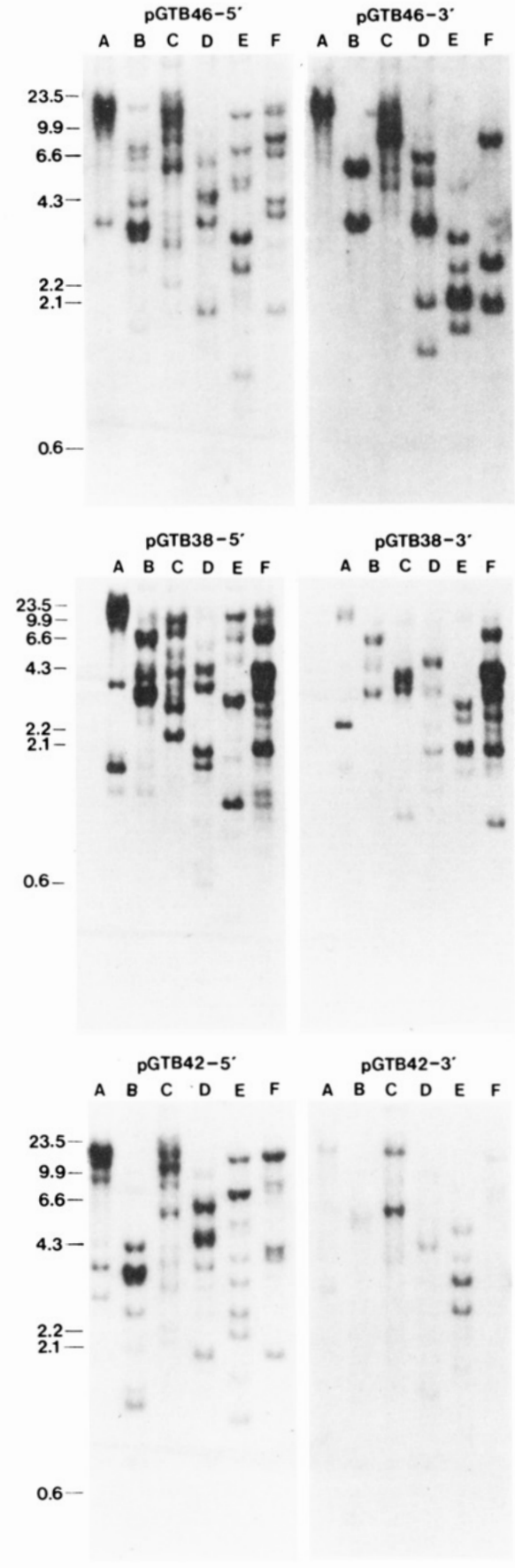

FIGURE 1: Southern blots of rat liver genomic DNA. In each panel, lanes A–F represent complete restriction digests of rat liver DNA: (lane A) *Bgl*I; (lane B) *Bgl*II; (lane C) *Eco*RI; (lane D) *Pst*I; (lane E) *Pvu*II; (lane F) *Sac*I. The cDNA probe is indicated above each panel. DNA molecular weight standards (*Hin*dIII digest of λ phage DNA) are shown in kilobases at the left margins of the panels.

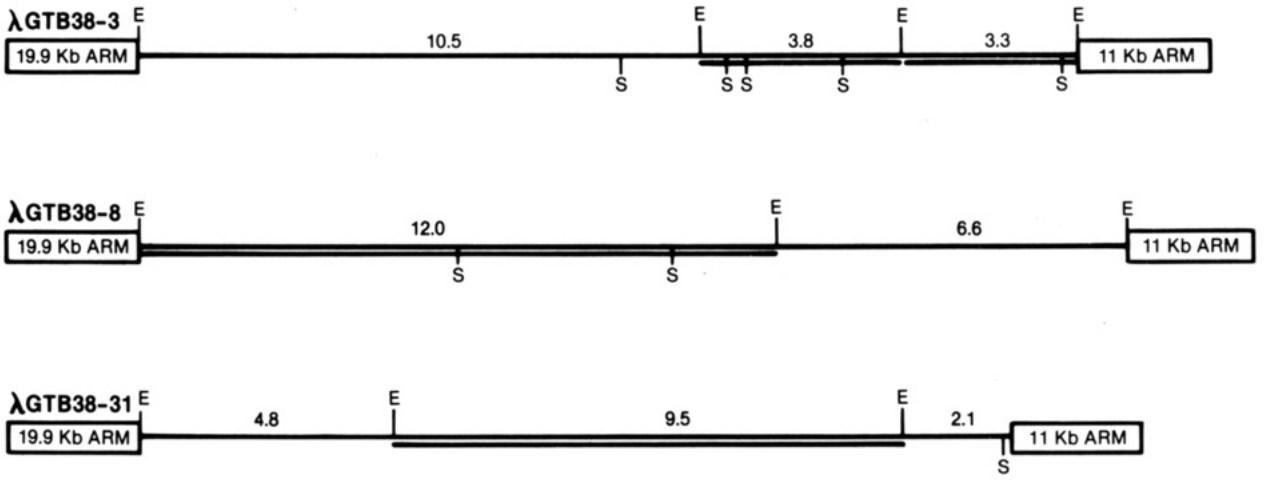

FIGURE 2: Restriction endonuclease maps of genomic clones isolated from the rat *Eco*RI library. Restriction sites are shown for *Eco*RI (E) and *Sma*I (S). Sizes in kilobases are indicated above each *Eco*RI fragment. Vector sequences are represented by boxes. *Eco*RI fragments that hybridize to pGTB38 are underlined.

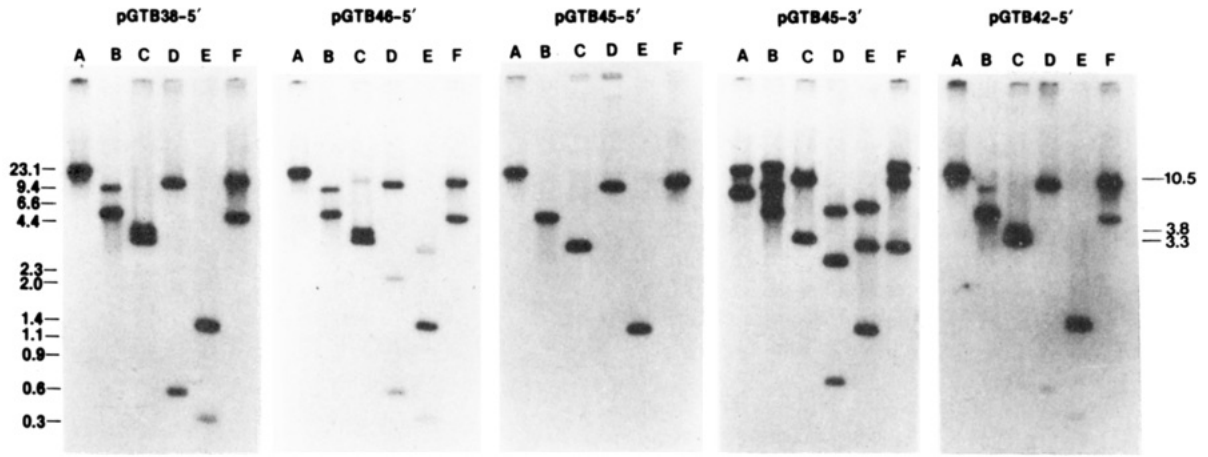

FIGURE 3: Hybridization of λGTB38-3 DNA blots with Ya/Yc cDNA probes. Genomic DNA cloned in λGTB38-3 was digested with six enzymes: *Bgl*I (lane A), *Bgl*II (lane B), *Eco*RI (lane C), *Pst*I (lane D), *Pvu*II (lane E), and *Sac*I (lane F). Southern blots of digests were prepared from agarose gel separations, and blots were hybridized with cDNA probes as indicated above each autoradiogram panel. DNA molecular weight standards (*Hin*dIII digest of λ DNA plus *Hae*III digest of *φX174* DNA) are shown in kilobases to the left of the panels; at the right are size markers in kilobases for the λGTB38-3 insert and *Eco*RI fragments.

genomic DNA was a continuous smear of hybridizing bands in each digest lane (not shown). The hybridization pattern of the shorter pGTB46-3′ probe is a subset of the hybridization pattern of pGTB38-3′ since the sequences of pGTB38 and pGTB46 in homologous regions of these clones (the 3′ protein coding region) are virtually identical.

The Yc 3′ probe prepared from pGTB42 contains almost no protein coding information and includes 134 bp of the 3′ untranslated region. Since there is only 23% homology with the 3′ untranslated region of pGTB38, the pattern of hybridization with pGTB42-3′ is clearly different and less complex than that with pGTB38-3′.

Since it is possible for only one digest fragment per gene to hybridize with both halves of a complementary sequence, the number of fragments hybridizing to both 5′ and 3′ probes in each genomic digest except *Bgl*II was used to estimate the number of genes in the Ya/Yc gene family. The highest estimate of the number of Ya/Yc genes is five to seven, calculated from the *Sac*I digest with probes derived from pGTB38, which encodes a full-length Ya sequence. In the Bg*l*I lane, only two or three large bands were detected with both 5′ and 3′ pGTB38 probes, indicating that at least some of these genes may be linked. The estimate with Yc probe pGTB42 was lower, two to three genes.

Isolation and Characterization of Genomic Clones Homolgous to Ya/Yc cDNA Clones. To isolate genomic sequences coding for glutathione *S*-transferase Ya/Yc subunits for studies of gene structure and to confirm the multiplicity of these genes suggested by results of genomic blots, a rat *Eco*RI genomic library was screened with pGTB38, a cDNA probe homolgous to both Ya and Yc sequences. This cDNA clone is 95% full length and codes for a Ya subunit (Pickett et al., 1984).

One million plaques were screened by plaque hybridization. Of the approximately 40 hybridizing plaques detected, 13 strongly hybridizing clones were isolated for further characterization. Restriction maps were prepared by single and double digests of phage DNA. A comparison of maps revealed three types of independent, nonoverlapping clones, represented by λGTB38-3, λGTB38-8, and λGTB38-31, shown in Figure 2. Since restriction maps are unique, these clones probably represent distinct genes. A blot hybridization of phage *Eco*RI digests confirmed their homology to the cDNA probe.

Three different transferase Ya cDNA clones (pGTB38, pGTB45, and pGTB46) and a transferase Yc cDNA clone (pGTB42) were used to probe λGTB38-3 in an attempt to identify the subunit encoded and the arrangement of protein-coding sequences of this genomic clone. Recombinant

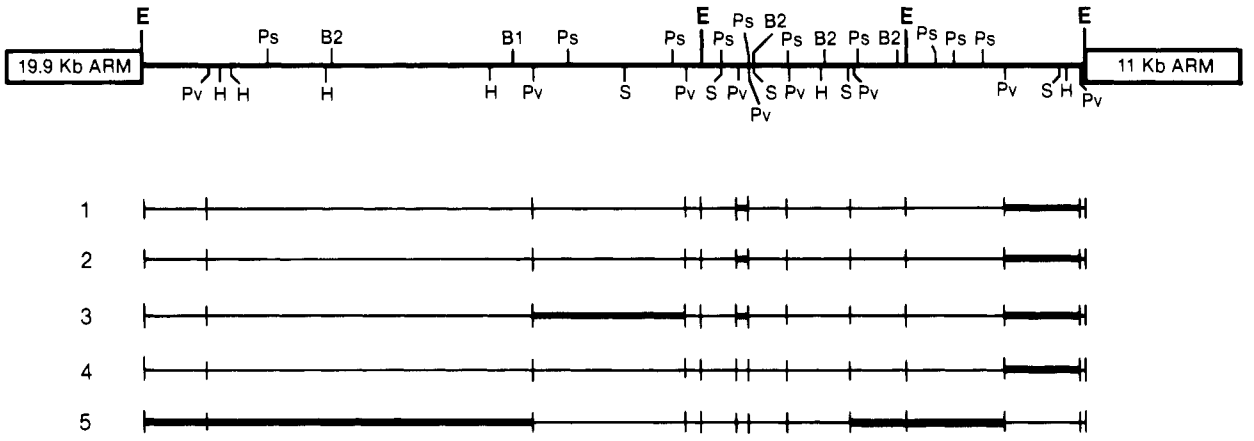

FIGURE 4: Detailed restriction endonuclease map and hybridization map for genomic clone λGTB38-3. Restriction sites in the insert are shown for *Bgl*I (B1), *Bgl*II (B2), *Eco*RI (E), *Hin*dIII (H), *Pst*I (Ps), *Pvu*II (Pv), and *Sma*I (S). Vector sequences are represented by boxes. Below are hybridization maps for *Pvu*II fragments with five different probes, identified by the number to the left of the maps: pGTB38-5′ (1), pGTB42-5′ (2), pGTB46′5′ (3), pGTB45-5′ (4), and pGTB45-3′ (5). *Eco*RI and *Pvu*II sites are indicated by vertical lines aligned with the restriction map. Hybridizing *Pvu*II fragments are indicated by bold lines. Hybridization maps are derived from data shown in Figure 3.

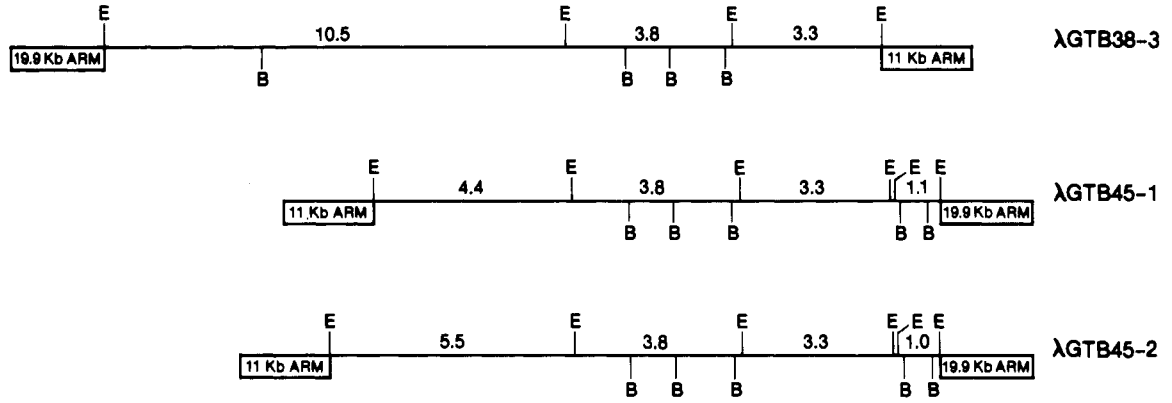

FIGURE 5: Restriction endonuclease maps of genomic clones λGTB45-1 and λGTB45-2 illustrating regions of overlap with λGTB38-3. Restriction sites are shown for *Eco*RI (E) and *Bgl*II (B). *Eco*RI fragment sizes are labeled in kilobases. Vector sequences are represented by boxes.

phage DNA was digested with six restriction enzymes, and a Southern blot of the digests was prepared. The blot was hybridized with 5′ and 3′ probes prepared by *Bgl*II/*Cla*I digests of each of the recombinant plasmids. Autoradiographs are shown in Figure 3. Hybridizing fragments were mapped on the basis of a detailed restriction map of λGTB38-3 prepared by single and double digests of recombinant phage DNA and subcloned *Eco*RI insert fragments. The restriction map for λGTB38-3 and hybridization maps for each of the probes are presented in Figure 4.

Two hybridizing fragments were detected with the Ya probe pGTB38-5′: a 0.3-kb *Pvu*II fragment within the 3.8-kb *Eco*RI fragment and a 1.2-kb *Pvu*II fragment within the 3.3-kb *Eco*RI fragment adjacent to the 11-kb vector arm. A Yc probe, pGTB42-5′, showed an identical hybridization pattern. The hybridization to these two *Pvu*II fragments is due to protein coding region homologies, since 5′ untranslated sequences for these two probes lack sequence homology.

The hybridization of the second Ya 5′ probe pGTB46-5′ to a third *Pvu*II fragment derived from the largest (10.5 kb) *Eco*RI fragment of λGTB38-3 can be attributed to the presence of 5′ untranslated sequences in this region. The pGTB46-5′ probe has a 5′ untranslated sequence 22 bp longer than that of pGTB38-5′. The pGTB46-5′ pattern also includes all other bands positive with pGTB38-5′ and pGTB42-5′ probes since its sequence includes 5′ protein coding sequences, which are largely conserved among Ya/Yc clones. The truncated Ya probe pGTB45-5′, lacking 5′ untranslated and 5′-most protein coding sequences, hybridizes to only one *Pvu*II fragment, which was also positive with the other three 5′ probes.

There was no hybridization of genomic clones λGTB38-8 or λGTB38-31 to any of the 3′ probes. However, λGTB38-3 hybridized with the 3′ probe prepared from pGTB45. Since the pGTB46-3′ probe completely overlaps the protein coding sequences of pGTB45-3′ and there is no hybridization to the pGTB46-3′ probe, the genomic sequences positive with this probe must be homologous only to the 3′ untranslated region of pGTB45. Probe pGTB45-3′ hybridizes with two λGTB38-3 *Pvu*II fragments: one upstream from the 5′ untranslated region identified with the pGTB46-5′ probe within the 10.5-kb *Eco*RI fragment and a second within what is most likely an intron sequence occurring between two 5′ hybridizing regions of this genomic clone detected within the 3.8- and 3.3-kb *Eco*RI fragments. This hybridization is most likely due to the presence of the type II Alu repetitive element in the 3′ untranslated region of pGTB45.

Since no 3′ protein coding region could be detected in λGTB38-3, an overlapping clone was sought. Attempts to isolate an overlapping clone from the rat *Eco*RI library with the pGTB38-3′ sequence or the 3.3-kb *Eco*RI fragment of λGTB38-3 as probe were unsuccessful. Consequently, a probe prepared from the *Bgl*II/*Hin*dIII fragment of pGTB45, a region homologous to the 3′ protein coding region of pGTB38, was used to screen a rat *Hae*III genomic library. Fifteen positive plaques were detected among the one-million plaques screened. Two new genomic clones, λGTB45-1 and -2, were isolated and found to overlap λGTB38-3 on the basis of the

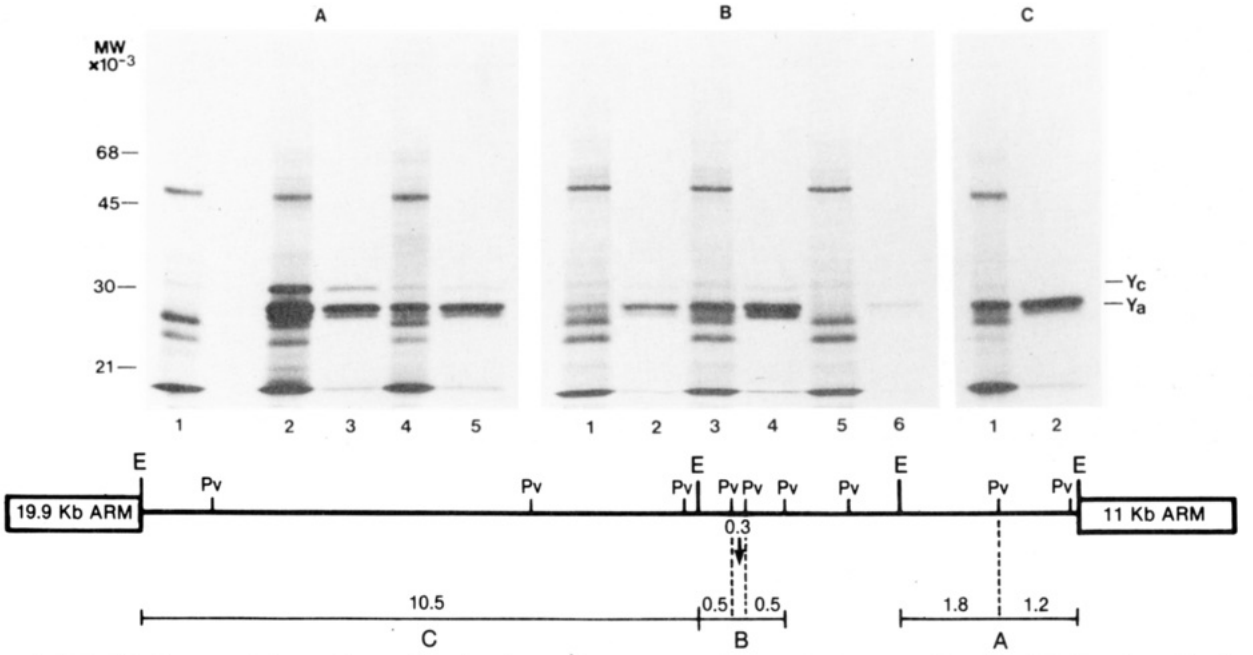

FIGURE 6: Hybrid select translation analyses using cloned genomic sequences. Top panels show one-dimensional SDS–polyacrylamide gel electrophoresis separations of polypeptides selected by genomic fragments cloned in λGTB38-3. Protein molecular weight markers shown to the left of panel A are bovine serum albumin (68 000), ovalbumin (45 000), carbonic anhydrase (30 000), and soybean trypsin inhibitor (21 000). Ya- and Yc-purified subunit standard markers are shown to the right of panel C. (Panel A) (lane 1) Endogenous in vitro translation products directed by the rabbit reticulocyte lysate in the absence of added mRNA; (lane 2) total translation products of mRNAs hybrid selected by the 1.2-kb *Pvu*II fragment of genomic DNA; (lane 3) immunoprecipitation of the translation products directed by mRNAs hybrid selected by the 1.2-kb *Pvu*II fragment of genomic DNA; (lane 4) total translation products of mRNAs hybrid selected by the 1.8-kb *Pvu*II fragment of genomic DNA; (lane 5) immunoprecipitation of the translation products directed by mRNAs hybrid selected by the 1.8-kb *Pvu*II fragment of genomic DNA. (Panel B) (lane 1) Total translation products of mRNAs hybrid selected by the 0.5-kb *Eco*RI/*Pvu*II fragment adjacent to the 10.5-kb *Eco*RI fragment of genomic DNA; (lane 2) immunoprecipitation of the translation products digested by mRNAs hybrid selected by the 0.5-kb *Eco*RI/*Pvu*II fragment adjacent to the 10.5-kb *Eco*RI fragment of genomic DNA; (lane 3) total translation products of mRNAs hybrid selected by the 0.3-kb *Pvu*II fragment of genomic DNA; (lane 4) immunoprecipitation of the translation products directed by mRNAs hybrid selected by the 0.3-kb *Pvu*II fragment of genomic DNA; (lane 5) total translation products of mRNAs hybrid selected by the 0.5-kb *Pvu*II fragment of genomic DNA; (lane 6) immunoprecipitation of the translation products directed by mRNAs hybrid selected by the 0.5-kb *Pvu*II fragments of genomic DNA. (Panel C) (lane 1) Total translation products of mRNAs hybrid selected by the 10.5-kb *Eco*RI fragment of genomic DNA; (lane 2) immunoprecipitation of the translation products directed by mRNAs hybrid selected by the 10.5-kb *Eco*RI fragment of genomic DNA. A restriction endonuclease map of sites for *Eco*RI (E) and *Pvu*II (Pv) in genomic clone λGTB38-3 is presented below. Fragment sizes are indicated in kilobases. Map sections labeled A, B, or C correspond to top panel A, B, or C.

comparison of *Eco*RI/*Bgl*II restriction maps (see Figure 5) and Ya/Yc probe hybridization patterns (not shown). These two new clones are very similar and provide a 1.2-kb extension of the insert of λGTB38-3 in the 3′ direction, on the basis of the arrangement of hybridizing fragments. The *Bgl*II/*Hin*dIII probe hybridized to the 200-bp *Bgl*II/*Eco*RI fragment adjacent to the 19.9-kb vector arms of λGTB45-1 and λGTB45-2. Since the 3′ end of pGTB38 does not hybridize to λGTB45-1 or λGTB45-2, the complete Ya gene may not be represented in these genomic clones.

The sequence encoded by these overlapping genomic clones belongs to a Ya gene since there was hybridization to a probe prepared from the 5′ untranslated region of pGTB46, which shows no homology to the 5′ untranslated region of the Yc clone pGTB42. The gene contains at least three intervening sequences, identified as the absence of hybridization to probes and the presence of sites for restriction enzymes that do not cleave cDNA sequences (*Eco*RI, *Pvu*II, and *Sac*I). One intron resides in the 5′ untranslated region of the gene, a second large (4 kb) intron interrupts the 5′ coding region, and a third intron is located in the central protein coding region that spans λGTB38-3 and λGTB45-1 (see Figure 4). Regions hybridizing to pGTB45-3′ detected upstream from the 5′ untranslated region of the gene and in the 4-kb intron contain a type 2 Alu repetitive sequence (see discussion of pGTB45 sequence data).

The *Eco*RI maps of λGTB45-1 and -2 provide a possible explanation for the failure to isolate a full-length clone from the *Eco*RI library. The central region of the gene cloned in these phages has two closely spaced (approximately 100 bp apart) *Eco*RI sites. These were detected in a blot-mapping experiment using a technique described by Schibler et al. (1982). Multiple-clustered *Eco*RI sites within a gene sequence would decrease the probability of packaging a complete gene during construction of an *Eco*RI library.

Comparison of sizes of cloned genomic fragments to genomic blots confirms that all clones isolated are authentic. However, it is clear that not all transferase genomic sequences have been isolated from the genomic libraries.

Hybrid Selection of Transferase mRNAs Using pGTB38-Negative Genomic Fragments. Preliminary hybridization experiments using pGTB38 as a probe with the *Eco*RI library clones showed no hybridization to the 3′ probe and the 4-kb nonhybridizing region separating 5′ hybridizing sequences of λGTB38-3. These results, along with the knowledge that 5′ sequences of transferase cDNA clones are conserved, while 3′ untranslated sequences have diverged, raised the possibility that *Eco*RI genomic clones may harbor 3′ transferase sequences that are not homologous to pGTB38. In order to detect genomic transferase sequences that are not homologous to this probe, a hybrid select assay was performed with cloned genomic DNA fragments. A map illustrating the relative positions of six genomic fragments used in hybrid select assays and autoradiographs of SDS–PAGE gels of in vitro translation products are presented in Figure 6. The results of this experiment demonstrate that several pGTB38-negative fragments select glutathione *S*-transferase B mRNAs.

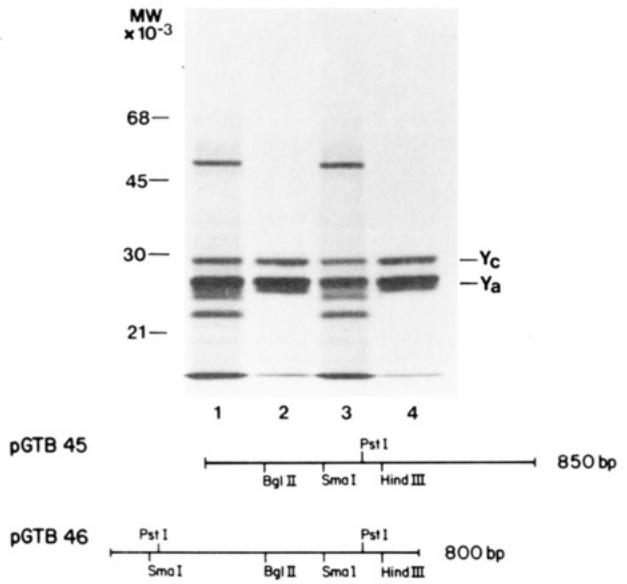

FIGURE 7: Characterization of cDNA clones, pGTB45 and pGTB46. Upper panel shows one-dimensional SDS–polyacrylamide gel electrophoresis separations of polypeptides selected by rat liver cDNA clones pGTB45 and pGTB46: (lane 1) total translation products of mRNAs hybrid selected by clone pGTB45; (lane 2) immunoprecipitation of the translation products directed by mRNAs hybrid selected by pGTB45; (lane 3) total translation products of mRNAs hybrid selected by clone pGTB46; (lane 4) immunoprecipitation of the translation products directed by mRNAs hybrid selected by pGTB46. Molecular size markers are identical with those presented in Figure 6. Shown below are restriction endonuclease maps for inserts of cDNA clones pGTB45 and pGTB46. Size of inserts is indicated to the right of each map.

The 0.3- and 1.2-kb *Pvu*II fragments (fragments labeled B and A, respectively, Figure 6), which hybridized to pGTB38, served as positive controls. These fragments selected Yα, Ya, and Yc mRNAs. The 1.2-kb fragment selected proportionately more Yc mRNA than any other fragment tested. The three fragments adjacent to control fragments, the 1.8-kb fragment and the two 0.5-kb fragments, also selected both Ya and Yα mRNAs and to a much lesser extent Yc mRNA although they did not hybridize with pGTB38. The region between map fragments labeled A and B in the map in Figure 6 selected insignificant amounts of Ya/Yc mRNAs (autoradiograph not shown). The 10.5-kb *Eco*RI fragment (map fragment C, also negative with pGTB38) was highly specific, selecting only the Yα and Ya mRNAs. No Yc band was visible on the autoradiograph even after prolonged exposure.

Isolation and Characterization of cDNA Clones Homologous to pGTB38-Negative Sequences of λGTB38-3. In order to identify and characterize the glutathione *S*-transferase Ya mRNAs hybrid selected by the genomic fragments that did not hybridize to pGTB38, cDNA clones other than pGTB38 were sought. Such clones were identified by screening a rat glutathione *S*-transferase B cDNA library previously described by our laboratory (Telakowski-Hopkins et al., 1985). Three nick-translated genomic fragments that were employed in the genomic hybrid select experiment described above served as probes in colony hybridization experiments. The 1.2-kb *Pvu*II genomic fragment that hybridized to pGTB38 was included in this experiment as a positive control. This fragment, which selected Yα, Ya, and Yc mRNAs, hybridized to 27 of the 64 clones in the cDNA library. Subsets of these 27 clones were positive with the other two genomic probes. The 1.8-kb *Pvu*II/*Eco*RI genomic fragment that was negative with pGTB38 and selected mostly Yα and Ya mRNAs was positive with six cDNA clones. The 10.5-kb *Eco*RI fragment, which hybrid selected only Ya and Yα mRNAs, was the most specific probe. Only one strongly hybridizing clone, pGTB45, was detected in the cDNA library. Clones pGTB45 and -46, positive with all three genomic probes, were chosen for further study.

Results of hybrid select assays, and restriction maps of pGTB45 and pGTB46 are presented in Figure 7. In hybrid select assays, these clones selected the Yα, Ya, and Yc mRNAs. Restriction maps were prepared by a series of single and double digests of the recombinant plasmids. A comparison of the maps of the inserts of these two clones reveals that they are overlapping in a region that corresponds to the central protein coding sequence of the composite full-length sequence. pGTB46 is extended in the 5′ direction, while pGTB45 is extended in the 3′ direction. Although the restriction map for the composite sequence is very similar to that published for Ya clone pGTR261 (Lai et al., 1984), clone pGTB45 has an unusually long 3′ untranslated region. As described earlier, the apparently indiscriminate hybridization of this 3′ probe with genomic DNA in Southern blots suggested the presence of a repetitive element in the 3′ sequence.

pGTB45 and nonoverlapping sequences of pGTB46 have been sequenced by the chemical method of Maxam and Gilbert (1980). The 5′ untranslated region and portions of the protein coding region of pGTB46 were found to be identical with homologous regions in pGTR261 (data not shown). pGTB46 includes the complete protein coding sequence since it includes both initiation (ATG) and termination (TAG) codons in positions expected on the basis of the comparison of this sequence with that of pGTR261 (data not shown).

Figure 8 illustrates the heterogeneity of 3′ sequences of Ya cDNA clones. It compares the 3′ sequence of the cDNA insert of pGTB45 from the *Hin*dIII cleavage site to its 3′ terminus with the 3′ sequences of Ya clones pGTR261 and pGTB38 and with a rat type 2 Alu repetitive sequence. pGTB45 overlaps the 3′ half of the coding region of pGTR261 and the complete 3′ untranslated region of pGTR261 up to the polyadenylation site. Within the 155-bp overlap shown in Figure 8, there are only two base substitutions and one base deletion. The 3′ coding sequence of pGTB45 shares less homology with pGTB38, showing 5 base substitutions in the last 53 bases of the coding sequence included in Figure 8. The degree of homology deteriorates markedly in the 110-bp untranslated region of pGTB38, revealing 59 base substitutions and 1 base deletion. The pGTB45 sequence continues for 231 bp beyond the polyadenylation site in pGTR261. This elongated 3′ untranslated region is unique in that it includes a full-length type 2 Alu repetitive sequence commencing 35 bp downstream from the site corresponding to the polyadenylation site of pGTR261. This type 2 Alu repetitive element shares 90% homology with a repetitive element found in the second intron of the rat growth hormone gene (Page et al., 1981). It has an A-rich region in its 3′ terminal sequence, which includes two overlapping polyadenylation signals.

Discussion

A family of glutathione *S*-transferase Ya/Yc genes has been detected in genomic blots of rat liver DNA. The members of this family share a high degree of homology in the protein coding region while the sequences corresponding to the 3′ untranslated region of the mRNA are more divergent. Although it is difficult to determine the exact number of Ya or Yc genes by Southern blot analysis, the hybridization of genomic fragments to cDNA probes corresponding to the 5′ and 3′ regions of pGTB38 suggests the presence of at least five Ya/Yc genes. Using a Yc-specific 3′ probe, we have deter-

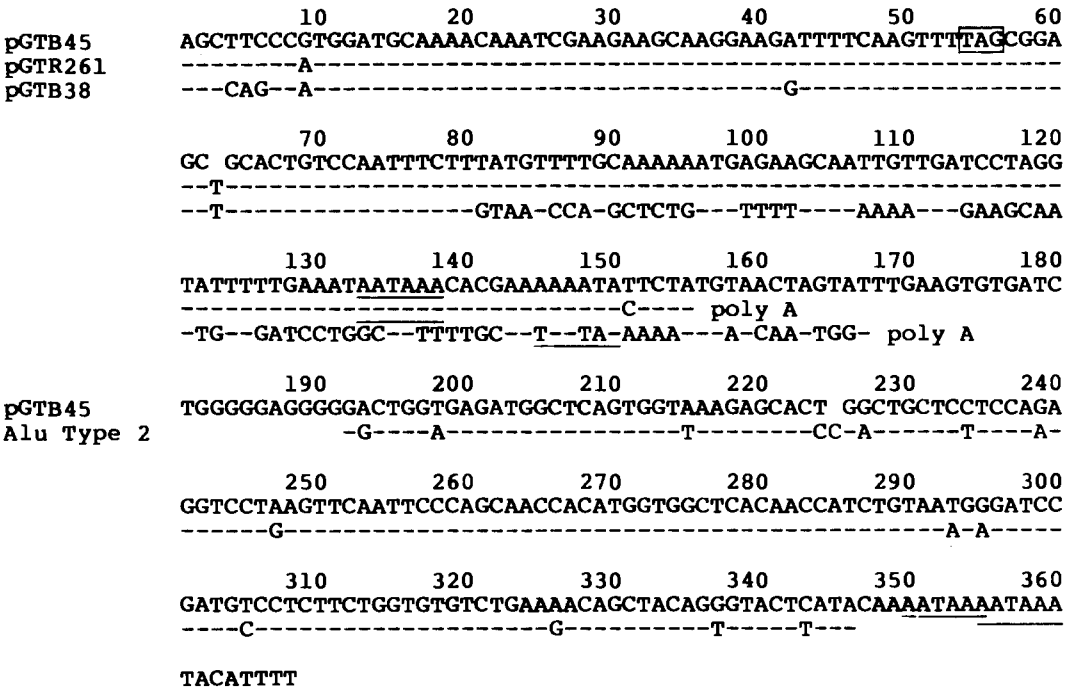

FIGURE 8: Comparison of partial nucleotide sequence of pGTB45 with homologous rat Ya cDNA clones and with a rat type 2 Alu repetitive sequence. The sequence of the pGTB45 insert from the *Hin*dIII cleavage site (position 1) to the 3′ terminus (position 368) is shown in the top line. Positions 1–155 are compared to the pGTR261 homologous region (Tu et al., 1984) (line 2), and positions 1–167 are compared with the pGTB38 homologous region (Pickett et al., 1984) (line 3). Positions 192–308 are shown in best alignment with a rat type 2 Alu sequence (Page et al., 1981). A dash in each sequence indicates base identity; base changes are shown. The pGTB45 termination codon (TAG) is boxed; polyadenylation signals (AATAAA) are underlined. A blank space in the pGTB45 sequence indicates a deleted base.

mined that at least two of these genes are Yc genes. The multiplicity of glutathione *S*-transferase genes was further supported by the isolation of three nonoverlapping genomic clones that hybridized to the Ya/Yc cDNA probes. We believe the complexity of the rat Ya/Yc gene family was underestimated in a recent study by Czosnek et al. (1984), who reported a minimum of two Ya/Yc genes on the basis of hybridization of genomic blots with a truncated (291 bp) Ya probe.

The organization of a nearly full length Ya gene represented by overlapping genomic clones λGTB38-3 and λGTB45-1 was examined. This genomic sequence spans approximately 10 kb and contains most of the structural gene encoding a Ya or Yα subunit. There are at least three intervening sequences in the structural gene. These have been identified by the absence of hybridization to the cDNA clones and the presence of restriction endonuclease sites not found in the cDNA inserts. The largest intervening sequence contains a type 2 Alu repetitive element.

The existence of multiple Ya/Yc genes is not surprising given the heterogeneity in the Ya and Yc subunits. Beale et al. (1982) have found microheterogeneity in the carboxyl terminal sequences of Ya and Yc subunits of glutathione *S*-transferase B. Recently, Sheehan and Mantle (1984) have isolated two Ya homodimers with different cyanogen bromide cleavage maps and activity toward 1-chloro-2,4-dinitrobenzene. The existence of two different Ya subunits is also supported by the recent characterization of two cDNA clones complementary to Ya mRNAs (Lai et al., 1984; Pickett et al., 1984). In the protein coding region of the two cDNA clones there are 15 nucleotide differences that account for 8 amino acid differences.

In the present study a third unique Ya cDNA clone, pGTB45, has been isolated and characterized. Although the protein coding region appears to be identical to that encoded by pGTR261 (Lai et al., 1984), its 3′ untranslated region is extended in length and contains a type 2 Alu repetitive element. It is likely that the presence of the repetitive element is due to alternative processing of the same Ya gene. This conclusion is based on the total sequence homology between pGTR261 and pGTB45 over identical regions of both clones.

An example of the generation of two mRNAs differing only in the presence or absence of a repetitive sequence from a single gene is found in a recent report by Kress et al. (1984). A type 2 Alu repetitive element in the mouse class I histocompatibility gene 3′ noncoding region introduces a novel polyadenylation signal that is functional. Depending upon which of the two functional polyadenylation signals in the gene is used, two different sizes of mRNAs are transcribed. Kress et al. speculate that the repetitive sequence may influence the subcellular compartmentalization of the protein product, since antigen proteins derived from the gene that includes the repetitive element are expressed to a lesser degree on the cell surface than are antigens that are encoded by closely related genes that do not contain repetitive sequences.

Although repetitive elements are commonly detected in association with introns and flanking regions of genes and in primary RNA transcripts, most repetitive sequences are spliced out during processing to the mature mRNA. Ryskov et al. (1984) have studied the organization of mouse B2 repetitive sequences (homologous to rat type 2 Alu sequences) in RNA transcripts. In mature poly(A+)-RNA, the repetitive elements are located exclusively at the 3′ end and occur in a universal orientation, the one in which signal regions are encoded. Type 2 Alu sequences contain an A-rich region at the 3′ end including multiple, overlapping polyadenylation signals (Jelenik & Schmid, 1982). Characteristically, the repetitive element in pGTB45 introduces two additional, overlapping polyadenylation signals (AAUAAA) in the 3′ sequence of the mRNA. The only other reports of specific mRNAs where repetitive elements have been found are the human LDL receptor mRNA, which contains multiple type 1 Alu sequences

(Yamamoto et al., 1984) and as mentioned previously the mouse class I histocompatibility antigen mRNA, which contains a type 2 Alu sequence (Hood et al., 1983; Kress et al., 1984). The 3′ location and orientation of repetitive sequences in these mRNAs also agree with the consensus organization reproted by Ryskov et al. (1984).

Although the function of repetitive elements is unknown, the relative abundance of mRNAs containing repetitive elements has been shown to be related to developmental and tissue-specific expression (Jelenik & Schmid, 1982). It is possible that this regulatory function is not mediated by the repetitive element per se but by the size heterogeneity of the 3′ untranslated regions it generates. Size heterogeneity of 3′ untranslated regions of mRNAs transcribed from the same gene with two polyadenylation signals independent of a repetitive element has also been linked to the subcellular fate of protein products (Parnes et al., 1983; Early et al., 1980).

It is interesting to compare the complexity of the glutathione *S*-transferase gene family to that found for the cytochrome P-450 multigene family. The cytochrome P-450s, like the glutathione *S*-transferases, are a complex group of closely related, multifunctional proteins with broad overlapping substrate specificities. They function in concert with the transferases in metabolism of xenobiotic compounds and are also induced upon administration of phenobarbital (PB) and 3-methylcholanthrene (3-MC) (Lu & West, 1980). In the rat, there are at least six members in a PB-inducible cytochrome P-450 gene family (Atchison & Adesnik, 1983; Mizukami et al., 1983). The pattern of conservation of 5′ sequences with divergence of 3′ untranslated sequences observed within this gene family is also seen within rat (Kawajiri et al., 1984) and mouse (Nakamura et al., 1983) 3-MC-inducible cytochrome P-450 gene families. In view of the parallels in protective function, inducibility, broad substrate specificity, and size and relatedness of gene family members, it is conceivable that these drug-metabolizing enzyme gene families share common regulatory mechanisms and that the families arose via similar evolutionary pathways.

The isolation and characterization of the genomic clones described in this study will allow for a detailed analysis of exon/intron structure as well as possible regulatory sequences in the 5′ flanking region of the glutathione *S*-transferase Ya/Yc structural genes. These studies should eventually aid in understanding the mechanisms by which the rat liver glutathione *S*-transferases are regulated by xenobiotics.

Registry No. Glutathione *S*-transferase, 50812-37-8.

REFERENCES

Arias, I. M., Fleischner, G., Kirsch, R., Mishkin, S., & Gatmaitan, Z. (1976) in *Glutathione: Metabolism and Function* (Arias, I. M., & Jakoby, W. B., Eds.) Raven Press, New York.

Atchison, M., & Adesnik, M. (1983) *J. Biol. Chem. 258*, 11285–11295.

Bass, N. M., Kirsch, R. E., Tuff, S. A., Marks, I., & Saunders, S. J. (1977) *Biochim. Biophys. Acta 492*, 163–175.

Beale, D., Ketterer, B., Carne, T., Meyer, D., & Taylor, J. B. (1982) *Eur. J. Biochem. 126*, 459–463.

Benton, W. D., & Davis, R. W. (1977) *Science (Washington, D.C.) 196*, 180–182.

Blattner, F. R., Williams, B. G., Blechl, A. E., Thompson, K. D., Faber, H. E., Furlong, L. A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L., & Smithies, O. (1977) *Science (Washington, D.C.) 196*, 161–169.

Blin, N., & Stafford, D. W. (1976) *Nucleic Acids Res. 3*, 2303–2308.

Boyer, T. D., Kenney, W. C., & Zakim, D. (1983) *Biochem. Pharmacol. 32*, 1843–1850.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Czosnek, H., Sarid, S., Barker, P. E., Ruddle, F. H., & Daniel, V. (1984) *Nucleic Acids Res. 12*, 4825–4833.

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R., & Hood, L. (1980) *Cell (Cambridge, Mass.) 20*, 313–319.

Grunstein, M., & Hogness, D. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3961–3965.

Hayes, J. D. (1984) *Biochem. J. 224*, 839–852.

Hood, L., Steinmetz, M., & Malissen, B. (1983) *Annu. Rev. Immunol. 1*, 529–568.

Jakoby, W. B. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 46*, 383–414.

Jelinek, W. R., & Schmid, C. W. (1982) *Annu. Rev. Biochem. 51*, 813–844.

Kawajiri, K., Gotoh, O., Sogawa, K., Tagashira, Y., Muramatsu, M., & Fujii-Kuriyama, Y. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1649–1653.

Ketterer, B., Beale, D., Taylor, J. B., & Meyer, D. J. (1983) *Biochem. Soc. Trans. 11*, 466–467.

Kress, M., Barra, Y., Seidman, J. G., Khoury, G., & Jay, G. (1984) *Science (Washington, D.C.) 226*, 974–977.

Kushner, S. R. (1978) in *Genetic Engineering* (Boyer, H. B., & Nicosia, S., Eds.) p 17, Elsevier/North-Holland, Amsterdam.

Lai, H.-C. J., Li, N., Weiss, M. J., Reddy, C. C., & Tu, C.-P. D. (1984) *J. Biol. Chem. 259*, 5536–5542.

Litwack, G., Ketterer, B., & Arias, I. M. (1971) *Nature (London) 234*, 466–467.

Lu, A. Y. H., & West, S. B. (1980) *Pharmacol. Rev. 31*, 277–295.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., & Efstratiadis, A. (1978) *Cell (Cambridge, Mass.) 15*, 687–701.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mannervik, B., & Jensson, H. (1982) *J. Biol. Chem. 257*, 9909–9912.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Meyer, D. J., & Ketterer, B. (1982) *FEBS Lett. 150*, 499–502.

Mizukami, Y., Fujii-Kuriyama, Y., & Muramatsu, M. (1983) *Biochemistry 22*, 1223–1229.

Nakamura, M., Negishi, M., Altieri, M., Chen, Y.-T., Ikeda, T., Tukey, R. H., & Nebert, D. W. (1983) *Eur. J. Biochem. 134*, 19–25.

Norgard, M. V., Keem, K., & Monahan, J. J. (1978) *Gene 3*, 279–292.

Page, G. S., Smith, S., & Goodman, H. M. (1981) *Nucleic Acids Res. 9*, 2087–2104.

Parnes, J. R., Robinson, R. R., & Seidman, J. G. (1983) *Nature (London) 302*, 449–452.

Pickett, C. B., & Lu, A. Y. H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 893–897.

Pickett, C. B., Wells, W., Lu, A. Y. H., & Hales, B. F. (1981) *Biochem. Biophys. Res. Commun. 99*, 1002–1010.

Pickett, C. B., Donohue, A. M., Lu, A. Y. H., & Hales, B. F. (1982a) *Arch. Biochem. Biophys. 215*, 539–543.

Pickett, C. B., Telakowski-Hopkins, C. A., Donohue, A. M., Lu, A. Y. H., & Hales, B. F. (1982b) *Biochem. Biophys. Res. Commun. 104*, 611-619.
Pickett, C. B., Telakowski-Hopkins, C. A., Donohue, A. M., & Lu, A. Y. H. (1983) *Arch. Biochem. Biophys. 221*, 89-98.
Pickett, C. B., Telakowski-Hopkins, C. A., Ding, G. J.-F., Argenbright, L., & Lu, A. Y. H. (1984) *J. Biol. Chem. 259*, 5182-5188.
Reddy, C. C., Li, N., & Tu, C.-P. D. (1984) *Biochem. Biophys. Res. Commun. 121*, 1014-1020.
Rigby, P. W. J., Dieckmann, M., Rhodes, C., & Berg, P. (1977) *J. Mol. Biol. 113*, 237-251.
Ryskov, A. P., Ivanov, P. L., Krzamerov, D. A., & Georgiev, G. P. (1984) *Mol. Biol. 18*, 74-83.
Sargent, T. D., Wu, J.-R., Sala-Trepat, J. M., Wallace, R. B., Reyes, A. A., & & Bonner, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3256-3260.
Schibler, U., Pittet, A.-C., Young, R. A., Hagenbuchle, O., Tosi, M., Gellman, S., & Wellaur, P. K. (1982) *J. Mol. Biol. 155*, 247-266.
Sheehan, D., & Mantle, T. J. (1984) *Biochem. J. 218*, 893-897.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503-517.
Telakowski-Hopkins, C. A., Rodkey, J. A., Bennett, C. D., Lu, A. Y. H., & Pickett, C. B. (1985) *J. Biol. Chem. 260*, 5820-5825.
Tu, C.-P. D., Lai, H.-C. J., Li, N.-Q., Weiss, M. J., & Reddy, C. C. (1984) *J. Biol. Chem. 259*, 9434-9439.
Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L., & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 3727-3731.
Wahl, G. M., Stern, M., & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3683-3687.
Woo, S. L. C. (1979) *Methods Enzymol. 68*, 389-395.
Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L., & Treiber, C. (1970) *Viology 40*, 734-744.
Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., & Russell, D. W. (1984) *Cell (Cambridge, Mass.) 39*, 27-38.

The Antitumor Agent Mitoxantrone Binds Cooperatively to DNA: Evidence for Heterogeneity in DNA Conformation†

Leonard S. Rosenberg, Mark J. Carvlin, and Thomas R. Krugh*
Department of Chemistry, University of Rochester, Rochester, New York 14627
Received August 14, 1985

ABSTRACT: The equilibrium binding of the antitumor compound DHAQ, or mitoxantrone [1,4-dihydroxy-5,8-bis[[2-[(2-hydroxyethyl)amino]ethyl]amino]-9,10-anthracenedione], to various DNAs has been examined by optical titration and equilibrium dialysis methods. At low r (bound drug/DNA base pair) values, $r < 0.03$, DHAQ binds, in a highly cooperative manner, to calf thymus and *Micrococcus lysodeikticus* DNAs. The binding isotherms for the interaction of DHAQ with *Clostridium perfringens* DNA and poly(dA-dT)·poly(dA-dT) exhibit a small positive slope at low r values, suggestive of cooperative binding. In contrast, the binding of DHAQ to poly(dG-dC)·poly(dG-dC) shows no evidence of cooperative binding even at very low r values. At higher r values ($r > 0.05$), the binding of DHAQ to all the DNAs studied is characterized by a neighbor-exclusion process. A model is proposed to account for the two modes of binding exhibited in the cooperative binding isotherms. The main feature of the proposed model is that local sequence and structural heterogeneity of the DNA give rise to sets of binding sites to which DHAQ binds in a highly cooperative manner, while the majority of the DNA sites bind DHAQ via a neighbor-exclusion process. This two-site model reproduces the observed binding isotherms and leads to the conclusion that DHAQ binds in clusters to selected regions of DNA. It is suggested that clustering may play a role in the physiological activity of drugs.

The anticancer compound DHAQ,[1] mitoxantrone [1,4-dihydroxy-5,8-bis[[2-[(2-hydroxyethyl)amino]ethyl]amino]-9,10-anthracenedione], shown in Figure 1, was first synthesized by Zee-Cheng and Cheng (1978) and independently synthesized by Murdock et al. (1979). The drug has proven pharmacological activity in a number of tumor model systems (Wallace et al., 1979; Von Hoff et al., 1981; Drewinko et al., 1983) and has undergone phase I (Von Hoff et al., 1980) and phase II clinical trials (Anderson et al., 1982). Although the reason for the effectiveness of DHAQ is unknown, physicochemical studies have shown that DHAQ is capable of binding to DNA in vitro and of binding preferentially to DNA and RNA in intact cells, of increasing the stability of DNA to thermal denaturation, of unwinding covalently closed circular DNA, and of causing the cooperative condensation of nucleic acids at high concentrations (Kapuscinski et al., 1981; Johnson et al., 1979). Additional evidence from cell culture studies (Nishio & Uyeki, 1983) shows that DHAQ is an effective

† This work was supported by Research Grants CA-17865 and CA-35251 (T.R.K.) from the National Cancer Institute, USDHEW, and by a postdoctoral fellowship (M.J.C.) from the National Institutes of Health.

[1] Abbreviations: DHAQ (mitoxantrone), 1,4-dihydroxy-5,8-bis[[2-[(2-hydroxyethyl)amino]ethyl]amino]-9,10-anthracenedione; TLC, thin-layer chromatography; EDTA, ethylenediaminetetraacetic acid.

0006-2960/86/0425-1002$01.50/0