Morgan, J. P., & Blinks, J. R. (1982) *Can. J. Physiol. Pharmacol. 60*, 524–530.
Neely, J. R., Rovetto, M. J., & Oram, J. F. (1972) *Prog. Cardiovasc. Dis. 15*, 289–329.
Nicholls, D. G., & Ackerman, K. E. O. (1982) *Biochim. Biophys. Acta 683*, 57–88.
Pearce, F. J., Forster, J., DeLeeuw, G., Williamson, J. R., & Tutwiler, G. F. (1979) *J. Mol. Cell. Cardiol. 11*, 893–915.
Peuhkurinen, K. J. (1984) *J. Mol. Cell. Cardiol. 16*, 487–495.
Reers, M., Kelly, R. A., & Smith, T. W. (1989) *Biophys. J. 55*, 32a (Abstract).
Robitaille, P.-M., Lew, B., Merkle, H., Path, G., From, A. H. L., Bache, R., & Ugurbil, K. (1990) *Magn. Reson. Med.* (in press).
Rutter, G. A., & Denton, R. M. (1988) *Biochem. J. 252*, 181–189.
Starnes, J. W., Wilson, D. F., & Erecinska, M. (1985) *Am. J. Physiol. 249*, H799–H806.
Stoner, C. D., & Sirak, H. D. (1979) *J. Bioenerg. Biomembr. 11*, 113–145.
Sundqvist, K. E., Peuhkurinen, K. J., Hiltunen, J. K., & Hassinen, I. E. (1984) *Biochim. Biophys. Acta 801*, 429–436.
Tager, J. M., Wanders, R. J. A., Groen, A. K., Kunz, W., Bohnensack, R., Kuster, U., Letko, G., Bohme, G., Duszynski, J., & Wojtczak, L. (1983) *FEBS Lett. 151*, 271–280.
Ugurbil, K., Petein, M., Maidan, R., Michurski, S., & From, A. H. L. (1986) *Biochemistry 25*, 100–107.
Unitt, J. F., McCormack, J. G., Reid, D., MacLachlan, L. K., & England, P. J. (1989) *Biochem. J. 262*, 293–301.
Vary, T. C., Reibel, D. K., & Neely, J. R. (1981) *Annu. Rev. Physiol. 43*, 419–429.
Williamson, J. R. (1965) *J. Biol. Chem. 240*, 2308–2321.
Williamson, J. R. (1966) *J. Biol. Chem. 241*, 5026–5036.
Zimmer, S. D., From, A. H. L., Foker, J. E., Michurski, S. P., & Ugurbil, K. (1987) *Ann. N.Y. Acad. Sci. 508*, 501–503.
Zimmer, S. D., Ugurbil, K., Michurski, S. P., Mohanakrishnan, P., Ulstad, V. K., Foker, J. E., & From, A. H. L. (1989) *J. Biol. Chem. 264*, 12402–12411.

Structure of the Rabbit Cytochrome P450IIC3 Gene, a Constitutive Member of the P450IIC Subfamily[†,‡]

Ging Chan[§] and Byron Kemper*
Department of Physiology and Biophysics, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801
Received September 29, 1989; Revised Manuscript Received December 19, 1989

ABSTRACT: Fragments of rabbit DNA have been cloned which encompass the gene for rabbit cytochrome P450IIC3. Cytochrome P450IIC3 is a constitutive member of the cytochrome P450IIC subfamily which contains both constitutive and phenobarbital-responsive genes. The cytochrome P450IIC3 gene spans at least 25 kbp and contains 8 introns which have 5′ GT's and 3′ AG's and are located in the same positions as in other family II genes. The 5′ flanking region contains a consensus TATA site about 25 bp from the RNA initiation site which was mapped by the primer extension method. Other potential regulatory sequences include a CCAAT sequence and sequences similar to binding sites for the liver-specific factor HNF-1 and the general transcription factors AP-1 and OCT. The gene sequence differs from the cytochrome P450IIC3 cloned cDNA sequence in only 2 of 1400 nucleotides, 1 of which results in an amino acid change. Since sequences coding for exon 1 and part of exon 2 were not present in the cloned cytochrome P450IIC3 cDNA, a 475-nucleotide fragment of the 5′ end of the cDNA was amplified by the polymerase chain reaction, cloned, and sequenced. The sequence of this cDNA was identical with the gene exon sequence. The protein sequence derived from the gene differs in nine positions from that determined directly for the protein. A comparison of the amino acid sequence by exons with other cytochrome P450IIC proteins revealed that, in general, the similarity was greater in C-terminal exons than in N-terminal ones. This trend is exaggerated in the comparison with P450IIC4/5, for which exon 8 had a high similarity of 87% compared to 65% overall and 58% for exon 6. However, divergence of nucleotides that do not result in amino acid changes was similar in exon 5 and exon 8, indicating that gene conversion was not responsible for the difference in similarity between these two exons and suggesting that conservation of protein sequence in this region may be important.

Cytochrome P450IIC3 is a constitutive form of cytochrome P450 (P450)[1] in the rabbit liver that is not detectably induced by phenobarbital (Johnson, 1980; Koop & Coon, 1979) and metabolizes a variety of substrates, including polycyclic hydrocarbons, alcohols, macrolide antibiotics, and steroids. The enzyme is present in two catalytic forms in liver, one of which efficiently catalyzes both 6β- and 16α-hydroxylation of progesterone while the other is less efficient and catalyzes only 16α-hydroxylation of progesterone (Dieter & Johnson, 1982). The low-activity form is observed in some outbred animals and in some inbred strains, notably IIIVO/J and B/J. The low 16α-hydroxylase activity is stimulated by 5β-pregnane-

[†] This work was supported by grants from the Air Force Office of Scientific Research (AFOSR-84-103) and by the National Institutes of Health (GM 39360).
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02901.
* Correspondence should be addressed to this author at the Department of Physiology and Biophysics, 524 Burrill Hall, 407 S. Goodwin Ave., Urbana, IL 61801.
[§] Present address: Lilly Research Labs, Department of Molecular Biology, Building 98C/3, Indianapolis, IN 46285.

[1] Abbreviations: P450, cytochrome P450; kbp, kilobase pair(s).

0006-2960/90/0429-3743$02.50/0

3β,20α-diol, and the stimulated activity as well as the high 16α-hydroxylase activity is inhibited by antibodies to P450IIC3 (Schwab & Johnson, 1985; Reubi et al., 1984; Johnson et al., 1983). Microsomes from an inbred strain, III/C, with high 16α-hydroxylase activity respond to allosteric effectors as if they contain two forms of P450 with either high or low 16α-hydroxylase activity. The low-activity form is kinetically similar to that in strain B/J. Since the inbred strains are presumably homozygous, Schwab and Johnson (1985) have proposed that two distinct genetic loci code for P450IIC3 proteins.

The complete amino acid sequence of P450IIC3 (P450 3b) from outbred animals has been determined (Ozols et al., 1985), and a partial cDNA has been cloned and sequenced which encodes the C-terminal 400 of the 489 amino acids in the protein (Leighton et al., 1984). The sequence derived from the cDNA differs in six positions from the protein sequence. P450IIC3 cDNA is about 70% similar in DNA sequence to the other rabbit members of the P450IIC gene subfamily and about 50% similar to the P450IIB subfamily genes. The members of the P450IIB subfamily and some of the P450IIC genes, P450IIC1, P450IIC2, and P450IIC4, are induced by phenobarbital (Leighton & Kemper, 1984; Zhao et al., 1987). It is likely, therefore, that the ancestral gene for these two subfamilies was inducible by phenobarbital-like compounds. Constitutive genes, like P450IIC3, presumably have lost the regulatory sequences responsive to phenobarbital. To begin to define the differences between the constitutive and phenobarbital-inducible genes, we have been isolating and characterizing genes in the rabbit P450IIC gene subfamily. We now report the structure for the constitutive P450IIC3 gene.

Materials and Methods

Materials. Oligonucleotides were synthesized by the University of Illinois Biotechnology Center. Double- and single-stranded DNA sequencing kits were obtained from Promega Biotech and United States Biochemical Corp.

Bacterial Strains and Propagation of Virus. λ phage Charon 4A and EMBL3 were propagated in *Escherichia coli* LE 392 (Borck et al., 1976; Murray et al., 1977). Large-scale isolation of λ phage was according to Silhavy et al. (1984). Plasmid DNA was grown in *E. coli* NM522 (Gough & Murray, 1983). Single-stranded DNA of phagemids was prepared by using phage IR1 (Enea & Zinder, 1982) or MK103 as helper phages as described by Mead et al. (1985). Plasmid DNA was prepared by the modified boiling method (Holmes & Quigley, 1981). For double-stranded DNA sequencing, DNA prepared by the boiling method was dissolved in 100 μL of 91% NaI and purified with GeneClean (Bio 101 Co.) as recommended by the manufacturer.

Screening of Rabbit Liver Genomic Libraries. A total of 2 million plaque-forming units of rabbit liver genomic libraries in λ phage Charon 4A (Maniatis et al., 1978) and EMBL 3 (Clontech, Inc.) were screened in triplicate, by in situ plaque filter hybridization (Maniatis et al., 1978). P450IIC1, P450IIC2, and P450IIC3 cDNAs were used as a mixed probe to screen the Charon 4A library. Two of the triplicate filters were probed by the mixed probe, and the third triplicate filter was probed with P450IIC3 cDNA. DNA fragments containing exons 2–6 (λC3.2) and exons 7–9 (λC3.1) were isolated from the Charon 4A library. For selection of the 5′ region of the gene, the EMBL3 library was screened with DNA fragments with intron 1 or exon 3–5 sequences and an oligonucleotide specific for exon 2. After development of the X-ray films, all the Hybond filters (Amersham) were stripped of radioactivity and screened with a P450IIC3 gene fragment containing the intron 6 sequence or P450IIC3 cDNA fragments containing the sequence from exons 2–3, exons 2–5, or exons 6–9. The hybridization and washing conditions were according to the manufacturer's specification. Clones hybridizing with intron 1 and exon 2 containing probes but not with those containing exon 3–8 or 6–9 sequences were selected for further analysis, and clone λC3.3 containing exons 1 and 2 was isolated.

Subcloning and Sequencing of P450IIC3 Gene Fragments. For clone λC3.1, 100 μg of phage DNA was sonicated, and DNA fragments from 400 to 600 bp were isolated after electrophoresis in a 1% agarose gel. The fragments were cloned into the *Hin*cII site of M13mp11, and about 1600 white plaques were picked into 16 Corning 96-well microtiter plates. Clone λC3.2 DNA was digested with *Eco*RI and subcloned into the *Eco*RI site of pTZ18R (Mead et al., 1986), and white colonies were transferred to 96-well microtiter plates. An *Hin*dIII/*Eco*RI fragment containing exon 1 and the 5′ flanking region of clone λC3.3 was subcloned into pTZ18R, and clones with sequential deletions produced by DNase I treatment (Lin et al., 1985) were isolated. Subclones for each of the λ phage that contained exon sequence were identified by colony filter hybridization with P450IIC3 cDNA probes. The DNA sequence was determined by the chain termination method (Sanger et al., 1977).

Southern Blot Analysis. Rabbit liver DNA was digested with one or two restriction enzymes. For double digestions, DNA was extracted with phenol/chloroform before the second digestion. Twenty-five micrograms of DNA was loaded into each lane of a 0.5% agarose gel, and the DNA fragments were separated by electrophoresis at 50 mA for 18 h. DNA was transferred to Hybond filters and hybridized with a 1.4 kbp *Eco*RI/*Nsi*I gene fragment containing intron 1 sequence (Southern, 1975).

Primer Extension. RNA was isolated from the livers of rabbits treated with 100 mg/kg phenobarbital or 0.9% NaCl for 17 h by the guanidinium thiocyanate method (Maniatis et al., 1982). For primer extension, the oligonucleotide CATGCTTAAGGATTTGTT, which is complementary to nucleotides 171–188 (see Figure 2) in exon 1 of P450IIC3 cDNA, was hybridized to 24 μg of total liver RNA isolated from control and phenobarbital-treated rabbits. The complex was incubated with 200 units of M-MLV reverse transcriptase for 60 min at 37 °C. Extended DNA fragments were analyzed on a 6% acrylamide DNA sequencing gel, and sequencing ladders primed from the same oligonucleotide with cloned genomic DNA as template were used as markers.

Polymerase Chain Amplification of P450IIC3 cDNA. Synthesis of cDNA corresponding to the 5′ region was primed with an oligonucleotide in exon 3, AAGGCTTCCTCTTGAAT, as described above in the primer extension section. The cDNA was added to a reaction mixture containing 75 mM Tris-HCl (pH 8.8), 7.5 mM $MgCl_2$, 18.5 mM $(NH_4)_2SO_4$, 11.3 mM mercaptoethanol, 7.5 μM EDTA, 225 μM deoxynucleotide triphosphates, 15% dimethyl sulfoxide, 1.5 units of *Thermus aquaticus* DNA polymerase, and 20 pM each of two oligonucleotides in a final volume of 100 μL (Roth et al., 1988). The oligonucleotide primers were the exon 3 oligonucleotide used to prime the reverse transcription and an oligonucleotide that included the initiator ATG codon of P450IIC3, TGCAGTTAACAGCCATGG. The underlined G and A were substitutions for T and G, respectively, in the gene sequence which created a *Hpa*I restriction site. The reaction was carried out in a Perkin-Elmer Cetus thermal cycler for 30 cycles of 2 min at 90 °C, 3 min at 50 °C, and

Table I: Comparison of Intron Splice Sites of P450IIC and P450IIB2 Genes

intron	P450	amino acid[a]	donor	intron size (bp)	acceptor
1	IIC1	Lys56	ACCAAAgtaagt		
	IIC2	Asn56	ACCAATgtaagt	1100	ttccagCTGTCA
	IIC3	Met56	AGCATGgtacgt	2200	ttccagCTAGCA
	IIC4	Lys56	ACTAAGgtaagt	2500	tcccagTTCTCA
	IIB2	Gln57	ATGCAGgtgaga	3200	ttgcagCTTCGC
2	IIC2	Gly111	GATTTGgtaggt	250	ctttagGAGTCA
	IIC3	Gly111	GATTAGgtatgc	1800	tgatagGAATTG
	IIC4	Gly111	GCCTTGgtaagt	168	tggcagGAATTG
	IIB2	Gly112	AATATGgtaaga	300	tgtcagGTGTGA
3	IIC2	Ala161	CCAATGgtgggt	2750	cttcagCCTCAC
	IIC3	Ala161	CTAATGgtatat	3600	cctcagCATCTC
	IIC4	Ala161	CCAATGgtggtg	3000	cttcagCCTTAC
	IIB2	Gly162	CCCAGGgtgaat	2300	ctccagGAGCCC
4	IIC2	Gln214	GTACAGgtgaga	1750	tttcagTTCTGC
	IIC3	Gln214	ATACAGgtaaaa	2500	atacagCTCTAC
	IIC4	Gln214	TTGCAGgtgaga	1500	ttttagGTCTAC
	IIB2	Gln215	AGCCAGgtccgt	800	ttacagGTGTTT
5	IIC2	Gln273	GAACAGgtaaag	3800	ttgtagGAAAAG
	IIC3	Lys272	GAAAAGgtagaa	1200	tgctagGAAAAA
	IIC4	Lys274	GAGAAGgtaaat	?	
	IIB2	Lys274	GAGAAGgtgagt	500	aaccagGAGAAG
6	IIC2	Ala321	TCATAGgtatgt	>3500	
	IIC3	Ala320	TTACAGgtatga	>6000	ttacagCTAAAG
	IIC5			>3500	tatcagCTAGGG
	IIB2	Glu322	TCACAGgtatat	1600	acgcagAGAAAG
7	IIC3	Lys382	CCAAAGgtgata	3500	tatcagGGCACA
	IIC5	Lys383	CCTAAGgtaatc	2000	atccagGGTACG
	IIB2	Lys384	CCCAAGgtgagg	300	tcccagAACACT
8	IIC3	Gly430	CAGCAGgtagca	2500	ttttagGAAAAA
	IIC5	Gly431	CAGCAGgtgaaa	2500	tctcagGAAAAC
	IIB2	Gly432	CCACAGgtgagg	3200	ccacagGAAAGC

[a] The amino acids are those at the exon/intron junction.

3 min at 70 °C. The ends of the amplified fragment were made blunt by incubation with T4 DNA polymerase, 5′-phosphates were added with T4 polynucleotide kinase, and the fragment was subcloned into pTZ19R. The DNA sequence was determined by double-stranded dideoxynucleotide sequencing (Sanger et al., 1977).

Results

Structure of the P450IIC3 Gene. The rabbit P450IIC3 gene contains 9 exons and spans at least 25 kbp (Figure 1). Of 44 λ clones containing P450IIC3 sequences, 3 clones were analyzed in detail. The clone λC3.1 and λC3.2 were isolated from the same rabbit DNA library in λ phage Charon 4A (Maniatis et al., 1978) and contained exons 2 through 9. Repeated attempts to isolate a fragment containing exon 1 from this library were unsuccessful. The third clone, λC3.3, was isolated from a second rabbit DNA library in EMBL3 and contained exons 1 and 2 and at least 8 kbp of 5′ flanking sequence. Multiple attempts to isolate a DNA fragment to close the gap in intron 6 from both rabbit DNA libraries were unsuccessful. Either this intron is extremely large or it contains sequences that caused it to be underrepresented in the two libraries.

For the sequences of exons 2–8, oligonucleotides at either end of the exon were used to sequence across the exon to the junction of the exon with the intron on the other side. The 5′ flanking region and exon 1 were sequenced on one strand by sequencing progressive deletions from the 5′ end generated by DNase I treatment, and specific oligonucleotides were used to sequence the second strand. DNA fragments corresponding to exon 9 were produced by sonication, subcloned, and sequenced (Figure 1). The lengths of the introns were determined by Southern analysis of the λ phage DNA using exon-specific probes. Compared with P450IIC3 cDNA, clone λC3.1 contained one silent change in exon 9 in 705 bp of exon sequence, and clone λC3.2 contained one replacement change at codon 257 in exon 5, resulting in a substitution of leucine for serine. The sequence corresponding to exon 1 and half of exon 2 was missing in the P450IIC3 cDNA clones so that it cannot be compared with the exon 1 sequence in λC3.3, but the derived amino acid sequence is 98% similar to that determined for the protein sequence of P450IIC3 (Ozols et al., 1985). In addition, the sequence of the 5′ 475 bp of the P450IIC3 cDNA produced by the polymerase chain reaction was identical with that of the gene exon sequences in λC3.3. The sequences of the 5′ flanking region and exons 1 and 2, which were missing from the P450IIC3 cDNA clone, are shown in Figure 2. It is not clear why full-length P450IIC3 cDNA clones were not isolated by Leighton et al. (1984). However, an 11 bp G-rich sequence in exon 2 (underlined in Figure 2) near the 5′ end of the cloned cDNA (indicated by the asterisk) is the complement of a sequence in exon 1 (underlined). Hybridization of this region in either the mRNA or the single-stranded cDNA may have prevented full-length synthesis of double-stranded cDNA.

The sequences at the intron–exon junctions for all the codons are shown in Table I. In each case, the introns contain the canonical 5′GT...AG3′ sequence, and the positions of the insertion of the introns are identical with those of other P450IIC (Govind et al., 1986; Zhao et al., 1990) and P450IIB genes (Mizukami et al., 1983; Suwa et al., 1985). The sizes of the P450IIC3 introns are similar to those of other P450IIC members except for intron 2 which is 5–10-fold larger in P450IIC3.

The 1850 bp of 5′ flanking region that were sequenced are remarkably AT rich with over 70% of the nucleotides either

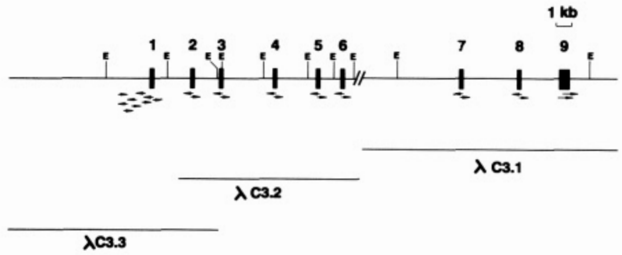

FIGURE 1: Schematic diagram of the structure of the rabbit P450IIC3 gene. Solid rectangles indicate exons. The locations of *Eco*RI sites (E) are shown. The arrows indicate the direction and regions that were sequenced. Below the diagram the positions of the three rabbit DNA fragments cloned in λ phage that were characterized are shown.

A or T and long strings of either A or T (Figure 2). The P450IIC3 sequence contains a canonical TATA sequence at −26 and a CCAAT sequence at −462. A search for about 20 sequences known to bind to regulatory factors revealed several sequences that match the minimum consensus sequence, GNTNNNTNNNNNNC, described for HP1 (Schorpp et al., 1988) which is similar to liver-specific sequences present in other liver genes also called HNF-1 (Courtois et al., 1988; Costa et al., 1989; Kugler et al., 1988; Maire et al., 1989) as indicated in Figure 2. Sequences similar to the AP-1 binding site are present at −367 and −1363, and multiple sequences with seven of eight matches with the octomer binding protein are present. A sequence similar to the 3′ portion of glucocorticoid response elements is present in intron 1. A number of imperfect inverted repeats are also present in the 5′ flanking region which are potential protein binding sites. Functional assays will be required to determine whether these sequences regulate transcription.

Initiation Site of RNA Synthesis. The major termination sites for fragments extended by reverse transcriptase with rabbit liver RNA as a template were clustered at a position about 25 bp 3′ of the canonical TATA sequence (Figure 3). The amount of extended fragment was about the same for RNA from control and phenobarbital-treated rabbits, consistent with the constitutive nature of the P450IIC3 gene. In parallel experiments, extended fragments for the phenobarbital-inducible P450IIC1 gene were about 15-fold greater in the phenobarbital-treated sample (not shown). The 5′ ends of the P450IIC3 mRNA appear to be heterogeneous with at least three bands occurring at the major start site. The A corresponding to the largest fragment of the three was designated +1. In addition, a second minor start site is observed about 30 bp 5′ of the main one. There are no obvious TATA sequences in the appropriate position to direct initiation at this minor site.

Genomic Southern Analysis. The rabbit P450IIC genes can be placed in subgroups of P450IIC1/C2/C14, P450IIC4/5, and P450IIC3 based on sequence similarity. The closely re-

```
                                                                                          HNF-1
CTTTTCTTTA TTTCTTTTTA AAAATATTTA TTTATTTTTA TTTGAAAAGC AGAGTTACAG AGAGAGCAAG AGATCTTCTA TCTGCTGGTT TTGGCCCCGA  -1750
                                                                                                  HNF-1
AATGGCACAA CAGCCAGAGC TGGGCAAGAC TGAAGCCAGG AGCCAGCTTC TAGATCCAGG TCTCCCACCT GAGTGCAGGG GCCCAAGAAC TTGGTCTTCT  -1650
                                                                                  HNF-1
TCTACTGCTT TTCCCAGGTA ATTAGCAGGA AGCTGAATCC CAAGTGGAGC AGCCAGGTAT TGAACCCATG CTCTTATATG GATGCTGGGG TCCCAAGTTG  -1550
                                  OCT                                 OCT
CAGCTTAACT CGAATACTAA CCCCTTTTTT TAATTTTCAT CATTTCTATT TTGCACCTAG TTTTTATACA AATGATTCCT TTTTAGTCTT TACATTTATT  -1450
                                    HNF-1                                                        AP-1
TAATTTATAC TGGTCATTTG TTAAAAATGC TCAAGATCCT TCTTTTCCTT TTTAAAATAA ACAAAAAGTA CCTTGTGTTT ATATTTTGAA TAATTAAAAA  -1350
TACATTATGT TTGCTTGTAG TCATTATGTT ATCAGGACAA TATGTTTGCC CACACTGTGC TGTAAAGACC AGAAAGTTAT TTCACCCATC TCACTTTAAT  -1250
TAATACCATT AATTAAACCT TTCCTGTGCT GCCTTAAGCC TCCTCCCCTC AGGATCCTCT GGTAACCATT ATTCTGCTCT CAATATCTAT GAAGTCCATG  -1150
                                   HNF-1
TAAGATTCCA CAAATATGAG CAAAACTATG TGGCTCTTGT CTTTCTCTAC AGTGCTTATT AAACTTAATA AATAAGAGAG GCACTTTTTT ATATTAATAT  -1050
TTATAATTTT TATTAATATG AGGTCAGATT TCATGTCACT AATACATATA AGTCTAAGAA TATAATTATA CTTCCCTCCC TTTATTCTTC ATTTTTTAAA  -950
AATTTTTGTG ATAATATATT TTTAATTTAG TTTATAGTCA CAGGCTAAAT GTTCCACTAC CTAAAGGTTT CAACAACTAA AAAGCATAGA GATCACTGTT  -850
CATCAGGAAT AGATAAAGGC TATAAACAAT AATCAAATCC CAAGATGTCA GAATCACTCA TATACATTAA TTCATTTGTG CTCTATCATT AGCTACCACA  -750
   OCT
AGTTATGCAA AACATGGGAT ATGTGCCTTT TTGGTTCTGG CTTATTTCAC TAAGTATAAT GGTCTCCAGT TGCAACAATT TTGTCACAAA AGATGGGAGG  -650
ATTTCATTCT TGTATGTCTG AGTAGCATTC CATTGTGCAC ATATACCACA TTTTTTTTAT CCAGTAATCA TTTGATGAAC ATCTAGGTTG ATTCCATAAA  -550
                                                                                                CBP
TCTTAGCTAT TCTGAATTGA ACTGCTATAA ACATAGGGTA CAGGTAATTT TTTCAGGTCA TTGTTAAATC ATTTATTTTA TACATTTCCA ATTTCTTGAT  -450
                                                                                             AP-1
ATAGGCACTA ATTACCATAA AATTTTCTTT TAATACTTCT TTTGCTATAT CCCATAAGTT TTGATGTGTT GTGCTATTAT TTTTACTCAT TCAATAATTT  -350
                                              OCT
TTTAATCTAT TTTGATTTCT TTTGTGATGC ACTGTTCATT CATTAGCATT GTTCAGTCTC AACATTTTTG TTTATTTTCT GGAGAGTTGG GATATCTATG  -250
                                              HNF-1
GTATTTTCTA AGACTGGCTT TTCTCCTCTA ATGCCTCTGG TTTTTCCATG ACACCTTTTA GAAAAAAAAA AAAAACCTTT TTTTGGAGAA GTGAAAGGTA  -150
GTATTTAGTG GGAATTTCAA CATTTAAAAT GATTGTTTTT TTTTCCTACC GTAAGACCAA AGTGCAATGT TTATTGAAGG TTCAAAGATT AGTACATATT  -50

                         TATA                    +1                        M  D  L  L   I  I  L  G  I  C    10
TGGTGATAAG ATTGTTGTGC ATTTATAAAG TCTTTGAATG CTCAACTTTA CACTGCATTTG ACACAGCCA TGGATCTCCT CATTATCTTG GGGATTTGTC  +51

L  S  C  V   V  L  L   S  L  W   K  K  T  H   G  K  G   K  L  P   P  G  P  T   P  L  P   V  V  G   N  L  L  Q    44
TTTCTTGTGT GGTTCTTCTT TCTTTGTGGA AGAAAACCCA TGGCAAAGGG AAGCTCCCAC CTGGACCCAC TCCTCTCCCA GTTGTTGGAA ATCTGCTACA  +151

 L  E  T   K  N  I   N  K  S  L   S  M                                                                         56
GCTAGAAACT AAGAATATCA ACAAATCCTT AAGCATGgta cgtgtgatca atttctctc agttgtttac aaagagtgaa taagacagtc tgataacagg  +251

aaaagattttcaaggga atattgaaga gaaggcttta aaggaaattg tccattggat acatagatgt aattttagaa tgtatatagc taggagcatt atg  +351
                       GRE
gggcccttgacctttca gcatttgttc tgaactcttg gattcacaca agaaaagaaa tcaaggagta gatatatgta agggagcagg cagccagatt tat  +451

                                                                                                L  A  K  E  Y   61
tgaaagcatgttagcat catatgtgag tgtaagccac ctctcctcat gagaggagga ctgctacctc ta..(1900 bp)..ttccagCTAG CAAAAGAGTA  +2450

 G  S  I   F  T  L   Y  F  G  M   K  P  A   V  V  L   Y  G  Y  E   T  V  K   E  A  L   I  D  R  G   E  E  F    94
TGGTTCTATA TTCACCTTGT ATTTTGGAAT GAAGCCCGCT GTGGTGTTGT ATGGATATGA AACAGTCAAA GAGGCCCTGA TTGATAGGGG AGAGGAGTTT  +2550
 S  G  R   G  I  F  P   V  F  D   R  V  T   K  G  L                                                           110
TCTGGCAGGG GAATTTTCCC AGTGTTTGAC AGAGTCACCA AAGGATTAGg tatgcttctg tttgggaata tcgatggttt.....                  +2630
```

FIGURE 2: Nucleotide sequence of the 5′ flanking region and exons 1 and 2 of the P450IIC3 gene. The major RNA initiation site is designated +1. Imperfect inverted repeats are indicated by arrows underlining the sequence with a dot at the center. Sequences similar to regulatory factor binding sites are shaded. Sequences that are similar to sequences conserved in the phenobarbital-inducible rabbit P450IIC genes are underlined at −269 and −92. Eleven base pair inverted repeats that could form the stem of a loop near the 5′ end of the cloned cDNA (asterisks) are underlined in exon 1 and exon 2. The single-letter amino acid code is indicated over the DNA sequence in the exon regions. The intron sequence is in lower case letters.

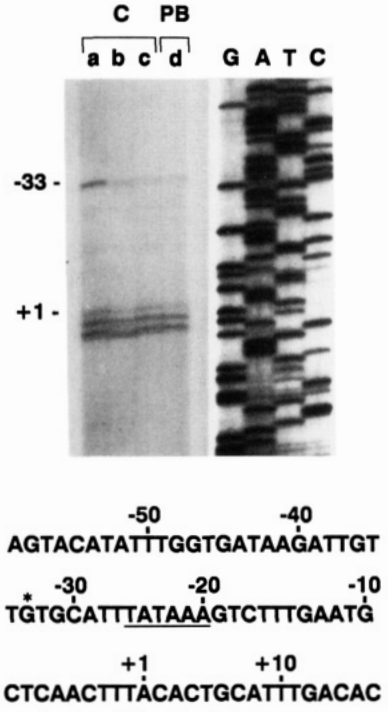

FIGURE 3: Determination of the 5′ end of P450IIC3 mRNA by primer extension. An oligonucleotide binding to the exon 1 region of the mRNA was extended with M-MLV reverse transcriptase, and the products were analyzed by gel electrophoresis as described under Materials and Methods. The sequence of the primer differed with sequences of other P450IIC cDNAs at 4–6 of 18 total nucleotide positions and was specific for P450IIC3. Preparations of RNA from three different control (C) rabbits (lanes a–c) and from a phenobarbital (PB)-treated rabbit (lane d) were analyzed. A sequencing ladder using the same oligonucleotide as primer was used as a marker. The DNA sequence surrounding the RNA initiation sites is shown at the bottom of the figure. +1 refers to the start site for the largest of the extended fragments in the major custer, and the minor start site at −33 is indicated by an asterisk in the sequence.

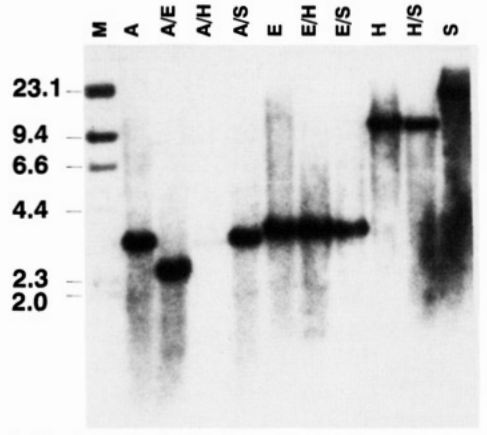

FIGURE 4: Southern analysis of the P450 gene. Rabbit DNA was digested with one or two enzymes as indicated, and the DNA was separated by electrophoresis in 0.5% agarose. For the hybridization, a DNA fragment from intron 1 was used as probe. *Hin*dIII-digested λ phage DNA was used as a marker (M). A = *Apa*I, E = *Eco*RI, H = *Hin*dIII, S = *Sal*I.

lated members of the subgroups often cause cross-hybridization when analyzed by Southern analysis. For example, Southern analysis using the 3′ untranslated region of P450IIC5 as a probe results in multiple bands (Tukey et al., 1985). To determine whether other closely related uncharacterized members of P450IIC3 might exist, rabbit liver DNA was analyzed by Southern hybridization with a 1.4 kbp *Eco*RI/*Nsi*I fragment containing intron 1 as probe (Figure 4). A

Table II: Discrepancies between the Amino Acid Sequences[a] Derived from the IIC3 Gene and cDNA and That Determined for the Protein (3b)[b] and Comparison with Other Rabbit P450IIC Proteins

position	gene	cDNA	3b	c1	c2	c4	c5	c14
49	Asn	Asn	Asp	Asp	Asp	Asp	Asp	Asp
	AAT	AAT[c]	GAT	GAC	GAC	GAC	GAC	GAC
82	Thr	Thr	Gly	Ala	Ala	Ala	Ala	Ala
	ACA	ACA[c]	GGA	GCA	GCA	GCA	GCA	GCA
84	Lys	Lys	Ile	Lys	Lys	Lys	Lys	Lys
	AAA	AAA[c]	ATA	AAA	AAG	AAG	AAG	AAG
89	Asp	Asp	Tyr	Asp	Asp	Asp	Asp	Asp
	GAT	GAT[d]	TAT	GAT	GAT	GAT	GAT	GAT
223	Gap	Gap	Gly	Ile	Phe	Leu	Leu	Met
	GAP	GAP[d]	GGG	ATT	TTC	CTT	CTT	ATG
256	Leu	Ser	Ser	Ile	Ile	Val	Val	Ile
	TTG	TCG[d]	TCG	ATT	ATT	GTT	GTT	ATT
339	Met	Met	Ser	Met	Met	Met	Met	Met
	ATG	ATG[d]	AGT	ATG	ATG	ATG	ATG	ATG
343	Thr	Thr	Ser	Ser	Ser	Ser	Ser	Ser
	ACC	ACC[d]	AGC	AGC	AGC	AGC	AGC	AGC
430	Ala	Ala	Thr	Thr	Thr	Ala	Ala	Thr
	GCA	GCA[d]	ACA	ACA	ACA	GCA	GCA	ACA

[a] Numbering of amino acids is according to the protein sequence of Ozols et al. (1985). [b] For amino acids in the P450(3b) sequence, possible codons are included in the table. [c] Sequence is from cDNA amplified by the polymerase chain reaction. [d] Sequence is from cloned cDNA (Leighton et al., 1984).

single band was observed for DNA digested with *Apa*I, *Eco*RI, *Sal*I, and *Hin*dIII alone or in combination, indicating that no other genes very closely related to P450IIC3 are present in the rabbit. This suggests that the gene cloned in this study is representative of the only P450IIC3 gene. In contrast, two P450IIC3 genes have been proposed based on enzymatic characteristics (Schwab & Johnson, 1985).

Protein Sequence of P450IIC3. The complete amino acid sequence of P450IIC3 could be derived from the gene sequence. Compared to the sequence derived either from the cloned cDNA (Leighton et al., 1984) or from the cDNA amplified by the polymerase chain reaction, the only difference was at amino acid 257 in exon 5 (Table II). At this position, the protein sequence agrees with the cDNA sequence. In contrast, the gene-derived sequence differs from the sequence determined for the corresponding protein (Ozols et al., 1985) in nine positions (Table II). These differences could be the result of sequencing or cloning errors or polymorphisms.

P450IIC Codon Similarities. The overall amino acid similarity of P450IIC3 with the other P450IIC members is about 65%, suggesting that a gene duplication separating P450IIC3 from the others occurred about 150 million years ago (Nelson & Strobel, 1987). However, there is considerable variation in the amount of similarity in different regions of the proteins. A comparison of the similarity of P450IIC3 by exons shows that overall C-terminal exons are more similar than the N-terminal ones (Figure 5). The highest similarity in all cases is exon 8, and the lowest is either exon 5 or exon 6 in the middle of the proteins. The similarity comparison with P450IIC4/5 has a sharp break between exon 6 and exon 7 increasing from a low of 58% similarity in exon 6 to 87% in exon 8. One possible explanation for these changes could be a gene conversion event that either increased divergence in exon 6 or decreased it in exon 8. To examine this possibility, the divergence of replacement and silent mutations were analyzed in the exons as described by Perler et al. (1980). If a gene conversion event occurred, the divergence of silent sites in the exons should parallel that of replacement sites. As expected, the number of replacement changes correlated with the amino acid changes observed (Table III). Although there was some

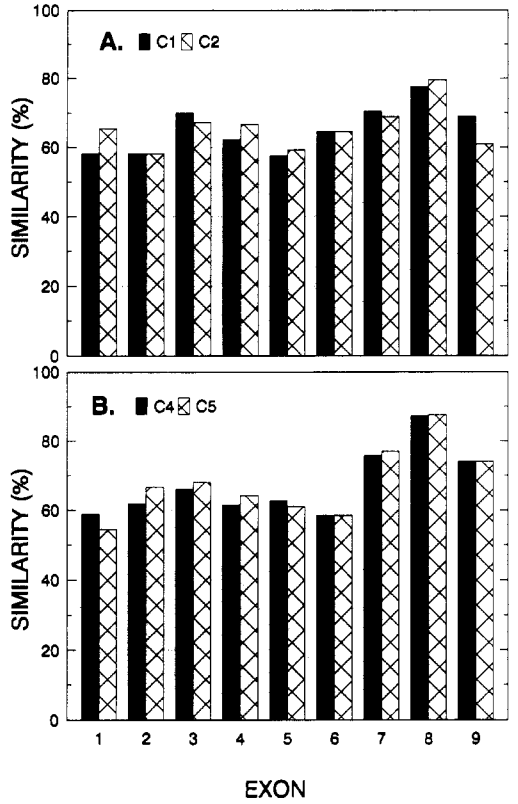

FIGURE 5: Comparison by codons of the amino acid similarity of P450IIC3 and other rabbit P450IIC proteins. The similarity for each codon was determined for the P450IIC1 and P450IIC2 subgroups in (A) and with the P450IIC4 and P450IIC5 subgroups in (B).

Table III: Divergence[a] of Silent (sil) Sites and Replacement (rep) Sites by Codon between the P450IIC3 Gene and Other Rabbit P450IIC Genes

	IIC1		IIC2		IIC4	
codon	rep	sil	rep	sil	rep	sil
1	0.21	0.38	0.18	0.42	0.24	0.42
2	0.21	0.31	0.22	0.35	0.21	0.38
3	0.15	0.43	0.16	0.45	0.17	0.45
4	0.21	0.37	0.20	0.44	0.20	0.35
5	0.24	0.31	0.24	0.33	0.23	0.33
6	0.17	0.31	0.18	0.37	0.25	0.45
7	0.15	0.28	0.17	0.27	0.11	0.24
8	0.10	0.32	0.09	0.19	0.06	0.38
9	0.16	0.34	0.21	0.36	0.13	0.40
all	0.18	0.34	0.18	0.35	0.18	0.37

[a] Divergence is not corrected for back-mutations.

variation in the estimated divergence of silent sites, there was no correlation between replacement changes and silent changes. For example, divergence between P450IIC3 and P450IIC4 of silent sites in exons 5 and 8, which have the greatest difference in similarities, was 0.45 and 0.35, respectively, while divergence of replacement sites was 0.25 and 0.06. Furthermore, P450IIC1 and P450IIC2 are 88% similar in amino acids, about equal to the exon 8 similarity between P450IIC3 and P450IIC4/5. This would suggest that if a gene conversion event is responsible for the increased similarity in exon 8, it occurred at about the same time as the gene duplication leading to P450IIC1 and P450IIC2 genes. However, silent site divergence for P450IIC1 and P450IIC2 was 0.15 (not shown), less than that for the exon 8 comparison, suggesting that the separation of the P450IIC1 and P450IIC2 genes occurred later than the separation of the exon 8 sequences of P450IIC3 and P450IIC4/5. Gene conversion, therefore, does not appear to be the mechanism responsible for the high similarity in exons 7, 8, and 9 of P450IIC3 and P450IIC4/5.

Discussion

The entire P450IIC3 gene, with the exception of a portion of intron 6, has been cloned. The gene spans >25 kbp, and, thus, as proposed on the basis of the partial structures of other rabbit P450IIC genes (Govind et al., 1986), these genes are larger than the P450IIB, P450I, and P450XXI genes that have been characterized. A constitutive rat P450IIC gene also was greater than 25 kbp in length (Morishima et al., 1987). The reason for the expansion of the intron sizes in this gene subfamily is not clear. However, the P450IIC subfamily contains more characterized genes than any other subfamily so that this region of the genome has undergone multiple gene duplications in the last 150 million years (Nelson & Strobel, 1987). Perhaps, duplications of smaller sequences within introns also were frequent, leading to the increased sizes. The larger introns may have decreased the potential for gene conversion events in the exon regions, permitting evolution of novel substrate specificities. Intron 6 may be particularly large, or alternatively difficult to clone, since multiple attempts to clone the missing intron 6 sequence were unsuccessful.

The 5′ flanking region contains the common promoter sequences TATA and CCAAT. A liver nuclear protein that binds to CCAAT and unrelated enhancer sequences has been described (Johnson et al., 1987). Other potential regulatory sequences, such as AP-1 and the octomer binding sequence, are also present, but have not yet been shown to the functional. Interestingly, a glucorticoid response element is present in intron 1, but induction of P450IIC3 by dexamethasone has not been described. Glucocorticoid response elements have also been reported in P450IA (Kawajiri et al., 1986; Hines et al., 1985) and a rat P450IIC gene (Morishima et al., 1987). In the former case, the element has been reported to be functional (Hines et al., 1988; Sherratt et al., 1989). Whether these sequences are functional in the P450IIC3 gene requires further study.

When compared by dot matrix analysis, overall similarity of the 5′ flanking region of the P450IIC3 gene with other rabbit P450IIC genes was not detectable (not shown). In contrast, dot matrix analysis among the phenobarbital-inducible P450IIC genes, C1, C2, and C4, revealed significant homology extending at least to −150 (Zhao et al., 1990). Two localized regions of high homology were present at similar positions in these three P450IIC genes and P450IIC3. The first of these was the TATA sequence at about −25 which is common to most RNA polymerase II genes. The second sequence is CAAAGTGCA present at −92. This sequence contains eight of nine identities with a nonamer present at about −100 in the P450IIC1, P450IIC2, and P450IIC4 genes. This region in the P450IIC3 gene contains 11/12, 10/12, and 9/10 matches with P450IIC1, P450IIC2, and P450IIC4 genes, respectively. This conservation suggests that this region may be functionally important. A third sequence, TTTATTTTC, at −279 also matches 8 of 9 bp of another nonamer conserved in the other sequences, but this similarity is less significant because the high AT content of the flanking region increases the probability of random occurrence.

A striking characteristic of the 5′ flanking sequence of P450IIC3 is a high percentage of A's and T's. The first 1.85 kbp of the sequence is 72% AT. Near the RNA initiation site, the region −1 to −300 contains 69% AT and 40% T compared to 56–61% and 25–30% for the phenobarbital-inducible P450IIC1, P450IIC2, and P450IIC4 genes. There are also long strings of A's and T's present in the P450IIC3 gene that are absent in the other P450IIC genes. In the first 1850 bp,

dA·dT tracts at least 9 bp long occur 21 times. The region from −166 to −195 in th P450IIC3 gene contains 27 of 30 A's or T's and an A string of 14 and T strings of 7 and 4. A string of nine T's is present at −114. Long poly[d(AT)] regions prevent nucleosome formation, and sequences 20 bp long decrease the probability of nucleosome formation (Kunkel & Martinson, 1981; Prunell, 1982). At least 9 but <12 bp of oligo(dA)·oligo(dT) comprise the minimal high-affinity binding site for a dA·dT binding protein (Winter & Varshavsky, 1989). These A and T strings in the P450IIC3 gene may be long enough to alter the nucleosome structure near the transcription initiation site to permit constitutive expression. In yeast, constitutive expression of the *pet56*, *his3*, and *ded1* genes has been correlated with AT-rich promoter regions (Struhl, 1985). dA·dT tracts can also augment the activity of phage T7 promoters in *Saccharomyces cerevisiae* cells that express T7 RNA polymerase (Chen et al., 1987). In addition to nucleosome effects, these sequences may be recognized directly by RNA polymerase or by dA·dT binding proteins, resulting in transcription activation (Struhl, 1985).

It has been proposed that proteins that have either high or low progesterone 16α-hydroxylase and are recognized by P450IIC3-specific monoclonal antibodies may be encoded at two separate genetic loci (Schwab & Johnson, 1985). In the cloning of the P450IIC3 genes, no evidence for a second genetic loci has been obtained. Partial sequence characterization or detailed restriction analysis of 10 genomic clones selected by P450IIC3 probes has indicated that all the clones correspond to the same gene as reported herein. In addition, only a single band was detected by Southern analysis using an intron 1 probe, suggesting that no other very closely related genes were present in the genome. It is possible that a second progesterone 16α-hydroxylase gene may be present with common P450IIC3 antigenic determinants but significantly different overall sequence or that the two genes are so closely related that they cannot be distinguished by the DNA analyses performed. Since the DNA sequence for the gene can now be fused with that of the incomplete cDNA to form a full-length cDNA, the expression of the cloned cDNA in cultured cells should be possible and may help resolve whether one or two genetic loci are required to produce the high- and low-activity P450IIC3s.

The amino acid sequence of P45IIC3 derived from the gene sequence differs from that determined for the protein in nine positions (Table II). If the differences are due to polymorphism, then the substitution probably would result from a single base change in the codon. Of the nine differences, only three could not result from a single nucleotide change: the gap at 223 in the gene-derived sequence instead of Gly, Thr_{82} instead of Gly, and Met_{339} instead of Ser (Table II). For the last of these differences, all the other rabbit P450IIC proteins have methionine in agreement with the gene- and cDNA-derived sequences (Table II). Although none of the other P450IIC proteins have a gap at position 223 as predicted by both the cDNA and gene sequences, it is difficult to explain the difference between a gap and an amino acid on the basis of polymorphism. At position 82, Gly, as determined for the protein sequence, is not present in other P450s, and Thr, as predicted by the P450IIC3 DNA sequences, is present in several other P450s but not other rabbit P450IIC proteins [see Nelson and Strobel (1987)]. For positions 84 and 89, the DNA-derived amino acids are identical with those in all the other rabbit P450IIC proteins. Therefore, for positions 82, 84, 89, and 339 and the gap at 223, the cDNA- and gene-derived sequences are probably more representative of the P450IIC3 amino acid sequence than that determined for the protein. On the other hand, at position 49, the Asp in the protein sequence but not the cDNA- or gene-derived Asn is identical with the other rabbit P450IIC proteins, suggesting an error in the derived sequences. However, in some human and rat P450IIC proteins, an Asn is present at this position so that the DNA-derived sequences may be correct. The remaining differences in the nucleic acid derived and protein sequences are consistent with polymorphic differences.

The similarity of P450IIC3 to the other P450IIC proteins varies considerably when individual exons are compared. Since gene conversion is likely to occur in the P450IIB genes and possibly occurred in exon 9 of P450IIC2 which has a region different from other P450IIC genes (Atchison & Adesnik, 1986), this mechanism could be responsible for the variance in similarity by exon. However, the relatively constant divergence of the silent sites in the exons suggests that gene conversion is not responsible. Exon 5 appears to be the most divergent of the exons, suggesting either little functional significance of this region or perhaps selection for mutations which could lead to new substrate specificity if this region is involved in substrate specificity. On the other hand, exon 8 is the most conserved region, and contains the heme binding region that is highly conserved among all P450s (Poulos et al., 1985). However, the conservation is higher between P450IIC4/5 and P450IIC3 than between P450IIC1/2 and P450IIC3, suggesting that a common function in this region may have been retained in the P450IIC3 and P450IIC4/5 isozymes. Identification of amino acids important in substrate recognition may clarify the difference in similarity among the isozymes.

Acknowledgments

We thank D. Kranz for assistance with the polymerase chain reaction method.

Registry No. Cytochrome P450IIC3, 9035-51-2.

References

Atchison, M., & Adesnik, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2300–2304.
Borck, K., Beggs, J. D., Brammar, W. J., Hopkins, A. S., & Murray, N. E. (1976) *Mol. Gen. Genet. 146*, 199–207.
Chen, W., Tabor, S., & Struhl, K. (1987) *Cell 50*, 1047–1055.
Costa, R. H., Grayson, D. R., & Darnell, J. E., Jr. (1989) *Mol. Cell. Biol. 9*, 1415–1425.
Courtois, G., Baumhueter, S., & Crabtree, G. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7937–7941.
Dieter, H. H., & Johnson, E. F. (1982) *J. Biol. Chem. 257*, 9315–9323.
Enea, V., & Zinder, N. Z. (1982) *Virology 122*, 222–226.
Gough, J. A., & Murray, N. E. (1983) *J. Mol. Biol. 166*, 1–19.
Govind, S., Bell, P. A., & Kemper, B. (1986) *DNA 5*, 371–382.
Hines, R. N., Levy, J. B., Conrad, R. D., Iverson, P. L., Shen, M.-L., Renli, A. M., & Bresnick, E. (1985) *Arch. Biochem. Biophys. 237*, 465–476.
Hines, R. N., Mathis, J. M., & Jacob, C. S. (1988) *Carcinogenesis 9*, 1599–1605.
Holmes, D. S., & Quigley, M. (1981) *Anal. Biochem. 114*, 193–197.
Johnson, E. F. (1980) *J. Biol. Chem. 255*, 304–309.
Johnson, E. F., Schwab, G. E., & Dieter, H. H. (1983) *J. Biol. Chem. 258*, 2785–2788.
Johnson, P. F., Landschulz, W. H., Graves, B. J., & McKnight, S. L. (1987) *Genes Dev. 1*, 133–146.

Kawajira, K., Watanabe, J., Gotoh, O., Tagashira, Y., Sogawa, K., & Fujii-Kuriyama, Y. (1986) *Eur. J. Biochem. 159*, 219–225.

Koop, D. R., & Coon, M. J. (1979) *Biochem. Biophys. Res. Commun. 91*, 1075–1081.

Kugler, W., Wagner, U., & Ryffel, G. U. (1988) *Nucleic Acids Res. 16*, 3165–3174.

Kunkel, G. R., & Martinson, H. G. (1981) *Nucleic Acids Res. 9*, 6869–6888.

Leighton, J. K., & Kemper, B. (1984) *J. Biol. Chem. 259*, 11165–11168.

Leighton, J. K., Debrunner-Vossbrinck, B. A., & Kemper, B. (1984) *Biochemistry 23*, 204–210.

Lin, H. C., Lei, S.-P., & Wilcox, G. (1985) *Anal. Biochem. 147*, 114–119.

Maire, P., Wuarin, J., & Schibler, U. (1989) *Science 244*, 343–346.

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

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

Mead, D. A., Szczesna-Skorupa, E., & Kemper, B. (1985) *Nucleic Acids Res. 13*, 1103–1118.

Mead, D. A., Szczesna-Skorupa, E., & Kemper, B. (1986) *Protein Eng. 1*, 67–74.

Mizukami, Y., Sogawa, K., Suwa, Y., Muramatsu, M., & Fujii-Kuriyama, Y. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3958–3962.

Morishima, N., Yoshioka, H., Higashi, Y., Sogawa, K., & Fujii-Kuriyama, Y. (1987) *Biochemistry 26*, 8279–8285.

Murray, N. E., Brammer, W. J., & Murray, K. (1977) *Mol. Gen. Genet. 150*, 53–61.

Nelson, D. R., & Strobel, H. W. (1987) *Mol. Biol. Evol. 4*, 572–593.

Ozols, J., Heinemann, F. S., & Johnson, E. F. (1985) *J. Biol. Chem. 260*, 5427–5434.

Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R., & Dodgson, J. (1980) *Cell 20*, 555–566.

Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., & Kraut, J. (1985) *J. Biol. Chem. 260*, 16122–16130.

Prunell, A. (1982) *EMBO J. 1*, 173–179.

Reubi, I., Griffin, K. J., Raucy, J. L., & Johnson, E. F. (1984) *Biochemistry 23*, 4598–4603.

Roth, M. E., Lacy, M. J., McNeil, L. K., & Kranz, D. M. (1988) *Science 241*, 1354–1358.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schorpp, M., Kugler, W., Wagner, U., & Ryffel, G. U. (1988) *J. Mol. Biol. 202*, 307–320.

Schwab, G. E., & Johnson, E. F. (1985) *Biochemistry 24*, 7222–7226.

Sherratt, A. J., Banet, D. E., Linder, M. W., & Prough, R. A. (1989) *J. Pharmacol. Exp. Ther. 249*, 667–672.

Silhavy, T. J., Berman, M. L., & Enquist, L. W. (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Southern, E. (1975) *J. Mol. Biol. 98*, 503–517.

Struhl, K. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8419–8423.

Suwa, Y., Mizukami, Y., Sogawa, K., & Fujii-Kuriyama, Y. (1985) *J. Biol. Chem. 260*, 7980–7984.

Tukey, R. H., Okino, S., Barnes, H., Griffin, K. J., & Johnson, E. F. (1985) *J. Biol. Chem. 260*, 13347–13354.

Winter, E., & Varshavsky, A. (1989) *EMBO J. 8*, 1867–1877.

Zhao, J., Leighton, J. K., & Kemper, B. (1987) *Biochem. Biophys. Res. Commun. 146*, 224–231.

Zhao, J., Chan, G., Govind, S., Bell, P., & Kemper, B. (1990) *DNA Cell Biol. 9*, 37–48.