cDNA Cloning, Sequence Analysis, and Induction by Aryl Hydrocarbons of a Murine Cytochrome P450 Gene, Cyp1b1

ZHIYUAN SHEN,¹ JINGMEI LIU,¹ ROBERT L. WELLS, and M.M. ELKIND

ABSTRACT

C3H mouse embryo fibroblast cells, designated 10T1/2, can be transformed by physical and chemical agents including polycyclic aromatic hydrocarbons. In a previous report (Shen *et al.*, Proc. Natl. Acad. Sci. USA 90, 11483–11487, 1993), we identified a cytochrome P450 gene induced by polycyclic aromatic hydrocarbons (PAHs) that is different from *1A1* or *1A2*, and which we tentatively named *P450CMEF*. Here, we report the entire cDNA sequence of *P450CMEF* (5,128 bp) and the amino acid sequence deduced from it (543 residues). A comparison of the latter sequence with known cytochrome P450s indicates that *P450CMEF* is in a new subfamily of family 1 of the P450 superfamily. Accordingly, the Committee on Standardized Cytochrome P450 Nomenclature designated the gene *Cyp1b1*. Exposure to various aryl hydrocarbons (2.5 hr) induced *Cyp1b1* mRNA in 10T1/2 cells to different degrees: 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, 7,12-dimethylbenz[*a*]anthracene, benz[*a*]anthracene, benzo[*a*]pyrene, and β -naphthoflavone were strong inducers; α -naphthoflavone and 3-methylcholanthrene, were moderate inducers; and benzo[*e*]pyrene was a weak inducer.

INTRODUCTION

THE CYTOCHROME P450 ENZYMES are heme-thiolate pro-L teins that reductively metabolize many endogenous as well as xenobiotic compounds. They are an important component of the mixed-function oxidase system. More than 200 P450 genes have been documented (Nelson et al., 1993), which have been grouped into 36 gene families according to their amino acid sequences and gene structure. All of the P450 proteins have a signature amino acid sequence, FXXGXXXCXG, in which cysteine is involved in heme binding (Nelson et al., 1993). In most cases, P450s from the same family of genes have more than a 40% amino acid identity and those in the same subfamily more than 55% (Nelson et al., 1993). The oxidases in the metabolism of aryl hydrocarbons are called aryl hydrocarbon hydroxylases (AHHs) because they catalyze hydroxylation reactions (Singer and Grunberger, 1983). Products of p450 genes in family 1-i.e., *IA1* and *IA2*-are involved in polycyclic aromatic hydrocarbon (PAH) metabolism (Nebert and Gonzalez, 1987; Gonzalez, 1989; Gonzalez et al.,

1991; Nebert et al., 1991) as, for example, in the case of 3methylcholanthrene (3MC) (Gonzalez et al., 1984; Kawajiri et al., 1984; Kimura et al., 1984).

The transformable cell line, C3H mouse embryo fibroblasts called 10T1/2, have been widely used in studying the induction of neoplastic transformation (focus formation) by physical and chemical agents (Reznikoff et al., 1973; Han and Elkind, 1979; Nesnow et al., 1981, 1989; Suzuki et al., 1981; Landolph, 1985a,b; Wells et al., 1992). Based on (i) the profiles of the metabolic products of 7,12-dimethylbenz-[a]anthracene (DMBA) and benzo[a]pyrene (BaP) in 10T1/2 cells (Pottenger et al., 1991; Wells et al., 1991); (ii) the lack of inducibility of AHH activity by 3MC in these cells (Gehly et al., 1979; Ho et al., 1983; Okey et al., 1983); and (iii) the observation that probes from Cyplal and Cyp1a2 did not yield detectable hybridizations with mRNA from 10T1/2 cells (Christou et al., 1990; Shen, unpublished data)-it was concluded that a different P450 gene was involved in PAH metabolism in these cells (Pottenger and Jefcoate, 1990; Shen et al., 1993b).

The new P450 gene to be described was tentatively named

Department of Radiological Health Sciences, Colorado State University, Fort Collins, CO 80523.

¹Present address: Life Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545.

P450CMEF (for P450 from C3H mouse embryo fibroblasts); the sequence of part of its cDNA was been reported (Shen *et al.*, 1993b). The latter cDNA fragment hybridized with a major \sim 5.3-kb and a minor \sim 4.2-kb transcript. Both transcripts were inducible by DMBA and benz[*a*]anthracene (BA); sequence comparisons suggested that the fragment came from a gene in family 1.

In this paper, we report the cDNA cloning, sequence analysis, and mRNA induction of *P450CMEF*. Based upon the amino acid sequence deduced from the foregoing, *P450CMEF* has been assigned to a new subfamily, *Cyp1b* (assignment by Committee on Standardized Cytochrome Nomenclature).

MATERIALS AND METHODS

Cell culture and chemical treatment

C3H mouse embryo fibroblasts, designated 10T1/2, were used. Details of their culture have been reported (Wells et al., 1991, 1992; Shen et al., 1993b). DMBA and BaP, were purchased from Aldrich Chemical Co. (Milwaukee, WI; α -naphthoflavone (α NF), β NF, and benzo-[e]pyrene (BeP) from Sigma Chemical Co. (St. Louis, MO); 3MC from Eastman Kodak (Rochester, NY) and 2,3,7,8-tetrachlorodibenz-p-dioxin (TCDD) from Midwest Research Institute, NCI Chemical Carcinogen Repository (Kansas City, MO). All chemicals were dissolved in acetone, HPLC-grade (Fisher Scientific, Fair Lawn, NY), and stored at stock concentrations of 10 mM, or 100 μM for TCDD only. For BaP, BeP, αNF , βNF , and 3MC, 15 μ l of the stock solutions were added to 15 ml of Eagle's basal medium with 10% fetal calf serum (BME-10), which resulted in 10 μM final concentrations. For DMBA and TCDD, the stock solutions were diluted with acetone before use to give final concentration of 0.3 μM and 10 nM, respectively. Control cells were treated with equal volumes of acetone alone.

Library construction and screening

From subconfluent, exponentially growing cells that were treated with 0.3 μ M DMBA for 3 hr, mRNA was extracted with the FastTrack mRNA Isolation Kit (Invitrogen, San Diego, CA). About 5 μ g of the mRNA was reverse-transcribed into cDNA using a combination of oligo(dT) and random primers in a first-strand cDNA synthesis. A cDNA library was constructed with the α -ZAP II cloning system (Stratagene, La Jolla, CA). A known fragment of *P450CMEF* (Shen *et al.*, 1993b) was labeled with [³²P]dCTP and used to screen the cDNA library by employing Duralon nylon membranes (Stratagene, La Jolla, CA) according to the manufacturer's manual.

DNA sequencing

A two-step cycle-sequencing protocol similar to that previously described (Shen *et al.*, 1993a) was used for sequencing. Specific sequencing primers, sense or antisense, were synthesized with an Applied Biosystems Synthesizer

(Model 380B, Applied Biosystems, CA) according to the sequence of known regions of the cDNA. Primers were labeled with [35S]dCTP and/or [35S]dATP, along with unlabeled dTTP and/or dGTP, by omitting at least one of the above four dNTPs, but not both ³⁵S-labeled compounds, depending on the downstream flanking sequences of a particular sequencing primer. In this way, adequate labeling of at least 4 bases and an extension of about 10 bases were obtained before the sequencing primers were arrested. The labeled and extended primers were subsequently used in dideoxy-termination, linear polymerase chain reaction (PCR) cycle sequencing at an elevated annealing temperature. Other procedures were as already described (Shen et al., 1993a,b) except that, in some instances, denaturing Long-Ranger Sequencing Gels (AT Biochem, Malven, PA) were run according to the suggestion of manufacturer. Single antisense strands were also sequenced to confirm the sequence of the coding region of the gene. GenBank's computer programs were used for sequence analysis.

Plasmid handling and other procedures

Positive phage plaques were purified and *in vivo* excision (EXASSIST/SOLR SYSTEM, Stratagene, La Jolla, CA) was used to make the pBluescript (SK⁻) plasmid containing the cDNA inserts. Plasmid DNA was isolated from host strain XL-1 Blue for direct PCR cycle sequencing and restriction endonuclease digestion by using QIAGEN Plasmid Kits (QIAGEN, Chatsworth, CA). Antisense singlestranded plasmid DNA was also purified following the manufacturer's procedure for the pBluescript EXO/ MUNG DNA sequencing system (Stratagene, La Jolla, CA). RNA isolation and Northern blot hybridization were done as previously described (Shen *et al.*, 1993b).

RESULTS

cDNA cloning and sequencing

A λ -ZAP II cDNA library was constructed (Stratagene, La Jolla, CA) from the mRNA of DMBA-treated cells (see Materials and Methods). Using the fragment of cDNA that we had already sequenced as as probe, 75 putative *P450*containing phage plaques were identified from among 4 × 10^s pfu in the first round of screening. Ten of these selected at random were picked for a second round of screening. Three purified plaques were obtained, and these were cloned into the pBluescript (Sk⁻) plasmid *via in vivo* excision of λ phage DNA (see Materials and Methods).

Phage DNA from these plaques was also PCR-amplified by using primers designed from the already known cDNA region of *P450CMEF* and the vector sequence to determine the length and orientation of the inserts *via* the protocol of Smith *et al.* (1990). Upon confirmation with restriction endonuclease digestions of the purified plasmids, it was determined that one of the three clones, #71, contained an insert of ~5.1 kb with its 5' end linked to the T3 promoter and its 3' end to the T7 promoter of the pBluescript (SK⁻) vector.

MURINE CYTOCHROME P450 GENE, CYP1B1

CAGACACCTCTGCGACGCCAGCTTGAAAACTGCCAGCGGGCTGGGATTTAACTCCA 1707

117	* TEACTTTGCGGGGGTCCCAGGTGCAAACTTGAGACACTTTTGTCCCTGTGCTCCCATCCTAA 1752 * DOTAL COTTTACGCGGGTCCCAGCTAACTTCAACACTTGACGACTACTCCCATCCTAAA 466													1752	GCC ANA CGG AGG TGC ATC GGT GAG GAA CTG TCT ANG ATG CTT CTG gly lys arg arg cys ile gly glu glu leu ser lys met leu leu		
177	теталогттловетствалтваютсяловатловское сососстеловате табетстевоваетсе сососстве сососств																
237	TAGG	alac	6666 TT(T)		1666	GACT	CTCA TGCT	GGTT	GAAG	CTTC	GCCA	GAAT	CTGG	NGCA(AGGC	481	phe leu phe ile ser ile leu ala bis gln cys asn phe lys ala
																1842	ANC CAN MAT GAG TOO TOA MAD ATG AGT TTO AGT TAT GGO CTG ACC
357	ATG met	ala	thr	AGC SOT	leu	AGT Ser	ala	asp	80T	pro	gln	gln	leu	841	TCG	496	asn gin asn glu ser ser asn met ser phe ser tyr gly leu thr
402	CTG	TCT	ACC	CNG	CAG	xcc	аст	СТТ	стс	ста	стс	TTC	TCC	: GT(CTG	1887	ATT ANG COC ANG TOG TIT AGA ATC CAT GTG TCT CTC AGA GAG TCG
16	leu	84I	thr	gln	gln	thr	thr	leu	leu	leu	leu	phe		Va]	leu	311	The TAR bio TAR mer bue wid the bis wat set led wid did set
447	GCC	GCC	GTG	CYC	TTA	GGC	CYG	TGO	сто	CTG	CGA	CNG	TGO	ယ	CGG	1932 526	ATG GAA CTC CTG GAT AAT GCT GTT AAA AAG CTG CAA ACT GAG GAA met glu leu leu asp asp ala val lys lys leu glp thr clu glu
31	ala	ala	wal	his	leu	gly	gln	trp	leu	leu	arg	gln	trp	glr	arg		
492	***	cca	TGG	TCC	TCG	CCC	CCA	GGT	CCC	TTT	CCT	TGG	CCA	CTO	ATC	1977	GGC TGC ANG TGA GAGGCCGGAGGGAGCTGGAATGTTTAAGGAATACCTATCTCAT gly cys lys ***
	-1-							4-1			P					2012	TAXCTGGGGAGGAAAAATAGATTTTTTCCCCCCAGCTTCATTTTTCGCCTTCAGCTCAGCAG
537 61	GGA gly	A S D	ala	GCG	GCT ala	GTT Val	GGC gly	CAG gln	GCG ala	TCG	CAC his	1eu	TAC tyr	: TTC : ph∉	ala	2092	TGAATGGAAACCAAGTGGTCTGAAGGTGGGGGCATGCTCACCAATTCATGGCTCCTCTTGG
587	cac	CTT	601	160	<u></u>	* 3*	GGC	GAC	GTT		C M G	3.77	CGT	CTO	aac	2152	ACCTGTGCTGGAGTTCCTGGAAGTATTTTGGAATTGAAGAGCAAAAGGGCCCAAGGAATT
76	arg	leu	414	arg	arg	tyr	gly	asp	val	phe	gln	11.	arg	leu	gly	2212	TGGAGCCTGTTGTTTTTTGGGTTTTTCAGCTAAACACATGCACACACA
627	AGC	TGT	ccc	GTG	GTG	GTG	ста	алт	GGA	GλG	AGT	GCC	ATC	CAC	CXG	2272	CACAACTATCTAAGAAAGTATTTCAGTAATTCTGCCTTTTTGGGTAGATTTGTGAGGGAA
91	84I	сув	pro	val	VAl	Val	leu	**1	gly	glu	ser	ala	11.	his	gln	2332	CTTCTATGTGCAGAAATTGACCCCATAGGAAACTGCAGTAAGCAAAGGCTTAGGATATAC
672	GCC	ста	GTG	CYG	CYG	aac	AGC	ATC	TTC	GCG	GAC	CGG	cca	ccc	TTC	2392	CCANGATTCAAAGACATGTGATTTCAGTGTAAAATGTAAAGACCAGAAGTCCTCCTACCA
106	414	leu	VAL	gin	gin	gly	8 41	11.	p he	414	asp	arg	pro	pro	phe	2452	AGAGACAGCTTGCTTGGAAAAATGCTTCAACTCCTTTATAGCCCTGGATGAGGCTTTCTG
717	occ	TCT	TTC	CGT	GTG	GTG	TCT	GGT	GGC	CGC	AGT	CTG	000	TTC	GGT	2512	CCTGCCTGTTGATGGGCCTCOCACTTTAGAATGGACCATAAAGTCAGTTGTCCCCTAAGA
			pue	.		•••		4-3	413	ary		194		pue	411	25/2	ATTIATOGCTUATTIAACAACTOCAGTAGGTTTTAGAGATTAGCGAATTIAAAGGTAAAT
762	CAC his	TAC tyr	TCT	GAG glu	CAC his	TGG trp	AAG 1ys	ACG thr	CAG gin	CGA	CGC ATT	TCG	GCC	TAT tyr	AGC SOL	2692	A GTA TACCA GTTA TA A GGGA GA BA BA BA BA CATTACTA COUNTRY A COUNTRY A COUNTRY A COUNTRY A COUNTRY A COUNTRY A
		-				-	-							-		2752	GGATTGTAATGAACTGGTTTCAGTAATAAGTCTGATAAACTAAAAAAGGAACCCAGTGTT
807 151	ACG thr	ATG met	CGT	GCT	TTC	TCC	ACG thr	CGC	CAC	CCG	CGC	AGT	CGC	GGT	CTT	2812	CCCALATAGTATTCAATGCATATAAAAATTTACCTAACTGCTCATAAACATCATATGAAT
	~	C1 <i>C</i>		~~~									,	1-1		2872	attaactattcaatcaggaaaacacttgaataatattatttcttatgtcatattgcaaacc
166	leu	glu	gly	his	ala	leu	ala	glu	ala	Arg	glu	leu	Val	GCA Ala	val	2932	AAAAAGTATAATCAGTTGAGTGAAATTTGCCCAATTCCAGGAAACGAGCCTGACTAAATC
897	ста	GTG	CGG	CGC	TGT	GCG	GGC	GGC	GCC	TTC	CTC	GAT	CC 3	ACG	CAG	2992	ANTOCGTTTTGACATCRGGGRACTGRGTANCRGRACRGTCTTTTTTTATTATRGTTCAT
181	leu	val	arg	arg	cys	a1a	gly	gly	ala	phe	leu	asp	pro	thr	gln	3052	ACAAGGGACAAATCTCAAAACTAAACCAGCGAAATATGTTCTGATCTGGTAATTGTAGTT
942	cca	GTC	ATT	GTG	GCG	GTG	GCC	л	GTC	ATG	λGC	GCT	GTG	TGC	TTC	3112	ACAACATGGATGGTTCTTTTGGAATTGACCCCATTAATATGTTAAGCTTACAAATTCTTG
196	pro	val	11.	val	ala	val	ala	890	val	met	ser	ala	val	суя	рье	31/2	CTICLCLCLATITITAAATTAGTTTAAAATCCATCTTGATTATTTCGTTCCAGTTGAAG
987	GGC	TGT	CGG	TAC	λλς	CAC	GAC	GAT	GCG	GAG	TTC	СТА	GAG	СТС	стс	3292	ANTGCCANATGTANANATANATANATGATANGTANATANAGCCATGAATGTCTGGGT
211	<u>d</u> ı A	cys	arg	tyr	åsn	018	asp	asp	ala	glu	phe	leu	glu	leu	leu	3352	GTTCATTAGAGACCATATGATAGGAGGCTCACTACAACCAGTAATTGGTAGTTCTTTTT
1032 226	λGC 8€Γ	CAC his	AAT	GAG Glu	GAG glu	TTC	GGG	CGC	ACA	GTG	GGT	GCG	GGC	AGC	CTG	3412	CCGAGTTCTTACTATTGTTGTTGTAGTTGTTTTTAAGTTGAGTATTAAAGTCTGTGTGT
1077	CTC	()7		0.000	,		3-1				4-1		4-1			3472	CTTTTGTTTTGTTTTTAAAGTCAGAATCTGAAGATTGTCTTGGGAGGTTGTTTAACAAC
241	val	asp	Val	leu	pro	trp	leu	gln	leu	phe	pro	asn.	pro	val	arg	3532	TGCTTCTTGAGACAGCAGAGATGCCTAAAATCAGTTTCTTTTATTTA
1122	ACC	ACC	ттс	CGC	лла	TTC	GNG	CAG	стс	ж	CGC	ж	TTC	AGC	ллс	3592	CGATGAGCAGCCACTTCCTGGAAGCTGAGAGAGGTTAGAAATACATGTATTCCGATTATG
256	thr	thr	phe	arg	1 7 #	phe	glu	gln	leu	4 8 h	arg		phe	ser		3652	TAAACAATCCGAAAGCCCTTTGAGAAGCGTTGGATGGGTGGAAAAGTGGGGACAGGGCCC
1167	TTC	GTT	стс	GAC	77 0	TTC	CTG	AGG	CAC	ccc	GAA	AGC	ста	GTG	ccc	3712	GTGGGGGGGTTTCCCATTTCTGTTTCATTTCGATTAGCCTGAACAAAAGAAGCCCTCAGC
2/1	pne	VAI	1 e u	asp	1 ys	рье	1eu	arg	bis	arg	glu	ser	leu	val	pro	3772	ACTCAGAGAGCTTACTGCTGGCTCTGTTTCATTAGGCTTCTTGGAATTTTCCTATCAAAA
1212 286	GGG gly	GCT Ala	GCT	CCT	CGY	GAC	ATG met	ACG thr	GAC	GCC	TTC	ATC	CTC	TCT	GCC	3832	TTACAAAGGAAGGGGAGTGCGATAGTCTCTGTGTGCAGAGGCTGCAAGACAGAGAAAGCA
1757		110														3952	ANALTITGAGGCATGAATGTTTAGACAAAAGGAGTTTTGAGTGGCAGGCTTTACCTTCAA
301	glu	175	lys	ala	ser	gly	ala	pro	g1y	4sp	asp.	Ser	ser	666	leu	4012	GTGTCTCTTAAGAGTGGCTGGAGCCCTGGCCCTTTCCTCCTATCTCCCCCCATCTATTTTG
1302	GAC	TTG	GAG	GAT	GTG	CCT	acc	ACT	ATT	ACG	GAC	ATC	TTC	GGA	GCC	4072	TGGTTGCAAGTTGAATTATAAAAAATTACAATGTAAACCTTATTTAAATTTCTCCACAAC
316	asp	leu	glu	asp	Val	pro	ala	thr	11.	thr	asp	11.	phe	gly	ala	4132	AGAACCCCTTGÄATGTGATACGCAGGTTGTGTCAGTATTTTAAAGGCCTTAATTTCAACA
1347	AGC	CAG	GAC	ACC	стт	TCC	ACC	GCG	стс	CTG	TGG	CTG	стс	ATC	CTC	4192	AAAATTAACTCTATGCAATAATTTTTAAGTTTTTCAGACATTTTAAAAGCATGCGCCTTA
331	ser	ğın	asp	thr	leu	84I	thr	ala	leu	leu	trp	leu	leu	11.	leu	4252	GTAAGGCTGGGACGGTGATTTGAGCTTCCATGGCCCTTTTACTATGAGAAATTTACAATA
1392	TTT	ACC	λGλ	TAC	CCG	GAT	GTT	CAG	GCC	CGC	GTG	CAG	GCT	GAG	TTG	4312 4	СТСАААСТСТТСАТСТАТСАСТСТОСАСТССТТАТТАААССААСС
340	Pue	Chr	ind	cyr	pro	шp	VAI	g1n		arg	VAI	gin	414	ğıu	Ten	4432	CTITETTTATTATATATATATATATATATATATATATATAT
1437 361	GAC	CAG gln	GTT Val	GTG Val	GGG Glv	AGG	GAC	CGC	CTG	DEC	TGC	λTG met	AGT BOT	GAC	CAG Gln	4492	ATAATTGTCATATTTTGGAAGCTCGAATAGATTACAGGCTTAGTCAGAAAGCCTGCTTTTG
1487			- -	~- ~-		- 3	ء معد	•	-					3.000	~~~	4552	TTTCTGCCACAGAGGACTTCTTTTGCTGCTCACCATGTTCCTCAGAGTATCTACTTCAG
376	pro	asn	leu	pro	tyr	ATC	met	ala	phe	leu	tyr	glu	80T	net	arg	4612	ATCAGTGCGGCAAAAGCATGTCTCATTAAATGCCTCATCTGTACAGTATAGCTTGGACTG
1527	TTC	TCC	AGC	TTT	TTG	сст	GTC	ACC	ATT	ссл	слс	GCC	лсс	vcc	GCC	4672	тетслаллееттестттессесскаласлалалсадатестслеетстватера
391	phe	30T	84I	phe	leu	pro	val	thr	il •	pro	hi.	ala	thr	thr	ala	4732	ттаатсясяадаяттсятоссссядяясття сссттва сстотояла татта сасаетс
1572	AAC	ACC	TTT	GTT	TTA	GGT	TAC	TAC	ATC	ccc	λλG	AAT	ACG	GTC	GTT	4792	TTATGCATTCCCACCAGGGCTGTACTACATTAAGCAGATAGAT
- 00	e s 1	LUL	Ъте	481	Ten	g t À	cyr	cyr	110	pro	тÀя	480	car	VAI	VAI	4832	MULAIWOUTTTUGCTGTGACACAARTTGTATGGTAAAAGGTGTATGCTTACCAGGTGATAAT
1617 421	TTT phe	GTT Val	AAC	CAG gln	TGG trd	TCT Ser	GTG Val	AAT ase	CAT his	GAC 48D	CCA pro	GCC ala	λλG lys	TGG tro	CCT pro	4972	TCTTAACATGTATATAGCTGATACAGTGCACACACACATATTGTGCACCACACACA
1662	330	-	- a1a	anc	-	a>=	~	a~~	000	-		a12	330			5033	
436	4.51	pro	glu	asp	phe	asp	pro	ala	arg	phe	leu	asp	lys	and and	gly	5092	univiane i letalutitigatataclaaatttäääätttäätteteteteattäteteteä TGAAACATTGTTACCTAAACAATAAAACCAATTTÄÄÄÄYY

FIG. 1. The cDNA sequence and amino acid sequence of the open reading frame of the mouse Cyp1b1 gene determined from a cDNA library made from actively growing C3H mouse 10T1/2 cells. The putative poly(A) addition signal (AAT-AAA), and the poly(A) addition site (CA) are identified (by underlining) according to the suggestions of Birnstiel et al. (1985).

1707 TTC ATT ANC ANG GCG CTA GCC AGC AGT GTG ATG ATA TTC TCA GTG 451 phe ile ann lys ala leu ala ser eer val met ile phe ser val

About 0.1–0.2 μ g of double-stranded plasmid DNA derived from clone #71 was used for direct DNA sequencing to obtain the whole cDNA sequence (see Materials and Methods). Parts of the inserts in clones #7 and #41 were also sequenced, but only revealed sequences identical to those in the corresponding regions of clone #71.

Sequence analysis

The cDNA sequence of clone #71 is shown in Fig. 1. It is 5,128 bases in length ending with CA. From 1,437-1,753 bp is the fragment that was previously reported (Shen *et al.*, 1993b). Ten base pairs upstream from the 3' CA is the location of a putative poly(A) addition signal sequence, AATAAA. Birnstiel *et al.* (1985) reported that most poly(A) signals are added to a CA located 9-30 bp downstream of the poly(A) addition signal. Accordingly, it is likely that the cDNA sequence in Fig. 1 contains the 3' end

of *P450CMEF* cDNA except for the poly(A) itself. Further, if we assume that the average size of a poly(A) site is 260–300 bp (Birnstiel *et al.*, 1985), we would estimate that the entire length of the mRNA corresponding to the cDNA in Fig. 1 is ~ 5.4 kb. The latter length agrees with what we have observed with Northern blot hybridization, ~ 5.3 kb (Shen *et al.*, 1993b). Hence, the sequence in Fig. 1 probably is the entire cDNA sequence of *P450CMEF*.

The regions corresponding to the degenerate primers that were used in the original identification of this gene (Shen *et al.*, 1993b) are upstream, 1,419-1,438 bp, and downstream, 1,753-1,774. Accounting for the degeneracies that were designed into the primers and the use of inosine, which is nonspecific in its base-pairing, the downstream primer has turned out to completely match the sequence. However, the upstream degenerate primer contained five mismatches (25%).

Analysis of the sequence in Fig. 1 revealed an open read-



FIG. 2. Phylogenetic tree generated by progressive alignment according to Feng and Doolittle (1987), which shows the relationship of P450CMEF with other P450 proteins. (\bigcirc) Node or branching point; (\rightarrow) a terminal sequence; and (number) distance to previous node in arbitrary units. The names on the right indicate the species and specific P450 sequences that were compared.

ing frame starting at position 375, which would code for a putative sequence of 543 amino acids having a molecular weight of ~60 kD. Accordingly, we propose that 357-359 bp (Met¹) should be the translation start codon because: (i) in the same reading frame, there is a stop codon at 237-239 bp (*i.e.*, 120 bp upstream of this assigned start codon); (ii) the next possible start codon would be 810-812 bp which would code for Met¹⁵²; and (iii) the proposed sequence of amino acids between Met¹ and Met¹⁵² is a region that is highly homologous to Cyp1a suggesting that it is unlikely that 810-812 bp is the start codon.

The sequence of the open reading frame has a larger degree of homology with P450s in subfamily 1A than with any other. By using the GenBank alignment program (Altschul and Erickson, 1986), with a Dayhoff cost matrix and a penalty factor of 2.5 for opening a gap and an incremental penalty factor of 0.5, the sequence was found to have a 38-39% amino acid identity with subfamily 1A P450s. However, when the FASTA program in GenBank was used (Pearson and Lipman, 1988), which only aligns certain primary regions, a homology of 40-41% with 1A P450s was found. The preceding comparisons suggest at least that *P450CMEF* should be in a different subfamily from that of 1A.

In view of the preceding uncertainties, the classification of *P450CMEF* to a specific *P450* family was guided by the phylogenetic tree-building program in GenBank of Feng and Doolittle (1987). By comparing the homologies of some 20 amino acid sequences in families 1, 2, and 3 with the putative sequence of P450CMEF, a phylogenetic tree was constructed (Fig. 2). It is evident from Fig. 2 that, although *P450CMEF* should be in a new subfamily, it still belongs in family 1. Based on the open reading frame discussed above, the Committee on Standardized Cytochrome P450 Nomenclature has assigned this gene to a new subfamily called *Cyplbl* (D. Nelson, D. Nebert, personal communications, 1993), a designation that we will also use in the remainder of this report.

Induction of Cyp1b1 by PAHs in 10T1/2 cells

The ability of BA and DMBA to induce *Cyp1b1* mRNA has been previously described (Shen *et al.*, 1993b). Other reports, as well as our own results (unpublished data), have shown that in 10T1/2 cells, 3MC and BeP do not induce AHH activity (Gehly *et al.*, 1979; Ho *et al.*, 1983). TCDD (Ho *et al.*, 1983), BaP (Gehly *et al.*, 1979), and β NF do induce AHH activity, and α NF frequently is an inhibitor.

With regard to the *Cyp1b1* mRNA, the action of a number of aryl hydrocarbons has been examined. Subconfluent 10T1/2 cells were exposed for 2.5 hr to 10 μ M of BaP, BeP, α NF, β NF, or 3MC, or to 10 nM of TCDD. Total RNA was used for Northern blot hybridization. The results from a single experiment are shown in Fig. 3a. Compared to acetone-treated controls, the averages (and standard deviations) of the relative mRNA levels induced by these treatments in several replicate experiments are shown in Fig. 3b (see legend for the details). It is evident that BA, BaP, β NF, and TCDD are strong inducers; α NF and 3MC are moderate inducers; and BeP induces to only a minor degree.



FIG. 3. a. Northern blot hybridization of P450CMEF (Cyplb1) with total RNA from TCDD-treated C3H-10T1/2 cells and cells treated with different PAHs. Subconfluent cells were incubated with 10 μM of αNF , βNF , BaP, BeP, 3MC, or 10 nM of TCDD for 2.5 hr. Acetonetreated cells (AC) were used as the controls. The label at the top of the figure indicates the treatment. Bases 979-1,404 (Fig. 1) was amplified from a cDNA plasmid, purified with the QIAGEN gel purification kit (QIAGEN, Chatsworth, CA); the antisense strand was labeled with [³²P]dCTP by linear PCR and used for Northern blot hybridization of P450CMEF (Cyp1b1) mRNA. After the radioactivity of the membrane had been scanned with an Ambis Radioanalytical Image System (San Diego, CA) and an X-ray film had been exposed, the membrane was stripped and rehybridized with a β -actin probe (Shen *et al.*, 1993b) for the normalization of RNA loading, and a second X-ray film was exposed. b. Induction of P450CMEF (Cyp1b1) mRNA in C3H mouse 10T1/2 cells. Quantitation was obtained from the relative counting rates of ³²Plabeled probes retained on the membranes. Averages of three to five hybridization experiments, including the results in Fig. 3a, and their standard deviation are shown.

DISCUSSION

The 5,128-bp sequence that we have described for P450-CMEF has an open reading frame on the basis of which this gene has been assigned to a new rodent subfamily with the designation Cyp1b1. This P450 gene was first detected by the application of highly degenerate primers and PCR amplification (Shen *et al.*, 1993b). This method proved to be successful even though the upstream degenerate, inosine-containing primer that was used proved to contain five mismatches (25%), illustrating the versatility of the method.

A 55-kD P450 protein, called P450-EF had been identified from BA-treated 10T1/2 cells (Pottenger *et al.*, 1991). Western blots showed that P450-EF has no cross-reactivity with P450 proteins in subfamily 1A. Although P450-EF might belong in a new family as suggested by Pottenger *et al.* (1991), because sequence information is not available, the relationship between the genes P450-EF and Cyp1b1 is not known.

In our previous report, two mRNA bands, ~5.3 kb and ~4.2 kb, were observed when a probe from bases 1,419-1,774 in Fig. 1 was used. The same two bands were observed (Fig. 3a) when a probe from bases 979-1,404 was used. In addition to the possible reasons for two mRNAs that have already been discussed (Shen *et al.*, 1993b), because 10T1/2 cells are immortal we note that a transcript from a mutated copy of the gene also is a possibility.

ACKNOWLEDGMENTS

This research was supported by the Department of Health and Human Services, U.S. Public Health Service, *via* Grant Number CA 47497 from the National Cancer Institute. The results reported are the responsibility of the authors and do not represent the official views of the National Cancer Institute.

Sequence reported in this paper has been deposited in GenBank under accession number U03283.

REFERENCES

- ALTSCHUL, S.F., and ERICKSON, B.W. (1986). A nonlinear measure of subalignment similarity and its significance levels. Bull. Math. Biol. 48, 617-632.
- BIRNSTIEL, M.L., BUSSLINGER, M., and STRUB, K. (1985). Transcription termination and 3' processing: The end is in site) Cell 41, 349-359.
- CHRISTOU, M., STEWART, P., POTTENGER, L.H., FAHL, W.E., and JEFCOATE, C.R. (1990). Differences in the modulation of P4501A1 and epoxide hydratase expression by benz-[a]anthracene and 2,3,7,8-tetrachlorodibenzo-p-dioxin in mouse embryo versus mouse hepatoma-derived cell lines. Carcinogenesis 11, 1691-1698.
- FENG, D.F., and DOOLITTLE, R.F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol. 25, 351-360.
- GEHLY, E.B., FAHL, W.E., JEFCOATE, C.R., and HEIDEL-BERGER, C. (1979). The metabolism of benzo(a)pyrene by cytochrome p-450 in transformable and nontransformable C3H mouse fibroblasts. J. Biol. Chem. 254, 5041-5038.
- GONZALEZ, F.J. (1989). The molecular biology of cytochrome p450s. Pharmacol. Rev. 40, 243-288.
- GONZALEZ, F.J., MACKENZIE, P.I., KIMURA, S., and NE-BERT, D.W. (1984). Isolation and characterization of fulllength mouse cDNA and genomic clones of 3-methylcholanthrene-inducible cytochrome P1-450 and p3-450. Gene 29, 281-292.
- GONZALEZ, F.J., CRESPI, C.L., and GELBOIN, H.V. (1991). cDNA-expressed human cytochrome P450s: A new age of mo-

lecular toxicology and human risk assessment. Mutat. Res. 247, 113-127.

- HAN, A., and ELKIND, M.M. (1979). Transformation of mouse C3H/110T1/2 cells by single and fractionated doses of X-rays and fission-spectrum neutrons. Cancer Res. **39**, 123-130.
- HO, D., GILL, K., and FAHL, W.E. (1983). Benz[a]anthracene and 3-methylcholanthrene induction of cytochrome p-450 in C3H/10T1/2 mouse fibroblasts, modulation role of cytotoxic 3-methylcholanthrene metabolites. Mol. Pharmacol. 233, 198-205.
- KAWAJIRI, K., GOTOH, O., SOGAWA, K., TAGASHIRA, Y., MURAMATSU, M., and FUJIIK-KURIYAMA, Y. (1984). Coding nucleotide sequence of 3-methylcholanthrene-inducible cytochrome p-450d cDNA from rat liver. Proc. Natl. Acad. Sci. USA 81, 1649-1653.
- KIMURA, S., GONZALEZ, F.J., and NEBERT, D.W. (1984).
 The murine Ah locus, comparisons of the complete cytochrome p₁-450 and p₃-450 cDNA nucleotide and amino acid sequences.
 J. Biol. Chem. 259, 10705-10713.
- LANDOLPH, J.R. (1985a). Mechanisms of chemically induced multistep neoplastic transformation in C3H 10T1/2 cells. In *Carcinogenesis*. E. Huberman and S.H. Barr, eds. (Raven Press, New York) vol. 10, pp. 211-223.
- LANDOLPH, J.R. (1985b). Chemical transformation in C3H 10T1/2 Cl8 mouse embryo fibroblasts: Historical background, assessment of the transformation assay, and evolution and optimization of the transformation assay protocol. In *Transformation Assay of Established Cell Lines: Mechanism and Application, IARC Scientific Publication No.* 67. T. Kakunaga and H. Yamasaki, eds. (International Agency for Research on Cancer, Lyon, France) pp. 185-198.
- NEBERT, D.W., and GONZALEZ, F.J. (1987). P450 genes: Structure, evolution, and regulation. Annu. Rev. Biochem. 56, 945-993.
- NEBERT, D.W., NELSON, D.R., COON, M.J., ESTABROOK, R.W., FEYEREISEN, R., FUJII-KURIYAMA, Y., GON-ZALEZ, F.J., GUENGERICH, F.P., GUNSALUS, I.C., JOHNSON, E.F., LOPER, J.C., SATO, R., WATERMAN, M.R., and WAXMAN, D.J. (1991). The p450 superfamily: Update on new sequences, gene mapping, and recommended nomenclature. DNA Cell Biol. 10, 1-14.
- NELSON, D.R., KAMATAKI, T., WAXMAN, D.J., GUEN-GERICH, F.P., ESTABROOK, R.W., FEYEREISEN, R., GONZALEZ, F.J., COON, M.J., GUNSALUS, I.C., GOTOH, O., OKUDA, K., and NEBERT, D.W. (1993). The p450 superfamily: Update on new sequences, gene mapping, accession number, early trivial-names of enzymes, and nomenclature. DNA Cell Biol. 12, 1993.
- NESNOW, S., LEAVITT, S., GARLAND, H., VAUGHAN, T.O., HYATT, B., MONTGOMERY, L., and CUBAK, C. (1981). Identification of cocarcinogens and their potential mechanism of action using C3H10T1/2CL8 mouse embryo fibroblasts. Cancer Res. 41, 3071-3076.
- NESNOW, S., GARLAND, H., and CURTIS, G. (1989). Inhibition of morphological transformation of C3H10T1/2CL8 mouse embryo cells by multiple carcinogen treatment. Cancer Lett. 47, 91-97.
- OKEY, A.B., MASON, M.E., GEHLY, E.B., and HEIDEL-BERGER, C. (1983). Defective binding of 3-methylcholanthrene to the Ah receptor within C3H/10T1/2 clone 8 mouse fibroblasts in culture. Eur. J. Biochem. 132, 219-227.
- PEARSON, W.R., and LIPMAN, D.J. (1988). Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444-2448.
- POTTENGER, L.H., and JEFCOATE, C.R. (1990). Characterization of a novel cytochrome P450 from the transformable cell line, C3H/10T1/2. Carcinogenesis 11, 321-327.

MURINE CYTOCHROME P450 GENE, CYP1B1

- POTTENGER, L.H., CHRISTOU, M., and JEFCOATE, C.R. (1991). Purification and immunological characterization of a novel cytochrome p450 from CeH/10T1/2 cells. Arch. Biochem. Biophys. 286, 488-497.
- REZNIKOFF, C.A., BERTRAM, J.S., BRANKOW, D.W., and HEIDELBERGER, C. (1973). Quantitative and qualitative studies of chemical transformation of cloned C3H mouse embryo cells sensitive to postconfluence inhibition of cell division. Cancer Res. 33, 3239-3249.
- SHEN, Z., LIU, J., WELLS, R.L., and ELKIND, M.M. (1993a). Cycle sequencing using degenerate primers containing inosines. BioTechniques 15, 82-89.
- SHEN, Z., WELLS, R.L., LIU, J., and ELKIND, M.M. (1993b). Identification of a cytochrome P450 gene by reverse transcription PCR using degenerate primers containing inosine. Proc. Natl. Acad. Sci. USA 90, 11483-11487.
- SINGER, B., and GRUNBERGER, D. (1983). Metabolic activation of carcinogens and mutagens. In *Molecular Biology of Mutagens and Carcinogens*. B. Singer and D. Grunberger, eds. (Plenum Press, New York) pp. 97-142.
- SMITH, D.P., JONSTONE, E.M., LITTLE, S.P., and HSIUNG, H.M. (1990). Direct DNA sequencing of cDNA inserts from plaques using the linear polymerase chain reaction. BioTechniques 9, 48-52.

- WELLS, R.L., SNEIDER, T.W., and ELKIND, M.M. (1991). Aryl hydrocarbon hydroxylase activity assayed in whole cell lysates using synchronous fluorescence spectroscopy. Anal. Biochem. **198**, 355-362.
- WELLS, R.L., CHEN, C.C., and ELKIND, M.M. (1992). Abrogation of killing and neoplastic transformation of C3H10T1/2 cells due to 7,12-dimethylbenz(a)anthracene. Carcinogenesis 13, 643-639.

Address reprint requests to: Dr. M.M. Elkind Department of Radiological Health Sciences Colorado State University Fort Collins, CO 80523

Received for publication January 5, 1994, and in revised form February 21, 1994.