Vol. 182, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1992 Pages 348-354

STRUCTURE OF THE MOUSE TYROSINE HYDROXYLASE GENE

Nakao Iwata, Kazuto Kobayashi[1], Toshikuni Sasaoka[2], Hiroyoshi Hidaka, and Toshiharu Nagatsu[1]*

Department of Pharmacology and [2]Biochemistry, Nagoya University School of Medicine, Nagoya 466, Japan

[1]Institute for Comprehensive Medical Science, School of Medicine, Fujita Health University, Toyoake, 470-11, Japan

Received November 29, 1991

SUMMARY: The mouse tyrosine hydroxylase (TH) gene was isolated from a genomic library by cross-hybridization with human TH cDNA probe. Nucleotide sequence analysis of two overlapping genomic clones showed that this gene is split into 13 exons distributed about 7.5 kb in length. The transcription initiation site was determined by primer extension analysis with mouse adrenal gland poly(A)$^+$RNA. The structure of the mouse TH gene was similar to that of the human TH gene, but it contained neither the alternative splice donor site around the 3'-end of the first exon nor an independent exon corresponding to the second exon of the human TH gene. There were the canonical TATA and GC boxes, cyclic AMP responsive element (CRE), and AP1 binding site in the 5'-flanking region of the mouse TH gene.

Tyrosine hydroxylase (tyrosine 3-monooxygenase; tyrosine, tetrahydropteridine: oxygen oxidoreductase (3-hydroxylating), EC 1.14.16.2.) (TH) catalyzes the first and rate-limiting step in catecholamine biosynthesis (1). This enzyme is selectively expressed in catecholaminergic neurons in the discrete regions of brain and sympathetic ganglia, as well as in chromaffin cells in adrenal medulla. Since catecholaminergic neurons are involved in a wide range of physiological actions, such as learning, memory, motor, and blood pressure (2), TH plays a central role in the brain and sympathetic functions of the mammalian nervous system.

Previously Grima *et al.* (3) and we (4, 5) reported the existence of multiple human TH isoforms which are different in their N-terminal coding regions. Characterization of the human TH gene showed that alternative splicing produces four kinds of mRNA from a single gene (6, 7). Linkage analysis showed that the human TH gene is localized to the short arm of chromosome 11, and is closely linked to insulin, insulin growth factor II, c-Ha-ras, and β-globin loci (8, 9).

Extensive data on genetic study of the mouse genome and recent improved techniques of genetic manipulation of the mouse embryo are important to investigate the development and

*To whom correspondence should be addressed.

0006-291X/92 $1.50

function of the complex mammalian nervous system. The mouse TH gene is mapped to the distal end of chromosome 7, and chromosomal region around this locus shares a number of loci homologous to human chromosome 11p15.5 (10). Furthermore our group recently described molecular cloning of the mouse TH cDNA and deduced its primary structure (11). In this paper, we subsequently report isolation and characterization of the mouse TH gene and compare its structure to the human gene.

MATERIALS AND METHODS

Library screening: A mouse genomic DNA library (stain DBA/2J) with EMBL-3 as a vector was purchased from Clontech. A total of 10^6 recombinant phages were plated out on NM539 lawn cells. Plaques were lifted onto nylon membranes (Hybond-N, Amersham). The human TH cDNA (type 4) fragments (4) were labeled with [α-^{32}P] dCTP (3,000 Ci / mmol) by Megaprime DNA labeling system (Amersham) and used as probes. The membranes were hybridized for 12 hrs at 42°C in a solution containing 6 x SSC (1 x SSC is 0.15 M NaCl and 15 mM sodium citrate), 0.3% SDS, 5 x Denhartdt's reagent, 50% formamide, salmon sperm DNA (100 μg/ml) and the labeled probe. Washing was carried out twice at 65°C with 2 x SSC containing 0.1% SDS.

DNA sequence analysis: Proper restriction fragments of the genomic clones were subcloned into Bluescript M13 (Stratagene) and pGEM7Zf (Promega) vectors. DNA sequences were determined by the dideoxy chain-termination method (12) with Sequenase DNA sequencing kit (Stratagene) with the T7 or SP6 promoter primer, or the synthetic oligonucleotide corresponding to the 5'- or 3'- terminal sequence of each exon. The oligonucleotides used were nt#1-20 (exon 1, sense), nt#61-80 (exon 1, antisense), nt#97-117 (exon 2, sense), nt#301-321 (exon 2, antisense), nt#322-342 (exon 3, sense), nt#475-495 (exon 3, antisense), nt#1057-1077 (exon 10, sense), nt#1093-1113 (exon 10, antisense), nt#1114-1134 (exon11, sense), nt#1209-1229 (exon 12, sense), and nt#1322-1342 (exon 12, antisense).

PCR (polymerase chain reaction): PCR was carried out in the reaction mixture of 25 μl containing 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 1.5 mM $MgCl_2$, 0.01% gelatin, 0.1% Triton X-100, 200 μM of each of dNTPs (dATP, dTTP, dCTP, and dGTP), and 1 unit of *Taq* DNA polymerase with 10 *p*g template DNA and 25 *p*moles of each of the primers. The following program was used: initial template denaturation step, 3 min at 95 °C, and DNA amplification step, 15 sec at 94°C, 30 sec at 55°C, 30 sec at 72°C, for 25 cycles, using GeneAmp PCR System (Perkin Elmer/Cetus).

Primer extension analysis: The 20-mer oligonucleotide, 5'-ACGGCTCTTCTGAAGCCCTT-3', that is complementary to the coding strand near the 5'-terminus of the mouse TH cDNA (nt#39-58) was synthesized and labeled at the 5'-end with [γ-^{32}P]ATP (6,000 Ci / mmol) using T4 polynucleotide kinase at 37°C for 30 min. The labeled primer was hybridized with 2 μg of mouse adrenal gland poly(A)$^+$RNA, incubated at 37°C for 1 hr in the reaction mixture of 80 μl containing 20 mM Tris-HCl (pH 8.3), 0.25 mM EDTA, 62.5 mM KCl, 10 mM $MgCl_2$, 10 mM DTT, 0.25 mM of each of dNTPs, 75 units of ribonuclease inhibitor, and 400 units of MMLV reverse transcriptase. The products were analysed by electrophoresis on 6% polyacrylamide-7 M urea gel. The DNA sequence of M13mp18 was used as a size maker.

RESULTS AND DISCUSSION

Isolation and characterization of the mouse TH gene: First we screened a mouse genomic DNA library with the full-length cDNA encoding human TH type 4 as a probe. Out of several positive clones, one was selected and named gMTH101. But this clone lacked the 5'-region of the mouse TH gene, therefore we rescreened the same library with the 0.3-kb *Eco*RI-

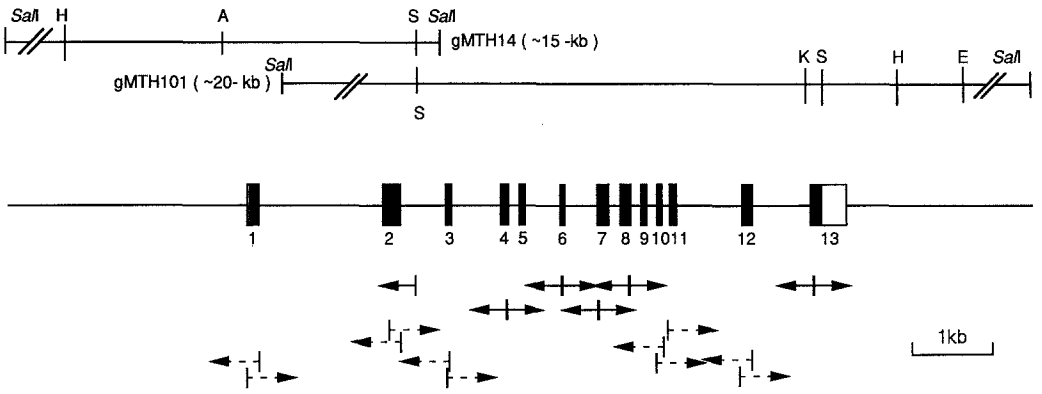

Fig. 1. Organization of the mouse TH gene. The upper panel shows the restriction enzyme maps of two overlapping genomic clones, gMTH101 (~20-kb) and gMTH14 (~15-kb). Restriction enzyme abbreviations: A, *Aat*II; S, *Sph*I; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I. The *Sal*I restriction sites are in the EMBL-3 vector arms. The midlle panel shows the exon/intron organization of the mouse TH gene. Exons are shown by boxes. Filled boxes indicate the protein-coding region and open boxes indicate the 5'- and 3'- untranslated regions. The lower panel shows sequencing strategy. Exons 1, 2, 3, 10, 11, and 12 were mapped by PCR with the T7 or SP6 promoter primer around the cloning site of plasmid vector, and the synthetic oligonucleotide corresponding to the 5'- or 3'- sequence of each exon. The nucleotide sequences of exons and their surrounding regions were determined with the synthetic oligonucleotide primers described above (dotted arrows). The remaining exons were localized by the restriction enzyme mapping and their sequences were determined from the proper restriction sites (arrows).

*Pst*I fragment containing 5'-terminal region of the human TH cDNA. Five positive clones were obtained, and one of them was named gMTH14 and used for further analysis. Restriction site mapping of gMTH101 and gMTH14 clones showed that they overlapped each other and covered an about 25-kb segment of the mouse genome (Fig. 1).

The 7.0-kb *Hin*dIII-*Sal*I fragment of gMTH101, and the 5.2-kb *Sph*I-*Sph*I fragment and the 2.0-kb *Sph*I-*Eco*RI fragment of gMTH14 were subcloned into plasmid vectors. We located all exons by the restriction site mapping or PCR method, as indicated in Fig. 1. The mouse TH gene spaned about 7.5-kb and contained 13 exons. All exon/intron junctional sequences were determined from the comparison with the mouse TH cDNA sequence (11). The sequences flanking splice doner and acceptor sites obeyed the GT-AG rule (Fig. 2) (13).

Determination of the transcription initiation site: The transcription start point of the mouse TH gene was identified by primer extension analysis. As show in Fig. 3, two bands of 89-bp and 87-bp were detected, and the 89-bp band was stronger than another. This result indicated two transcription start sites at the nucleotides 34-bp and 36-bp upstream from the initiation methionine codon. We numbered the nucleotide sequence begining with the major transcription initiation site at nucleotide 36-bp upstream from the initiation codon.

Comparison between the mouse and human TH gene structures: The organization of the mouse TH gene was compared to that of the human TH gene (Fig. 4A). Human TH is encoded by a single gene consisted of 14 exons (6, 7). Alternative splicing from the human TH gene generates four different mRNAs, which are constant for the major part, but are distinguishable from one another as to the insertion/deletion of 12-bp and 81-bp sequences near the N-terminus (3-5). Our analysis of the mouse TH gene showed that this gene is composed

Exon (size)	Exon 3'-boundary	Intron (size)	Exon 5'-boundary
Exon 1 (96-bp)	GTC ACG Val Thr	gtgaggatgaca·····Intron 1 (~1.6-kb)····· ttgactctcag	TCC CCA Ser Pro
Exon 2 (225-bp)	TTT GAG Phe Glu	gtgggttggctt·····Intron 2 (~0.6-kb)·····ttgctcctgaag	ACA TTT Thr Phe
Exon 3 (75-bp)	AAG G Lys Val	gtgaggatggat·····Intron 3 (~0.6-kb)·····ttgcttgtgaag	TT CCC Val Pro
Exon 4 (89-bp)	CAT CCG His Pro	gtgagcttgtgt·····Intron 4 (115-bp)····· ttggccatcag	GGC TTC Gly Phe
Exon 5 (68-bp)	AAG CA Lys Gln	gtaaggagcct·····Intron 5 (~0.4-kb)·····tcccctccacag	G GGT Gln Gly
Exon 6 (51-bp)	ACC TG Thr Trp	gtaagaccctgc·····Intron 6 (~0.4-kb)·····tgtggcctacag	G AAG Trp Lys
Exon 7 (146-bp)	AAG G Lys Glu	gtgagatgcaga·····Intron 7 (68-bp)·····aggtccttacag	AA CGG Glu Arg
Exon 8 (136-bp)	CAG CC Glu Pro	gtgagtatgctg·····Intron 8 (~0.1-kb)·····ggcctcacacag	A GAC Pro Asp
Exon 9 (70-bp)	TCC CAG Ser Gln	gtatgtttag·····Intron 9 (~0.1-kb)·····ctctacccacag	GAC ATT Asp Ile
Exon 10 (57-bp)	TCC ACG Ser Thr	gttcgttttcag·····Intron 10 (~0.1-kb)·····agtatctggcag	GTG TAC Val Tyr
Exon 11 (96-bp)	CTC CTG Leu Leu	gtgagattaccc·····Intron 11 (~0.8-kb)·····gcctctttctag	CAC TCC His Ser
Exon 12 (134-bp)	CTC AG Leu Arg	gtgggtggagc·····Intron 12 (~0.7-kb)·····cacacag	G AAC Arg Asn
Exon 13 (415-bp)	AGC TAA Ser term	---AATAAAGGAAGGAAAGGTCTCCaggagtgttcctggc···3'	

Fig. 2. Exon/intron junctions of the mouse TH gene. The splice boundaries were determined by comparing the genomic DNA sequence with the mouse cDNA sequence (11). The polyadenylation signal is underlined.

of 13 exons, and does not contain an exon corresponding to the second exon of human TH gene. Exons 2-13 of the mouse TH gene corresponded to exons 4-14 of the human TH gene, respectively, and the positions of the exon/intron junctions were identical between both genes. Interestingly, there was not the alternative splice doner site around the 3'-end of exon 1 of the

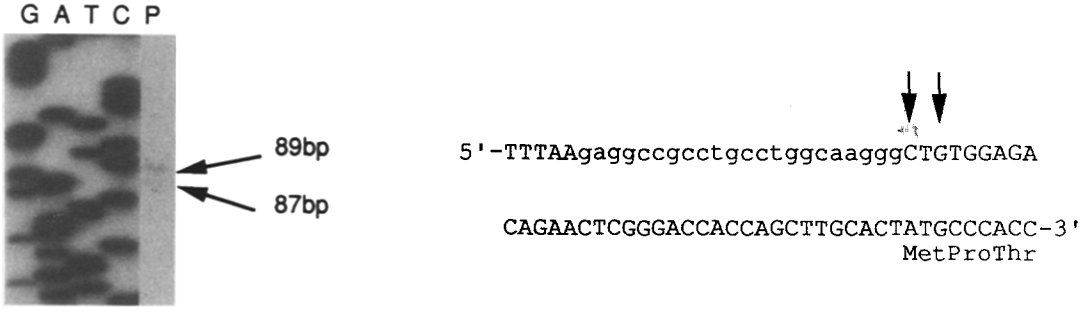

Fig. 3. Identification of the transcription initiation site of the mouse TH gene. Primer extension analysis was carried out with mouse adrenal gland poly(A)$^+$ RNA and the 20 mer synthetic oligonucleotide complementary to a part of TH mRNA, as described in MATERIALS AND METHODS. Two bands of 89-bp and 87-bp are shown by arrows (lane P). Lanes G, A, T, and C indicate the DNA sequence of M13mp18 determined with M13 sequencing primer (-40). The right panel shows the sequence around the transcription initiation site.

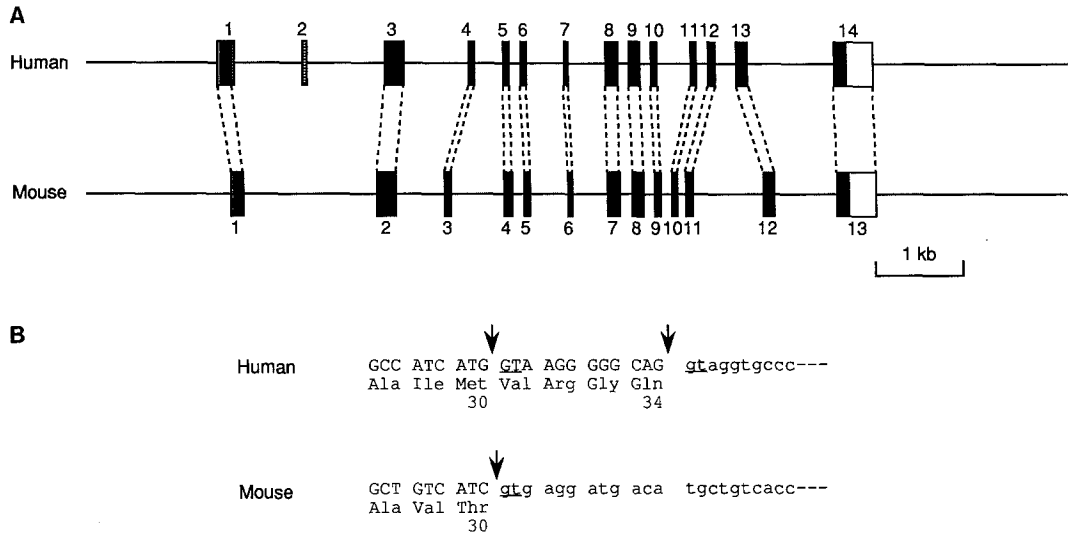

Fig. 4. Comparison between the mouse and human TH genes. A: Comparison of the structures between the mouse and human TH genes. Hatched boxes indicate the exon regions corresponding to the 12-bp and 81-bp insertion sequences of the human TH gene. B: The nucleotide sequence around the 3'-end of exon 1 of mouse and human TH genes. Arrows indicate the splice donor sites.

mouse gene (Fig. 4B). Recently, RT-PCR analysis of mouse adrenal RNA suggested the existence of a single species of the mouse TH mRNA (11). These results suggested that the mouse TH gene does not express the multiple forms of TH mRNA.

Potential cis-acting elements in the 5'-upstream region of the mouse TH gene: Fig. 5 shows the comparison of the 5'-flanking regions of the mouse, rat, and human TH genes. The nucleotide sequence (~260-bp) of the mouse TH gene had the homology of 94% and 74% to those of the rat (14) and human (7) TH genes, respectively. We found several consensus sequences of the promoter elements in the mouse TH gene. The typical TATA box (13), TTTAA, was found from -27-bp to -23-bp upstream from its transcription start site. The GC box, CCCGCC (15), were located from -115-bp to -110-bp. The consensus sequences of cyclic AMP responsive element (CRE), TGACGTCA (16, 17), and AP1 binding site , TGATTCA (18, 19), were localized from -42-bp to -35-bp and -204-bp to -198-bp, respectively. These consensus sequences were conserved among the mouse, rat, and human TH genes.

Many studies have reported the promoter analysis of the TH gene using cell culture and DNA transfection assay. The nucleotide sequence of 212-bp of the rat TH gene promoter is sufficient for the high-level expression in pheochromocytoma cell line (20). On the other hand, Gandelman *et al.* suggested the presence of multiple elements in the 5'-upstream and 3'-downstream regions of the human TH gene, to direct its cell-type specific expression (21). Moreover, the Fos protein family is suggested to bind AP-1 site of the TH gene and contribute to the regulation of this gene *in vivo* (22, 23). We recently reported transgenic mice which carry the whole human TH gene containing the 5'-upstream region of 2.5-kb, the entire exon-intron structure, and the 3'-downstream of 0.5-kb (24). Our results indicated the tissue-specific

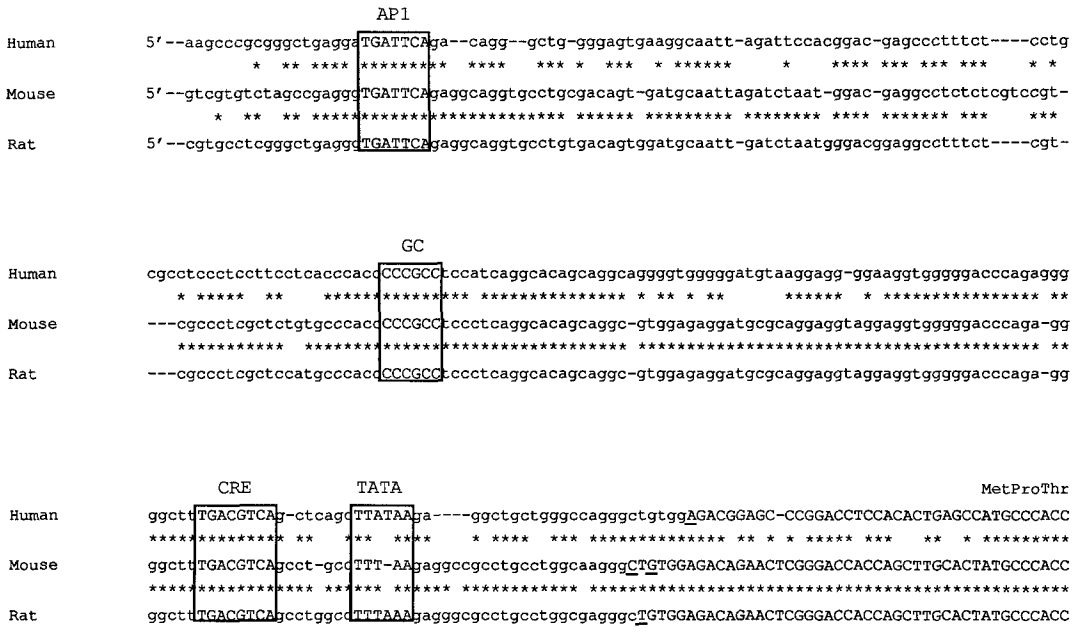

Fig. 5. Comparison of the 5'-flanking regions of the mouse, rat, and human TH genes. The matching bases are shown by asterisks. Putative cis-acting elements, TATA box, GC box, CRE, and AP1 binding site are boxed. The transcription initiation sites are underlined. The human and rat TH sequences are from references (7) and (14), respectively.

and high-level expression of the human TH gene, which contained the information sufficient for the transcriptional regulation of its gene in transgenic mice. Our studies with transgenic mice could be useful to understand the gene expression and function in catechalaminergic neurons.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research, Ministry of Education, Science and Culture, Japan. We would like to thank Drs. M. Katsuki and M. Kimura of Tokai University for helpful discussion and encouragement. We also thank Dr. Ichikawa of Nagoya University for valuable technical instruction.

REFERENCES

1. Nagatsu, T., Levitt, M., and Udenfriend, S. (1964) *J. Biol. Chem.* 239, 2910-2917.
2. *Catecholamines:Basic and Peripheral Mechanisms (eds. Usdin, E., Carlsson, A., Dahlström, A., and Engel, J.)* (1984) Alan R. Liss, Inc., New York.
3. Grima, B., Lamouroux, A., Boni, C., Julien, J.-F., Javoy-Agid, F., and Mallet, J. (1987) *Nature* 326, 707-711.
4. Kaneda, N., Kobayashi, K., Ichinose, H., Kishi, F., Nakazawa, A., Kurosawa, Y., Fujita, K., and Nagatsu, T. (1987) *Biochem. Biophys. Res. Commun.* 146, 971-975.
5. Kobayashi, K., Kaneda, N., Ichinose, H., Kishi, F., Nakazawa, A., Kurosawa, Y., Fujita, K., and Nagatsu, T. (1987) *Nucleic Acids Res.* 15, 6733.
6. O'Malley, K.L., Anhalt, M.J., Martin, B.M., Kelsoe, J.R., Winfield, S.L., and Ginns, E.I. (1987) *Biochemistry* 26, 6910-6914.
7. Kobayashi, K., Kaneda, N., Ichinose, H., Kishi, F., Nakazawa, A., Kurosawa, Y., Fujita, K., and Nagatsu, T. (1988) *J. Biochem.* 103, 907-912.

8. Moss, P.A.H., Davies, K.E., Boni, C., Mallet, J., and Reeders, S.(1986) *Nucleic Acids Res.* 14, 9927-9932.
9. O'Malley, K.L., and Rotwein, P. (1988) *Nucleic Acids Res.* 16, 4437-4445.
10. Brilliant, M.H., Niemann, M.M., and Eicher, E.M. (1987) *J. Neurogenet.* 4, 259-266.
11. Ichikawa, S., Sasaoka, T., and Nagatsu, T. (1991) *Biochem. Biophys. Res. Commun.* 176, 1610-1616.
12. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
13. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986) *Ann. Rev. Biochem.* 55,1119-1150.
14. Lewis, E.J., Harrington, C.A., and Chikaraishi, D.M. (1987) *Proc. Natl. Acad. Sci. USA* 84, 3550-3554.
15. Breathnach, R., and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349-383.
16. Montminy, M.R., Sevarino, K.A., Wagner, J.A., Mandel, G., and Goodman, R.H. (1986) *Proc. Natl. Acad. Sci. USA* 83, 6682-6686.
17. Montminy, M.R., and Bilezikjian, L.M., (1987) *Nature* 328, 175-178.
18. Lee, W., Haslinger, A., Karin, M., and Tjian, R. (1987) *Nature* 325, 368-372.
19. Lee, W., Mitchell, P., and Tjian, R. (1987) *Cell* 49, 741-752.
20. Cambi, F., Fung, B., and Chikaraishi, D. (1989) *J. Neurochem.* 53, 1656-1659.
21. Gandelman, K.Y., Coker III, G.T., Moffat, M., and O'Malley, K.L. (1990) *J. Neurochem.* 55, 2149-2152.
22. Gizang-Ginsberg, E., and Ziff, E.B. (1990) *Genes Dev.* 4, 447-491.
23. Stachowiak, M.K., Goc, A., Hong, J.S., Kaplan, B.B., and Stachowiak, E.K. (1990) *Mol. Cell. Neurosci.* 1, 202-213.
24. Kaneda, N., Sasaoka, T., Kobayashi, K., Kiuchi, K., Nagatsu, I., Kurosawa, Y., Fujita, K., Yokoyama, M., Nomura, T., Katsuki, M., and Nagatsu, T. (1991) *Neuron* 6, 583-594.