MOUSE DOPAMINE β -HYDROXYLASE: PRIMARY STRUCTURE DEDUCED FROM THE cDNA SEQUENCE AND EXON/INTRON ORGANIZATION OF THE GENE

Tamotsu Nakano^{1,2}, Kazuto Kobayashi², Shigeru Saito¹, Keisuke Fujita², and Toshiharu Nagatsu^{2*}

¹Department of Oral Biochemistry, Kanagawa Dental College, Yokosuka 238, Japan

²Institute for Comprehensive Medical Science, School of Medicine, Fujita Health University, Toyoake 470-11, Japan

Received October 22, 1992

SUMMARY: Genomic clones for mouse dopamine β-hydroxylase (DBH) were isolated from two genomic libraries derived from DBA/2J and 129/SV mouse strains, by plaque hybridization with the human DBH cDNA probe. Subsequently, cDNA encoding mouse DBH was amplified with reverse transcription-polymerase chain reaction (RT-PCR) method using primers corresponding to 5'- and 3'-portions of the mouse DBH mRNA, subcloned into a plasmid vector, and subjected to nucleotide sequence analysis. The clone encoded a protein of 621 amino acids with a calculated molecular mass of 70,186 daltons. The predicted amino acid sequence of mouse DBH showed 87%, 80% and 79% identities with the rat, bovine and human enzymes, respectively. Several potential amino acid sequences that are involved in the posttranslational modification and catalytic function of DBH were identified in mouse DBH protein. Nucleotide sequence analysis of the overlapping genomic clones showed that the mouse DBH gene was composed of 12 exons about 17 kb in length. Typical TATA and CCAAT boxes were observed in the 5'-upstream region of the gene. Northern blot analysis of adrenal gland RNA detected a single size species of the mouse DBH mRNA.

Dopamine β -hydroxylase (DBH; EC 1. 14. 17. 1) that catalyzes hydroxylation of dopamine to norepinephrine in the catecholamine biosynthesis pathway, is a copper-containing glycoprotein consisting of four identical subunits with molecular weight of 75 kDa (1,2). The enzyme exists as soluble and membrane-bound forms, in synaptic vesicles of norepinephrine and epinephrine neurons in brain and sympathetic ganglion, as well as in storage granules of adrenal medulla chromaffin cells. Several groups have reported the cDNA and predicted amino acid sequences of human (3,4), bovine (5-9) and rat (10) DBH. Actually, the recombinant protein expressed in cultured cells, exhibited the DBH enzyme activity (11,12). Moreover, we reported that the human DBH gene is composed of 12 exons, spanning ~23 kb of human

^{*} To whom correspondence should be addressed.

genome, and that alternative use of two poly A signals generates different mRNA types in 3'-untranslated region (4).

Since the DBH gene is expressed specifically in norepinephrine- and epinephrine-producing cells in the central and peripheral nervous systems, the regulatory elements involved in gene expression supply the information to understand molecular mechanism of neurotransmitter phenotype determination and the method to drive the foreign gene to the selected neuron subtype in transgenic mice. Recently, Mercer et al. (13) have reported that the 5.8-kb region of the human DBH gene promoter is sufficient to drive expression of the E. coli lacZ gene to norepinephrine neurons in brain, sympathetic ganglion, and adrenal chromaffin cells of trangenic mice. Our previous reports also showed that the 4-kb 5'-flanking region of human DBH gene promotes expression of human phenylethanolamine N-methyltransferase (PNMT) cDNA specifically in norepinephrine and epinephrine neuron subtypes of the trangenic mouse brain (14,15). As a part of our continuing investigation on the regulation of the DBH gene expression in transgenic mice, we present here cDNA cloning of mouse DBH and also charaterization of the mouse DBH gene structure, and compare its structure to the human gene.

MATERIALS AND METHODS

Screening of genomic DNA library: Two genomic DNA libraries used in this experiment were derived from DBA/2J (Clontech) and 129/SV (Stratagene) mouse strains, which were constructed with EMBL-3 and Lambda FIX II as vectors, respectively. A total of 1.5 x 10^6 recombinant phages of each library were plated out on lawn cells for screening. The phage plaques were lifted onto nylon membranes (Hybond-N, Amersham). The human DBH cDNA fragment (4) was labeled with $[\alpha^{-32}P]dCTP$ (110TBq/mmol) by the Megaprime DNA labeling system (Amersham) and used as a probe. The membranes were prehybrized for at least 2 hours at 65°C in 6 x SSC containing 5 x Denhardt's solution, 0.5% SDS and 0.1 mg/ml of salmon sperm DNA. Hybridization followed at 65°C for 17 hours in the above solution with a labeled probe. The membranes were washed with 2 x SSC containing 0.05% SDS at 65°C for 20 minutes twice.

Reverse transcription-polymerase chain reaction (RT-PCR) and cDNA cloning: Two oligonucleotides 5'-ATCCCTGCCATGCAAGCTCATCTCAGCCAC-3' (forward primer) which corresponds to the nucleotides from -9 to 21, and 5'-GCCAGCACTAGTCTATGGA-AGTGAGTCTTCACACAGGAAGG-3' (reverse primer) complementary to the nucleotides from 1,951 to 1,990, were used as RT-PCR primers. The first strand of cDNA was yielded by reverse transcription from 1 µg of mouse adrenal gland total RNA using oligo(dT) as a primer. Subsequently, the segment including the mouse DBH cDNA was amplified using RT-PCR primers. The cDNA fragment was electrophoresed on 0.8% agarose gel, eluted from gel slices, phosphorylated at the 5'-end with T4 polynucleotide kinase, and subcloned into the *Sma* I site of pUC119 vector.

Nucleotide sequence analysis: Proper restriction fragments of the genomic and cDNA clones were subcloned into pGEM-7Zf(+) and pUC119 vectors. Several deletion mutants of the appropriate clones were constructed according to the unidirectional deletion method (16). DNA sequences were determined by the dideoxy chain-termination method (17) with Sequenase DNA Sequencing kit (Stratagene) with T7 or SP6 promoter primer, or the synthetic oligonucleotide corresponding to the 5'- or 3'- terminal sequence of each exon.

Northern blot hybridization: Poly (A)⁺RNA (3 μ g) was electrophoresed on 2% agarose-formaldehyde gel, and transferred to a nylon membrane in 20 x SSC according to Thomas (18). The membrane was hybridized with ³²P-labeled mouse DBH cDNA fragment.

RESULTS AND DISCUSSION

Isolation of the gene and cDNA encoding mouse DBH: A DBA/2J mouse genomic DNA library was screened with the human DBH cDNA fragment (4) as a probe. Two positive phages designated as gMDBH(D)2 and gMDBH(D)3, were obtained (Fig. 1). Southern blot analysis with some restriction fragments of the human DBH cDNA showed that gMDBH(D)2 and gMDBH(D)3 contained the 5'- and 3'-regions of the mouse DBH gene, respectively, and that they did not overlap with each other. Then, we screened a 129/SV mouse genomic DNA library and obtained five overlapping positive clones (Fig. 1). Proper restriction fragments digested with BamH I or BamH I/Sal I, were subcloned into the pGEM-7Z(+) or pUC119 vector.

To isolate the mouse DBH cDNA with RT-PCR, we initially identified the first and last exons of the mouse DBH gene in the isolated phage clones, and determined the nucleotide sequences

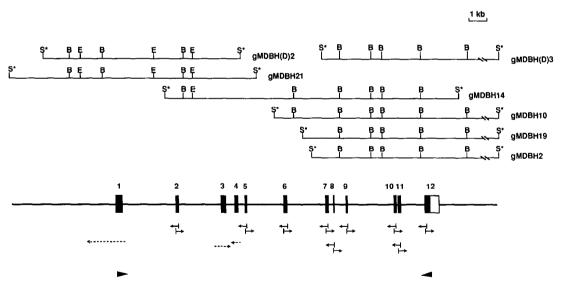


Fig.1. Organization of the mouse DBH gene. The upper panel shows the restriction enzyme maps of seven overlapping genomic clones, gMDBH (D)2 and 3 obtained from a DBA/2J mouse genomic library, and gMDBH2, 10, 14, 19 and 21 from a 129/SV mouse genomic library. Restriction enzyme abbreviations: B, BamH I; E, EcoR I; and S*, Sal I site that was artificially produced to construct the genomic DNA library. The exon/intron structure of mouse DBH gene is illustrated, indicating exons with boxes. Sequencing strategy is shown on the bottom. The nucleotide sequences of exons and their surrounding regions were determined with the synthetic oligonucleotide primers (arrows). Some exons were localized by the restriction enzyme mapping and their sequences were determined from the proper restriction sites (dotted arrows). Arrowheads indicate the location and the direction of the oligonucleotide primers used in RT-PCR described in MATERIALS AND METHODS.

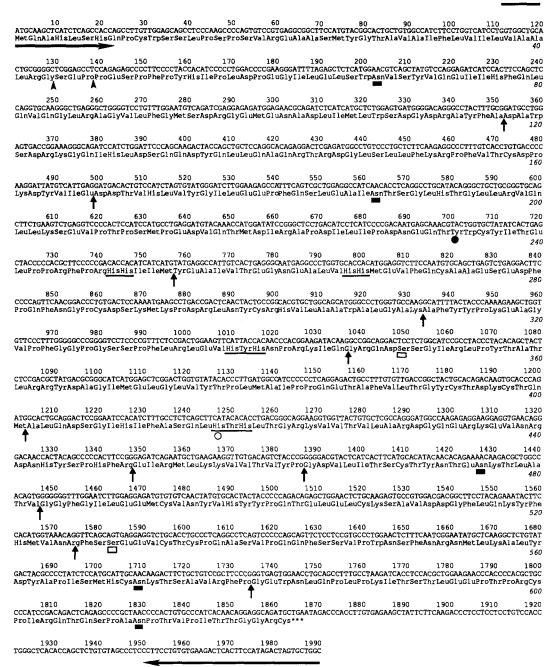
of their surrounding regions. As described in MATERIALS AND METHODS, the mouse DBH cDNA was amplified by RT-PCR method with adrenal gland total RNA using two synthetic oligonucleotides corresponding to 5'- or 3'- region of the mouse DBH mRNA as primers. The PCR product of about 2.0 kb was subcloned into the pUC119 vector and subjected to nucleotide sequence analysis (Fig. 2). The predicted amino acid sequence showed that this cDNA encoded the mouse DBH protein.

Subsequently, we determined the complete exon-intron structure of the mouse DBH gene. Positions of the exon/intron junctions were estimated in the mouse DBH cDNA sequence, based on the comparison with the human DBH gene structure (4). As shown in Fig. 1, 5'- and 3'-terminal sequences of each exon region were synthesized and used as primers, to localize each exon in subcloned restriction fragments and to determine the exon/intron junctional sequences. The mouse DBH gene spanned about 17 kb and consisted of 12 exons interrupted by 11 introns which ranged from about 0.1 kb to 2.8 kb. The sequences flanking splice donor and acceptor sites obeyed the GT-AG rule (19,20) (Fig. 3). The exon/intron junctions of the mouse DBH gene were identical with that of human DBH gene.

Primary structure of mouse DBH: As shown in Fig. 2, the cDNA clone encoded an open reading frame consisting of 621 amino acids with the calculated molecular weight of 70,186, that corresponded to the mouse DBH precursor protein including signal peptide. The sequence surrounding initiation methionine codon, CTGCCATGC, matched well the consensus sequence by Kozak (21). Previously, analysis of the bovine soluble DBH protein detected two N-terminal sequences that were of about equal stoichiometry and differed only in the presence or absence of the first three amino acids, Ser-Ala-Pro (22-24). We located two putative signal peptide cleavage sites in the mouse DBH precursor protein, based on the homology with the protein sequence data of the mature bovine DBH. The sequence of signal peptide showed high hydrophobicity and contained the N-terminal extension of 11 amino acids as compared to the bovine DBH signal peptide. Cleavage of signal peptide generates the mature protein of mouse DBH with molecular weight of ~65 kDa, without posttranslational modification.

The predicted amino acid sequence of mouse DBH was compared to the sequence data of DBH from other species (Fig. 4). It showed 87%, 80%, and 79% identities with the rat (10), bovine (5-9), and human (3,4) DBH sequences, respectively. We found several amino acid sequences that are involved in the posttranslational modification and catalytic function of DBH protein. Five potential N-linked glycosylation sites that are identical to the consensus sequence, Asn-X-Ser/Thr (25), were localized in the mouse DBH sequence. Three of these sites, Asn-68, Asn-188, and Asn-569, were conserved among other three species. Wang *et al.* (6) have shown that two potential glycosylation site, Asn-170 and Asn-552 of bovine DBH, which correspond to Asn-188 and Asn-569 of mouse DBH, respectively, are actually glycosylated. There were two putative calmodulin-dependent protein-kinase II phosphorylation sites, that correspond to the consensus sequence, Arg-X-Y-Ser/Thr (26), at residues Ser-350 and Ser-508 in the mouse





<u>Fig. 2.</u> Nucleotide sequence and deduced primary structure of mouse DBH. The amino acid sequence is indicated below the nucleotide sequence with three-letter code, and the nucleotides are numbered from the putative intitiation methionine codon. The vertical arrowheads indicate the predicted cleavage sites of signal peptide. Exon/intron junctions are represented with vertical arrows. Potential N-glycosylation sites and calmodulin-dependent protein kinase II phosphorylation sites are marked with filled and open boxes, respectively. Putative β -cresol and β -ethyryltyramine binding sites are indicated with the filled and open circles, respectively. The paired histidines and His-X-His residues are underlined. Horizontal arrows indicate the location and the direction of the oligonucleotide primers used in RT-PCR (see MATERIALS AND METHODS).

Exon	3'-junction			5'-junction		
1	TAC TTT Tyr Phe	GCG gtgagt(Intron : Ala 117	1, ~2.8kb)	P	GAT GCC Asp Ala 118	
2	GTC ATT Val Ile	GAG gtaggc(Intron : Glu 166	2, ~2.3kb)	_ P	GAT GAC Asp Asp 167	
3	ATC ATC Ile Ile	ATG gtaaac(Intron : Met 252	3, ~0.6kb)	r	TAT GAG Tyr Glu 253	
4	GGT GCC Gly Ala	AAG gtgtgt(Intron - Lys 311	4, ~0.3kb)	P	GCA TTT Ala Phe 312	
5	ATA CAA Ile Gln	G gtagga(Intron ! Gly 346	5, ~2.0kb)	_ 0	GC CGG Gly Arg 346	
6	ACC CAG Thr Gln	ATG gtgagt(Intron Met 401	5, ~2.2kb)	P	GCA CTG Ala Leu 102	
7	CAC TTC His Phe	CGG gtgaga(Intron ´ Arg 449	7, ~0.2kb)		GAG ATC Glu Ile 450	
8	GTC TAC Val Tyr	CCG gtgagt(Intron † Pro 462	3, ~0.6kb)	, G	GGG GAC Gly Asp 463	
9	GCC ACA Ala Thr	GTG gtaa(Intron ! Val 482), ~2.6kb)	Ĩ (GGG GGG Gly Gly 483	
10	GTA AAC Val Asn	AG gtgaag(Intronlo Arg 525), ~0.1kb)	P	G TTC Arg Phe 525	
11	CGC TTC Arg Phe	CCG gtatga(Intron1: Pro 578	1, ~1.4kb)	- G	GGT GAG Gly Glu 579	
12	AGA TGC Arg Cys 621	TGA atagacccaccttgtgaga ***	à			

<u>Fig. 3.</u> Exon/intron junctions of the mouse DBH gene. The splice boundaries were determined by comparing the genomic DNA sequence with the mouse cDNA sequence.

DBH sequence. The sequence of human DBH also contained one potential site of the same protein kinase at the position corresponding to Ser-350 of mouse DBH. However, in the rat and bovine DBH sequences, the consensus sequence around this potential phosphorylation site was converted to that of cyclic AMP-dependent protein kinase by the amino acid substitution of Gln-348 of mouse DBH to Arg, at the position -2 preceding the phosphorylation site. Furthermore, the paired histidines and His-X-His residues, which are thought to be related to copper binding, were located at positions, 248-249, 266-267, 337-339, and 416-418. Two

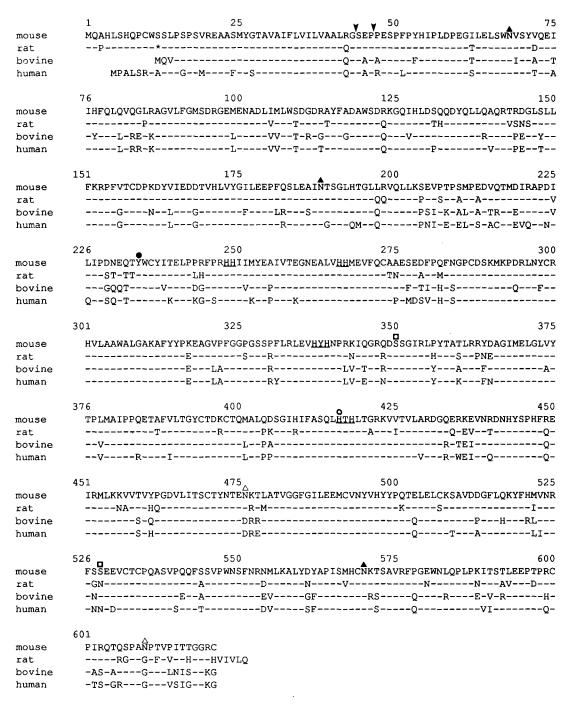


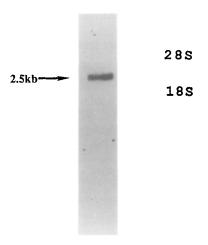
Fig. 4. Comparison of the deduced primary structures of mouse, rat, bovine and human DBH. The amino acid residues identical to mouse DBH are indicated as dashes, and gap (*) has been inserted to achieve a maximum homology. The vertical arrowheads indicate the predicted cleavage sites of signal peptides. Putative β -cresol and β -ethyryltyramine binding sites are indicated with the filled and open circles, respectively. Triangles represent the N-glycosylation sites conserved (\triangle) or not conserved (\triangle) among four species. Underlined, conserved HH or HXH residues indicate likely copper binding sites, and open squares show potential phosphorylation sites. The rat, bovine and human DBH profiles are derived from the reports of McMahon *et al.* (10), Taljanidisz *et al.* (5), and Kobayashi *et al.* (4), respectively.

catalytic active sites that were identified in bovine DBH with mechanism-based inhibitors, β cresol (27) and β -ethyryltyramine (28), were present at Tyr-234 and His-416 in the mouse
DBH sequence, respectively. These characteristic amino acid residues were conserved among
the sequences from four species.

Regulatory elements of the mouse DBH gene promoter: The nucleotide sequence of the 5'-region of mouse gene was compared to that of human DBH gene (4) (Fig. 5). Their nucleotide sequences showed the homology of 67%. Several transcription regulatory elements were found in the 5'-upstream region of the mouse DBH gene. The typical TATA (20) and CCAAT (29) boxes were found at nucleotides -38 bp to -33 bp, and -159 bp to -155 bp of the mouse DBH gene, respectively. These sequences were conserved between the mouse and human DBH genes. However, the sequences homologus to cyclic AMP response element



<u>Fig. 5.</u> Comparison of the 5'-flanking regions of the mouse and human DBH genes. Identical resudies are marked by (:). Gap (-) has been inserted to achieve a maximum homology. Putative cis-acting elements, TATA, CCAAT and CACCC boxes, cyclic AMP responsive element (TGACGTCC) are underlined. Human data are from reference (4).



<u>Fig. 6.</u> Northern blot hybridization of RNA from mouse adrenal gland probed with the 2.0-kb mouse DBH cDNA. Poly (A)⁺RNA profile (3.0 μg/lane) was electrophoresed in a 2% agarose gel in the presence of formaldehyde and transferred to a nylon membrane. The 28S and 18S ribosomal RNA bands are indicated as solid lines.

(CRE) (30), and CACCC box (31), which were observed in the 5'-region of the human DBH gene, were not found in the corresponding region of the mouse DBH gene.

Detection of the mouse DBH mRNA: Northern blot analysis of mouse adrenal gland RNA was carried out with the mouse DBH cDNA as a probe (Fig. 6). The result indicated the existence of a single mRNA species of 2.5 kb. We previously reported the generation of two different mRNAs (3.0 kb and 2.7 kb) encoding human DBH in pheochromocytoma. They were distinct only in the 3'-untranslated region and were produced by alternative use of the polyadenylation signals from a single human DBH gene (4). McMahon *et al.* (10) described that there were two mRNA species of 2.5 and 2.7 kb in rat DBH, and that the relative abundance of two mRNAs was different in rat adrenal gland and PC12 cells. On the other hand, Wu *et al.* (9) detected a single DBH mRNA in bovine adrenal medulla. Multiplicity of the DBH mRNA seems to be different in species or in tissues, and its physiological significance is still unclear.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, and Culture of Japan, and by a Grant-in-Aid from Fujita Health University to T. Nagatsu.

REFERENCES

- 1. Friedman, S. and Kaufman, S. (1965) J. Biol. Chem. 240, 4763-4773.
- 2. Stewart, L.C. and Klinman, J.P. (1988) Annu. Rev. Biochem. 57, 551-592.

- 3. Lamouroux, A., Vigny, A., Faucon Biguet, N., Darmon, M.C., Franck, R., Henry, J.-P. and Mallet, J. (1987) *EMBO J.* 6, 3931-3937.
- 4. Kobayashi, K., Kurosawa, Y., Fujita, K. and Nagatsu, T. (1989) Nucleic Acids Res.17, 1089-1102.
- Taljanidisz, J., Stewart, L., Smith, A.J. and Klinman, J.P. (1989) *Biochemistry* 28, 10054-10061.
- 6. Wang, N., Soithan, C., DeWolf, Jr., W.E., Wells, T.N.C., Kruse, L.I. and Leatherbarrow, R.J. (1990) *Biochemistry* 29, 6466-6474.
- 7. Robertson, J.G., Desai, P.R., Kumar, A., Farrington, G.K., Fitzpatrick, P.F. and Villafranca, J.J. (1990) J. Biol. Chem. 265, 1029-1035.
- 8. Lewis, E.J., Allison, S., Fader, D., Claflin, V. and Baizer, L. (1990) J. Biol. Chem. 265, 1021-1028.
- 9. Wu, H.J., Parmer, R.J., Koop, A.H., Rozansky, D.J. and O'Connor, D.T. (1990) *J. Neurochem.* 55, 97-105.
- 10. McMahon, A., Geertman, R. and Sabban, E.L. (1990) J. Neurosci. Res. 25, 395-404.
- 11. Ishii, A., Kobayashi, K., Kiuchi, K. and Nagatsu, T. (1991) Neurosci. Lett. 125, 25-28.
- 12. Lewis, E.J. and Asnani, L.P. (1992) J. Biol. Chem. 267, 494-500.
- 13. Mercer, E.H., Hoyle, G.W., Kapur, R.P., Brinster, R.L. and Palmiter, R.D. (1991) *Neuron* 7, 703-716.
- 14. Kobayashi, K., Sasaoka, T., Morita, S., Nagatsu, I., Iguchi, A., Kurosawa, Y., Fujita, K., Nomura, T., Kimura, M., Katsuki, M. and Nagatsu, T. (1992) *Proc. Natl. Acad. Sci. USA* 89, 1631-1635.
- 15. Morita, S., Kobayashi, K., Mizuguchi, T., Yamada, K., Nagatsu, I., Titani, K., Fujita, K., Hidaka, H. and Nagatsu, T. (1992) *Mol. Brain Res.* in press.
- 16. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103-119.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 18. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
- 19. Mount, S.M. (1982) Nucleic Acids Res. 10, 459-472.
- 20. Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349-383.
- 21. Kozak, M. (1984) Nucleic Acids Res. 12, 857-872.
- 22. Skotland, T., Ljones, T., Flatmark, T. and Knut, S. (1977) *Biochem. Biophys. Res. Commun.* 74, 1483-1489.
- 23. Joh, T.H. and Hwang, O. (1986) Annu. N.Y. Acad. Sci. 493, 343-350.
- 24. Taylor, C.S., Kent, U.M. and Fleming, P.J. (1989) J. Biol. Chem. 264, 14-16.
- 25. Hubbard, S.C. and Ivatt, R.J. (1981) Annu. Rev. Biochem. 50, 555-583.
- 26. Taylor, S.S., Buechler, J.A. and Yonemoto, W. (1990) *Annu. Rev. Biochem.* 59, 971-1005.
- 27. DeWolf, Jr., W.E., Carr, S.A., Varrichio, A., Goodhart, P.J., Mentzer, M.A., Roberts, G.D., Southan, C., Dolle, R.E., and Kruse, L.I. (1988) *Biochemistry* 27, 9093-9101.
- 28. DeWolf, Jr., W.E., Chambers, P.A., Southan, C., Sauders, D. and Kruse, L.I. (1989) *Biochemistry* 28, 3833-3842.
- Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 653-668
- Roesler, W.J., Vandenbark, G.R. and Hanson, R.W. (1988) J. Biol. Chem. 263, 9063-9066.
- 31. Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and Weissmann, C. (1983) *Cell* 32, 695-706.