Vol. 195, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 31, 1993 Pages 158-165

INCREASED HETEROGENEITY OF TYROSINE HYDROXYLASE IN HUMANS+

Hiroshi Ichinose[1], Tamae Ohye[1], Keisuke Fujita[1], Mitsuo Yoshida[2], Shintaroh Ueda[3] and Toshiharu Nagatsu[1]*

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

[2]Department of Neurology, Jichi Medical School, Tochigi 329-04, Japan

[3]Department of Anthropology, Faculty of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

Received July 11, 1993

Humans produce four different forms of tyrosine hydroxylase (TH) mRNA via alternative splicing of the gene. Here we demonstrate that New- and Old-World monkeys and the gorilla produce only two of the TH isoforms. Comparison among the genomic DNA sequences of various primates revealed that mutations that had accumulated in the genomic DNA created a new exon, resulting in the appearance of two new TH isoforms in man. These findings offer new insight into the sequence of events leading to the evolution of the higher primates into separate species. They also represent what may be the first evidence of a genetic difference between man and primates with respect to a specific brain function.

Catecholamines, such as dopamine, norepinephrine, and epinephrine, are neurotransmitters and hormones. They are deeply involved in various functions of the brain, including emotion, behavior, and learning. Tyrosine hydroxylase [tyrosine 3-monooxygenase, L-tyrosine, tetrahydropteridine: oxygen oxidoreductase (3-hydroxylating); EC 1.14.16.2] (TH) catalyzes the first and rate-limiting step of catecholamine biosynthesis (1,2). Recent studies on the cDNA structure of human TH revealed the presence of four isoforms (3-5), designated as type-1 to -4. Differences in these four isoforms are the insertion and/or deletion of 12- and 81-bp sequences between the 90th and the 91st nucleotides of type-1. These different mRNAs are generated through an alternative mRNA-splicing from a single primary transcript (Fig. 1) (6-8). Haycock demonstrated the presence of all four TH isoform proteins in the human adrenal medulla and in neuroblastoma cell lines utilizing

+Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. L14789-L14804.

*To whom all correspondence should be addressed.

0006-291X/93 $4.00

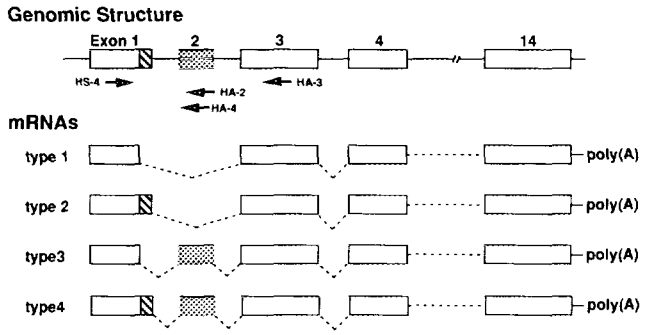

Fig. 1. A schematic illustration of the genomic structure for human TH, and the mechanism for alternative splicing that generates four isoforms of human TH. Locations and directions of primers used for the PCR are shown below the genomic structure.

isoform-specific antibodies (9). However there are no isoforms of TH in lower animals, such as rats (6), mice (10), an insectivore (*Sunkus murinus*) (11). It is important for understanding of physiological role(s) of the heterogeneity to explore when and how the heterogeneitiy of TH originated. In this paper, we determined the number of TH isoforms in Old-World and New-World monkeys, and analyzed genomic DNA sequences for TH in the higher primates.

MATERIALS AND METHODS

Reverse Transcription-Polymerase Chain Reaction (RT-PCR). Total cellular RNAs were extracted from human and animal tissues with guanidinium thiocyanate followed by centrifugation in cesium chloride solutions. RNAs were quantified by absorbance at 260 nm. Oligonucleotides were synthesized by a DNA synthesizer model 392 (Applied Biosystems). Sequences of synthetic oligonucleotide-primers were as follows: HS-4, 5'-dTGTCTGAGCTGGACGCCAAGCAG-3' (identical with nucleotides 53 to 75); HA-3, 5'-dCTCCTCAAAGGCCACAGCCTCCA-3' (complementary to nucleotides 296 to 318); HA-2, 5'-dTGCAGTTCCAGGCCACGGAGAGC-3' (complementary to nucleotides 128 to 150); HA-4, 5'-dGGCCACGGAGAGCCTGTGAGGCT-3' (complementary to nucleotides 118 to 140). Nucleotides are numbered based on human TH type-4 cDNA with the first base of the ATG initiation codon designated as +1. Total RNAs (1 μg) from various animals were reverse-transcribed with Moloney murine leukemia virus-reverse transcriptase (Bethesda Research Laboratories) using random hexamer. One fourth of the transcribed cDNA was used for each amplification with the primer sets indicated in the text. PCR was performed in 25 μl of reaction mixture composed of 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% gelatin, 0.1 mM concentration of each dNTP (dATP, dTTP, dCTP, and dGTP), 20 pmol of sense- and antisense-primers, 1 U of Perfect Match Polymerase Enhancer (Stratagene), and 0.6 U of *Thermus aquaticus* DNA polymerase (Perkin-Elmer Cetus). The following program was used for 30 cycles of amplification in a GeneAmp PCR System 9600 (Perkin Elmer Cetus): 30 sec at 94°C and 1 min at 60°C. The amplified DNA fragments were electrophoresed in 4% Nusieve GTG agarose gel (FMC Bioproduct) and stained with ethidium bromide.

Genomic DNA Analysis. DNAs of chimpanzee (*Pan troglodytes*), orangutan (*Pongo pygmaeus*), and gibbon (*Hylobates lar*) were prepared from lymphocytes of peripheral blood. DNA of gorilla (*Gorilla gorilla*) was obtained from lymph nodes. The other genomic DNAs were extracted from the human placenta, the macaque (*Macaca fuscata*) kidney, and the marmoset (*Callithrix jacchus*) kidney. Genomic DNAs of macaques were extracted from two individual monkeys and analyzed separately. Genomic DNA was amplified using the PCR, which was performed under

the same conditions as described in RT-PCR, with 100 ng of each genomic DNA used. Primers used for the amplification was shown in Fig. 3. Amplified DNA fragments were separated in 4% NuSieve agarose gel, extracted, ligated into a pUC119, and sequenced. The nucleotide sequences were determined by the dideoxynucleotide chain-termination method using Sequenase (United States Biochemical Corp., Cleveland, OH). All fragments were sequenced on both strands. The sequences around exon 2 of gorilla and macaque genomic DNA were confirmed several times.

RESULTS AND DISCUSSION

By using the RT-PCR, we examined tissues from Old-World monkeys (macaques), New-World monkeys (marmoset), and man. As shown in Fig. 2a, we obtained two bands using the HS-4 and HA-3 primer set, indicating the presence of both type-1 and -2 TH mRNAs, in the brain and the adrenal gland from various monkeys (macaques and marmoset) as well as in the postmortem human brain and adrenal gland. When we examined the rat brain, we detected only one band (Fig. 2a, lane 10), confirming that rats have only one type of TH molecule.

We amplified mRNAs of type-3 and -4 using two primer sets, HS-4 and HA-2, or HS-4 and HA-4, on the same samples used for the detection of type-1 and -2. As shown in Figs 2b and 2c, type-3 and -4 TH cDNAs were amplified only in the

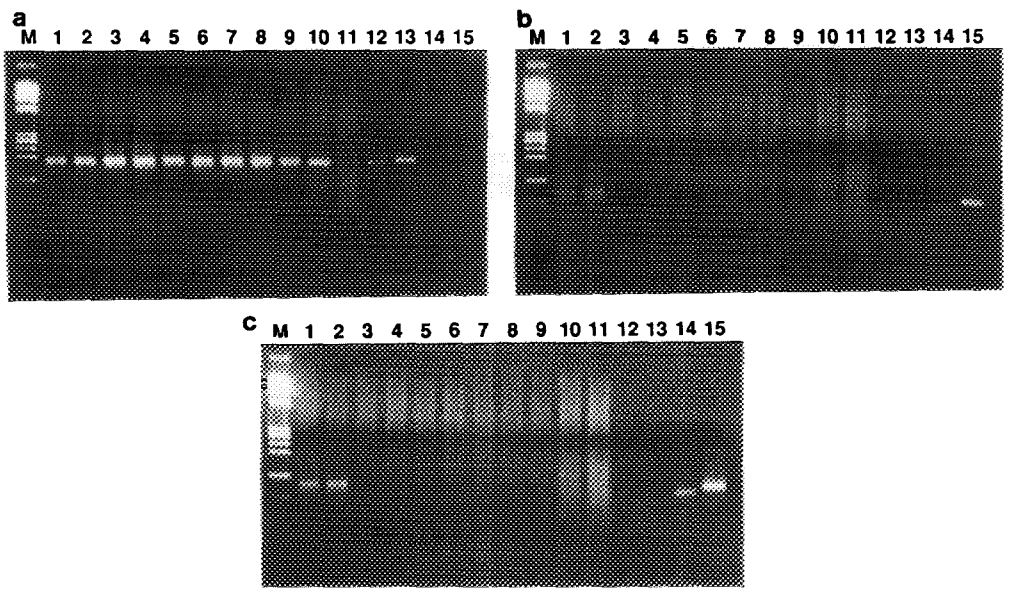

Fig. 2. Detection of TH mRNAs by the reverse transcription-PCR (RT-PCR). (a) Amplification of type-1 and -2 mRNAs with HS-4 and HA-3 used as primers. Species and tissues of RNAs are as follows: lane 1, human substantia nigra; 2, human adrenal gland; 3, macaque (*Macaca fuscata*: Japanese monkey) adrenal; 4, macaque (*Macaca fuscata*) substantia nigra; 5, macaque (*Macaca fuscata*) locus ceruleus; 6, macaque (*Macaca fascicularis*: crab-eating monkey) pons and medulla oblongata; 7, macaque (*Macaca fascicularis*) substantia nigra; 8, marmoset (*Callithrix jacchus*) adrenal; 9, marmoset (*Callithrix jacchus*) midbrain; 10, rat brain; 11 macaque (*Macaca fuscata*) liver; 12-15, human TH cRNAs (1 pg) types 1-4. (b) Amplification of type-3 and -4 mRNAs with HS-4 and HA-2 used as primers. (c) Amplification of type-3 and -4 mRNAs with HS-4 and HA-4 used as primers. Species and tissues of RNAs are the same as in Fig. 2a.

human tissues. RNAs from the macaque and marmoset tissues showed no bands. These results indicate that monkeys do not contain type-3 and -4 TH mRNAs, while the human brain and adrenal gland contain both TH isoforms. As shown later, the macaque has the same sequence as the complementary one of HA-2, and the same sequence except one-base in the middle portion as the complementary one of HA-4, in its genomic DNA. Thus we can deny that a mispairing of primers with cDNA templates is the cause of its inability to amplify types 3 and 4 mRNAs in macaques.

Next we examined genomic DNA sequences of man, chimpanzee, gorilla, orangutan, white-handed gibbon, macaque, and marmoset. Genomic DNA fragments around exon/intron junctions of intron 1 and intron 2 were amplified by a PCR technique using primers based on the sequence of the human TH gene that we previously reported (8).

Type-1 and -2 mRNAs are produced through a competition of two alternative 5' splice sites in the end of exon 1. Sequences around the alternative 5' splice sites were completely conserved among all primates examined (Fig. 3a), and were in good agreement with the consensus sequence (C/A)AG!GT(A/G)AGT (12). The 3' splice sites of the intron 2 were also well conserved [Fig. 3c], and well matched the consensus sequence $(C/T)_nN(C/T)AG!G$ (12). The 12-bp inserted sequences in type-2 TH were completely identical with that of the human TH. These observations confirmed the presence of type-1 and -2 mRNAs in the macaque and marmoset as shown by RT-PCR, and suggested the presence of both types of TH mRNAs in other apes.

Type-3 and -4 mRNAs are formed through the inclusion of exon 2, which is composed of 81 bases. We amplified the entire region of exon 2 (Fig. 3b).

When we compare the sequence of the macaque with that of man, we can note several short deletions, as shown by asterisks in Fig. 3b. At position 46, there was a one-base deletion. This deletion causes a frame-shift in the presumable exon 2 in the macaque, and results in generating a stop codon after 15 residues. This means that the macaque cannot synthesize active type-3 and -4 TH proteins, even if mRNAs corresponding to these types were produced.

At the 5' splice site of intron 2, the gorilla had a four-base duplication composed of a sequence, TAAG, similar to the human sequence. However, since the essential sequence at the 5' splice site, GT, which is located just before the duplication, has mutated to GA, the gorilla cannot process the pre-mRNA at the same position as man. There is a putative sequence for the splicing, A!GTAAGA, at just 4 bases downstream of the splice site in man. Although a splicing at this site may occur in the gorilla, active TH protein cannot be produced because of an alteration of the reading frame. Thus we can conclude that the gorilla also cannot produce TH proteins corresponding to type-3 and -4.

Nucleotide sequences in and around the presumable exon 2 of the chimpanzee, orangutan, and gibbon were very close to that of man (Fig. 3b). Although they do not have the 4-base (TAAG) duplication at the 5' splice site of

a

```
                 EXON 1                                    INTRON 1
            1                                                                               80
    man-JB  TCTGAGCTGG ACGCCAAGCA GGCAGAGGCC ATCATGGTAA GAGGGCAGGT AGGTGCCCGG CGGCCGCAGT GGACCGGAGC
       man  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- ---------- ---------- ----------
chimpanzee  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- ---------- ------T--- ----------
   gorilla  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- ---------- ---------- ----------
 orangutan  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- ---------- ---------- ----------
    gibbon  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- ---------- ----T----- ----------
   macaque  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- ---------- ----T----- ----G-----
  marmoset  TCTGAGCTGG ACGCCAAGCA G--------- ---------- ---------- --------A- G---TT---- ----------

                 INTRON 1
            81                                                                   153
    man-JB  CCAGGGCTGG TGCCAGCTGC CTCTGCTACT CCCCAGCCTG GCTGGCAGCC CCAGGCTCAG GGTCCATGCA AAC
       man  ---------- ---------- ---------- ---------- ---------- CCAGGCTCAG GGTCCATGCA AAC
chimpanzee  ---------- ---------- ------C--- ---------- ---------- CCAGGCTCAG GGTCCATGCA AAC
   gorilla  ---------- ---------- ------C--- ---------- ---------- CCAGGCTCAG GGTCCATGCA AAC
 orangutan  ---------- ---------- ------C--- -----AT--- ---------- CCAGGCTCAG GGTCCATGCA AAC
    gibbon  --------A- ---------- ------C--- -----A---- ---------- CCAGGCTCAG GGTCCATGCA AAC
   macaque  ---------- ---------- ---------- ---------- --C----A-- CCAGGCTCAG GGTCCATGCA AAC
  marmoset  ---------- -----CT--- -------G-- ---------A ---------T CCAGGCTCAG GGTCCATGCA AAC
```

b

```
                       INTRON 1                     EXON 2
            1                                                                               80
    man-JB  CCCCACCAGG CTCCCCATCA GGCATCCCCT CCCCAGGGCG CCCCGGGGCC CAGCCTCACA GGCTCTCCGT GGCCTGGAAC
       man     CACCAGG CTCCCCATCA GGCAT----- ---------- ---------- ---------- ---------- ----------
chimpanzee     CACCAGG CTCCCCATCA GGCAT----- --------G- ---------- ---------- ---------- ----------
   gorilla     CACCAGG CTCCCCATCA GGCAT----- --------G- ---------- ---------- -------T-- ----------
 orangutan     CACCAGG CTCCCCATCA GGCAT----- --------G- ---------- ---------- --------A- ----------
    gibbon     CACCAGG CTCCCCATCA GGCAT---G- --------G- A---A----- ---------- ---------- ----------
   macaque  CCCCACCAGG CTCCCCATCA G-****---- --------G- A---A*---- ---------G ---------- ----------

                       EXON 2                     INTRON 2
            81                                                                              160
    man-JB  TGCAGCCCCA GCTGCATCCT ACACCCCCAC CCCAAGGGTA AGTAAGAGGG GACTCTGGGA GGGGCTTCTG CTGCTCCCCT
       man  ---------- ---------- ---------- ---------- ---------- ---------- -----C---- ----------
chimpanzee  ---------- ---------- ---------G --------** **-------- ---------- ---------- ----------
   gorilla  ---------- ---------- ---------- -------A-- ---------- ---------- ---------- ----------
 orangutan  ---------- ---------- ---------- --------** **-------- --T------- ---------- ----------
    gibbon  ---------- ---------- ---------- ---G----** **-------- A--------- ---------- ----------
   macaque  --------T- ---------- --------G- --------** **-------C A----CA--- ---------- ----------

                       INTRON 2
            161                                                                             240
    man-JB  TCATGTTCCA CAACCCTGGA AGCTCAGGAT GAAGCTGATT CTTCTCTTAC AAGGGGCCCA GAGCCTTCTT GGGAGTTCAG
       man  ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---AGTTCAG
chimpanzee  -----C---- ---------- ---------- ---------- ---------- -C-------- ---------- ---AGTTCAG
   gorilla  ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---AGTTCAG
 orangutan  ---CA----- -G-------- ---------- ---------- ---------- ---------- --------C- ---AGTTCAG
    gibbon  ---CA----- -G-------- ---C------ ---------- ---------- ---------- ---------- --TAGTTCAG
   macaque  --G------- TG-------- --A-----G- ---------- ---------- ---A------ ---------- ---AGTTCAG

            241             256
    man-JB  CTCCAAGGGA TGAGCC
       man  CTCCAAGGGA TGAGCC
chimpanzee  CTCCAAGGGA TGAGCC
   gorilla  CTCCAAGGGA TGAGCC
 orangutan  CTCCAAGGGA TGAGCC
    gibbon  CTCCAAGGGA TGAGCC
   macaque  CTCCAAGGGA TGAGCC
```

Fig. 3. Nucleotide sequences of human, chimpanzee, gorilla, orangutan, gibbon, macaque, and marmoset TH genes. The exons/introns covered by the sequences are indicated. The human sequence at the top is quoted from Kobayashi *et al.* (8). Sequences of amplified DNA fragments from genomic DNAs of various primates are indicated below the human sequence. A dash (-) indicates an identical nucleotide with the top sequence, and an asterisk (*) indicates a gap in the sequence. Sequences of PCR primers used for the amplification are written in the figure with underlines. GT and AG sequences located at the 5' and 3' splice sites are indicated by double-underlines. (a) Sequences around alternative 5' splice sites of intron 1. (b) Sequences around exon 2 including 3' splice site of intron 1 and 5' splice site of intron 2. (c) Sequences around 3' splice site of intron 2.

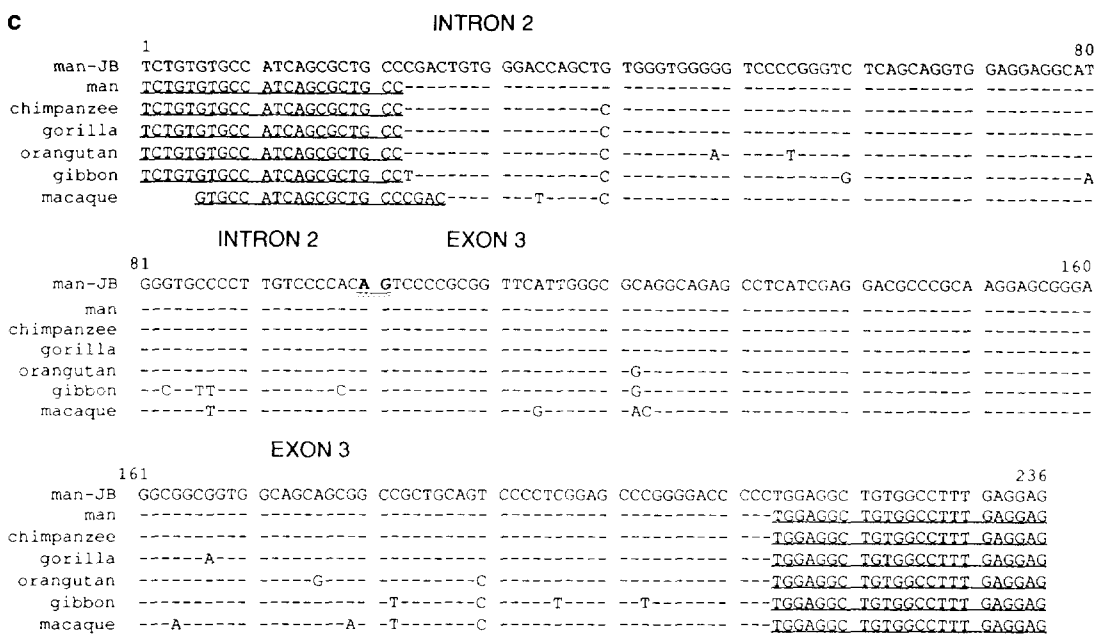

Fig. 3 - Continued

intron 2, alteration in the sequence of the 5' splice site is minor, from AGG!GTAAGT in man to AGG!GTAAGA in the chimpanzee, orangutan, and gibbon. Thus these higher apes except gorilla may have the capability to produce type-3 and -4 mRNAs. Direct analysis of mRNAs, however, would be required to determine the existence of these types.

We found the 4-base (TAAG) duplication at the 5' splice site of intron 2 only in man and gorilla, but not in the chimpanzee. Phylogenetic trees of hominoids suggest that the gibbon was split off from a common ancestor first, followed sequentially by the orangutan and gorilla, and that finally chimpanzee and man were separated, while distances between gorilla and man, and between chimpanzee and man are very close (13). Our observation of a 4-base duplication in gorilla and man, but not in the chimpanzee, might be explainable in either of two ways. One is that the duplication occurred independently in the human and gorilla lineages after the split of gorilla, chimpanzee, and man about 4-5 million years ago. The other is that the duplication occurred before the divergence of the three species, and that the four bases were deleted only in the chimpanzee. In any case, it would be of interest to investigate the time of the duplication.

There are two proposed mechanisms for the creation of two new isoforms. One mechanism is an exon/intron shuffling. Translocation of a gene can produce a new chimeric protein. The other mechanism is generation of a new splice site in an intron by mutation. Here we have demonstrated that a region that is not recognized as an exon in the macaque became an alternative exon in man. The sequence of exon 2, composed of 81 bases, shows no obvious homology with reported sequences. Whether the sequence in the macaque is derived from another region of the gene or

generated from mutations accumulated in an intron is not clear. O'Malley *et al.* (6) have suggested the existence of a sequence homologous to that of human exon 2 in the rat.

Generation of heterogeneity in the TH isoforms in primates may alter the *in vivo* concentrations of catecholamines. Accumulating evidence suggests that neurotransmitters themselves act as neurotrophic factors, and affect the growth of neurites (see review ref. 14). Alteration in the concentrations of catecholamines due to the presence of TH isoforms in primates may thus possibly affect the neural circuitry in the brain. Furthermore, cell bodies of dopamine and norepinephrine neurons contain neuromelanin, especially in primates. Neuromelanin is thought to be synthesized through a polymerization of 3,4-dihydroxyphenylalanine (DOPA), dopamine, or norepinephrine, which are products of TH. The heterogeneity of TH may be responsible for the presence of neuromelanin only in primates. The existence of neuromelanin in primates is thought to be a cause of the selective degeneration of the nigro-striatal dopaminergic neurons in Parkinson's disease (15).

Why is man "Man"? What are the crucial differences between primates and the other animals, or between man and the other primates? These are very interesting questions. The most obvious phenotypic difference between the two is intelligence. One might, therefore, expect to find some genetic differences in molecules involved in brain function. Of course, the heterogeneity of TH cannot account for all of the difference, but it may be one factor. Many investigations will be required in the future for answering these questions.

This is the first report, to our knowledge, about differences between brains of primates and of other mammals, and between monkeys or gorilla and man at the molecular level.

ACKNOWLEDGMENTS

We are grateful to Dr. Sidney Udenfriend (Roche Institute of Molecular Biology, Nutley, New Jersey, USA) for his valuable suggestions on this manuscript. We wish to thank Dr. Motoharu Hayashi of Kyoto University (Japan) for giving us monkey tissues. This work was supported by a grant-in-aid for Specially Promoted Research from the Ministry of Education, Science and Culture of Japan to T. N. and by a grant-in-aid from the Fujita Health University.

REFERENCES

1. Nagatsu, T., Levitt, M. and Udenfriend, S. (1964) *J. Biol. Chem.* **239**, 2910-2917.
2. Levitt, M., Spector, S., Sjoerdsma, A., and Udenfriend, S. (1965) *J. Pharmacol. Exp. Ther.* **148**, 1-8.
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. Le Bourdellès, B., Boularand, S., Boni, C., Horellou, P., Dumas, S., Grima, B. and Mallet, J. (1988) *J. Neurochem.* **50**, 988-991.
8. Kobayashi, K., Kaneda, N., Ichinose, H., Kishi, F., Nakazawa, A., Kurosawa, Y., Fujita, K. and Nagatsu, T. (1988) *J. Biochem.* **103**, 907-912.
9. Haycock, J. W. (1993) *J. Neurochem.* **60**, 493-502.
10. Iwata, N., Kobayashi, K., Sasaoka, T., Hidaka, H. and Nagatsu, T. (1992) *Biochem. Biophys. Res. Commun.* **182**, 348-354.
11. Ichikawa, S., Ichinose, H. and Nagatsu, T. (1990) *Biochem. Biophys. Res. Commun.* **173**, 1331-1336.
12. Green, M. R. (1991) *Annu. Rev. Cell Biol.* **7**, 559-599.
13. Ueda, S., Watanabe, Y., Saitou, N., Omoto, K., Hayashida, H., Miyata, T., Hisajima, H., and Honjo, T. (1989) *J. Mol. Biol.* **205**, 85-90.
14. Schwartz, J. P. (1992) *Int. Rev. Neurobiol.* **34**, 1-23.
15. Hirsch, E., Graybiel, A. M. and Agid, Y. A. (1988) *Nature* **334**, 345-348.