Genomic Structure and Expression of the Human Gene Encoding Cytochrome b$_{561}$, an Integral Protein of the Chromaffin Granule Membrane*

Meera Srivastava‡

From the Laboratory of Cell Biology and Genetics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

Cytochrome b$_{561}$ is an electron transfer protein unique to neuroendocrine secretory vesicles. The Southern blot hybridization shows that it is a single copy gene highly conserved throughout phylogeny. The transcription unit spans approximately 11 kilobases, and heterogeneous transcription sites are located 404 bases 5' to the translation initiation codon. The sequence of the 5'-flanking region is GC-rich and lacks a typical TATA box at the usual position. However, it has a CAAT sequence that is 132 and potential recognition sequences for several transcription factors including SP1, GR-PR-MMTV, AP4, gERE, J CV repeat, AP2, and NF-κB. Each of the five transmembrane segments are encoded by five consecutive exons. This corroborates the five-transmembrane model proposed for human, mouse, and Xenopus rather than six proposed for bovine. The cytochrome was found to be highly expressed in colon cancer cell lines, T cell lymphomas, and K-562 cell lines. However, in B-cell lymphomas such as Burkitt's and Daudi, the cytochrome b$_{561}$ expression was completely shut down. The results in this report are the first to demonstrate the structural organization and regulatory sequences of the cytochrome b$_{561}$ gene encoding an integral membrane protein of neuroendocrine storage vesicles of neurotransmitters and peptide hormones. Unexpected results on cytochrome b$_{561}$ expression in cells of lymphocytic origin and its complex regulation in tumor cells provide new insights into cytochrome b$_{561}$ gene regulation.

The integral membrane protein, cytochrome b$_{561}$ was first discovered in catecholamine storage granules (1) and was subsequently found to be specific to catecholamine and neuropeptide secretory vesicles of the adrenal medulla, pituitary gland, and other neuroendocrine tissues (2-8). This 30-kDa protein comprises 15% of the granule membrane protein (9) and is present in both small synaptic vesicles and large dense core vesicles (10, 11) where it plays a central role with ascorbic acid in the biosynthesis of several catecholamine (6, 12) and peptide neurotransmitters (13, 14). Among the cytochromes, cytochrome b$_{561}$ is unique in its localization to neuroendocrine tissues and appears to be unusual in its enzymatic mechanisms. Unlike other electron carriers which need other proteins for its oxidation-reduction cycle, the cytochrome b$_{561}$ does not interact with any other protein. Instead, ascorbic acid serves as the extravesicular electron donor for cytochrome b$_{561}$, and semidehydroascorbic acid acts as the intravesicular electron acceptor, which replenishes the high intravesicular ascorbic acid concentrations. Ascorbic acid serves as a cofactor for intravesicular dopamine β-hydroxylase activity in catecholamine storage granules (15) and peptidyl α-amidating monooxygenase activity in neuropeptide storage vesicles (14). Cytochrome b$_{561}$ appears to be a simple, functionally symmetric electron transfer protein and acts as a pure electron channel.

Understanding the mechanism of this long range electron transfer process has been facilitated by the cloning of bovine cytochrome b$_{561}$ which revealed that it has six transmembrane helices with N- and C-terminal ends of the protein facing the cytoplasm (16). Subsequent biochemical analysis using peptide antibodies raised against N-terminal and C-terminal ends of the protein and protease digestion suggested that the N terminus was unavailable to both (17). In addition, the human cytochrome b$_{561}$ lacked the first putative 22 amino acids at the N terminus predicted from the bovine sequence to be in the cytoplasmic side. Based on the sequence comparison and low hydrophobicity of the second transmembrane domain, combined with the biochemical evidence, a revised hypothesis was proposed in which the gene product is likely to be distributed within the membrane with five transmembrane helices (18). Subsequent cloning of the mouse and Xenopus cytochrome b$_{561}$ confirmed this structure and further revealed that it was colocalized to specific tissues with either DBH or peptidylglycine α-amidating monoxygenase. In addition, we found that cytochrome b$_{561}$ expression was developmentally regulated in mouse and Xenopus. Furthermore, more recent data on cytochrome b$_{561}$ expression in different tissues exhibited uniform distribution in all the neuroendocrine tissues tested.

In order to better understand the molecular basis of developmental and tissue-specific cytochrome b$_{561}$ expression, it is necessary to characterize the genomic structure and regulatory elements of the gene. This approach could also help establish the relationship between this gene and those of other membrane proteins, thereby providing clues concerning the evolution of the eukaryotic chromaffin granule. Cloning and sequence analysis of the regulatory region of the cytochrome b$_{561}$ gene would provide reagents and information necessary to examine the expression of a major structural protein of the chromaffin granule. For these reasons, we have determined the structural organization and promoter sequence of human cytochrome b$_{561}$, and studied gene expression in different cell lineages.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: LCBG, NIDDK, NIH, Bldg. 8, Rm. 405, Bethesda, MD 20892. Tel.: 301-594-5185; Fax: 301-402-3299.

‡ To whom correspondence should be addressed: LCBG, NIDDK, NIH, Bldg. 8, Rm. 405, 9000 Rockville Pike, Bethesda, MD 20892. Tel.: 301-594-5185; Fax: 301-402-3299.

† M. Srivastava, C. Flores, J. Oldes, H. B. Pollard, and P. J. Fleming, submitted for publication.

‡ M. Srivastava, submitted for publication.

Received for publication, April 19, 1995, and in revised form, July 11, 1995}

Printed in U.S.A.
Materials and Methods

Library Screening—A genomic DNA library from human peripheral blood leukocytes in Lambda DASH (Stratagene) was plated, and replicas containing 1 \times 10^6 clones were screened using the full-length human cytochrome b_{561} cDNA from the HCR1 A clone (18), which had been radiolabeled by nick translation (Life Technologies, Inc.). Hybridization was performed at 42 °C in 6x SSC, 5x Denhardt’s solution, 0.1% SDS, 50% formamide, and 0.1 mg/ml salmon sperm DNA. Positive cell colonies were washed at 65 °C in 1x SSC and 0.1% SDS and exposed to XAR-5 film (Kodak) with an intensifying screen at -80 °C. λ-DNA was purified using a Qiagen kit (Qiagen, Inc.). Four positive colonies were sequentially plaque-purified, analyzed, and compared to each other by a combination of restriction mapping and Southern blot analysis. Hybridization with different labeled oligonucleotides corresponding to various regions of the cDNA led to the identification of two clones containing the entire cytochrome b_{561} gene.

Determination of the 5'-Flanking and Exon 1 Sequences—The 6.5 kb XbaI fragment hybridized with the 5'-region of the cDNA was isolated and digested with XhoI, and the 2.8-kb, 2.3-kb, and 1.4-kb fragments were subcloned into the pGX2627 vector. These subclones were sequenced using either two primers homologous to sequences flanking the multiple cloning site of pGX2627 or 24 oligonucleotides based on cDNA or derived genomic sequences. The entire gene and the flanking sequences were sequenced by Sanger’s dideoxy method (21) using cloned T7 DNA polymerase (Sequenase, U.S. Biochemical Corp.). Oligonucleotides were synthesized in an Applied Biosystems DNA synthesizer (Model 380B). DNA sequencing was performed with the PCGene program (IntelliGenetics) in order to obtain the 5'-flanking sequences and the first exon sequences. The 5'-flanking sequences were analyzed for regulatory elements using a computer search routine.

PCR Determination of Exon 2-5 and Intron Sizes—Using different sets of primers (Table I) and λ-DNA as template, the polymerase chain reaction was performed as described (22) using the GeneAmp kit (Perkin Elmer). With the following parameters: initial denaturation at 94 °C for 2 min followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 50 °C for 2 min, and extension at 72 °C for 3 min with a final extension at 72 °C for 7 min. Amplified products were electrophoresed on agarose gels, gel-purified using the GeneClean Kit (Bio 101), and subcloned into the TA cloning vector (Invitrogen) according to the manufacturer’s instructions. The plasmids containing the genomic fragments were selected by restriction endonuclease digestion and were sequenced using the M13 universal sequencing primer at the 5'-end and the T7 promoter sequence to obtain the 3'-end of the genomic fragment’s sequence. Some of the fragments containing the exon sequences were further sequenced using the oligonucleotides derived from the cDNA sequences to confirm the cDNA sequence, as well as to find if there were any small intron sequences in between the exons. The sizes of the introns were determined by comparing the PCR amplified products with the migrations of the DNA fragments of known sizes. The combination of these procedures identified five introns.

Mapping of the Transcription Start Sites—The ribonuclease protection assay was performed with a run-off transcript initiated on the 5'- upstream region and less than 200 bp of all introns are shown in Fig. 2. Exons flanked by introns ranged in size from 99 bases (exon 2) to 1374 bases (exon 5) (Fig. 2, Table I). The intron-exon junction sequences are consistent with the reported 5'-donor (GTGAG) and the 3'-acceptor (CTAG) consensus sequence (21) (Table I). The intron splice phase is type 0 (the intron occurs between codons) for intron 3, type 1 (the intron interrupts the first and the second bases of the codon) for introns 1 and 2, and type 2 (the intron interrupts the second and the third base of the codon) for intron 4 (22).

Sequence Determination and Exon-Intron Organization—Through a combination of restriction mapping, PCR, and sequence analysis, the human cytochrome b_{561} gene was determined to contain 6 exons interrupted by 5 introns, defining a gene of approximately 11 kb in size. Fig. 1 shows the organization of the cytochrome b_{561} gene in relation to the cDNA sequence and transmembrane domains. The XbaI fragment of the λ clone identified with the primer from exon 1 was subcloned and sequenced. The DNA sequence of the rest of the exons, the exon/intron splice junctions and the introns were obtained using the λ clones as templates and the oligonucleotide primer sets A → H listed in Table I. The location of each splice junction is shown in Fig. 1. The length of each intron was then determined by PCR using pairs of oligonucleotides from the exons flanking each intron, followed by analysis on agarose gels.

Partial nucleotide sequences, including the sequence of the 5'-upstream region and less than 200 bp of all introns are shown in Table I. The intron-exon junction sequences are consistent with the reported 5'-donor (GTGAG) and the 3'-acceptor (CTAG) consensus sequence (21) (Table I). The intron splice phase is type 0 (the intron occurs between codons) for intron 3, type 1 (the intron interrupts the first and the second bases of the codon) for introns 1 and 2, and type 2 (the intron interrupts the second and the third base of the codon) for intron 4 (22).

Relationship of Splice junctions to Protein Domains—Exons often represent the functional units of the protein. Indeed, the first, fourth, and fifth transmembrane domains are entirely within exons 1, 4, and 5, respectively. However, the second and third transmembrane domain sequences are interrupted at the end by introns 2 and 3. Exon 1 has the mRNA leader sequence and the first transmembrane domain. Exon 2 contains the conserved His-91 and the putative cyclic AMP binding site and two protein kinase C (PKC) sites. Exon 4 has the conserved His-160 which ligands to the heme moiety and is adjacent to the heme-binding hydrophobic pocket. Exon 5 contains the fifth transmembrane domain with one (PKC) site. Hence, the five transmembrane segments of the protein are structurally dis-
cretes and apparently have different functional roles. Because this is the first report on the genomic sequence of cytochrome \( b_{561} \), from any species to be cloned, the degree of evolutionary conservation of the cytochrome \( b_{561} \) gene organization is yet to be learned.

Repetitive DNA—The human cytochrome \( b_{561} \) gene was examined for the presence of Alu repeats. Two complete Alu repeats were identified with 82% and 93% identity to published Alu consensus sequences (25). They are found at position 2986 in the 5'-flanking sequence and at position 8051 in intron 5. The dinucleotide repeat CACACACACACA was found at position 2949 and at position 3276 which might be a useful microsatellite marker for mapping the cytochrome \( b_{561} \) gene defects linked to a specific disease on chromosome 17.

Mapping the Transcription Initiation Site—Primer extension analysis and reverse transcriptase-PCR experiments using human brain mRNA identified two major transcription sites and two minor transcription sites (18). To confirm the validity of the above analysis, a ribonuclease protection assay was carried out utilizing human brain RNA and a 415-base pair RNA probe that was generated using the T7 promoter (see "Materials and Methods"). Corresponding to the primer extension study, two protected RNA species of approximately 131 and 139 nucleotides were detected which were not observed in the lane containing yeast tRNA control. Two minor bands were also observed by this assay, suggesting that additional transcripts of the cytochrome \( b_{561} \) gene might be expressed although at much lower levels (Fig. 3). This result was confirmed by replicate analysis.

Analysis of the 5'-Flanking Sequences—To identify sequences that potentially control cytochrome \( b_{561} \) gene expression, a dynamic computer program (26) was utilized to locate the regulatory elements. The region 5' to the identified transcription start sites is GC-rich and contains a possible CAAT box at ~132 nucleotides and potential recognition sequences for several transcription factors (Fig. 4). However, there is no TATA-like element in close proximity. An examination of the 5'-flanking sequences revealed the presence of six \( C^3 \) repeats, six SP-1, two AP-2, MRE, GR-PR-MMTV, AP-4, PEA-3, E-alpha H box, and HTNF-1 hist, all within 1000 nucleotides of the transcription start site. Imperfect sequences, which only closely resemble the reported binding sites, have not been listed in Fig. 4. However, they might have a potential regulatory function, as it has been shown that many binding sites deviate to a certain extent from reported consensus sequences.

Differential Expression of Cytochrome \( b_{561} \) in Cancer Cell Lines—Despite extensive studies on the ubiquitous expression of cytochrome \( b_{561} \) in normal neuroendocrine tissues, little is known about its expression in tumor tissues or regulation of cytochrome \( b_{561} \) expression underlying its induction or suppression at the molecular level, which will provide the molecular basis of the physiological function of cytochrome \( b_{561} \).

Therefore, Northern blot analyses were performed using poly(A) RNA from tumor cell lines derived from melanoma (G 361), lung carcinoma (A 549), colorectal carcinoma (SW 480), Burkitt’s lymphoma (Raji) T cell lymphoblastic leukemia cell line (MOLT-4) and chronic myelogenous leukemia cell line (K-562), HeLa cell, 53 and promyelocytic leukemia (HL-60), and peripheral blood leukocyte as normal cells. The colon carcinoma cells (SW 480), T cell lymphoma (HL-60) and K-562, expressed high levels of cytochrome \( b_{561} \) (Fig. 5A, a) compared to peripheral blood leukocytes (data not shown). Again in Burkitt’s which is a B cell lymphoma cell type, the cytochrome \( b_{561} \) expression was not detected, although the amount of mRNA was almost same in all the cell lines tested using actin as the probe (Fig. 5A, b). To further verify the results obtained by Northern blot analysis, reverse transcriptase-PCR analysis was carried out using mRNA from diverse tumor cell lines.
lines and tumor tissues, including HL-60, HeLa, Daudi, K-562, melanoma, aorta, ovary, Burkitt's, colon, lung, and breast.

Using primers from the coding region of human cytochrome b561, the expected 727-bp band was detected in all the cancer cell lines except Burkitt's and Daudi where the cytochrome b561 expression was completely absent (Fig. 5B, a). In order to verify this observation further, an oligonucleotide from the internal sequence was end-labeled with 32P and used for Southern blot hybridization. The results in Fig. 5B, b, also show that there is no signal detected in Burkitt's and Daudi cell lines, although the amount of mRNA used was the same as confirmed by the amplification of mRNAs from all the cancer lines using primers from a different gene (Fig. 5B, c). This limited analysis suggests that expression of cytochrome b561 can be selective in hematopoietic cells.

**DISCUSSION**

To understand the molecular basis of the tissue-specific expression of cytochrome b561, we have now cloned and charac-

**Fig. 2.** Partial nucleotide sequence of the human cytochrome b561 gene. Nucleotides in exons are shown in bold letters with the deduced amino acid shown below. The transcription initiation sites and the end of the human cytochrome b561 mRNA is shown in bold type uppercase. The asterisks indicate the two strong transcription initiation sites. Numbering of the nucleotides begins at 4572 bp before the start of the transcription site and is located on the left side, whereas the amino acids are numbered under the protein sequence and on the right side. Intron boundaries having the consensus GT and AG dinucleotides are underlined. Also, two direct Alu repeats (positions 2986 and 8051), ATG, stop codon, and two dinucleotide CA repeats are underlined.

[Image of nucleotide sequence and protein sequence]
terized the genomic sequence of cytochrome \textit{b} \textsubscript{561} including the promoter sequences possibly regulating the specific expression of cytochrome \textit{b} \textsubscript{561}. Although the structure, function, and expression of cytochrome \textit{b} \textsubscript{561} protein has been studied in detail, the biological significance of its uniform expression in neuroendocrine tissues has not been understood. The results described here, therefore, are of particular significance in as much as they represent first characterization of the gene encoding the integral membrane protein of the chromaffin granule and its differential expression in hematopoietic cells.

Southern hybridization analysis established that cytochrome \textit{b} \textsubscript{561} gene is highly conserved (data reviewed, but not included) and occurs as a single copy of the cytochrome \textit{b} \textsubscript{561} gene in the human genome, spanning approximately 11 kb, and is split into five exons. The coding sequence of the gene is 753 bp, representing 7.1\% of the gene, and perfect agreement between the nucleotide sequences of the chromosomal gene exons and the cDNA sequences was observed. The introns interrupt the cytochrome \textit{b} \textsubscript{561} protein coding sequence in such a way that many of the protein segments are revealed as products of individual exons. A similar correspondence has been observed in other highly hydrophobic proteins such as human myelin (27), human T cell-specific proteolipid protein (28), and human vacuolar ATPase (29). Studies on the structure of the bovine cytochrome \textit{b} \textsubscript{561} have been interpreted to indicate that the protein might traverse the membrane six times with N- and C-terminal membrane-spanning domains.

### Table II

| Exon size (bp) | Exon-intron Length (bp) | Intron-exon |
|---------------|-------------------------|-------------|
| 608           | gttgagt (1275) ggacag   | 609         |
| 608           | AAATG                   | CCCCTG      |
| 707           | yAsnA                   | lal eu      |
| 99            | gttgagt (270) gcccag    | 708         |
| 811           | TGGTGT                  | GCTTG       |
| 104           | gttgagt (350) tcctag    | 812         |
| 158           | gttgagt (440) tcctag    | TryLe       |
| 1374          | gttct (1100) ggacag     | 970         |
|               | ACCC                   | GGGCA       |
|               | 2665                   | yGl yL      |
|               | ATCTG                  | 2344        |
|               |                        | GCTTC       |

**Fig. 3.** Site of transcription initiation as determined by ribonuclease protection analysis. Human total brain RNA (20 μg) or yeast tRNA was annealed to a \textsuperscript{32}P-labeled antisense RNA probe (prepared by T7 polymerase transcription) and digested with RNase A and T1 as described under "Materials and Methods." The probe fragments protected by total RNA from human brain was analyzed on 6% polyacrylamide-8 M urea gel electrophoreses and were compared with the adjacent sequencing ladder.

**Fig. 4.** Nucleotide sequence of the 5'-flanking sequence of the human cytochrome \textit{b} \textsubscript{561}. Amino acids encoded by the first exon are indicated below the appropriate codons. The transcription start sites are indicated in bold, and the last one is numbered 1. Numbers to the left refer to the nucleotide position relative to the transcription initiation site. Sequences similar to consensus sequences for transcription factors (e.g. SP-1, AP-2, AP-4, GR-PR, MMTV, NF-κB, JCV repeat, PEA-3, H4TF-1, hist, E-alpha box, and HinF-hist) are underlined.
C-terminal ends facing the cytoplasm. However, subsequent biochemical analysis using antibodies raised against N- and C-terminal peptides along with topological orientation suggested that N-terminal is not facing the cytoplasm. Comparative amino acid sequences derived from human, mouse, and Xenopus, along with low hydrophobicity and unusual amino acid composition of the second transmembrane segment, suggested that bovine is 22 amino acids shorter, and a model with acid composition of the second transmembrane segment, suggested that N-terminal is not facing the cytoplasm. Comparative amino acid sequences derived from human, mouse, and Xenopus, along with low hydrophobicity and unusual amino acid composition of the second transmembrane segment, suggested that bovine is 22 amino acids shorter, and a model with acid composition of the second transmembrane segment, suggested that N-terminal is not facing the cytoplasm.

Fig. 5. A, a, Northern blot analysis of poly(A)⁺ RNA from different cancer lines and tissues from human. The blot was hybridized to the human cytochrome b₅₆₁ coding region. A, b, the same blot was hybridized to β-actin after stripping. B, reverse transcriptase-PCR and Southern blot analysis of human cytochrome b₅₆₁ mRNA. Poly(A)⁺ RNA from different cancer cell lines was reverse-transcribed and run on a 1.5% agarose gel followed by blotting onto nitrocellulose paper (see "Materials and Methods"). B, a, agarose gel electrophoresis of reverse transcriptase-PCR amplified mRNA from different cancer lines from human. B, b, agarose gel electrophoresis of reverse transcriptase-PCR amplified mRNA from cancer cell lines using oligonucleotide primers from a different gene. B, c, Southern blot of the same gel probed with ³²P-labeled antisense primer from set D in Table I. Control lane has the markers.

Acknowledgments—I am indebted to Dr. Harvey B. Pollard for his valuable advice and critical evaluation of the manuscript and for his support during the course of this work. I also thank Dr. Patrick J. Fleming for his helpful suggestions.

REFERENCES
1. Spiro, M. J., and Ball, E. G. (1961) J. Biol. Chem. 236, 225–30
2. Duong, L. T., and Fleming, P. J. (1982) J. Biol. Chem. 257, 8561–8564
3. Pruss, R. M., and Shephard, E. A. (1987) Neuroscience 22, 149–157
4. Hertngale, H., Winkler, H., and Lohs, H. (1973) J. Neurochem. 20, 977–985
