roc. Natl. Acad. Sci. USA
Vol. 86, pp. 2587–2591, April 1989
Biochemistry

Characterization of the human 5-lipoxygenase gene

(DNA sequence/intron/exon junctions/promoter/protein domains)

COLIN D. FUNK, SHIGERU HOSHIKO, TAKASHI MATSUMOTO, OLOF RÅDMARK, AND BENGT SAMUELSSON

Department of Physiological Chemistry, Karolinska Institutet, S-104 01 Stockholm, Sweden

Contributed by Bengt Samuelsson, December 27, 1988

ABSTRACT **The gene for human 5-lipoxygenase has been isolated from three different bacteriophage genomic libraries and a genomic cosmid library. The gene spans >82 kilobases and consists of 14 exons. The size range for the exons is 82–613 base pairs, whereas that for the introns is ≈200 bp to >26 kb. A major site of transcription initiation in leukocytes was mapped to a thymidine residue 65 base pairs upstream of the ATG initiation codon by nuclease S1 protection and primer extension experiments. Other potential minor initiation sites were found. The putative promoter region contains no TATA and CCAAT sequences in the expected positions upstream of the major transcription initiation site but contains multiple GC boxes within a (G + C)-rich region, as does the immediate 5′ region of the first intron. Characteristics common to the 5′ end of the human 5-lipoxygenase gene and the promoter regions of the housekeeping genes raise important questions concerning the regulation of 5-lipoxygenase gene expression.**

The enzyme 5-lipoxygenase (arachidonate:oxygen 5-oxidoreductase, EC 1.13.11.34) catalyzes the formation of 5-hydroperoxy-6,8,11,14-icosatetraenoic acid (5-HPETE) from arachidonic acid as well as the subsequent conversion of 5-HPETE to 5,6-oxido-7,9,11,14-icosatetraenoic acid (leukotriene A_4, LTA_4) (1–4). LTA_4 is metabolized further to the various leukotrienes that have biological activities related to immediate hypersensitivity and inflammation (5, 6).

5-Lipoxygenase has been purified from a number of sources, including human (1, 7) and porcine (3) leukocytes, rat basophilic leukemia cells (4, 8), and mouse mastocytoma cells (2). Ca^{2+} and ATP are required for maximal enzyme activity (1–4, 7). The human enzyme is also stimulated by certain unidentified cytosolic and membrane-bound proteins (7, 9). 5-Lipoxygenase is a cytosolic enzyme; however, it was shown that in the presence of Ca^{2+} there is a reversible membrane association of the enzyme (10). The membrane binding was proposed to be involved in the activation of 5-lipoxygenase (11). Cellular stimulation of leukotriene synthesis was recently shown to result in a shift of the soluble 5-lipoxygenase to a membrane-bound form (12, 13).

Recently, cDNAs for human placenta (14), differentiated HL-60 cell (15), and rat basophilic leukemia cell (16) 5-lipoxygenase have been isolated. The mature human enzyme consists of 673 amino acids (14, 15). A search of the sequence for the "EF-hand" calcium-binding domain found in many calcium-binding proteins failed to find any significant homologies (14). However, some homology, albeit weak, to a 17-amino acid consensus sequence of a group of Ca^{2+}-dependent membrane-binding proteins was observed (15, 16). No ATP-binding sites were easily predicted from the primary and secondary structural data (14, 15).

5-Lipoxygenase appears to be expressed primarily in cells of myeloid lineage. The regulation of this tissue-specific 5-lipoxygenase gene expression is poorly understood. To address this situation we have begun characterization of the human 5-lipoxygenase gene. Clones spanning the gene and putative promoter region have been isolated and characterized.*

MATERIALS AND METHODS

Screening of Human Genomic DNA Libraries. Four genomic DNA libraries were screened in these studies and are as follows: (*i*) a human leukocyte library in the cloning vector EMBL-3 (Clontech); (*ii*) a human leukemia cell library in the cloning vector EMBL-4 (gift from J. Sumegi); (*iii*) a human fetal liver library in the cloning vector λCharon 4A (17) (obtained from the American Type Culture Collection); and (*iv*) a human fetal liver library in the cosmid cloning vector pHC79-2cos/tk (18) (gift from J. Collins). Screening was carried out with various 5-lipoxygenase cDNA probes (14) labeled with ^{32}P by the random priming method (19). Screening, hybridization, and washing procedures were carried out as described (20) and by standard methods (21, 22).

Clone Characterization. DNA from phage and cosmid clones and from specific subcloned fragments was characterized by standard restriction endonuclease mapping and Southern blot analysis (23). Phage clones 1x9A, 1x15A, 1x22A, and 1x27A were also mapped by the method of Rackwitz *et al.* (24). DNA sequencing of specific restriction fragments subcloned into phage M13 vectors (25) was performed by the chain-termination method (26).

Characterization of the 5′ End of Human Leukocyte mRNA. For nuclease S1 protection analysis, the 307-nucleotide *Pst* I/*Bst*EII fragment (−292 to +15) of the λ1x12A phage DNA was labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP at the *Bst*EII site. The purified labeled fragment was hybridized to total and poly(A)$^+$ RNA from human leukocytes, and nuclease S1 analysis was carried out essentially according to the method of Berk and Sharp (27).

For primer extension analysis, the plasmid, pPC50, containing the changed 5-lipoxygenase cDNA sequence at position −2 (dC to dG; unpublished data) was first digested with *Nae* I, 5′ end-labeled, and subsequently digested with *Sph* I. The 49-nucleotide fragment (−1 to +48) was purified by polyacrylamide gel electrophoresis and hybridized with total and poly(A)$^+$ RNA from human leukocytes for 4 hr at 42°C in 30 μl of hybridization buffer (50 mM Pipes, pH 6.4/0.1 mM EDTA/0.4 M NaCl/80% formamide). After precipitation, the hybridized primer/RNA was dissolved in 20 μl of reaction buffer (50 mM Tris·HCl, pH 8.0/4 mM $MgCl_2$/10 mM dithiothreitol/1 mg of gelatin per ml), 1 μl of dNTPs and 4 units of murine Moloney leukemia virus reverse transcriptase were added, and the mixture was incubated for 30 min at 37°C. The extracted products were analyzed by electrophoresis in an 8% polyacrylamide/7.5 M urea gel.

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

*The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04520).

RESULTS

Isolation and Characterization of 5-Lipoxygenase Genomic Clones. Four genomic libraries were screened with several different 5-lipoxygenase cDNA probes. A map of the 5-lipoxygenase gene was constructed from the overlapping genomic DNA fragments (Fig. 1). Restriction fragments from total leukocyte DNA that were detected by various ^{32}P-labeled 5-lipoxygenase cDNA probes in Southern blots (data not shown) could be accounted for in this map, suggesting that the locus exists as a single copy. The gene consists of 14 exons distributed over >82 kb. Clones λ1x12A (exon 1 and part of exon 2), λ1x15A (exons 5–7), λ1x9A (exons 7 and 8), λ1x17A (exons 9–14), and λ1x3A (exon 14) were obtained from a human leukocyte library. From a human leukemia cell library two additional clones were isolated, λ1x22A (exons 1 and 2) and λ1x27A (exon 2), which extended further downstream from λ1x12A. Further screening with a third library, a human fetal liver library, yielded two more clones, λ1x43A (exon 2) and λ1x70C1 (exon 4). Since it was evident from screening the three genomic phage libraries that the 5-lipoxygenase gene was very large and that segments were still missing we turned to a human cosmid library to isolate further clones. Two clones, clx5B-1 (exons 2 and 3) and clx2C-1 (exons 4–10), were isolated and mapped.

Partial Sequence and Organization of the Human 5-Lipoxygenase Gene. The intron/exon boundaries and all exons were sequenced (see Fig. 2). All introns conform to the GT-AG rule (28), and surrounding sequences are closely related to the consensus sequences surrounding splice junctions (29). The exons ranged in size from 82 bp (exon 3) to 613 bp (exon 14). All nucleotides in the coding region for human 5-lipoxygenase are the same in the genomic and cDNA clones (14, 15). Two minor differences involving extra base insertions (thymine at position 2188 and guanine at position 2240; numbering assignments according to ref. 14) are observed in the 3′ noncoding region of the cDNA sequence of Dixon *et al.* (15), on the one hand, and the genomic sequence and our previously published cDNA sequence (14), on the other hand. A few differences in the 5′ noncoding region of exon 1 and the cDNA sequences are also observed (see below).

Introns ranged in size from 192 bp (intron K) to >26 kb (intron C). Even after screening four genomic libraries, a fragment of unknown length within intron C remains uncloned.

The poly(A) signal AATAAA (30) was found in exon 14 of clones λ1x3A and λ1x17A and was followed by 15 nucleotides that corresponded to the cDNA sequence, including four adenosine residues at the 3′ end. Thereafter, the sequence diverged from the poly(A) stretch of the cDNA. Twenty-nine base pairs after the poly(A) signal the sequence TGTGTTAT appears, which closely resembles the consensus YGTGTTYY sequence found in many mammalian genes and is thought to be functionally important (31, 32).

The 5′ End of the Human 5-Lipoxygenase Gene. The nucleotide sequence of a 532-bp region upstream from the ATG start codon, encompassing the putative promoter, was determined (Fig. 3). The region closest to the initiation codon is very (G + C)-rich (≈80%) and contains several interesting features. There are six tandem repeats of the sequence CGGGGG, two repeats of CCCGCCC, and four CCGGG sequences. Regions from eight of these (G + C)-rich areas correspond to the core consensus for Sp1 binding (GGGCGG or CCGCCC) (33, 34). Three additional "GC boxes" occur within intron A close to the junction boundary (Fig. 3). Neither a "TATA box" nor a "CCAAT box" exists in the putative promoter region, within 250 bp of the transcription initiation sites (see below). An 11-bp inverted repeat is found at positions −26 to −36 and −141 to −131.

We had previously sequenced 34 bp of 5′ noncoding DNA from a placenta 5-lipoxygenase cDNA clone (14). All nucleotides, except the three end nucleotides (−32 to −34), are the same in the genomic DNA, which suggests that there was a cloning artifact at the 5′ end of the cDNA clone. There are three differences with the other reported human 5-lipoxygenase cDNA sequence (15) (deletion of guanine at −11, cytosine instead of guanine at −27, and insertion of an extra cytosine at −30; numbering assignments according to Fig. 3 and ref. 14).

The sites of transcription initiation in leukocytes were determined by nuclease S1 protection and primer extension experiments (Fig. 4). A major initiation site at a thymidine residue (position −65) was found by both techniques. Other potential minor initiation sites determined by nuclease S1 mapping and primer extension analysis occurred at positions −107, −97, −89, and −62 and at positions −66, −35, and −34, respectively (Figs. 3 and 4).

DISCUSSION

Gene Organization and Putative Protein Domains. The structure and organization of the human 5-lipoxygenase gene have been investigated in the present study. The gene consists of 14 exons divided by 13 introns (Fig. 2 and Fig. 1) and is relatively long (>82 kb) considering the length of the coding sequence. There are 5 introns with lengths of >8 kb. Exons 1–7, encoding the amino-terminal half of 5-lipoxygenase, are spread out over >65 kb, whereas the carboxyl-terminal encoding exons (exons 8–14) are clustered in a 6-kb segment of DNA. At present it is difficult to say if certain exons correspond to structural or functional domains of 5-lipoxygenase. Putative Ca^{2+}- and ATP-binding sites could not be easily predicted from the cDNA sequence (14, 15). However, exon 7, which encodes amino acids 278–326, corresponds exactly to one of the most hydrophobic segments of 5-lipoxygenase (14). Thus, this segment could represent an important domain structure.

The homology of 5-lipoxygenase to soybean lipoxygenase isozyme 1 (35) has been previously recognized (15). More recently, rat basophilic leukemia cell 5-lipoxygenase (16), soybean lipoxygenase isozymes 2 (36) and 3 (37), and pea

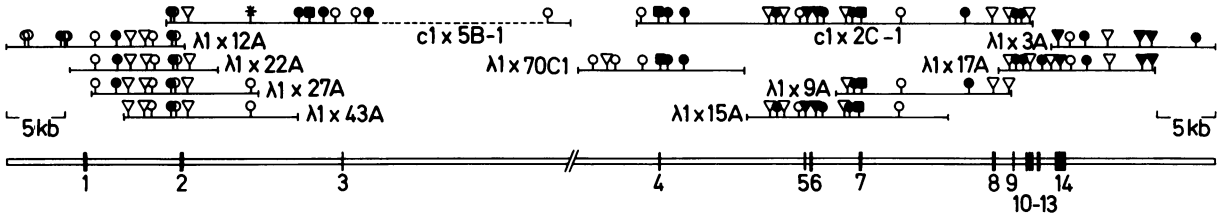

FIG. 1. The human 5-lipoxygenase gene. The gene locus is represented by the long box and exons are shown as black rectangles. Inserts from recombinant phage and cosmid clones are shown at the top along with the restriction sites for *Bam*HI (♦), *Eco*RI (♀), *Sma* I (▽), *Xho* I (▼), and *Cla* I (▮; *Cla* I mapping of clones λ1x17A and λ1x3A was not performed). The dotted line within clone clx5B-1 contains two *Bam*HI sites and four *Eco*RI sites that were not mapped. The asterisk above the site of the clx5B-1 clone indicates a potential restriction site polymorphism. kb, Kilobases.

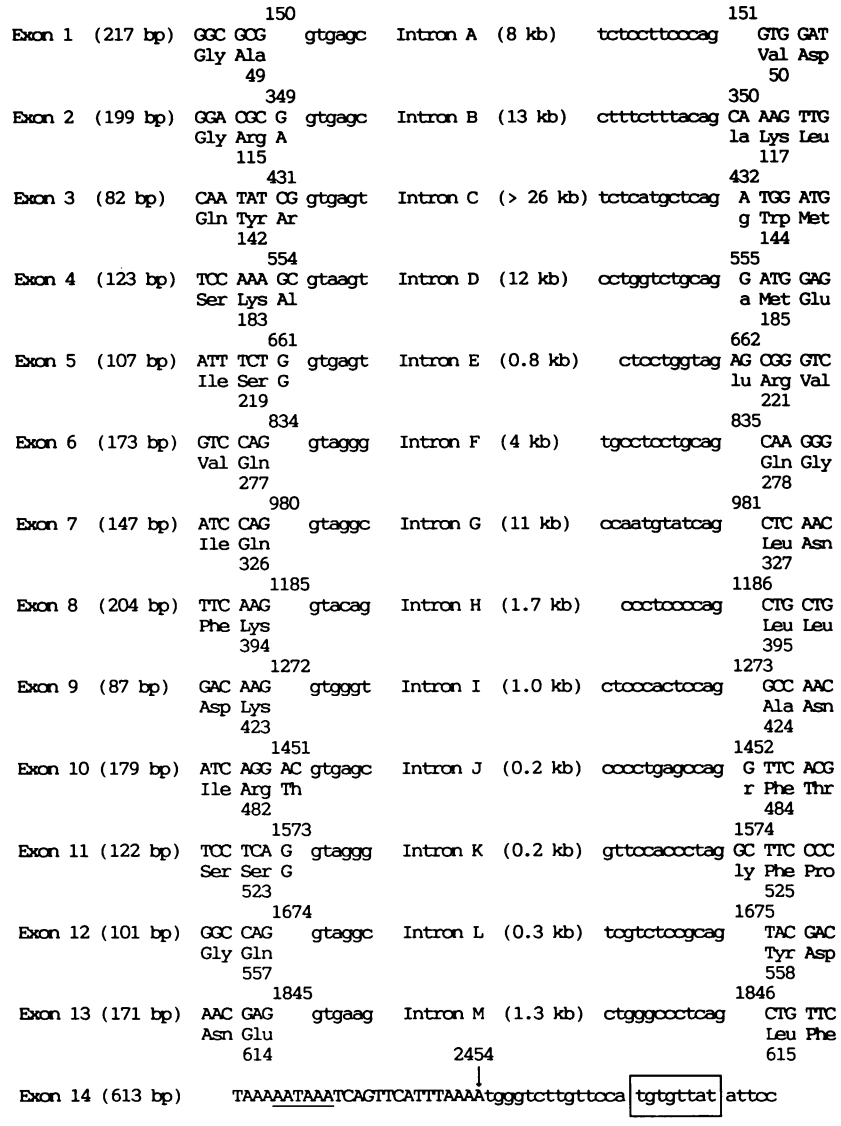

FIG. 2. Intron/exon organization of the human 5-lipoxygenase gene. Nucleotide sequences of intron/exon junctions were determined. Exon sequences are shown in uppercase letters; intron sequences are shown in lowercase letters. Nucleotides that border the 5′ and 3′ ends of the introns are numbered above the DNA sequence with the first base of the ATG initiation codon designated as +1 (nucleotides within introns are not numbered). Amino acids bordering the splice junctions are numbered with the first amino acid of mature 5-lipoxygenase designated as +1. Exon sizes and approximate intron sizes are indicated. Exon 1 size is determined from the major transcriptional initiation site. The 3′ end of exon 14 was not precisely mapped. Its size is determined up to and including the adenosine residue marked by an arrow. The poly(A) signal is underlined and a putative 3′ end mRNA processing signal is boxed. bp, Base pairs.

seed lipoxygenase (38) sequences have been reported. There are striking homologies in the carboxyl-terminal half of these proteins (Fig. 5). Between amino acid residues 546 and 558 of human 5-lipoxygenase there are 12 of 13 residues identical in all six lipoxygenases, and between residues 351 and 401 there are five conserved histidine residues and six more conserved acidic and basic residues. The latter region lies within a segment postulated by Shibata *et al.* (36) to be involved in binding iron. Within the proposed iron-binding domain lies a short sequence with homology to the interface binding

FIG. 3. DNA sequence of the 5′ end of the human 5-lipoxygenase gene. The first base of the ATG initiation codon is designated +1. The 5′ ends of fragments protected by S1 nuclease digestion and generated by primer extension are indicated by ○ and ●, respectively. The major transcription initiation site is indicated by ↓. Sites corresponding to the core consensus for Sp1 binding are boxed. An 11-bp inverted repeat is underlined. The DNA was sequenced entirely on both strands.

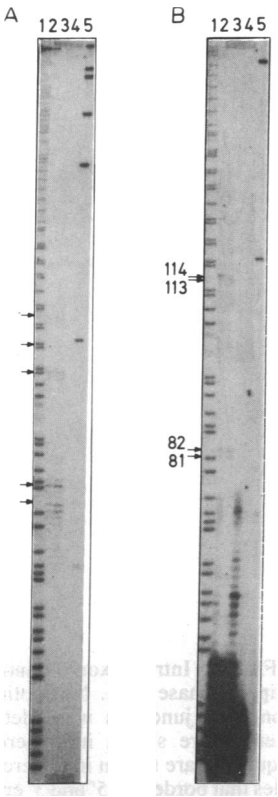

FIG. 4. Mapping the sites of transcription initiation in the human leukocyte 5-lipoxygenase gene by nuclease S1 and primer extension analysis. (*A*) The *Pst* I/*Bst*EII (−292 to +15) fragment of the λ1x12A phage DNA was labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP at the *Bst*EII site. A 307-nucleotide ^{32}P-labeled fragment was isolated by 5% polyacrylamide gel electrophoresis. The labeled DNA was hybridized to 5 μg of poly(A)$^+$ RNA from human leukocytes (lane 2), 20 μg of total leukocyte RNA (lane 3), or 20 μg of yeast tRNA (lane 4) according to the method of Berk and Sharp (27). The RNA·DNA hybrids were digested with 200 units of nuclease S1 at 37°C for 1 hr and analyzed by 8% polyacrylamide/7 M urea gel electrophoresis and autoradiography. To align the nuclease S1-protected fragment with the genomic sequence, a parallel Maxam–Gilbert (46) chemical reaction (A + G) was performed with the ^{32}P-labeled template used for the nuclease S1 analysis (lane 1). ΦX174 *Hin*cII-digested fragments were also used as fragment length markers (lane 5). The arrows indicate the protected fragments by nuclease S1 digestion. (*B*) The 49-nucleotide primer (−1 to +48) was hybridized with 5 μg of poly(A)$^+$ RNA from human leukocytes (lane 2), 20 μg of total leukocyte RNA (lane 3), or 20 μg of yeast tRNA (lane 4) for 4 hr at 42°C (27), and the hybridized primer/RNA was extended with 4 units of reverse transcriptase for 30 min at 37°C. The products were analyzed by electrophoresis in an 8% polyacrylamide/7.5 M urea gel followed by autoradiography. The Maxam–Gilbert chemical reaction (A + G) of the 307-nucleotide *Pst* I/*Bst*EII fragment and ΦX174 *Hin*cII-digested fragments (lanes 1 and 5, respectively) were used as fragment length markers. The numbers on the left indicate the fragment length generated by primer extension.

domain of lipases (39). This segment could be involved in arachidonic acid binding and is ideally contacted with the putative ferric catalytic center. This interesting region of the protein is encoded primarily by exon 8 and by part of exon 9.

The 5-lipoxygenase gene is likely represented by a single copy in the human haploid genome. Total genomic DNA hybridization analysis (data not shown) and the fact that clones from four different genomic libraries exhibited near-identical restriction maps (see below) tend to confirm this. However, the presence of very similar functional genes or pseudogenes cannot be ruled out. The only other cloned lipoxygenase gene is from soybeans (isozyme 3) and its structure is totally unrelated to the 5-lipoxygenase gene.

After restriction digest mapping and Southern blot analysis of one of the cosmid clones (clx5B-1) we noticed that one *Eco*RI site was absent when compared to the corresponding overlapped regions of two phage clones (λ1x27A and λ1x43A) obtained from two different libraries (Fig. 1). Detailed mapping of a 4-kb region 3′ to the *Eco*RI site in question (data not shown) revealed identical restriction maps in the phage clone λ1x43A and the cosmid clone. It is, therefore, probable that the absence of the *Eco*RI site within intron B of clone clx5B-1 represents a restriction site polymorphism.

The 5′ End of the 5-Lipoxygenase Gene. The putative promoter region of the 5-lipoxygenase gene is lacking typical TATA and CCAAT boxes in close proximity to the transcription initiation sites and exhibits features common to the promoter regions of the housekeeping genes (40). The 5′ flanking region of the 5-lipoxygenase gene contains eight potential sites for Sp1 binding (33, 34) in addition to similar short GC repeating units (Fig. 3). The presence of an 11-bp inverted repeat unit could also have some relevance to transcription factor binding and thus to transcriptional regulation. In addition, at least three GC boxes, close to the junction border of exon 1, within intron A, are found (Fig. 3). Sequences within the first intron of several genes are known to modulate transcriptional activity in either a positive or negative fashion (41). It is possible that the GC boxes in the first intron of the 5-lipoxygenase gene could play a similar role (cf. ref. 42).

There is one major transcription initiation site of the 5-lipoxygenase gene in human leukocytes that was determined by nuclease S1 protection and primer extension experiments (Fig. 4). It occurs at a thymidine residue 65 bp upstream of the ATG initiation codon. However, consistent with other genes that lack TATA and CCAAT sequences, the 5-lipoxygenase gene appears to have multiple transcription initiation sites. It is possible, though, that some of these sites could represent incomplete digestion by nuclease S1 (positions −89, −97, −107) or premature termination by reverse transcriptase (positions −34, −35).

The 5′ flanking sequence (positions −292 to −1) shows some sequence similarity (65%) to a region of intron 2 of the human ζ-globin pseudogene (43). The latter sequence contains multiple copies of the sequence CGGGG, similar to the repeating CGGGGG units in the 5-lipoxygenase gene. The 177-bp intron 2 sequence of the bovine arginine vasopressin-neurophysin II gene (44) displays similar homology.

The resemblance of the putative promoter region of the human 5-lipoxygenase gene to those of the housekeeping genes is rather surprising. The products of housekeeping genes usually display wide tissue distribution and perform essential metabolic functions. Their genes are often constitutively expressed with little regulation (40). However, 5-lipoxygenase gene expression has been assumed to take place primarily in cells of myeloid lineage. In addition, HL-60 cells in the undifferentiated state do not express 5-lipoxygenase but when induced to differentiate by such stimuli as dimethyl sulfoxide 5-lipoxygenase gene expression is turned on (15). Obviously, the 5-lipoxygenase gene does not fit the housekeeping category. A similar situation is found with the promoter that directs expression of the nerve growth factor receptor gene (45). Further studies to investigate the promoter activity and transcriptional regulatory elements of the 5-lipoxygenase gene will prove invaluable in defining factors involved in the control of 5-lipoxygenase gene expression.

We thank Prof. Håkan Persson for helpful discussion and Dr. Osamu Makabe for synthesis of oligonucleotides. This study was supported by a fellowship from Fonds de la Recherche en Santé du Québec (to C.D.F.) and by grants from the Swedish Medical Research Council (03X-217).

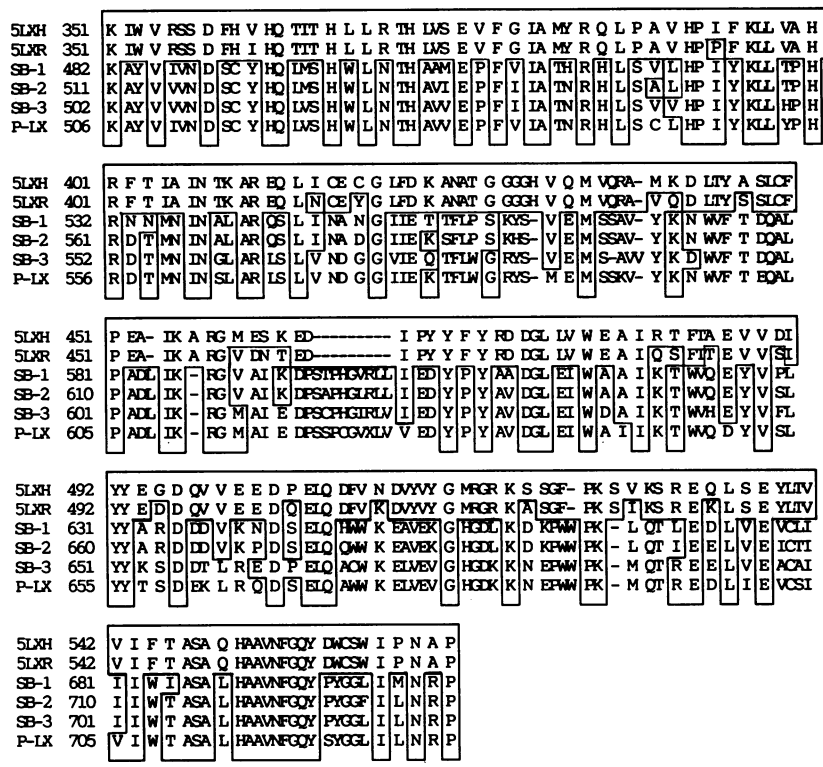

FIG. 5. Comparison of the amino acid sequence of human 5-lipoxygenase with the sequences of other lipoxygenases. Standard one-letter codes are used for amino acids. Boxed regions indicate identity to human 5-lipoxygenase residues. Spaces inserted for optimal alignment of the sequences are represented by dashes. Numbers indicate the amino acid position within the corresponding sequence. 5LXH, human 5-lipoxygenase (14); 5LXR, rat basophilic leukemia cell 5-lipoxygenase (16); SB-1, SB-2, and SB-3, soybean lipoxygenase isozymes 1, 2, and 3, respectively (35–37); P-LX, pea seed lipoxygenase (38).

1. Rouzer, C. A., Matsumoto, T. & Samuelsson, B. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 857–861.
2. Shimizu, T., Izumi, T., Seyama, Y., Tadokoro, K., Rådmark, O. & Samuelsson, B. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4175–4179.
3. Ueda, N., Kaneko, S., Yoshimoto, T. & Yamamoto, S. (1986) *J. Biol. Chem.* **261,** 7982–7988.
4. Hogaboom, G. K., Cook, M., Newton, J. F., Varrichio, A., Shorr, R. G. L., Sarall, H. M. & Crooke, S. T. (1986) *Mol. Pharmacol.* **30,** 510–519.
5. Samuelsson, B. (1983) *Science* **220,** 568–575.
6. Samuelsson, B., Dahlén, S.-E., Lindgren, J. Å., Rouzer, C. A. & Serhan, C. N. (1987) *Science* **237,** 1171–1176.
7. Rouzer, C. A. & Samuelsson, B. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6040–6044.
8. Goetze, A. M., Fayer, L., Bouska, J., Bornemeier, D. & Carter, G. W. (1985) *Prostaglandins* **29,** 689–701.
9. Rouzer, C. A., Shimizu, T. & Samuelsson, B. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7505–7509.
10. Rouzer, C. A. & Samuelsson, B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7393–7397.
11. Samuelsson, B., Rouzer, C. A. & Matsumoto, T. (1987) *Adv. Prostaglandin Thromboxane Leukotriene Res.* **17,** 1–11.
12. Rouzer, C. A. & Kargman, S. (1988) *J. Biol. Chem.* **263,** 10980–10988.
13. Wong, A., Hwang, S. M., Cook, M. N., Hoogaboom, G. K. & Crooke, S. T. (1988) *Biochemistry* **27,** 6763–6769.
14. Matsumoto, T., Funk, C. D., Rådmark, O., Höög, J.-O., Jörnvall, H. & Samuelsson, B. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 26–30, and correction (1988) **85,** 3406.
15. Dixon, R. A. F., Jones, R. E., Diehl, R. E., Bennett, C. D., Kargman, S. & Rouzer, C. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 416–420.
16. Balcarek, J. M., Theisen, T. W., Cook, M. N., Varrichio, A., Hwang, S.-M., Strohsacker, M. W. & Crooke, S. T. (1988) *J. Biol. Chem.* **263,** 13937–13941.
17. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15,** 687–702.
18. Lindenmeier, W., Hauser, H., de Wilke, I. G. & Schutz, G. (1982) *Nucleic Acids Res.* **10,** 1243–1256.
19. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
20. Funk, C. D., Rådmark, O., Fu, J. Y., Matsumoto, T., Jörnvall, H., Shimizu, T. & Samuelsson, B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6677–6681.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
22. DiLella, A. G. & Woo, S. L. C. (1988) *Methods Enzymol.* **152,** 199–212.
23. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
24. Rackwitz, H.-R., Zehetner, G., Frischauf, A.-M. & Lehrach, H. (1984) *Gene* **30,** 195–200.
25. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
27. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12,** 721–732.
28. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
29. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
30. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
31. McLauchlan, J., Gaffney, D., Whittor, J. L. & Clements, J. B. (1985) *Nucleic Acids Res.* **13,** 1347–1368.
32. Gil, A. & Proudfoot, N. J. (1984) *Nature (London)* **312,** 473–474.
33. Dynan, W. S. & Tjian, R. (1985) *Nature (London)* **316,** 774–778.
34. Kadonaga, J. T., Jones, K. A. & Tjian, R. (1986) *Trends Biochem. Sci.* **11,** 20–23.
35. Shibata, D., Steczko, J., Dixon, J. E., Hermodson, M., Yazdanparast, R. & Axelrod, B. (1987) *J. Biol. Chem.* **262,** 10080–10085.
36. Shibata, D., Steczko, J., Dixon, J. E., Andrews, P. C., Hermodson, M. & Axelrod, B. (1988) *J. Biol. Chem.* **263,** 6816–6821.
37. Yenofsky, R. L., Fine, M. & Liu, C. (1988) *Mol. Gen. Genet.* **211,** 215–222.
38. Ealing, P. M. & Casey, R. (1988) *Biochem. J.* **253,** 915–918.
39. Guidoni, A., Benkowka, F., DeCaro, J. & Rovery, M. (1981) *Biochim. Biophys. Acta* **660,** 148–150.
40. Dynan, W. S. (1986) *Trends Genet.* **2,** 196–197.
41. Killen, P. D., Burbelo, P. D., Martin, G. R. & Yamada, Y. (1988) *J. Biol. Chem.* **263,** 12310–12314.
42. Nakanishi, T., Kohno, K., Ishiura, M., Ohashi, H. & Uchida, T. (1988) *J. Biol. Chem.* **263,** 6384–6391.
43. Proudfoot, N. J., Gil, A. & Maniatis, T. (1982) *Cell* **31,** 553–563.
44. Ruppert, S., Scherer, G. & Schutz, G. (1984) *Nature (London)* **308,** 554–557.
45. Sehgal, A., Patil, N. & Chao, M. (1988) *Mol. Cell. Biol.* **8,** 3160–3167.
46. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.