Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 6836–6840, September 1988
Genetics

Organization of the human α_2-plasmin inhibitor gene

(fibrinolysis/serine protease inhibitors/serpin gene superfamily/human genomic clones)

SHINSAKU HIROSAWA*, YUICHI NAKAMURA*, OSAMU MIURA*, YOSHIHIKO SUMI†, AND NOBUO AOKI*‡

*The First Department of Medicine, Tokyo Medical and Dental University, Yushima, Bunkyo-Ku, Tokyo 113, Japan; and †Department of Biochemistry, University of Tokyo School of Medicine, Hongo, Bunkyo-Ku, Tokyo 113, Japan

Communicated by Earl W. Davie, May 25, 1988 (received for review October 28, 1987)

ABSTRACT **We have isolated overlapping phage genomic clones covering an area of 26 kilobases that encodes the human α_2-plasmin inhibitor. The α_2-plasmin inhibitor gene contains 10 exons and 9 introns distributed over ≈16 kilobases of DNA. To our knowledge, the number of introns is the highest yet reported for a member of the serine protease inhibitor (serpin) superfamily. All introns are located in the 5′-half of the corresponding mRNA. The 5′-untranslated region and the leader sequence are interrupted by 3 introns totaling ≈6 kilobases. A "TATA box" sequence is located 17 nucleotides upstream from the proposed transcription initiation site. Multiple "GC box" sequences, G+C-rich sequences, and "CCAAT box"-like sequence, the hepatitis B virus enhancer element-like sequence and the human immunodeficiency virus enhancer-like sequence appear in the 5′-flanking region. The NH_2-terminal region, which implements factor XIII-catalyzed cross-linking of α_2-plasmin inhibitor to fibrin, is encoded by the 4th exon. The reactive site and plasminogen-binding site, both located in the COOH-terminal region, are encoded by the 10th exon. When similar amino acids of α_2-plasmin inhibitor and other members of the serpin gene superfamily are aligned, the position of the 7th intron of the α_2-plasmin inhibitor gene aligns precisely with that of the second intron of the genes for rat angiotensinogen and human α_1-antitrypsin genes and is misaligned by only one nucleotide with that of the third intron of antithrombin III, suggesting that the α_2-plasmin inhibitor gene originates from the common ancestor of these serine protease inhibitors.**

α_2-Plasmin inhibitor (α_2PI; α_2-antiplasmin) is a plasma glycoprotein that functions crucially in the regulation of fibrinolysis (1–3). Human α_2PI is one of the major serine protease inhibitors (serpin superfamily) and is highly structurally similar to the other serpin superfamily members (4–6). However, α_2PI contains an extra ≈50-residue peptide beyond the COOH-terminal ends of the other family members (4). This extra peptide contains a plasminogen-binding site (4, 7) that endows the inhibitor with high affinity for plasminogen and enables the inhibitor to compete with fibrin for binding to plasminogen (8–10). During blood coagulation, α_2PI is cross-linked by activated factor XIII to the α chain of fibrin at the glutamine residue proximal to the NH_2-terminal end (11–13). The cross-linked α_2PI inhibits *in situ* plasmin generation on the fibrin surface by physiologically occurring fibrin-associated plasminogen activation (14, 15). These properties peculiar to α_2PI enable it to be a much more specific and effective inhibitor of plasmin-catalyzed fibrinolysis than any other major protease inhibitors, such as α_2-macroglobulin (2, 9, 16, 17). In individuals with a congenital deficiency of α_2PI, hemostatic plugs are dissolved prematurely by physiologically occurring fibrinolytic processes before the restoration of injured vessels, resulting in a severe hemorrhagic tendency (18, 19). The role of α_2PI in modulating fibrinolytic reactions has been reviewed recently (2, 3).

Studies from our laboratories (4) and those of others (5, 6) have led to the isolation of the cDNA coding for human α_2PI. Subsequently, the chromosomal localization of the α_2PI gene was demonstrated (20). In this investigation, the cDNA for human α_2PI was used for the isolation of overlapping genomic clones from a λ phage library. Organization of the gene was then analyzed§ and compared with those of the genes for other serine protease inhibitors.

MATERIALS AND METHODS

cDNA for α_2PI. A partial cDNA clone for α_2PI, pPI 39, has been described (4). A longer cDNA, covering the regions coding for the COOH-terminal 6 amino acids of the signal peptide and the whole plasma protein plus the 3′-noncoding region up to the poly(A) sequence was subsequently assembled from clonal members of a new human hepatoma cell cDNA library. The nucleotide sequence of the region coding for the mature plasma protein was completely accordant with those of the cDNA already reported (5, 6).

Screening of the Human Genomic DNA Library. The human genomic library was provided by H. Matsushime and M. Shibuya (Medical Institute, University of Tokyo, Japan) (21). The library was prepared from human placenta DNA by partial digestion with *Alu* I and *Hae* III and subsequent cloning in the bacteriophage vector Charon 4A with *Eco*RI linker. The library was screened by *in situ* hybridization of 1.2×10^6 phage plaques (22) with two α_2PI cDNA fragments corresponding to amino acids 31–130 and 179–429 as probes (4, 6). A 15-mer synthetic oligonucleotide, 5′-ACTCCCC-TGCCAGCC-3′, that is the complementary sequence to bases −15 to −5 of the cDNA (6) plus the donor signal at 5′ (AC) and *Eco*RI linker at 3′ (CC), was used as a probe to obtain a fragment containing the 5′-untranslated region. The probes of cDNA fragments were labeled by nick-translation, and the 5′-end of the oligonucleotide was labeled by T4 polynucleotide kinase. Fragments of human genomic DNA were mapped with the restriction endonucleases *Eco*RI, *Bam*HI, *Hin*dIII, *Dra* I, and *Xba* I. Subcloning of the genomic DNA fragments in the plasmid pUC-18 and -19 was done.

Southern Blotting (23). The plasmid containing α_2PI gene was isolated and subjected to restriction endonuclease digestions. The DNA fragments were then separated on agarose gels, transferred to a nitrocellulose filter, and hybridized as described (24) using cDNA and oligonucleotide probes, which correspond to several regions of the α_2PI gene.

DNA Sequencing. Appropriate DNA fragments, isolated and digested with various restriction endonucleases, were

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.

Abbreviations: α_2PI, α_2-plasmin inhibitor; nt, nucleotide(s).

‡To whom reprint requests should be addressed.

§The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03830).

sequenced by the dideoxy nucleotide chain-termination method of Sanger *et al.* (25) using plasmid as described by Hattori and Sakaki (26).

Primer Extension. One nanogram of a 5'-end-labeled synthetic oligonucleotide (5'-ACCAGGAGCCCCCAGAGCAGCGCCATGTTC-3'), complementary to nucleotides (nt) −4 to +26 as shown in Fig. 2, was hybridized with 50 μg of total RNA from the hepatoma cell line Hep G2 or 2 μg of poly(A)$^+$ RNA from normal human liver. The total RNA was prepared by the guanidine thiocyanate extraction method. Poly(A)$^+$ RNA was prepared by oligo(dT)-cellulose chromatography. The hybridization occurred in 80% (vol/vol) formamide/0.4 M NaCl/40 mM Pipes, pH 6.8, 1 mM EDTA by heating to 80°C for 5 min and then incubating at 45°C for 24 hr. The hybrids were recovered, and primer-extension reactions were done at 42°C for 1 hr with 40 units of reverse transcriptase purified from Rous-associated virus 2 (Takara Shuzo, Kyoto, Japan) and analyzed on 12% sequencing gels (27).

RNA-Blot Hybridization. Poly(A)$^+$ RNA prepared from human normal liver cells was separated by electrophoresis on a formaldehyde/agarose gel, transferred to a nitrocellulose filter, and hybridized as described (24).

RESULTS AND DISCUSSION

Three bacteriophage clones (designated as λPI1, λPI2, and λPI6) containing sequences of α_2PI gene were isolated from the genomic DNA library. By restriction endonuclease mapping, these clones were found to overlap (Fig. 1). λPI1 carried the DNA insert of the *Eco*RI fragment of 13 kilobases (kb), and λPI2 carried *Eco*RI fragments of 13, 0.5, and 1.8 kb. λPI6 had 11-kb and 1.8-kb *Eco*RI fragments. Both λPI2 and λPI6 contained the 1.8-kb fragments with identical sequence, indicating that both clones (λPI2 and λPI6) overlap each other. The synthetic 15-mer oligonucleotide probe, corresponding to the 5'-untranslated region of the cDNA, and the cDNA probe, corresponding to the 5'-region of the mRNA for α_2PI, hybridized to the 11-kb fragment contained in λPI6. The cDNA probe that corresponds to the 3'-region of the mRNA of α_2PI hybridized to the *Eco*RI fragment of 13 kb contained in λPI1 and λPI2. These results show that these clones contain the entire region of the α_2PI gene (Fig. 1).

The gene structure of α_2PI was further characterized by subcloning appropriate fragments in the plasmid pUC-18 and -19. Southern blotting analysis of restriction enzyme-digested plasmid DNAs containing exons and exon–intron boundaries were used to deduce the overall gene organization. The genomic DNA sequence of selected regions of the α_2PI gene was compared with the cDNA sequences, and this comparison allowed a precise definition of the exon–intron boundaries (Fig. 2). All boundaries were consistent with the "GT-AG" rule formulated by Breathnach and Chambon (28). The α_2PI gene was found to be ≈16 kb in length and to consist of 10 exons and 9 introns.

The sequence of the exons agrees perfectly with the sequence of the entire coding region of the cDNA earlier reported (4–6). However, a small section of the 5'-untranslated region (nt −17 to −5 in Fig. 2) differed slightly as compared with the cDNA reported by Tone *et al.* (6)—CTGGCAGGGGA for the cDNA and CTGGGCAGGGAGG for the genomic DNA. The sequence difference might either have been caused by polymorphism or the origin of the libraries—hepatoma cell line for the cDNA and placenta for the genomic library.

To identify the transcription initiation site, we constructed a 5'-end-labeled, antimessage sequence 30-base oligonucleotide primer complementary to a sequence of nt −4 to 26 including the initiator methionine codon in the cDNA. This primer was hybridized to human hepatoma cell line RNA or human normal liver cell poly(A)$^+$ RNA and extended with reverse transcriptase. The products of the reaction were sized by denaturing PAGE and migrated as 48- and 65-base pair (bp) fragments (Fig. 3). Although the band corresponding to 65-bp transcript was very faint from hepatoma cell line RNA, the results suggest two major transcription initiation sites; 22 nt and 39 nt upstream of the initiator methionine codon in the cDNA. Therefore, one transcription initiation site may be located at nt −22, suggesting sequence nt −22 to −5 is exon 1. To further define the initiation site, RNA blotting was done with a synthetic oligonucleotide probe, corresponding to proposed exon 1 (nt −22 to −5) or to its immediate upstream sequences (nt −39 to −23), and the cDNA probe. The probe corresponding to the proposed exon 1 hybridized to a band that corresponds to the mRNA of α_2PI (≈2.4 kb), which was identified by the cDNA, whereas the other probe (nt −39 to −23) failed to hybridize. These results together with the presence of mRNA (G)GGT sequence (consensus sequence at the intron–exon junctions) indicate that the region from nt −22 to −5 is the first exon. Another possible transcription initiation site is not known, but the longer transcript may represent crosshybridization to another mRNA.

The 1120-nt sequence of the 5'-flanking region was determined. The result reveals the presence of a TATA box (29) and 4 GC box sequences (5'-GGGCGG-3' and its inverted complement sequence 5'-CCGCCC-3') (30) (Fig. 2). Three of these GC box sequences are present in the ≈350-nt region upstream of the transcription initiation site. In this region are also several G+C-rich sequences in addition to the segments containing GC boxes (Fig. 2). McKnight and Kingsbury (30) stressed the importance of the GC box and G+C rich sequences upstream of the TATA box for maintaining transcription efficiency in eukaryotes. They further reported the

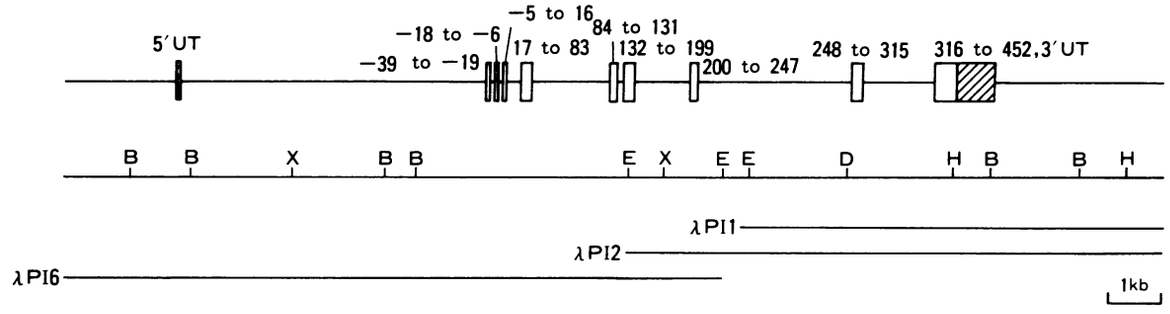

FIG. 1. Organization of the human α_2-plasmin inhibitor gene. The first line shows the positions of exons as rectangles, and the numbers above the line indicate the amino acids at which intron–exon junctions occur. Untranslated regions (UT) are shown as hatched areas. The second line indicates the positions of restriction endonuclease-recognition sites. Straight lines at bottom indicate the region of the three phage clones (λPI1, λPI2, and λPI6). B, *Bam*HI; D, *Dra* I; E, *Eco*RI; H, *Hin*dIII; and X, *Xba* I. Note from Fig. 2 that a small 5'-untranslated region exists in the second exon.

FIG. 2. (*Figure continues on the opposite page.*)

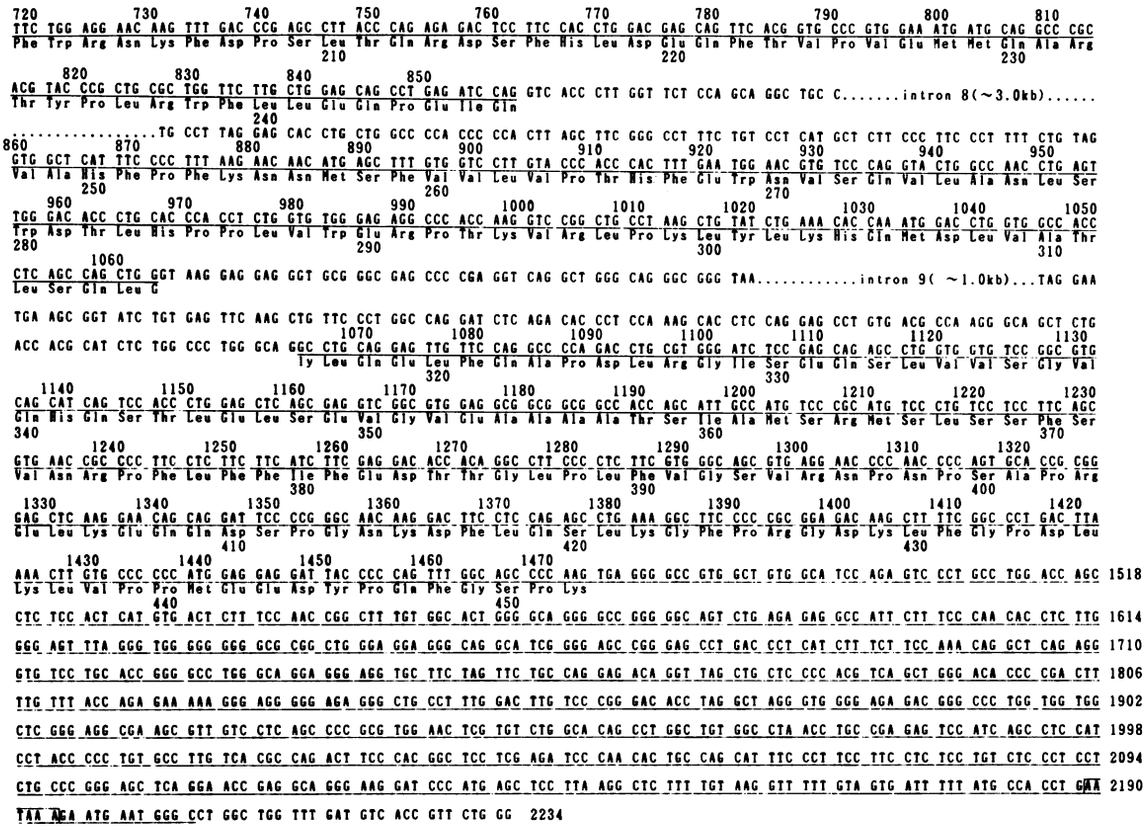

FIG. 2. Nucleic acid sequence of the α_2-plasmin inhibitor gene. Exons are underlined with solid lines. Bases in the exons and the 5'- and 3'-flanking regions are numbered relative to the translation initiation site. Amino acids are numbered from the NH_2-terminal residue in the plasma protein. Regions corresponding to a potential TATA box, the GC boxes, a potential transcriptional start site (−22) and a polyadenylylation recognition site (2189–2194) are boxed. The direct repeats of CCAAT box-like sequence are indicated by dots. G + C-rich sequences are indicated by the dashed underlines. The sequence (−123 to −108), similar to the hepatitis B virus enhancer sequence, is indicated by a waved line. The sequence (−809 to −800), similar to the human immunodeficiency virus enhancer sequence or κ-immunoglobulin light-chain gene enhancer sequence, is bracketed.

CCAAT box homology (31) downstream of the GC box (27, 31). In our study, we found the two direct repeats of the CCAAT box homology sequence, 5'-GCCATCA-3', separately located in the downstream regions of the two different GC boxes (Fig. 2). The TATA box may determine the position of the start of transcription, whereas the GC box may be the site interacting with a cellular transcription factor necessary for transcriptional activity (32).

The first base of the most proximal GC box sequence or the CCAAT box homology sequences is located 88 or 74 bases upstream, respectively, from the proposed transcription initiation site (Fig. 2). The relative positions of these sites are accordant with that usually found in eukaryotes (28). The first thymine of the TATA box is located 17 bases upstream from the proposed transcription initiation site (Fig. 2). The TATA box is usually found between 20 and 30 bases upstream from the transcription initiation site on most eukaryotic protein-coding genes. Therefore, the distance between the TATA box and the transcription initiation site here proposed might be an exceptional case among eukaryotes.

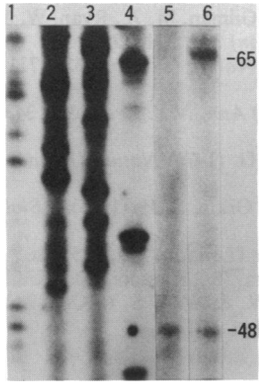

FIG. 3. Primer-extention reactions by reverse transcriptase. An end-labeled oligonucleotide probe from the gene for the α_2PI was used. Lanes: 1–4, DNA sequencing ladder for size comparison; 5 and 6, primer-extension reactions with liver cell line Hep G2 and normal liver cell RNAs, respectively. Figures at right are the lengths of the primer-extension products, corresponding to 48 and 65 nt.

It is interesting to note that the 16-bp sequence (nt −123 to −108 in Fig. 2) is 88% similar to the 17-bp sequence (nt 1193–1209) in the hepatitis B virus enhancer element (33), which displays tissue-specific activity (34, 35) and shows high homology with sequences in the promoter region of several liver-specific genes; α-fetoprotein, α_1-antitrypsin, and albumin (33). Also interesting is the presence of a 10-bp sequence, GTGACTTTCC, between nt −799 and −810 (Fig. 2). This sequence differs only by one base from the human immunodeficiency virus enhancer sequence, GGGACTTTCC (36), that is 100% similar to an enhancer sequence in the κ-immunoglobulin light-chain gene (37). It is quite interesting to see whether these sequences are also functional elements for the enhancement of the transcriptional activity of α_2PI gene.

The lengths of exons 1–10 were 17, 67, 39, 63, 202, 144, 204, 143, 205, and 1169 bp, respectively. Exon 1 is located 6 kb upstream from exon 2 that contains initiation codon ATG. The signal peptide is encoded by exons 2, 3, and a part of exon 4 (Figs. 1 and 2). The signal peptide consists of 39 amino acids, of which 23 are hydrophobic and form hydrophobic cores (Fig. 2). This agrees with the characteristic features of signal peptides (38). One base difference was noted in the sequence coding for the signal peptide as compared with the sequence of the cDNA reported by Tone *et al.* (6). The nucleotide (97 in their numbering system) was thymine in

their cDNA, but the nucleotide at this position (nt 97) was cytosine in our study and also in the cDNA reported by Holmes *et al.* (5). Consequently, the predicted amino acid at position −7 was arginine in our study and the study by Holmes *et al.* (5), whereas tryptophan was predicted by the cDNA sequence reported by Tone *et al.* (6). The difference may have been caused by the difference of cell types from which the cDNA was derived. Tone *et al.* (6) used the cDNA derived from a liver carcinoma cell line for sequencing the 5′-region, whereas normal cells were used in our study and the study by Holmes *et al.* (5).

The sequence of the 3′-noncoding region, including the consensus polyadenylylation signal AATAAA (39), is identical with that of the cDNA reported by Tone *et al.* (6), except for one substitution (T → G) at nt 1547. When compared with the cDNA sequence of the 3′-noncoding region reported by Holmes *et al.* (5), however, five minor differences including one deletion, three insertions, and one substitution were noticed (6). The poly(A) addition site, determined from the cDNA sequences reported by Sumi *et al.* (4) and Tone *et al.* (6), is cytosine at nt 2207 in Fig. 2. Another possible polyadenylylation site, determined from the cDNA sequences reported by Holmes *et al.* (5), is thymine at nt 2212. These differences are probably due to the origins used for the construction of the cDNA libraries.

α_2PI contains three functional domains—the reactive site, the plasminogen-binding site, and the cross-linking site for the fibrin α chain (2, 3). The plasminogen-binding site and the cross-linking site are peculiar to α_2PI among serine protease inhibitors and make α_2PI the most specific and effective one in inhibiting plasmin-catalyzed fibrinolysis (2, 3). The cross-linking site domain is located in the NH_2-terminal region (12) and is encoded by exon 4. The plasminogen-binding site domain is located in the COOH-terminal region (4, 7) and is encoded by exon 10. The reactive-site peptide bond that is cleaved by the reaction with plasmin has been postulated to be Met-362 to Ser-363 (4) or Arg-364 to Met-365 (5), and the reactive site domain containing these peptide bonds is encoded by exon 10, like the plasminogen-binding site domain.

Homologous amino acid sequences of human α_2PI and other serpin superfamily members (antithrombin III, α_1-antitrypsin, and rat angiotensinogen) were aligned as previously reported (6), and the positions of the introns were compared. Only one intron of nine introns of α_2PI, intron 7, was located at the position equivalent to those of the other serpin members. When the positions of these introns are compared at the nucleotide level, the intron of α_2PI aligns precisely with those of α_1-antitrypsin and angiotensinogen (40). However, the intron of antithrombin III is misaligned by only one nucleotide as shown by Prochownik *et al.* (40).

Although the serpin gene superfamily may originate from the same ancestor, explaining the discrepancies in intron positions of its members is difficult. Cornish-Bowden (41) has suggested that random losses of most introns occur during evolution from an ancestral gene. Others (42) have suggested that introns have been introduced into a particular family after the divergence of its members from an ancestral gene. The former proposal suggests that α_2PI may be evolutionally primitive because the number of introns in the α_2PI gene is the highest among the serpin gene superfamily members. The latter proposal suggests, on the contrary, that α_2PI may be evolutionally new. The former proposal agrees with the phylogenetic tree of the serpins constructed by Tone *et al.* (6), which suggested that α_2PI was the first gene to branch from the common ancestor of the serpins.

We thank Drs. Masami Muramatsu and Masaharu Sakai, Department of Biochemistry, University of Tokyo School of Medicine, for valuable advice during the course of this work, and Dr. Yataro Ichikawa, Central Research Laboratories, Teijin Ltd., for synthesizing the oligonucleotide, and Dr. Yoshiyuki Sakaki, Kyushu University School of Medicine, for critical reading of the manuscript. This research was supported, in part, by grants from the Ministry of Education, Science and Culture of Japan (62480260), Teijin Ltd., and the Mitsubishi Foundation.

1. Moroi, M. & Aoki, N. (1976) *J. Biol. Chem.* **251,** 5956–5965.
2. Aoki, N. & Harpel, P. C. (1984) *Semin. Thromb. Hemostasis* **10,** 24–41.
3. Aoki, N. (1986) *J. Protein Chem.* **5,** 269–277.
4. Sumi, Y., Nakamura, Y., Aoki, N., Sakai, M. & Muramatsu, M. (1986) *J. Biochem.* **100,** 1399–1402.
5. Holmes, W. E., Nelles, L., Lijnen, H. R. & Collen, D. (1987) *J. Biol. Chem.* **262,** 1659–1664.
6. Tone, M., Kikuno, R., Kume-Iwaki, A. & Hashimoto-Gotoh, T. (1987) *J. Biochem.* **102,** 1033–1041.
7. Sasaki, T., Morita, T. & Iwanaga, S. (1986) *J. Biochem.* **99,** 1699–1705.
8. Moroi, M. & Aoki, N. (1977) *Thromb. Res.* **10,** 581–586.
9. Aoki, N., Moroi, M. & Tachiya, K. (1978) *Thromb. Haemostasis* **39,** 22–31.
10. Wiman, B., Lijnen, H. R. & Collen, D. (1979) *Biochim. Biophys. Acta* **579,** 142–154.
11. Sakata, Y. & Aoki, N. (1980) *J. Clin. Invest.* **65,** 290–297.
12. Tamaki, T. & Aoki, N. (1982) *J. Biol. Chem.* **257,** 14767–14772.
13. Kimura, S. & Aoki, N. (1986) *J. Biol. Chem.* **261,** 15591–15595.
14. Sakata, Y. & Aoki, N. (1982) *J. Clin. Invest.* **69,** 536–542.
15. Aoki, N., Sakata, Y. & Ichinose, A. (1983) *Blood* **62,** 1118–1122.
16. Aoki, N., Moroi, M., Matsuda, M. & Tachiya, K. (1977) *J. Clin. Invest.* **60,** 361–369.
17. Aoki, N. (1979) *Prog. Cardiovasc. Dis.* **21,** 267–286.
18. Aoki, N., Sakata, Y., Matsuda, M. & Tateno, K. (1980) *Blood* **55,** 483–488.
19. Aoki, N. (1984) *Semin. Thromb. Hemostatis* **10,** 42–50.
20. Kato, A., Nakamura, Y., Miura, O., Hirosawa, S., Sumi, Y. & Aoki, N. (1988) *Cytogenet. Cell Genet.*, in press.
21. Matsushime, H., Wang, L. H. & Shibuya, M. (1986) *Mol. Cell. Biol.* **6,** 3000–3004.
22. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
23. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 199–206.
25. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
26. Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152,** 232–238.
27. Sollner-Webb, B. & Reede, R. H. (1979) *Cell* **18,** 485–499.
28. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
29. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4853–4857.
30. McKnight, S. L. & Kingsbury, R. (1982) *Science* **217,** 316–324.
31. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8,** 127–142.
32. McKnight, S. L. & Tjian, R. (1986) *Cell* **46,** 795–805.
33. Shaul, Y. & Ben-Levy, R. (1987) *EMBO J.* **6,** 1913–1920.
34. Jameel, S. & Siddiqui, A. (1986) *Mol. Cell. Biol.* **6,** 710–715.
35. Tur-Kaspa, R., Burk, R. D., Shaul, Y. & Shafritz, D. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1627–1631.
36. Franza, B. R., Jr., Josephs, S. F., Gilman, M. Z., Ryan, W. & Clarkson, B. (1987) *Nature (London)* **330,** 391–395.
37. Nabel, G. & Baltimore, D. (1987) *Nature (London)* **326,** 711–713.
38. Jackson, R. C. & Blobel, G. (1980) *Ann. N.Y. Acad. Sci.* **343,** 391–403.
39. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
40. Prochownik, E. D., Bock, S. C. & Orkin, S. H. (1985) *J. Biol. Chem.* **260,** 9608–9612.
41. Cornish-Bowden, A. (1982) *Nature (London)* **297,** 625–626.
42. Leicht, M., Long, G. L., Chandra, T., Kurachi, K., Kidd, V. J., Mace, M., Jr., Davie, E. W. & Woo, S. L. C. (1982) *Nature (London)* **297,** 655–659.