Structure of the Human Plasminogen Activator Inhibitor 1 Gene: Nonrandom Distribution of Introns†

David J. Loskutoff,‡,§ Marijke Linders,§ Jaap Keijer,§ Harry Veerman,§ Harm van Heerikhuizen,‖ and Hans Pannekoek*,§

Department of Molecular Biology, Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, and Biochemical Laboratory, Free University of Amsterdam, Amsterdam, The Netherlands

Received March 25, 1987; Revised Manuscript Received April 17, 1987

ABSTRACT: Plasminogen activator inhibitor 1 (PAI-1) is the primary inhibitor of tissue-type plasminogen activator and thus performs an essential role in the regulation of the fibrinolytic process. It is a member of a large family of serine protease inhibitors (serpins). We determined the structure of the PAI-1 gene in order to more completely investigate the relationship of PAI-1 to other serpins and, at the same time, to begin to delineate structure–function relations in PAI-1 itself. A human genomic cosmid DNA library was screened and found to contain two independent clones, each harboring the entire PAI-1 gene. Restriction site mapping, electron microscopic inspection of heteroduplexes, and nucleotide sequence analysis all demonstrate that the PAI-1 gene is approximately 12.2 kilobase pairs in length and consists of nine exons and eight introns. All intron–exon boundaries are in accord with the "GT-AG" rule, including a cryptic acceptor splice site found in intron 7. The intron–exon pattern of the PAI-1 gene is distinct from that of most other serpins except that intron 3 of PAI-1 occupies an identical position as intron E of ovalbumin. Comparison of our data with the proposed subdomain structure of serpins suggests that seven of the eight introns may occupy a nonrandom position in the gene. These introns either delineate boundaries of individual structural subdomains or are located in random coil regions of the protein. Transcription of the PAI-1 gene in cultured vascular endothelial cells results in two distinct mRNA species. Our data suggest that these two transcripts arise by alternative polyadenylation.

Plasminogen activator inhibitor 1 (PAI-1)[1] is a glycoprotein present in blood and platelets and is synthesized by a variety of cells including endothelial cells (Erickson et al., 1985; Sprengers & Kluft, 1987). This fast-acting inhibitor functions as the physiological counterpart of tissue-type plasminogen activator (t-PA) and thus acts as a regulator of the fibrinolytic process. Recently, full-length PAI-1 cDNAs have been constructed and their complete nucleotide sequence has been determined (Ny et al., 1986; Pannekoek et al., 1986; Ginsburg et al., 1986; Andreasen et al., 1986). These cDNAs were derived from two distinct PAI-1 mRNA species that harbor an identical codogenic region but differ in the length of their 3′ untranslated region, suggesting that these different PAI-1 mRNA species arise either by alternative polyadenylation or by alternative splicing. Experiments devised to map the chromosomal location of the PAI-1 gene (Ginsburg et al., 1986) suggest that these mRNAs are transcribed from a single gene.

Alignment of the inferred amino acid sequence of PAI-1 with that of other proteins clearly showed that PAI-1 belongs to the gene family of the serine protease inhibitors (serpins) (Ny et al., 1986; Pannekoek et al., 1986). This family of molecules evolved from a common ancestral gene 200–500 million years ago (Hunt & Dayhoff, 1980) and functions to regulate the major protease systems of the body, including the fibrinolytic, coagulation, and complement cascades (Travis & Salvesen, 1983). Proteins belonging to this gene family share a highly homologous amino acid sequence (Hunt & Dayhoff, 1980), and it has been suggested that they have a similar ternary structure (Loebermann et al., 1984; Carrell & Boswell, 1986).

The elucidation of the intron–exon distribution of a gene may provide insights into its evolutionary background and at the same time provide important criteria for establishing structure–function relationships in its gene product (Patthy, 1985). In this paper, we have established the number and the precise location of introns in the human PAI-1 gene and have compared these data with those reported on other serpins. In addition, we have attempted to correlate the position of the introns in this gene with specific structural features of the protein (Gilbert, 1978; Blake, 1978; Go, 1981; Craik et al., 1983). Our studies suggest that seven of the eight introns in the PAI-1 gene may be located in segments of the gene that encode either boundaries of presumed subdomain structures or random coil regions of the protein. These findings provide support for the hypothesis that the position of introns within genes is not random.

Experimental Procedures

Screening of a Human Genomic DNA Library. The human placenta derived genomic cosmid DNA library was kindly provided by Dr. F. Baas (Department of Pediatric Endocrinology, Academical Medical Center, Amsterdam). This li-

† This work was supported in part by a grant from MEDIGON (900-526-70) to H.P. and in part by grants from the NIH to D.J.L.

* Address correspondence to this author c/o Publication Secretariat.

‡ On leave of absence from the Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, CA.

§ Central Laboratory of the Netherlands Red Cross Blood Transfusion Service.

‖ Free University of Amsterdam.

[1] Abbreviations: PA, plasminogen activator; PAI, plasminogen activator inhibitor; t-PA, tissue-type plasminogen activator; serpin, serine protease inhibitor; bp, base pair(s); kbp, kilobase pair(s); PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); EDTA, ethylenediaminetetraacetate; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/87/0426-3763$01.50/0

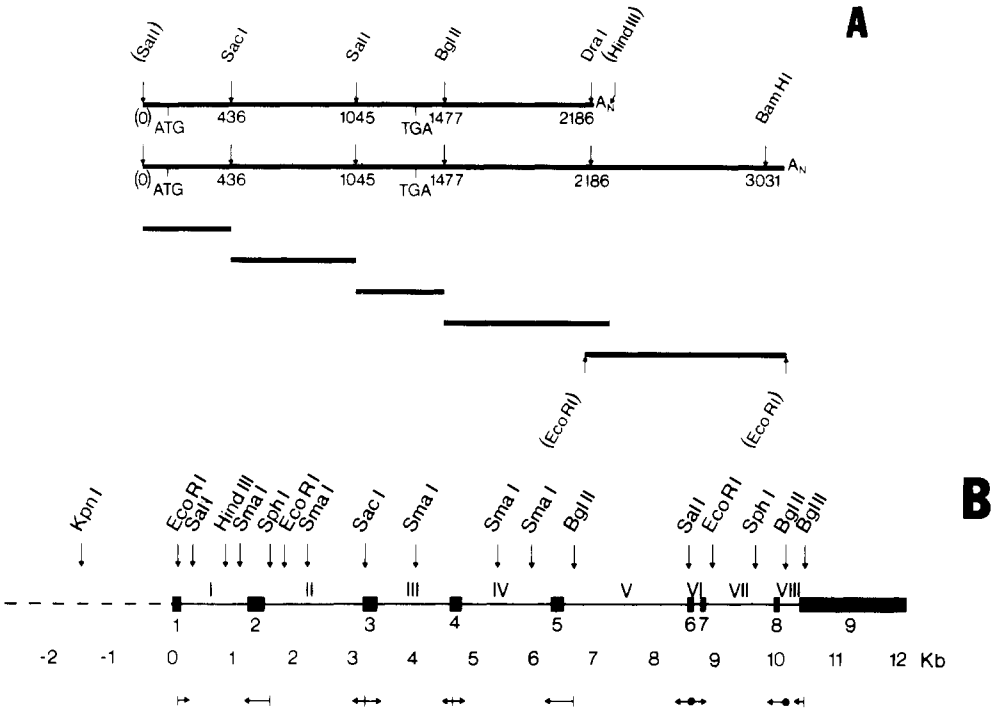


FIGURE 1: Restriction map of PAI-1 cDNA and of the PAI-1 gene. (A) The relevant restriction sites in PAI-1 cDNA (Ny et al., 1986; Pannekoek et al., 1986) employed for the generation of probes are indicated as well as the positions of the initiation and the termination codons. The numbering of the sites is according to Pannekoek et al. (1986). The bracketed *Sal*I site upstream from PAI-1 cDNA was derived from the multilinker of the pUC9 vector that had been digested with *Pst*I and G-tailed (Pannekoek et al., 1986). The bracketed *Hin*dIII site originates from the multilinker of vector pSP65 (Pannekoek et al., 1986). The probes used are (*Sal*I)–*Sac*I, *Sac*I–*Sal*I, *Sal*I–*Bgl*II, *Bgl*II–(*Hin*dIII) and a 1-kbp fragment extending 5′ from the poly(A) tail of the 3-kbp cDNA. The latter fragment was derived from a deletion library constructed by using the single-strand M13 method of Dale et al. (1985). It thus covers the 3′ end of the long variant of PAI-1 cDNA (Ny et al., 1986). (B) Schematic representation of the intron–exon distribution of the PAI-1 gene. The nine exons are indicated on the scale with the solid black bars and the eight introns by the horizontal line running through them. The restriction sites within the PAI-1 gene that were used for nucleotide sequencing of the junctions of introns and exons are also indicated. In most instances, sequencing was performed according to the dideoxy chain-termination method and the universal M13 primer (Sanger et al., 1977) (indicated with a vertical bar) by cloning fragments spanning one or two intron–exon boundaries into M13mp18 or M13mp19 (Yanisch-Perron et al., 1985). In two cases it was necessary to synthesize oligonucleotide primers. Thus, the primer 5′-TGGACCAGCTGACACGGCTG and the complementary primer from position 665 to position 684 in the cDNA were used to determine the boundaries of exon 4–intron IV and exon 4–intron III, respectively. In three cases, the chemical degradation procedure was employed (Maxam & Gilbert, 1977) (indicated with a black dot). In these instances, the indicated 5′ termini of *Sal*I and *Bgl*II digestions were labeled with ^{32}P and polynucleotide kinase. The length of the individual introns was determined by restriction site mapping and estimated to be approximately 1100, 1640, 1200, 1480, 2060, 113, 1160, and 340 bp (intron I–VIII, respectively). The precise length of the exons is indicated in Figure 3.

brary is a collection of approximately 340 000 independent colonies harboring cosmid pJBF [a 3.8 kilobase pair (kbp) derivative of cosmid pJB8 (5.4 kbp) that lacks a DNA segment covering its unique *Sal*I site (Ish-Horowicz & Burke, 1981)]. The library was constructed by partially digesting human placental DNA with *Sau*3A and inserting the resulting genomic fragments (35–50 kbp) in the unique *Bam*HI site of pJBF DNA. High-density in situ colony screening was done as described (Maniatis et al., 1982) by hybridization with a "nick-translated" PAI-1 cDNA fragment of approximately 2.2 kbp (Pannekoek et al., 1986). Positive colonies were isolated, and cosmid DNAs were purified by CsCl/ethidium bromide density centrifugation (Maniatis et al., 1982; Birnboim & Doly, 1979).

Restriction Site Mapping, Southern Blotting, and Nucleotide Sequence Analysis. Restriction enzymes and DNA-modifying enzymes were purchased from New England Biolabs (Beverly, MA) and were used as recommended by the manufacturer. Restriction fragments were isolated from low-melting agarose gels by repeated extractions with phenol and phenol/chloroform as described (Maniatis et al., 1982). Blot hybridization of genomic DNA fragments was done by the procedure of Southern (1975). Probes employed for hybridization included a 2.2-kbp PAI-1 cDNA (Pannekoek et al., 1986), and specific fragments of it, and a 1-kbp fragment derived from the 3′-terminal region of the larger 3-kbp cDNA (Ny et al., 1986). A schematic presentation of the length and location of these probes is given in Figure 1A. Nucleotide sequencing was performed either by the chemical degradation method after 5′ end labeling of isolated restriction fragments (Maxam & Gilbert, 1977) or by the dideoxy chain-termination procedure (Sanger et al., 1977) after subcloning genomic fragments in M13mp18 or M13mp19 DNA (Yanisch-Perron et al., 1985). In the latter cases, either the universal M13 primer (New England Biolabs) or primers that were synthesized on an Applied Biosystems type 381A DNA synthesizer were employed.

In Vitro Transcription of PAI-1 cDNA. The 2.2-kbp *Bam*HI–*Hin*dIII fragment containing the PAI-1 cDNA was inserted into pSP65 DNA (Melton et al., 1984) digested with *Bam*HI and *Hin*dIII. The resulting plasmid pSP65/PAI-1 contained PAI-1 cDNA downstream from the SP6 promotor. This plasmid was linearized with *Hin*dIII and transcribed in vitro with SP6 RNA polymerase according to the manufacturer's instructions (New England Nuclear, Dreieich, FRG).

Heteroduplex Mapping by Electron Microscopy. DNA/RNA hybrids were prepared as follows. *Bam*HI-linearized cosmid DNA (8 μg/mL) was denatured for 90 s at 100 °C in a buffer containing 70% (v/v) formamide, 0.1 M PIPES (pH 7.4), 0.4 M NaCl, and 0.01 M EDTA. After being cooled to 63 °C, SP6-generated PAI-1 RNA was added to a final concentration of 0.8 μg/mL, and the mixture was incubated

at a temperature linearly decreasing from 63 to 48 °C over a period of 6 h. The hybrids were brought to 65% (v/v) formamide, 0.08 M NaCl, 0.02 M PIPES, 6 mM EDTA, 90 mM Tris-HCl (pH 8.0), and 0.003% cytochrome *c* and then spread on a hypophase of doubly distilled water. Plasmid pBR322 DNA and single-stranded DNA from phage M13mp18 were used as internal standards for length measurements. Photographs were taken with a Zeiss EM109 electron microscope at a magnification of approximately 9000-fold. Contour lengths were measured on 10-fold enlarged photographic prints with a Hewlett-Packard 9874A digitizer coupled to a Hewlett-Packard 9825A calculator.

Results

Isolation of Cosmids Containing PAI-1 DNA. The human cosmid DNA library was screened by colony hybridization with radiolabeled 2.2-kbp PAI-1 cDNA as a probe. This cDNA is 2198 base pairs (bp) in length, not including the poly(A) tail, and consists of a 5′ untranslated region of 126 bp, a codogenic part of 1206 bp, and a 3′ untranslated region of 866 bp (Pannekoek et al., 1986). Two independent positive clones were isolated from this library, and restriction enzyme analysis revealed DNA inserts of approximately 42 and 47 kbp, respectively. Southern blotting of *Bam*HI-digested cosmid DNA demonstrated a restriction fragment of approximately 23 kbp that was common to both of these clones and hybridized to all of the PAI-1 cDNA probes (i.e., derived from any part of 3-kbp full-length PAI-1 cDNA) (Figure 1A). A second fragment of approximately 90 bp was recognized by the 1-kbp probe (i.e., the 3′ probe).

In order to determine the location, orientation, and length of the PAI-1 gene contained within the 23-kbp fragment, the cosmid (not shown) and the isolated 23-kbp *Bam*HI fragment (Figure 1B) were digested with various restriction enzymes and analyzed by Southern blotting and hybridization with PAI-1 cDNA probes. These studies revealed that the *Bam*HI site defining the 3′ end of the 23-kbp fragment is in fact the *Bam*HI site located only 89 bp from the poly(A) tail of the 3-kbp full-length PAI-1 cDNA (Ny et al., 1986). The cosmid DNA extends a minimum of 10 kbp in the 3′ direction from this site and approximately 10.5 kbp in the 5′ direction from the *Eco*RI site located at position 52 in the cDNA (Pannekoek et al., 1986). Analysis of the restriction fragments suggests that the cosmid harbors the entire PAI-1 gene and that the gene itself comprises about 12.2 kbp. The 23-kbp *Bam*HI fragment also contains the majority of the gene. It lacks only the 89-bp 3′ fragment located at the very 3′ end of the molecule.

Estimation of the Number of Introns by Heteroduplex Mapping. The SP6 in vitro transcription system (Melton et al., 1984) was employed to synthesize PAI-1 RNA corresponding to the 2.2-kbp cDNA. This RNA was hybridized with the denatured 23-kbp *Bam*HI fragment, and the resulting PAI-1 DNA and PAI-1 mRNA heteroduplexes were analyzed by electron microscopy (Figure 2). Single-stranded and double-stranded stretches on 20 different intact heteroduplex molecules were measured, and these data, together with the restriction map (Figure 1), were employed both to estimate intron number and location and, as guidelines for DNA sequencing, to determine the intron–exon junctions more precisely. Inspection of these heteroduplexes revealed at least seven introns, which varied in length from about 300 bp to 2 kbp. Restriction analysis (not shown) indicated that the heteroduplex is oriented in a 5′ to 3′ direction (from left to right in the photograph). Intron 1 is indicated. The length of the exons was estimated to be between 100 and 300 bp.

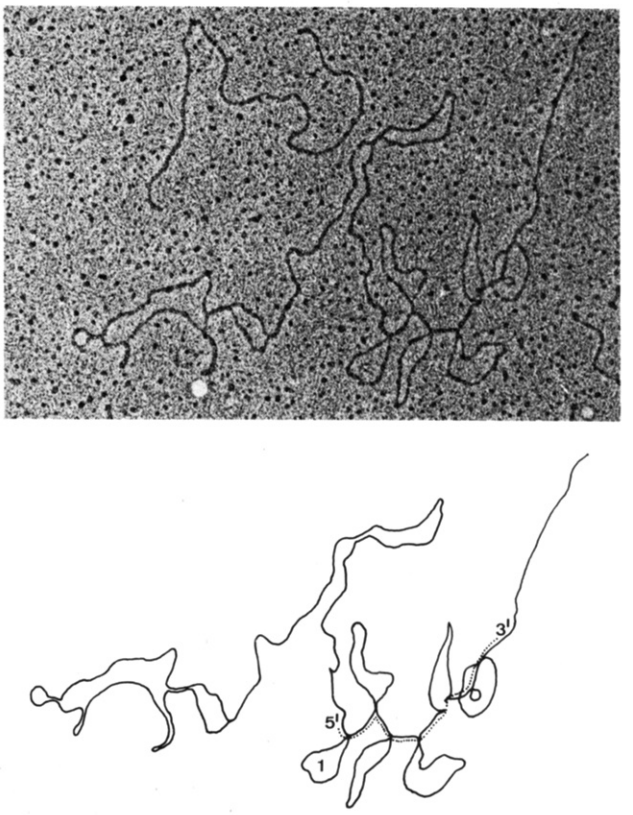
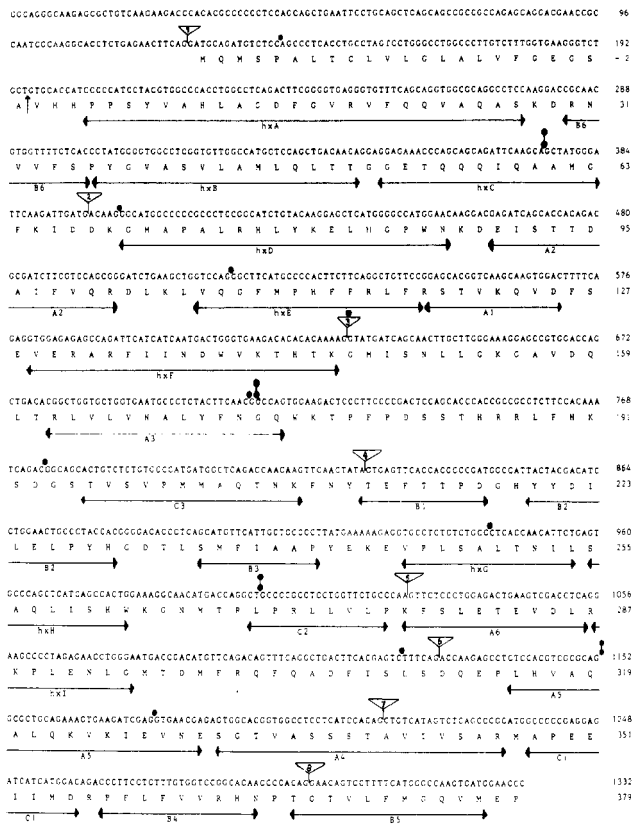


FIGURE 2: Heteroduplex map of the PAI-1 gene. (Top) Photograph of heteroduplexes formed between *Bam*HI-digested cosmid DNA and PAI-1 mRNA synthesized in vitro by employing the SP6 system. The heteroduplexes were prepared and visualized as described under Experimental Procedures; 9000× magnification. (Bottom) Artist's interpretation of the photograph. The dotted line corresponds to the position of the mRNA. The positions of intron 1 and both the 5′ and 3′ ends of the mRNA are indicated.

DNA Sequencing of Intron–Exon Junctions of the PAI-1 Gene. The DNA sequence of each exon, and of both sides of each intron–exon junction, was determined as described under Experimental Procedures. The strategy for sequencing is outlined in Figure 1B. The results, together with the amino acid sequence derived from PAI-1 cDNA, are presented in Figure 3. In addition to the seven introns detected by restriction analysis (Figure 1B) and heteroduplex mapping (Figure 2), DNA sequencing revealed the presence of another intron of 113 bp. This intron was located between the sixth and seventh exons shown in Figure 2. This small intron could not be visualized with certainty by electron microscopy. Thus, this portion of the PAI-1 gene consists of eight introns and nine exons. The location of the different introns is indicated in the DNA sequence of PAI-1 cDNA (Figure 3).

It has been suggested that the ternary structure of the various serpins is similar (Loebermann et al., 1984; Carrell & Boswell, 1986). On the basis of this assumption, we have also indicated in Figure 3 the portions of the protein that presumably constitute the common β-sheet, α-helix, and random coil structures. It can be seen that seven of the eight introns are located outside of the subdomain regions, either at their borders or in random coil structures.

We previously reported the isolation of a full-length PAI-1 cDNA plasmid that harbors an insertion of 21 bp, positioned between nucleotides 1213 and 1214 (Pannekoek et al., 1986). The insertion encodes seven amino acid residues, which are "in-frame" with the remainder of the protein and located near the reactive center of PAI-1 (P_1–P_1'; Arg_{346}-Met_{347}; Andreasen et al., 1986). It was noted that the sequence of this insertion,



No.	exon	intron		exon
1	TCAG	GTAG...	...CTTGCTCTAG	GATG
2	GATG	GTGA...	...CCCGGTGCAG	ACAA
3	AAAG	GTGA...	...TTTCCCATAG	GTAT
4	TATA	GTAA...	...CTCTGTACAG	CTGA
5	CCAA	GTAA...	...CCCGTTGCAG	GTTC
6	TCAG	GTAA...	...TCTGTTTTAG	ACCA
7	ACAG	GTGA...	...TCTCTTGCAG	CTGT
7a	ACAG	GTGA...	...TCTGTCTAAG	ACAT
8	ACAG	GTGA...	...CCACCCTCAG	GAAC

FIGURE 3: Intron positions of the PAI-1 gene and the presumed subdomain structure of the PAI-1 protein. (Top) The position of the eight introns in the nucleotide sequence of PAI-1 cDNA has been determined for both sides of the intron–exon junctions. The introns are localized between nucleotides 125–126, 397–398, 631–632, 826–827, 1025–1026, 1126–1127, 1213–1214, and 1297–1298 and are indicated with a triangle. The amino acid sequence (single letter symbols) derived from the nucleotide sequence is separately numbered. The subdomain structure of α_1-antitrypsin has been determined from its three-dimensional structure (Loebermann et al., 1984; Carrell & Boswell, 1986). We have aligned the α_1-antitrypsin subdomain structure (horizontal lines, arrowheads) onto the amino acid sequence of PAI-1 using the Microgenie program (Beckman Inc.). β-Sheets are indicated by arrows overlining the subdomains A1–A6, B1–B6, and C1–C3, while α-helices are indicated by arrows overlining hxA–hxI. The black dots above the nucleotide sequence indicate the corresponding position of introns found in other serpins (i.e., rat angiotensinogen, chicken ovalbumin or the related gene Y, human antithrombin III, and human α_1-antitrypsin). (Bottom) The DNA sequence of the exon–intron boundaries of the PAI-1 gene illustrating the GT-AG rule for the splice junctions. The cryptic acceptor splice site (7a) results in an insertion of 21 bp (Pannekoek et al., 1986).

i.e., 5′-ACATTGTCCCTTCTCTTGCAG, matches that of an acceptor splice site (Mount, 1982). Indeed, this sequence is found in the genomic PAI-1 DNA and is preceded by a pyrimidine-rich stretch and an AG dinucleotide [5′-TCTGTCTAAG(ACATT...)-3′] (Figure 3). This finding provides additional evidence that alternative splicing may create another protein that lacks PAI-1 activity (Pannekoek et al., 1986). At present, it is not known whether this phenomenon is an artifact due to incomplete splicing or represents a functional mechanism to vary the specificity or the activity of PAI-1.

Discussion

In this paper, we have investigated the structure of the human PAI-1 gene and compared it to the gene structure of other serpins. Our results demonstrate that the PAI-1 gene is approximately 12.2 kbp in length and consists of nine exons and eight introns. It should be noted, however, that the 5′ end of the cDNA used in these experiments (Pannekoek et al., 1986) does not contain the entire 5′ end of the molecule (i.e., it lacks at least 16 nucleotides; Andreasen et al., 1986). Thus, our results do not exclude the possibility that one or more introns may exist in this portion of the 5′ untranslated region. However, we have sequenced the genomic PAI-1 DNA beyond the 5′ end of the cDNA. These studies not only revealed the 16 additional nucleotides detected by Andreasen et al. (1986) but also indicated the presence of a "TATA box" at −30 from the 5′ end of their cDNA (A. J. van Zonneveld and D. J. Loskutoff, unpublished observation). The finding of this characteristic transcriptional regulatory element (Bucher & Trifonov, 1986) at the proper distance from the 5′ end of the PAI-1 mRNA argues against the presence of another intron in this region.

To investigate whether additional introns exist in the 3′ untranslated region, cosmid DNA and the 3-kbp cDNA containing the poly(A) tail (Ny et al., 1986) were digested with *Bgl*II (position 1477) and *Bam*HI (position 3031) and analyzed by Southern blotting using the 3′-terminal 1-kbp fragment (Figure 1A) as the hybridization probe. Identical fragments of 1.6 kbp were detected (data not shown). Thus, this part of the 3′ untranslated region is devoid of introns. This observation suggests that the two descrete PAI-1 mRNA species synthesized by human cells (Ny et al., 1986; Pannekoek et al., 1986) result from alternative polyadenylation and not from alternative splicing. DNA sequencing of the remaining 89 bp of the genomic PAI-1 DNA (3′ to the *Bam*HI site) showed that this segment is free of introns.

It was originally predicted that genes belonging to the serine protease inhibitor superfamily would contain a similar intron–exon pattern (Hunt & Dayhoff, 1980) and thus provide a record of the evolutionary creation of this gene family and the structure of the primordial serpin gene (Cornish-Bowden, 1982). The gene structure of a number of serpins has now been determined (Catterall et al., 1978; Breathnach et al., 1978; Leicht et al., 1982; Heilig et al., 1982; Tanaka et al., 1984; Prochownik et al., 1985; Kidd & Woo, 1986; this paper), and it is evident that this hypothesis is an oversimplification. For example, there is little or no correspondence between the locations of the introns of chicken ovalbumin and human α_1-antitrypsin (Leicht et al., 1982), but the four introns of α_1-antitrypsin and angiotensinogen appear to be equivalently placed (Tanaka et al., 1984). The antithrombin III gene contains five introns, only one of which is homologously located (Prochownik et al., 1985). Analysis of the PAI-1 gene (Figure 3) indicates that intron 3 occupies an identical position as intron E of ovalbumin and ir.tron E of the ovalbumin-related gene Y (Heilig et al., 1982). The other seven introns of the PAI-1 gene are at positions unrelated to those within other serpin genes. These observations indicate that introns within this gene family occur both at equivalent positions and at unrelated positions.

Prochownik et al. (1985) compared the known gene structure of four serpins (ovalbumin, angiotensinogen, α_1-antitrypsin, and antithrombin III) and suggested two models that would account for the placement of introns in serpins. The

first and simplest model assumed that there were at least 15 introns in the ancestral gene. These 15 introns represent the 15 different introns found in the four serpins. The selective loss of introns with time was then suggested to account for the variable gene structure among the four serpins examined. The alternative model assumed that the primordial gene contained only two introns (i.e., the introns that occupy equivalent positions in the separate members). The subsequent introduction and loss of introns was suggested to have created the present diversity of the gene structures. In general, the positions of the introns in the PAI-1 gene do not correspond to those of introns in other genes. Thus, the inclusion of our data into the first model would increase the number of introns in the primordial gene to 22. The average length of the exons in such a gene would be about 50 bp. This option is unlikely since it assumes that exon length would have been significantly smaller in primordial genes than in present genes. Intron 3 of PAI-1 does occupy an equivalent position with intron E of ovalbumin. Thus, incorporation of our data on the position of introns in the PAI-1 gene into the second model would predict that the primordial gene would contain three introns, instead of the proposed two. It is clear that this value also represents a minimum estimate. A more precise definition of the structure of the primordial serpin gene must obviously await the elucidation of the chromosomal structure of several other members of the serpin family.

The data in Figure 3 also suggest that there are constraints placed on the distribution of introns in the PAI-1 gene. This conclusion follows from consideration of the introns in terms of the presumed subdomain structure of the serpins. Carrell and Boswell (1986) have argued that the general folding structure of serpins will be similar because of the corresponding distribution and size of α-helices, β-sheets, and random coil segments in these molecules. Specifically, the three-dimensional structure of these proteins should be comparable with the one established for α_1-antitrypsin (Loebermann et al., 1984). Alignment of the amino acid sequence of PAI-1 with those of α_1-antitrypsin and other serpins allowed us to extrapolate the general subdomain structure of the serpins onto the primary amino acid sequence of PAI-1 (see Figure 3). As can be observed, seven introns are positioned in a gene segment either encoding a boundary of a particular subdomain (α-helix or β-sheet) or located within a nonstructured part of the protein. Only intron 7 is clearly located within a subdomain (β-sheet A4). This intron is unusual in that two acceptor splice sites are optional for it, yielding either PAI-1 or a protein with seven extra amino acids that has poor PAI-1 activity (Pannekoek et al., 1986). In Figure 3 we also analyzed the position of introns in other serpin genes in relation to the suggested subdomain structure of serpins. Although not as obvious as for PAI-1, the introns of rat angiotensinogen, chicken ovalbumin (and gene Y), human antithrombin III, and human α_1-antitrypsin show a preference for a position close to the boundary of a subdomain or for a nonstructured region. In this respect, it is noteworthy to mention that Craik et al. (1983) reported that introns in some gene families coincide with gene segments encoding surface loops. These authors argued that, by a mechanism called "intron sliding", variability of the amino acid sequence can be introduced without affecting the overall folding structure of the protein. Our deductions on the nonrandom distribution of introns in the PAI-1 gene and other serpin genes are compatible with the notion that variability of the amino acid sequence among serpins, due to possible abberations of splicing, would not affect the general structure of these proteins.

An evolutionary model has been advanced to account for the apparent nonrandom placement of introns in some genes of higher organisms (Gilbert, 1978, 1985; Blake, 1978). The model proposes that introns delimit exons that specify autonomous structural and/or functional domains. This hypothesis, commonly called "exon shuffling", implies that exons are inserted by transposition-like events into a preexisting gene in a manner that is concordant with the translational phase of the gene (Patthy, 1985; Rogers, 1985). Examination of the precise location of the introns both within the PAI-1 gene and within other serpin genes (Figure 3) reveals that there is no preference for a particular translational phase. Unless a rather complicated model is conceived for the creation of genes, we feel that it is unlikely that the exon shuffling theory fits for the PAI-1 gene or for other serpin genes.

ACKNOWLEDGMENTS

We acknowledge P. Tayman for typing the manuscript and A. J. van Zonneveld and M. Sawdey for helpful discussions.

Registry No. PAI-1, 105913-11-9.

REFERENCES

Andreasen, P. A., Riccio, A., Welinder, K. G., Douglas, R., Sartorio, R., Nielsen, L. S., Oppenheimer, C., Blasi, F., & Dano, K. (1986) *FEBS Lett. 209*, 213–218.

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.

Blake, C. C. F. (1978) *Nature (London) 273*, 267.

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., & Chambon, P. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4853–4857.

Bucher, P., & Trifonov, E. N. (1986) *Nucleic Acids Res. 14*, 10009–10026.

Carrell, R. W., & Boswell, D. R. (1986) in *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G., Eds.) pp 403–420, Elsevier Science Publishers, Amsterdam.

Catterall, J. F., O'Malley, B. W., Robertson, M. A., Staden, R., Tanaka, Y., & Brownlee, G. G. (1978) *Nature (London) 275*, 510–513.

Cornish-Bowden, A. (1982) *Nature (London) 297*, 625–626.

Craik, C. S., Rutter, W. J., & Fletterick, R. (1983) *Science (Washington, D.C.) 220*, 1125–1129.

Dale, R. M. K., McClure, B. A., & Houchins, J. P. (1985) *Plasmid 13*, 31–40.

Erickson, L. A., Schleef, R. R., Ny, T., & Loskutoff, D. J. (1985) *Clin. Haematol. 14*, 513–530.

Gilbert, W. (1978) *Nature (London) 271*, 501.

Gilbert, W. (1985) *Science (Washington, D.C.) 228*, 823–824.

Ginsburg, D., Zeheb, R., Yang, A. Y., Rafferty, U. M., Andreasen, P. A., Nielsen, L., Dano, K., Lebo, R. V., & Gelehrter, T. D. (1986) *J. Clin. Invest. 78*, 1673–1680.

Go, M. (1981) *Nature (London) 291*, 90–92.

Heilig, R., Muraskowsky, R., Kloepfer, C., & Mandel, J. L. (1982) *Nucleic Acids Res. 10*, 4363–4382.

Hunt, L. T., & Dayhoff, M. O. (1980) *Biochem. Biophys. Res. Commun. 95*, 864–871.

Ish-Horowicz, D., & Burke, J. F. (1981) *Nucleic Acids Res. 9*, 2989–2998.

Kidd, V. J., & Woo, S. L. C. (1986) in *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G., Eds.) pp 421–440, Elsevier Science Publishers, Amsterdam.

Leicht, M., Long, G. L., Chandra, T., Kurachi, K., Kidd, V. J., Mace, M., Davie, E. W., & Woo, S. L. C. (1982) *Nature (London) 297*, 655–659.

Loebermann, H., Tokuoka, R., Diesenhofer, J., & Huber, R. (1984) *J. Mol. Biol. 177*, 531–556.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, pp 1-545, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560-564.
Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., & Green, M. R. (1984) *Nucleic Acids Res. 12*, 7035-7056.
Mount, S. M. (1982) *Nucleic Acids Res. 10*, 459-472.
Ny, T., Sawdey, M., Lawrence, D., Millan, J. L., & Loskutoff, D. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6776-6780.
Pannekoek, H., Veerman, H., Lambers, H., Diergaarde, P., Verweij, C. L., Van Zonneveld, A. J., & Van Mourik, J. A. (1986) *EMBO J. 5*, 2539-2544.
Patthy, L. (1985) *Cell (Cambridge, Mass.) 41*, 657-663.
Prochownik, E. V., Bock, S., & Orkin, S. H. (1985) *J. Biol. Chem. 260*, 9608-9612.
Rogers, G. (1985) *Nature (London) 315*, 458-459.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Southern, E. (1975) *J. Mol. Biol. 98*, 503-517.
Sprengers, E. D., & Kluft, C. (1987) *Blood 69*, 381-387.
Tanaka, T., Ohkubo, H., & Nakanishi, S. (1984) *J. Biol. Chem. 259*, 8063-8065.
Travis, J., & Salvesen, G. S. (1983) *Annu. Rev. Biochem. 52*, 655-709.
Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103-119.

Articles

Role of Calcium Ions and the Heavy Chain of Factor XIa in the Activation of Human Coagulation Factor IX†

Dipali Sinha,*,‡ Frances S. Seaman,‡ and Peter N. Walsh‡,§

Thrombosis Research Center, Department of Medicine, and Department of Biochemistry, Temple University School of Medicine, Philadelphia, Pennsylvania 19140

Received October 27, 1986; Revised Manuscript Received January 13, 1987

ABSTRACT: Since optimal rates of factor IX activation by factor XIa require the presence of calcium ions and the heavy chain of the enzyme as well as the active-site-containing light chain, we have studied the effects of calcium ions and the heavy chain on the reaction kinetics. Whereas the amidolytic activities of factor XIa and of its active-site-containing light chain were almost indistinguishable, the two enzymes behaved quite differently when factor IX was the substrate. Factor XIa was 100-fold more potent in the presence of Ca^{2+} than in its absence. On the contrary, the presence or absence of Ca^{2+} made very little difference in the case of the isolated light chain of factor XIa. Moreover, the enzymatic activity of the light chain was almost identical with that of intact factor XIa when Ca^{2+} was absent. Using an optimal concentration of Ca^{2+}, we studied the activation in the presence of various concentrations of two monoclonal antibodies, one (5F4) directed against the light chain of factor XIa and the other (3C1) against its heavy chain. Analysis of $1/V$ vs. $1/S$ plots showed that whereas inhibition by 5F4 was noncompetitive, 3C1 neutralized the enzyme in a classical competitive fashion. We conclude that in the calcium-dependent activation of factor IX by factor XIa the heavy chain of the enzyme is involved in the binding of the substrate and this is essential for optimal reaction rates.

Factor XI is a coagulation protein present in human plasma in zymogen form at a concentration of 4-6 μg/mL. Since a hemorrhagic state can result from its deficiency, it undoubtedly has a key role in the regulation of blood coagulation (Rosenthal et al., 1953). Factor XI migrates with an apparent molecular weight of 160 000 on sodium dodecyl sulfate ($NaDodSO_4$)[1] gels and consists of two identical disulfide-linked polypeptide chains (Koide et al., 1977; Kurachi & Davie, 1977; Bouma & Griffin, 1977). When activated by factor XIIa, an internal peptide bond in each of its two polypeptide chains is cleaved giving rise to a pair of disulfide-linked heavy and light chains with molecular weights of 50 000 and 30 000, respectively, each light chain containing one active site (Kurachi & Davie, 1977; Bouma & Griffin, 1977). Factor XIa is the activator of factor IX in the intrinsic pathway of blood coagulation (Davie et al., 1979; Fujikawa et al., 1974). The rate of activation of factor IX by factor XIa is greatly increased in the presence of calcium ions. Factor IX activation by factor XIa involves a two-step

† This research was supported by Health and Human Services Grants HL14217, HL25661, and HL36579, by Council for Tobacco Research, Inc., Grant CTR 1389, by a grant from The W. W. Smith Charitable Trust, by a grant from the American Heart Association, Pennsylvania Affiliate, and by Biomedical Research Support Grant S07 RO5417 from the Division of Research Resources, NIH.

* Address correspondence to this author c/o Dr. Peter N. Walsh, Thrombosis Research Center, Temple University School of Medicine.
‡ Thrombosis Research Center, Department of Medicine, Temple University School of Medicine.
§ Department of Biochemistry, Temple University School of Medicine.

[1] Abbreviations: $NaDodSO_4$, sodium dodecyl sulfate; TBS, Tris-buffered saline (50 mM Tris/100 mM NaCl, pH 7.5); BSA, bovine serum albumin; EDTA, ethylenediaminetetraacetate; Tris, tris(hydroxymethyl)aminomethane; pNA, *p*-nitroanilide; TCA, trichloroacetic acid; RIA, radioimmunoassay.

0006-2960/87/0426-3768$01.50/0