

GEN 02757

Cloning and sequencing of cDNA for the rat plasminogen activator inhibitor-1

(Recombinant DNA; hepatoma cells; serine protease inhibitor; serpin; repeated sequence; phage λ library)

Ron Zeheb and Thomas D. Gelehrter

Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109-0618 (U.S.A.)
Tel. (313)763-3460

Received 15 June 1988

Accepted 8 August 1988

Received by publisher 6 September 1988

SUMMARY

A cDNA encoding rat plasminogen activator-inhibitor (PAI-1) has been isolated from an HTC rat hepatoma cell cDNA library constructed in phage λ gt10. The cDNA contains 118 bp of 5'-untranslated sequence, 1206 bp encoding a 402-amino acid (aa) protein and 1747 bp of 3'-untranslated sequence. The protein-coding sequence and the derived amino acid sequence share 82% and 81% identity, respectively, with human PAI-1 cDNA and protein. The rat cDNA encodes a preprotein with a 23-aa leader peptide and a predicted N-terminal serine for the mature protein. Three of four potential N-glycosylation acceptor sites as well as the active site of rat PAI-1 are identical to the human protein. The 3'-untranslated region contains a number of unusual regions, including 80 bp of tandemly repeated G_pA dinucleotides, a 115-bp stretch which shares greater than 90% sequence identity with a region within the 3'-untranslated cDNA of human PAI-1, and two 70-bp stretches of highly T-rich sequence located close to the 3'-terminus of the cDNA.

INTRODUCTION

Plasminogen activator-inhibitors (PAIs) are key regulatory proteins in those physiological and pathological processes which involve plasminogen activator (PA)-catalyzed conversion of plasminogen to plasmin. These processes include fibrinolysis, ovulation, blastocyst implantation, macrophage migra-

tion during inflammation and tissue remodeling (Verstraete and Collen, 1986; Canipari and Strickland, 1985; Strickland et al., 1976; Vassalli et al., 1976; Allen et al., 1986; Dano et al., 1985). Plasminogen activation has also been implicated in tumor invasion and metastasis (Dano et al., 1985). Two types of PAIs, designated PAI-1 and PAI-2, appear to be the major physiological inhibitors of PA

Correspondence to: Dr. R. Zeheb, Department of Human Genetics, University of Michigan Medical School, Medical Science II M4708, 1301 Catherine Rd., Ann Arbor, MI 48109-0618 (U.S.A.).

Abbreviations: aa, amino acid(s); A, adenosine; bp, base pair(s); C, cytidine; cDNA, DNA complementary to RNA; EDTA, ethylenediaminetetraacetic acid; G, guanosine; HTC, hepatoma

tissue culture; IL, interleukin; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); PA, plasminogen activator; PAI, plasminogen activator inhibitor; pfu, plaque-forming units; SSC, 150 mM NaCl/15 mM Na₃ citrate, pH 7; SSPE, 150 mM NaCl/10 mM NaH₂PO₄·H₂O/1 mM EDTA, pH 7; T, thymidine; TBE, 89 mM Tris base/89 mM boric acid/2 mM EDTA; t-PA, tissue-specific PA; uPA, urokinase; UV, ultraviolet light.

activity *in vivo*, and can be distinguished on the basis of biochemical, immunological and molecular biological properties.

PAI-1 is encoded by a single-copy gene on human chromosome No. 7 (Ginsburg et al., 1986). It is the major PA-inhibitor found in the blood, and has been isolated from human endothelial cells (Philips et al., 1984), platelets (Erickson et al., 1984), plasma (Kruithof et al., 1984), and fibrosarcoma cells (Andreasen et al., 1987), and from rat hepatoma cells (Zeheb et al., 1987) and a number of other cultured cell lines (Dano et al., 1985). PAI-1 is a glycoprotein of approximately 50 kDa which rapidly inhibits both tissue-type plasminogen activator (t-PA) and urokinase (uPA), and whose activity is acid- and denaturant-stable (Zeheb et al., 1987). PAI-1 activity can be regulated by a variety of hormones and other effectors including glucocorticoids (Gelehrter et al., 1983, 1987), thrombin (Gelehrter and Szyner-Laszuk, 1986), bacterial endotoxins, IL-1 and tumor necrosis factor (Sprengers and Kluft, 1987; Schleef et al., 1988), transforming growth factor-beta (Lund et al., 1987), and cyclic nucleotides (Barouski-Miller and Gelehrter, 1982; J.H. Heaton and T.D.G., unpublished). Abnormal expression of PAI-1 has been associated with cardiovascular disease, thromboembolic disease, and liver disorders in man (Sprengers and Kluft, 1987; Kruithof et al., 1988). Recently, the nucleotide sequences of the human PAI-1 cDNA (Ginsburg et al., 1986; Ny et al., 1986; Pannekoek et al., 1986; Wun and Kretzmer, 1987) and gene (Bosma et al., 1988; Loskutoff et al., 1987) have been published.

PAI-2 is encoded on human chromosome No. 18 (Webb et al., 1987). It occurs in both a glycosylated (60 kDa) and nonglycosylated (47 kDa) form and has been found in placenta and cells of the monocyte line (Sprengers and Kluft, 1987; Astedt et al., 1985; Wun and Reich, 1987). Serum levels of PAI-2 increase markedly during pregnancy (Kruithof et al., 1987). The cDNA for human PAI-2 has been sequenced (Ye et al., 1987; Schleuning et al., 1987; Antalis et al., 1988).

We have previously demonstrated that HTC rat hepatoma cells synthesize and secrete PAI-1 (Coleman et al., 1982; Cwikel et al., 1984; Coleman et al., 1986). Incubation of HTC cells with the synthetic glucocorticoid dexamethasone induces a rapid, five to ten-fold increase of PAI-1 activity,

protein and mRNA (Coleman et al., 1982; Cwikel et al., 1984; Gelehrter et al., 1987; J.H. Heaton and T.D.G. unpublished). In order to investigate the structural and functional properties of rat PAI-1, its relationship with other serine protease inhibitors (serpins) and its mode of regulation, we have cloned and sequenced the cDNA encoding HTC rat hepatoma cell PAI-1.

MATERIALS AND METHODS

(a) cDNA library construction

A cDNA library was constructed in the bacteriophage vector λ gt10 using mRNA extracted from HTC rat hepatoma cells which had been incubated for 16 h in the presence of 1 μ M dexamethasone (Sigma Chemical Co.). Total RNA was isolated from approximately 5×10^8 cells using a modified guanidine HCl procedure (Strohman et al., 1977), and the poly(A)⁺ RNA purified by affinity chromatography using oligo(dT)-cellulose (Pharmacia Inc.) (Maniatis et al., 1982). cDNA was synthesized using oligo(dT)-primers and reverse transcriptase utilizing a cDNA synthesis kit (Amersham Corp.). The cDNA was methylated to protect internal *Eco*RI sites, then ligated to *Eco*RI linkers according to the instructions provided by the manufacturer. The ligated DNA was annealed to phage λ gt10 arms and the resultant concatamers were packaged, *in vitro*, using Gigapack Gold (Stratagene) packaging extracts. The complete recombinant phage were then used to infect *E. coli* strains L87 (nonselective host) and NM514 (Hfl⁺, selective host), supplied with the kit. The library was characterized on both bacterial strains; subsequent plating and screening was performed using only the selective host.

(b) Screening of cDNA library

The unamplified HTC cDNA library, plated onto the selective bacterial host NM514, was screened by standard plaque hybridization techniques (Maniatis et al., 1982) utilizing the previously cloned human PAI-1 cDNA as a probe (Ginsburg et al., 1986). The probe was radiolabeled using the random primer method (Feinberg and Vogelstein, 1983) and used to

screen duplicate filters under stringent conditions (40% formamide, 5 × SSPE at 42°C). Filters were washed to a final stringency of 0.5 × SSC, 42°C and positive phage were plaque-purified.

(c) Characterization of cloned inserts

Plaque-purified phage were grown as minipreps, and phage DNA isolated as described (Maniatis et al., 1982). The purified DNA was digested with *EcoRI* (Boehringer-Mannheim) in order to release the cloned inserts, and analysed on 1% agarose (BRL) gels using TBE (Sigma, Fisher) as the electrode buffer (Maniatis et al., 1982). After staining with ethidium bromide (Sigma), bands were visualized under UV irradiation (Transilluminator, UVP Inc.) and the size of the cDNA inserts determined relative to size standards (*HindIII*- or *BstEII*-digested λ DNA, New England Biolabs, Inc.). Insert DNA was subcloned into the *EcoRI* sites of the multifunctional plasmid vectors pBSM13(-) and pBluescript SK(-) and used to transfect competent host *E. coli* strain AG-1 (Stratagene, Inc., La Jolla, CA). Plasmids were grown as minipreps and DNA was isolated and used to construct restriction maps.

(d) Nucleotide sequence analysis

Sequencing of cloned DNA inserts was facilitated by construction of nested deletions using exonuclease III/mung bean nuclease digestion, according to the procedure described by Stratagene. For construction of the nested deletions, DNA was isolated from large-scale preparations and purified by CsCl (Boehringer-Mannheim) banding (Maniatis et al., 1982). Sequence analysis was performed directly from double-stranded plasmid DNA according to the method of Sanger et al. (1977), as modified by Hattori and Sakaki (1986), and using deoxyadenosine 5'-[α -³⁵S]thiotriphosphate (Amersham). Both the coding and noncoding strands were sequenced and the data analysed with the IBI-Pustell sequence analysis programs (International Biotechnologies, Inc.) (Pustell and Kafatos, 1986), and with the Sequence Analysis Software Package (version 5) of the Genetics Computer Group at the University of Wisconsin (Devereux et al., 1984).

(e) Cell cultures

HTC rat hepatoma cells were cultured as previously described and the PAI-1 activity in the medium of the cultured cells was determined by a two-step esterolytic assay (Gelehrter et al., 1987).

RESULTS AND DISCUSSION

(a) Characterization and screening of the HTC cDNA library

By titring the HTC cDNA library on both the nonselective (L87) and selective (NM514) bacterial hosts, the library was characterized with respect to background (3.7×10^3 pfu/ml), number of recombinants (5.1×10^4 pfu/ml), percent recombinants (93%) and cloning efficiency (approximately 1×10^6 pfu/ μ g of DNA). In addition, purification of phage DNA from 18 randomly selected clones revealed that 17 of the 18 (94%) contained insert DNA ranging in size from approximately 400 bp to greater than 3.5 kb and averaging 2 kb.

The library (6×10^4 plaques) was screened using the 2-kb human PAI-1 cDNA (PAIB6) as a probe (Ginsburg et al., 1986). This probe does not contain either a poly(A) tail or the partial *Alu* repeated sequence which is present in the 3'-untranslated region of the 3-kb human PAI-1 cDNA (PAIB4). On primary screening, the probe hybridized to 26 plaques, 24 of which remained positive through secondary and tertiary rounds of plaque-purification and rescreening. These results suggest that the abundance of the message is approximately 1:2000 in the library. Studies from this laboratory (J.H. Heaton and T.D.G., unpublished) have shown that dexamethasone causes approximately a five-fold increase in the steady-state accumulation of PAI-1 mRNA; therefore, the abundance of PAI-1 mRNA in untreated HTC cells is approximately 0.01% of the total poly(A)⁺ RNA.

Phage DNA was purified from the positive clones and the size of the insert DNA determined. One clone, harboring an insert of approximately 3 kb, was selected for further analysis based on the close correspondence of its size to that of HTC PAI-1 mRNA, determined by Northern blot analysis of

PARTIAL RESTRICTION MAP OF HTC RAT PAI-1 cDNA

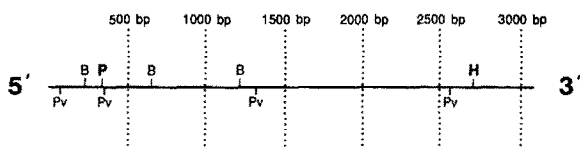


Fig. 1. Partial restriction map of rat PAI-1 cDNA. The map was constructed by digesting purified plasmid pBSM13(-), containing the PAI-1 cDNA insert, with the indicated restriction endonucleases singly and in combination. Digestion fragments were analysed by electrophoresis in 1% agarose gels, stained with ethidium bromide and visualized under UV irradiation. The cDNA clone is oriented 5' to 3' with respect to the coding strand. B, *BglII*; H, *HindIII*; P, *PstI*; Pv, *PvuII*.

total HTC cell RNA probed with the human PAIB6 cDNA (J.H. Heaton and T.D.G., unpublished). The insert DNA was subcloned into pBSM13(-) and into pBluescript SK(-) and a partial restriction map determined (Fig. 1). The presence of unique *PstI* and *HindIII* sites at the 5'- and 3'-ends of the insert, respectively, were used to orient the clone.

(b) Nucleotide sequence determination

Nested deletions of the subcloned DNA, overlapping by 150–300 bp, were constructed and both strands sequenced as described in MATERIALS AND METHODS, section d. The nucleotide sequence of the cDNA encoding rat PAI-1 is shown in Fig. 2. The HTC PAI-1 cDNA is 3071 bp in length and consists of 118 bp of 5'-untranslated sequence with an in-frame stop codon at position -51, 1206 bp encoding a 402-aa protein, and 1747 bp of 3'-untranslated sequence terminating in a poly(A) tail. Comparison of the nucleotide sequence of the rat PAI-1 cDNA with the previously characterized human PAI-1 cDNA (Ginsburg et al., 1986; Ny et al., 1986; Pannekoek et al., 1986; Wun and Kretzmer, 1987) shows that these three regions share 57%, 82%, and 59% sequence identity, respectively. The cDNAs are 68% identical, overall.

(c) Analysis of the 5'-untranslated region

The rat PAI-1 cDNA has 118 bp of 5'-untranslated sequence and is estimated to be missing 25–35 nt from the 5'-end. This estimate is supported

by a comparison of the sequence and organization of the rat and human (Bosma et al., 1988) PAI-1 genes. Analysis of the rat PAI-1 gene (C.J. Bruzdinski and T.D.G., unpublished) has identified a TATA box starting 174 nt upstream of the translational start codon. In addition, there is an intron/exon boundary at position -1 from the translation-start codon. The human gene has a TATA box which starts at position -173, and the boundary between exons 1 and 2 occurs at position -1. The high degree of similarity in the organization of these two genes suggests that the transcriptional start point of human PAI-1 (position -145) and that of the rat message may be similar. Finally, a full-length rat PAI-1 mRNA, starting at position -145, would result in a 29-bp distance between the TATA box and the transcriptional start point; a number which is in close agreement with the 30-bp distance generally observed in other eukaryotic genes (Bucher and Trifonov, 1986).

(d) Analysis of the protein-coding region

The rat PAI-1 protein shares considerable sequence similarity with other members of the serine protease inhibitor family (Table I). The rat PAI-1 cDNA encodes a preprotein with a 23-aa hydrophobic leader peptide, the cleavage of which results in a 379-aa mature protein having an N-terminal serine residue (Fig. 3). This is consistent with primary amino acid sequence data obtained on a sample of purified HTC PAI-1 protein (data not shown). The N-terminal regions of rat and human PAI-1 show the greatest extent of sequence divergence. Ten out of the first 14 aa of the mature rat protein differ from those of human PAI-1, whereas overall they share 81% sequence identity (89% sequence similarity). Divergence of the N-terminal region is a feature of other serpins and may represent an area where specific functional differences occur, or a relatively nonessential area which is able to tolerate a high frequency of amino acid substitutions without compromising the function of the protein. Since rat and human PAI-1 appear to function in the same way, it seems unlikely that the N-terminal sequence divergence reflects a functional difference between these two proteins. Four potential *N*-glycosylation acceptor sites are predicted (Fig. 3), three of which are identical to those predicted for human PAI-1. We have shown previously that full

-118 CCCCAGAG CTTTGTGAAG GAGGAACGCT GCACACCCGC CTCGCGCAGC ACACAGCCAA -59 1443 ACTCACAGAA TCTGGAATCG CCCAGCAATC AGCCTTTGGA CCCAGTTCOC ACCAAGGCC 1502
 -58 CCACAGCTCA GGCACAGCA ACAAGAGCCA ATCACRAGGC ACTTCCGAAA GCTCCAGGAT 2 1503 CAGCAGGGCC AACCCACTAC GCCTTCACTC AGCAAAGTCT TTTGTGTTC ACACACTC 1562
 3 GCAGATGCT TCAGCCCTCA CTTGCCCTAC CTTGGCCCTG GTTCTGGTCT TTGGGAAGG 62 1563 TGGGTACCTC TTGTATTCGC CTCCAITGCT ATGAAGGATG ACCCAGGCCA AAGGAAGAAG 1622
 63 GTTCGCTTCA CCCCTTCAG AGTCCCATAC AGCCACAGC GCCACCAACT TCGAGTAAA 122 1623 CACTGTCTA TCTCAAGCT CACTGTGAA ATGAACACT TGCCCATCCC CAAGGGGAG 1682
 123 AGTGTTCAG CATGTGTCC AGCCCTCCA AGACCGAAT GTGGTCTTCT CTCCTTAGG 182 1683 CAGATAGACA GATCGAATGA TCGCCGATA TCAAGCCTTC TCCCAGCTCC CGTCTGCCC 1742
 183 CGTGTCTCG GTGCTGGCTA TGCTGCAGCT GACCACAGCA GGAACACCC GGCAGCAGAT 242 1743 TCCCTTCCCT GACAGCCGCC TTGTGTATT TCAGATGCTA GATCACTGT TTACAGCTTT 1802
 243 CCAAGATCT ATGGATTCA ATATCAGTGA GAGGGGACA GCTCTGCCC TCGAAGCT 302 1803 TTTCCAGCCA CAACCTTTC TCAITTTGAA ACGGTGAAG AAGGTCAGA TGTGCAGTG 1862
 303 CTCGAAGG CTCATGGGT CATGGAACAA GAATGATC AGTACTGGG ACCCACTT 362 1863 CTTTCTCTT TATCTGGT CTCCCTGTA GGGGAGGG GTCTGGGA GATTCAGGG 1922
 363 TGTCCAGG GACCTAGAG TGTTCAGG CTTCAAGCC CACTTCTCA AGCTCTCG 422 1923 GTGTGATGA ATATTTATCT TGTATCTT ATACGTTTGT TGGGAGAG AAGCACTATT 1982
 423 GACCAGGT AAGCAGTGG ACTTCTAGA GGTGGAAGA GCCAGATTCA TCAACAACA 482 1983 AAGGAGAAAG CCTTTTATT AAACCATGG ATATGTTGTC CCAITTTGGG TCTGTATCCC 2042
 483 CTGGTGGAG AGGCACACA AAGGTATGAT CAGTACTTA CTGGCAAGG GGCCTGAAA 542 2043 TGTATGTCAG GAGGCACTA CTCACAAAC CCGCCCTTCG GGTGGCCCG CGTCGGGCT 2102
 543 TGAGCTGACA GGCCTGTGC TGTGAAAC CCTGATTTT AACGGCCAA TGGAAACCC 602 2103 CACTGTCCG CTTAGTGA GCGAAACAG CCGTTGCCCC ATCCCTCCC CGCATCTTC 2162
 603 CTTCTTAGG GCAACACC ACCAGGCCCT GTTCCACAAG TCTGATGTA GCACATCTC 662 2163 CCCGTGGCTC TTTTCTTAG GATCTTGGC AAGGTATGC TTGGCAGCCC ACGGTAAAAG 2222
 663 CGTCCCATG ATGGCTAGA ACAAAGTT CAACTACT GATTCACA CTCGGATGG 722 2223 AAGGGGAAA AAGATTAGT GGGAGAGA GAGAGAGA GAGAGAGA GAGAGAGA 2282
 723 GCAGGATC GACATCTGG AACTGCCCTA CCAGGGGAA ACCCTCAGCA TGTCTATTG 782 2283 GAGAGAGA GAGAGAGA GAGAGAGA GAGAGAGA GATCTTTGAG AGAGGCAAA 2342
 783 AGCACCTTT GAARAAGT TGCCCTCTC CGCATCACC AACATTTGG ACGCTAGCT 842 2343 GTGCTTCAA ATTTTTCAA TACATTCAGA ACCCAGTGG GAATGGGG TGTGTGACCT 2402
 843 CATCAGAAA TGAAGAGCA ACATGACCAG GCTGCCCCG CTCCTATCC TGCTAAGTT 902 2403 AACAGGACAG AACTTTCTC AATPACTGG TGACTGGT GCACAGCT GCACCTGGTA CTCACITCAA 2462
 903 CTCTCTGG ACTGAATGG ACCTCAGAG GCCCTGGAG AAGCTGGCA TGACTGACAT 962 2463 TGTCTAAT CCGGCTGCTG TAAGTGAACA GNGACAGT GGGGGGGGG GGTGAGGAT 2522
 963 CTTCACTCA ACCCAGGCC ACTTCAAG TCTTTCCGAC CAGAGACAG TCTCTGTAGC 1022 2523 GAAAGAAC GCCAGTCTT GGTCAACCAC CTTAGTTAGA TAAATTTTT TGAAGCTTC 2582
 1023 ACAAGCTCA CAAAAGTCA AGATCGAGT GAAGCAGGC GGCACAGTG 1082 2583 CTAGCTGGG GTATGATCAG AAAACCAAT TACTGAAAA CTGCACAGCA AGGTAACGTG 2642
 1083 CACAGCAAT CTAGTCTAG CCGGATGG CCCACAGGC ATGGTTTAG ACGATCTTT 1142 2643 AATCTAAT CATAGGGGC CGCTGTCAT CGTTACATC TCCACTGGAA AAAATPATC 2702
 1143 TCTCTTTGT GTTCGGCACA ATCCACAGA GACAATCTC TTCATGGGC AGTGTATGA 1202 2703 ATTTCTTT TGTGTGTG TGTGTGTTT ACGTTTCTC CCTCTCCC TTTCTCTCT 2762
 1203 GCCTTAGAG TGGATGAGA AGCTTTCT TTGGACAAA ACTGGACGTG TTATAGCAG 1262 2763 AATTCATAT GCACCTGATA ACCATACC GGTACCACA GGGGCCAAA TGTGGGTCA 2822
 1263 AGACTCTGA GAAAAGATT GTTTAAGGA CTCCTTGGG AGAAGAGAA GGCCTTTCT 1322 2823 CATGGCTTG AATTTTGG GGTACATAG CTTTGTGTTG TTTGTTTCA CTTTGTAT 2882
 1323 TCTTACCCG GCAGTGTAA ATCTTCCA ACAGCTCCC AGACCTCAGA CTCGGAAGA 1382 2883 ATAAACAGT AATGCTGTT TAAAAATA AAAAAATAG AGAATATGA GACAAAAA 2942
 1383 GAAAAGAGT TAACTCCCT ACTAGGGACC TATCTACTA AGGTCTATC CAACATAGA 1442 2943 AAAAAAAAAA A 2953

Fig. 2. Nucleotide sequence of the cDNA encoding rat PAI-1. The nucleotides are numbered beginning with the first nt of the cDNA insert. The translation-start codon is shown by rightward arrowhead. The translation-termination codon and an upstream, in-frame stop codon are shown by leftward arrowheads. T-rich stretches of sequence and a G_pA dinucleotide repeated sequence are shown by single- and double-line overlays, respectively. A region having greater than 90% sequence identity with a segment in the 3'-untranslated cDNA of human PAI-1 is boxed. The polyadenylation signal is indicated by five heavy dots.

TABLE I

Amino acid sequence similarities between rat PAI-1 and related proteins

Protein	Sequence identity ^a %	Sequence similarity ^b %
PAI-1 (human)	81	89
PAI-2 (human)	30	52
Antithrombin III (human)	28	52
α_1 -Antitrypsin (human)	28	53
α_1 -Antichymotrypsin (human)	29	51
Ovalbumin (chicken)	27	51
Gene Y protein (chicken)	31	55

^a Related proteins were identified by scanning the Genbank peptide file using the WordSearch program (Devereux et al., 1984). Optimal alignment of rat PAI-1 to each related protein was determined using the Gap alignment program with a match display threshold of 1.5. At this setting only exact matches are used in calculating the percent amino acid sequence identity.

^b Sequence similarity was determined as in ^a except that the match display threshold was set to 0.5, a setting which scores 'conservative' amino acid changes as a match (Devereux et al., 1984).

glycosylation of the protein is not needed for activity (Zeheb et al., 1987). By homology to human PAI-1, the active site P1 and P1' residues are predicted to be arginine 369 and methionine 370, respectively (numbering based on the 402-aa preprotein). The amino acids at the active site are identical to those identified for the active site of human PAI-1, and present an appropriate site for cleavage by arginine-specific serine proteases.

The human PAI-1 amino acid sequence is more closely related to rat PAI-1 (81% sequence identity) than it is to human PAI-2 (30% sequence identity (Ginsburg et al., 1986)), suggesting that PAI-1 and PAI-2 diverged prior to the separation of these two species.

(e) Analysis of the 3'-untranslated region

The rat PAI-1 cDNA has a long 3'-untranslated region (1747 bp) within which are located a number

of unusual and interesting sequences. Unlike human PAI-1, for which there are two size classes of message, only one size mRNA is found in Northern blots of rat HTC hepatoma cells (J.H. Heaton and T.D.G., manuscript in preparation) and normal hepatocytes (J.H. Heaton, V.L. Nebes, L.G. O'Dell, S.M. Morris Jr. and T.D.G., Molecular Endocrinology, submitted). The size of the mRNA (approximately 3 kb) corresponds well to the size of the cDNA. It has been suggested that the smaller human PAI-1 mRNA is formed by polyadenylation in response to a weak upstream poly(A) addition signal which does not match the consensus sequence. Candidate poly(A) addition sequences have been proposed including CCTAAA (located at position 2000 bp of the cDNA), AAATAA (position 2049 bp), TTTAAA (position 2060 bp), (Ginsburg et al., 1986) and AATAAT (position 11 174 bp of the gene sequence, position 2050 bp of the cDNA) (Bosma et al., 1988). Polyadenylation in response to any of these signals could result in the formation of a smaller message whose size would correspond to the observed size (approximately 2 kb) of the smaller human mRNA. Comparison with the rat cDNA sequence reveals that there is one consensus poly(A) addition signal, AATAAA, located 18 bp 5' to the start of the poly(A) tail, and that of the four possible alternative sites found in the human sequence, only one (TTTAAA) is preserved intact in the rat sequence. Since only one size message is formed in the rat, this would suggest that this site is not an alternative poly(A) addition signal.

The 3'-untranslated region contains an 80-bp segment of polypurines organized in the form $(GA)_{34}A_2(GA)_5$. Tandem repeats of the dinucleotide G_pA have been found within the introns of other genes including the murine immunoglobulin μ - δ heavy-chains (Richards et al., 1983) and IgG3 constant region (Wels et al., 1984), a murine pro-opiomelanocortin pseudogene (Notake et al., 1983; Uhler et al., 1983), and the 3'-end spacer of mouse ribosomal RNA genes (Kominami and Muramatsu, 1983). In addition, another serpin, mouse α_1 -antitrypsin, has a long G_pA dinucleotide repeat approximately 300 bp upstream from the start of transcription (Krauter et al., 1986). The rat PAI-1 cDNA appears to be the first example of such a nucleotide sequence appearing in the exon of a gene. As yet, no clear-cut function has been associated with any of

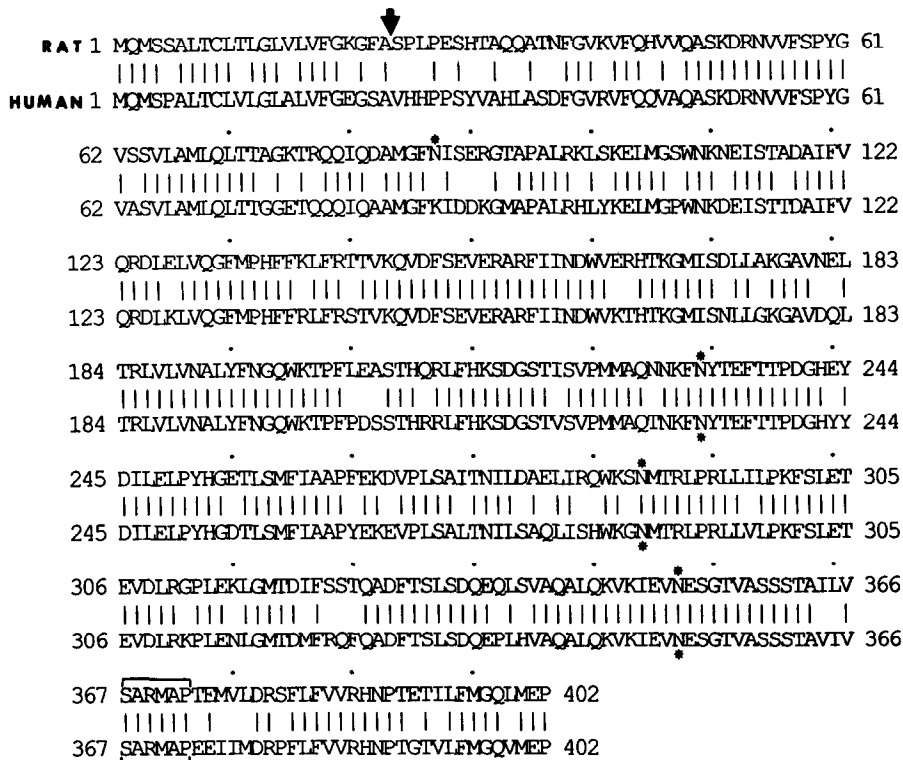


Fig. 3. Alignment of rat and human PAI-1 peptide sequences. Alignment was determined as described in footnote a to Table I. The amino acids are numbered beginning with the initiator methionine. Identical amino acids are connected by vertical lines. The downward arrow indicates the leader peptide cleavage site. Potential *N*-linked glycosylation acceptor sites are marked by asterisks. The active sites are enclosed by brackets.

these sequences, though it has been proposed that they may serve as 'hot spots' for homologous recombination. The occurrence of such a sequence in the 5' 'switch' region of the murine immunoglobulin gene (Wels et al., 1984) has prompted speculation that it may be involved in class-switching. It is interesting to note that the human PAI-1 gene contains eleven sequences which are 77% to 90% identical to the consensus for the *Alu* repeated sequence; four of these are contiguous on their 3'-ends with long stretches of polypurine repeats which basically conform to an organization of $(GA_3)_n$ (Bosma et al., 1988). None of the human PAI-1 gene *Alu*-like sequences or polypurine sequences occur within exons.

Sixty bp 3' to the end of the G_pA dinucleotide repeat is a 115-bp region which has greater than 90% sequence identity to a stretch of nucleotide sequence in the 3'-untranslated region of human PAI-1 (Fig. 2). This extremely high degree of sequence

identity is unusual in that it occurs in non-protein coding regions of the two cDNAs. The preservation of a high degree of nucleotide sequence similarity within the 3'-untranslated regions of mRNAs from different animal species has been reported for a number of highly conserved proteins including actin (Ordahl and Cooper, 1983), tubulin (Cowan et al., 1983), myosin (Saidapet et al., 1984) and the B subunit of creatine kinase (Billadello et al., 1986; Papenbrock and Wille, 1986). Recently, the cDNA encoding human multilineage-colony-stimulating factor was identified by hybridization with highly conserved regions in the 3'-untranslated region of mouse IL-3 (Dorssers et al., 1987). Nucleotide sequence preservation within the 3'-untranslated region implies a specific and essential function. As yet, the nature of that function is unknown.

Finally, close to the 3'-end of the cDNA are two, approximately 70-bp stretches of highly T-rich sequence, one of which is located in a position

analogous to the AT-rich region previously identified in the human PAI-1 cDNA (Ginsburg et al., 1986). Studies of message stability have shown that mRNA from some genes are degraded at a much faster rate than others (Ross and Kobs, 1986). The presence of AU-rich sequences in the 3'-untranslated regions, in particular the nucleotide sequence UUAUUUAU, may be associated with message lability (Shaw and Kamen, 1986). The highly T-rich regions of the rat PAI-1 cDNA may be examples of other sequences which may play a role in the regulation of mRNA stability.

(f) Conclusions

(1) A nearly full-length cDNA encoding rat PAI-1 from HTC cells has been cloned and sequenced. The cDNA is 3071 bp in length and consists of 118 bp of 5'-untranslated sequence, a 1206-bp open reading frame, and 1747 bp of 3'-untranslated sequence. The HTC PAI-1 cDNA shares 68% sequence identity with the cDNA encoding human endothelial cell PAI-1.

(2) The cDNA encodes a 402-aa preprotein having a 23-aa leader peptide and four potential acceptor sites for *N*-linked glycosylation. The rat PAI-1 protein shares considerable amino acid similarity with other members of the serine protease inhibitor family. Rat and human PAI-1 share 81% amino acid sequence identity (89% sequence similarity).

(3) The 3'-translated region of the cDNA contains a number of unusual regions which may be involved in the regulation of PAI-1 gene expression in HTC cells. These regions include 80 bp of (G_pA)-repeated dinucleotide sequence; 115 bp which share greater than 90% sequence identity with a region within the 3'-untranslated cDNA of human PAI-1; and two 70-bp stretches of highly T-rich sequence located close to the 3'-terminus of the cDNA.

ACKNOWLEDGEMENTS

This work was supported by National Institute of Health research grant CA22729 (T.D.G.) and by a University of Michigan Gastrointestinal Hormone Research Center Grant (DK34933) pilot award (R.Z.). The authors are grateful to Dr. J.L. Slightom

for sharing the human PAI-1 gene sequence manuscript prior to publication. We thank Miguel A. Rodriguez for technical assistance and Ms. Joanne Heaton and Dr. Patrick Venta for critical review of the manuscript.

REFERENCES

- Allen, E.H., Hamilton, J.A., Medcalf, R.L., Kubota, M. and Martin, T.J.: Cyclic AMP-dependent and -independent effects on tissue-type plasminogen activator activity in osteogenic sarcoma cells; evidence from phosphodiesterase inhibition and parathyroid hormone antagonists. *Biochim. Biophys. Acta.* 888 (1986) 199–207.
- Andreassen, P.A., Pyke, C., Riccio, A., Kristensen, P., Nielsen, L.S., Lund, L.R., Blasi, F. and Dano, K.: Plasminogen activator inhibitor type 1 biosynthesis and mRNA level are increased by dexamethasone in human fibrosarcoma cells. *Molec. Cell. Biol.* 7 (1987) 3021–3025.
- Antalis, T.M., Clark, M.A., Barnes, T., Lehrbach, P.R., Devine, P.L., Schevzov, G., Goss, N.H., Stephens, R.W. and Tolstoshev, P.: Cloning and expression of a cDNA coding for a human monocyte-derived plasminogen activator inhibitor. *Proc. Natl. Acad. Sci. USA* 85 (1988) 985–989.
- Astedt, B., Lecander, I., Brodin, T., Lundblad, A. and Low, K.: Purification of a specific placental plasminogen activator inhibitor by monoclonal antibody and its complex formation with plasminogen activator. *Thromb. Haemostasis* 53 (1985) 122–125.
- Barouski-Miller, P.A. and Gelehrter, T.D.: Paradoxical effects of glucocorticoids on regulation of plasminogen activator activity of rat hepatoma cells. *Proc. Natl. Acad. Sci. USA* 79 (1982) 2319–2322.
- Billadello, J.J., Kelly, D.P., Roman, D.G. and Strauss, A.W.: The complete nucleotide sequence of canine brain B creatine kinase mRNA: homology in the coding and 3' noncoding regions among species. *Biochem. Biophys. Res. Comm.* 138 (1986) 392–398.
- Bosma, P.J., Van den Berg, E.A., Kooistra, T., Siemieniak, D.R. and Slightom, J.L.: Human plasminogen activator inhibitor-1 gene. *J. Biol. Chem.* 263 (1988) 9129–9141.
- Bucher, P. and Trifonov, E.N.: Compilation and analysis of eukaryotic PolII promoter sequences. *Nucleic Acids Res.* 14 (1986) 10009–10026.
- Canipari, R. and Strickland, S.: Plasminogen activator in the rat ovary. *J. Biol. Chem.* 260 (1985) 5121–5125.
- Coleman, P.L., Barouski, P.A. and Gelehrter, T.D.: The dexamethasone-induced inhibitor of fibrinolytic activity in hepatoma cells. *J. Biol. Chem.* 257 (1982) 4260–4264.
- Coleman, P.L., Patel, P.D., Cwikel, B.J., Rafferty, U.M., Szyner-Laszuk, R. and Gelehrter, T.D.: Characterization of the dexamethasone-induced inhibitor of plasminogen activator in HTC hepatoma cells. *J. Biol. Chem.* 261 (1986) 4352–4357.

- Cowan, N.J., Dobner, P.R., Fuchs, E.V. and Cleveland, D.W.: Expression of human alpha-tubulin genes: interspecies conservation of 3' untranslated regions. *Mol. Cell. Biol.* 3 (1983) 1738-1745.
- Cwikel, B.J., Barouski-Miller, P.A., Coleman, P.L. and Gelehrter, T.D.: Dexamethasone induction of an inhibitor of plasminogen activator in HTC hepatoma cells. *J. Biol. Chem.* 259 (1984) 6847-6851.
- Dano, K., Andreasen, P.A., Grondahl-Hansen, J., Kristensen, P., Nielsen, L.S. and Skriver, L.: Plasminogen activators, tissue degradation, and cancer. *Adv. Cancer Res.* 44 (1985) 139-266.
- Devereux, J., Haerberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12 (1984) 387-395.
- Dorssers, L., Burger, H., Bot, F., Delwel, R., Geurts van Kessel, A.H.M., Löwenberg, B. and Wagemaker, G.: Characterization of a human multilineage-colony-stimulating factor cDNA clone identified by a conserved noncoding sequence in mouse interleukin-3. *Gene* 55 (1987) 115-124.
- Erickson, L.A., Ginsberg, M.H. and Loskutoff, D.J.: Detection and partial characterization of an inhibitor of plasminogen activator in human platelets. *J. Clin. Invest.* 74 (1984) 1475-1472.
- Feinberg, A.P. and Vogelstein, B.: A technique for radiolabeling cDNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132 (1983) 6-13.
- Gelehrter, T.D. and Szynger-Laszuk, R.: Thrombin induction of plasminogen activator-inhibitor in cultured human endothelial cells. *J. Clin. Invest.* 77 (1986) 165-169.
- Gelehrter, T.D., Barouski-Miller, P.A., Coleman, P.L. and Cwikel, B.J.: Hormonal regulation of plasminogen activator in rat hepatoma cells. *Mol. Cell. Biochem.* 53-54 (1983) 11-21.
- Gelehrter, T.D., Szynger-Laszuk, R., Zeheb, R. and Cwikel, B.J.: Dexamethasone inhibition of tissue-type plasminogen activator (tPA) activity: paradoxical induction of both tPA antigen and plasminogen activator inhibitor. *Mol. Endocrinol.* 1 (1987) 97-101.
- Ginsburg, D., Zeheb, R., Yang, A.Y., Rafferty, U.M., Andreasen, P.A., Nielsen, L., Dano, K., Lebo, R.V. and Gelehrter, T.D.: cDNA cloning of human plasminogen activator-inhibitor from endothelial cells. *J. Clin. Invest.* 78 (1986) 1673-1780.
- Hattori, M. and Sakaki, Y.: Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* 152 (1986) 232-238.
- Kominami, R. and Muramatsu, M.: A mouse type 2 Alu sequence (M2) is mobile in the genome. *Nature* 301 (1983) 87-89.
- Krauter, K.S., Citron, B.A., Hsu, M.-T., Powell, D. and Darnell, J.D. Jr.: Isolation and characterization of the α_1 -antitrypsin gene of mice. *DNA* 5 (1986) 29-36.
- Kruihof, E.K.O., Gudinchet, A. and Bachmann, F.: Plasminogen activator inhibitor 1 and plasminogen activator inhibitor 2 in various disease states. *Thromb. Haemostasis* 59 (1988) 7-12.
- Kruihof, E.K.O., Tran-Thang, C., Gudinchet, A., Hauert, J., Nicoloso, G., Genton, C., Welti, H. and Bachmann, F.: Fibrinolysis in pregnancy: a study of plasminogen activator inhibitors. *Blood* 69 (1987) 460-466.
- Kruihof, E.K.O., Tran-Thang, C., Ransijn, A. and Bachmann, F.: Demonstration of a fast-acting inhibitor of plasminogen activators in human plasma. *Blood* 64 (1984) 907-913.
- Loskutoff, D.J., Linders, M., Keijer, J., Veerman, H., van Heerikhuizen, H. and Pannekoek, H.: Structure of the human plasminogen activator inhibitor 1 gene: nonrandom distribution of introns. *Biochemistry* 26 (1987) 3763-3768.
- Lund, L.R., Riccio, A., Andreasen, P.A., Nielsen, L.S., Kristensen, P., Laiho, M., Saksela, O., Blasi, F. and Dano, D.: Transforming growth factor-beta is a strong and fast acting positive regulator of the level of type-1 plasminogen activator inhibitor mRNA in WI-38 human lung fibroblasts. *EMBO J.* 6 (1987) 1281-1286.
- Maniatis, T., Fritsch, E.F. and Sambrook, J.: *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.
- Notake, M., Tobimatsu, T., Watanabe, Y., Takahashi, H., Mishina, M. and Nujma, S.: Isolation and characterization of the mouse corticotropin- β -lipotropin precursor gene and a related pseudogene. *FEBS Lett.* 156 (1983) 67-71.
- Ny, T., Sawdey, M., Lawrence, D., Millan, J.L. and Loskutoff, D.J.: Cloning and sequence of a cDNA coding for the human β -migrating endothelial-cell-type plasminogen activator inhibitor. *Proc. Natl. Acad. Sci. USA* 83 (1986) 6776-6780.
- Ordahl, C.P. and Cooper, T.A.: Strong homology in promoter and 3'-untranslated regions of chick and rat β -actin genes. *Nature* 303 (1983) 348-9.
- Pannekoek, H., Veerman, H., Lambers, H., Diergaarde, P., Verweij, L., Van Zonneveld, A.-J. and Van Mourik, J.A.: Endothelial plasminogen activator inhibitor (PAI): a new member of the Serpin gene family. *EMBO J.* 5 (1986) 2599-2544.
- Papenbrock, T. and Wille, W.: The 3' noncoding region of the mouse brain B creatine kinase mRNA: a sequence with exceptional homology among species. *Nucleic Acids Res.* 14 (1986) 86-90.
- Philips, M., Juul, A.-G. and Thorsen, S.: Human endothelial cells produce a plasminogen activator inhibitor and a tissue-type plasminogen activator-inhibitor complex. *Biochim. Biophys. Acta* 802 (1984) 99-110.
- Pustell, J. and Kafatos, F.C.: A convenient and adaptable micro-computer environment for DNA and protein sequence manipulation and analysis. *Nucleic Acids Res.* 14 (1986) 479-488.
- Richards, J.E., Gilliam, A.C., Shen, A., Tucker, P.W. and Blattner, F.R.: Unusual sequences in the murine μ - δ heavy-chain region. *Nature* 306 (1983) 483-487.
- Ross, J. and Kobs, G.: H4 histone messenger RNA decay in cell-free extracts initiates at or near the 3' terminus and proceeds 3' to 5'. *J. Mol. Biol.* 188 (1986) 579-593.
- Saidapet, C., Khandekar, P., Mendola, C. and Siddiqui, M.A.Q.: Tissue specificity of 3'-untranslated sequence of myosin light chain gene: unexpected interspecies homology with repetitive DNA. *Arch. Biochem. Biophys.* 233 (1984) 565-572.
- Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74 (1977) 5463-5467.
- Schleuning, W.-D., Medcalf, R.L., Hession, C., Rothenbuhler, R., Shaw, A. and Kruihof, E.K.O.: Plasminogen activator

- inhibitor 2: regulation of gene transcription during phorbol ester-mediated differentiation of U-937 human histiocytic lymphoma cells. *Mol. Cell. Biol.* 7 (1987) 4564–4567.
- Shaw, G. and Kamen, R.: A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* 46 (1986) 659–667.
- Schleef, R.R., Bevilacqua, M.P., Sawdey, M., Gimbrone Jr., M.A. and Loskutoff, D.J.: Cytokine activation of vascular endothelium. *J. Biol. Chem.* 263 (1988) 5797–5803.
- Sprengers, E.D. and Kluft, C.: Plasminogen activator inhibitors. *Blood* 69 (1987) 381–387.
- Strickland, S., Reich, E. and Sherman, M.I.: Plasminogen activator in early embryogenesis: Enzyme production by trophoblast and parietal endoderm. *Cell* 9 (1976) 231–240.
- Strohman, R.C., Moss, P.S., Micou-Eastwood, J., Spector, D., Przybyla, A. and Paterson, B.: Messenger RNA for myosin polypeptides: isolation from single myogenic cell cultures. *Cell* 10 (1977) 265–273.
- Uhler, M., Herbert, E., D'Eustachio, P. and Ruddle, F.D.: The mouse genome contains two nonallelic pro-opiomelanocortin genes. *J. Biol. Chem.* 258 (1983) 9444–9453.
- Vassalli, J.-D., Hamilton, J. and Reich, E.: Macrophage plasminogen activator: modulation of enzyme production by anti-inflammatory steroids, mitotic inhibitors, and cyclic nucleotides. *Cell* 8 (1976) 271–281.
- Verstraete, M. and Collen, D.: Thrombolytic therapy in the eighties. *Blood* 67 (1986) 1529–1541.
- Webb, A.C., Collins, K.L., Snyder, S.E., Alexander, S.J., Rosenwasser, L.J., Eddy, R.L., Shows, T.B. and Auron, P.E.: Human monocyte Arg-Serpin cDNA. Sequence, chromosomal assignment, and homology to plasminogen activator-inhibitor. *J. Exp. Med.* 166 (1987) 77–94.
- Wels, J.A., Word, C.J., Rimm, D., Der-Balan, P., Martinez, H.M., Tucker, P.W. and Blattner, R.: Structural analysis of the murine IgG3 constant region gene. *EMBO J.* 3 (1984) 2041–2046.
- Wun, T.-C. and Kretzmer, K.K.: cDNA cloning and expression in *E. coli* of a plasminogen activator inhibitor (PAI) related to a PAI produced by Hep G2 hepatoma cell. *FEBS Lett.* 210 (1987) 11–16.
- Wun, T.-C. and Reich, E.: An inhibitor of plasminogen activation from human placenta. *J. Biol. Chem.* 262 (1987) 3646–3653.
- Ye, R.D., Wun, T.-C. and Sadler, J.E.: cDNA cloning and expression in *Escherichia coli* of a plasminogen activator inhibitor from human placenta. *J. Biol. Chem.* 262 (1987) 2718–2725.
- Zeheb, R., Rafferty, U.M., Rodriguez, M.A., Andreasen, P. and Gelehrter, T.D.: Immunoaffinity purification of HTC rat hepatoma cell plasminogen activator-inhibitor-1. *Thromb. Haemostasis* 58 (1987) 1017–1023.

Communicated by J.L. Slightom.