Volume 12 Number 24 1984 Nucleic Acids Research

cDNA and gene nucleotide sequence of porcine plasminogen activator

Yoshikuni Nagamine*, David Pearson, Michael S.Altus and E.Reich

Friedrich Miescher-Institut, P.O. Box 2543, CH-4002 Basel, Switzerland

Received 18 September 1984; Revised and Accepted 5 November 1984

ABSTRACT

We have isolated cDNA and genomic clones coding for porcine plasminogen activator (urokinase, uPA). The cDNA is 2375 nucleotides long: it consists of a 5'-non-coding region (104 nucleotides), an open reading frame of 1329 nucleotides, and 3'-non-coding region of 942 nucleotides apart from the poly A tail. The genomic segment corresponding to the transcribed sequence is 5.85 kb long; it is composed of 11 exons and 10 introns. The 5'-flanking genomic region contains a number of sequences of potential regulatory significance, including possible hormone receptor binding sites and a sequence which we tentatively propose may be involved in activation of transcription by cAMP.

The full sequence of both cDNA and genomic clones, the latter including 1.3 kb of flanking region, is presented and discussed, and the deduced amino acid sequence compared with that of human uPA.

INTRODUCTION

The plasminogen activator (PA) genes are interesting subjects for study owing both to the biological contexts in which they are expressed and to the complex patterns of regulatory phenomena that govern their expression. PA secretion appears to be a general mechanism used by cells for inducing localized extracellular proteolysis; and enzyme production has been conspicuously associated with numerous examples of tissue remodeling and cell migration, processes in which extracellular proteolysis is likely to play an important role. Examples include ovulation (1), trophoblast implantation (2), mammary involution (3), inflammation (4), spermatogenesis (5), and stem cell migration (6). High levels of PA production are also generally associated with primary malignant tumors in man (7) and experimental animals (8), and a requirement for PA in metastasis has been demonstrated in an experimental system based on a human tumor (9).

In a minority of cases, PA production appears to be governed by temporal programming, but PA gene expression is generally regulated by cell and/or tissue-specific patterns of response to hormones. Thus, depending on cell type, PA synthesis can be modulated by one or more polypeptide hormones (8,10,11), sex or adrenocortical steroids (12,13), hormone-like agents such as retinoids (14), tumor promoters (15,16), neural transmitters (17) and glucose (17), or some combination of these. The questions posed by this variety of responses invite analysis at the molecular level.

Two genetically and immunologically distinct forms of PA are known to be produced by mammals - urokinase (uPA) and tissue plasminogen activator (tPA). We have previously described the isolation of an incomplete cDNA clone for porcine uPA and its use in defining some aspects of the hormonal modulation of PA synthesis (18). As a next step in the study of PA regulation we report here the isolation of full length cDNA and genomic clones for porcine uPA, and their nucleotide sequences.

MATERIALS AND METHODS

Materials

Restriction endonucleases were purchased from Bethesda Research Laboratories, New England Biolab and Biotec; bacteriophage λ packaging kits and RNase inhibitor from Biotec; reverse transcriptase from Stehelin (Basel); nuclease S1 from Boehringer Mannheim; Escherichia coli RNase H, DNA polymerase I large fragment, polynucleotide kinase, and terminal deoxynucleotide transferase from Behtesda Research Laboratories; α-^{32}P-ATP from New England Nuclear, and γ-^{32}P-ATP and a-$_{32}$P-ddATP from Amersham. All chemicals used in this work were either of enzyme grade or of the highest grade commercially available.

The computer program used for sequence analysis was purchased from SciSoft. The porcine kidney cell line, LLC-PK$_1$, was that described previously (11).

Methods

Isolation of full length cDNA clones. The preparation of poly(A)$^+$ RNA from calcitonin-treated cultures of LLC-PK$_1$ cells, its size fractionation by sucrose density gradient centrifugation, and

synthesis of the first strand of cDNA were performed as reported elsewhere (18). Secondary strand synthesis using the large fragment of DNA polymerase I, RNase H and DNA ligase was by the method of Okayama and Berg (19) as modified by Gubler and Hoffman (20). Double stranded DNA was tailed with poly dC according to Deng and Wu (21), annealed to pBR322 previously tailed with poly dG at the Pst I site (22), and transfected into E. coli HB101. Ten thousand recombinant clones were examined by means of high density screening procedures (1000 clones per filter)(23) using pPK59 (18) as a probe.

Isolation of genomic clones. DNA from LLC-PK_1 cells was prepared by the method of Blin and Stafford (24), and used to construct a genomic library according to Karn et al. (25). The isolated DNA was digested partially with the restriction enzyme Sau3A and size-fractionated by sucrose density gradient centrifugation. DNA fragments in the size range 14-20 kb were collected, and 1.0 µg was ligated to 1.5 µg of BamHI-digested EMBL4 λ phage DNA (26) and packaged in vitro (27). The primary library of $6x10^5$ clones was screened by plaque hybridization (28) using pPK79 (see Results) as a probe.

S1 nuclease mapping. Two µg of size fractionated poly(A)$^+$RNA prepared from calcitonin-treated LLC-PK_1 cells, $7x10^4$ cpm of end labeled DNA fragment and 10 µg of yeast tRNA were lyophilized and dissolved in 30 µl of 80% formamide, 40 mM Pipes, pH 6.4, 1 mM EDTA and 0.4 M NaCl. The solution was first treated at 90°C for 1 min and then incubated at 51°C for 3 hours. After this the solution was diluted with 900 µl of S1 buffer containing 50 mM sodium acetate, pH 6.4, 0.28 M NaCl, 4.5 mM $ZnCl_2$, 0.1% SDS and 2000 units/ml nuclease S1 and incubated at either 37°C or 45°C for 60 min. The protected radioactive fragment was precipitated with ethanol after extracting the solution with phenol-chloroform, and analyzed by polyacrylamide gel electrophoresis in parallel with other samples containing reaction mixtures used for determining the nucleotide sequence of the radioactive fragment that was hybridized.

DNA sequence determination and sequence analysis. DNA sequencing was performed mainly by the chemical digestion method of Maxam and Gilbert (29), but the chain terminating method based on di-

deoxynucleotides (30) was used for a few fragments. DNA fragments were labeled either at the 5' end with polynucleotide kinase and γ-^{32}P-ATP, or at the 3' end with terminal transferase and α-^{32}P-ddATP. Sequence analysis was performed according to Queen and Korn (31).

RESULTS

Isolation of full length cDNA clones

We have previously described a plasmid, pPK59 (18), which contained a 650 bp insert corresponding to a portion of porcine uPA-cDNA. With the aim of obtaining full-sized cDNA clones two additional libraries were constructed from size-fractionated poly(A)$^+$RNA isolated from calcitonin-treated cultures. One library was prepared according to a previously described method (18), slightly modified by the inclusion of RNase inhibitor during the synthesis of the first cDNA strand. Ten thousand recombinant clones were screened by colony hybridization (24) using the insert from pPK59 as a probe, and sixty positive colonies were taken for plasmid isolation. The plasmid (pPK79) containing the largest insert was selected for study. From the size of its insert - 2 kb - it was not expected to be full length; nucleotide sequence analysis and comparison with the known amino acid sequence of human uPA showed that a small part of the 5'-protein coding and all of the 5'-untranslated sequences were missing. A second cDNA library was then prepared as outlined in Methods and twelve thousand clones screened by colony hybridization using the insert from pPK79 as a probe. Ninety-two positive clones were obtained, and one plasmid containing the longest insert - pYN15- was studied further.

pYN15 was completely, and pPK79 was partially sequenced (Fig. 1). Except for potentially ambiguous segments, most of the cDNA sequences were read from only one strand because they could be verified by comparison with genomic sequences that were being determined concurrently. From the sequence data pYN15 contained 3' and 5' non-coding domains, and also included the entire coding region. The nucleotide sequence of pYN15 cDNA is given in Fig. 2, together with the deduced amino acid sequence which is compared with that of human uPA (32, 35) The two proteins show extensive homology, including the leader sequence: of 442 amino acids in

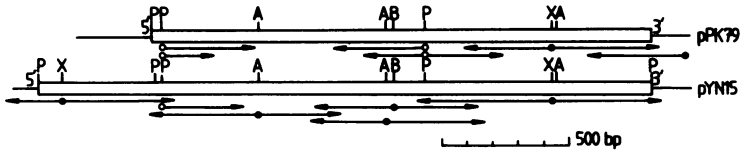


FIG. 1. Restriction maps of cDNA clones and sequencing strategy. Acc I (A), Pst I (P), BamHl (B), and Xba I (X) sites were determined. The straight lines and open boxes are pBR322 regions and cDNA inserts, respectively; open and filled circles designate the 3' and 5' ends, respectively, and the arrows indicate the direction and extent of sequence reading.

porcine uPA, 350 (79%) are identical with the human enzyme. The degree of homology is rather even throughout the length of the coding sequence, and all of the cysteine residues in the mature proteins are in comparable positions, indicating that the three-dimensional structures are probably very similar. Significant differences in amino acid sequence are found at only two sites. These are represented by the insertion of two and nine amino acid residues, respectively, at positions 9 and 135 in the porcine sequence. As expected from the similarity in amino acid sequence, dot matrix analysis (36) showed strong homology in nucleotide sequence between the two enzymes. This homology extended into both 3' and 5' non-coding regions, although at a slightly lower level than in the coding segment. Porcine uPA closely resembles human uPA also in a number of specific polypeptide domains including the leader sequence (residues -20 - -1), region of EGF homology (residues 1-46) (32), kringle structure (residues 52-133) (37), and the active site residues common to serine proteases, his-215,asp-266, and ser-367 (33).

Two poly(A) addition signals, AATAAA, separated by a sequence of 7 nucleotides, are located at the 3' end of porcine uPA cDNA.

Isolation and sequencing of genomic clones

A genomic DNA library was prepared (Methods) and $6x10^5$ recombinant phage plaques were screened with the pPK79 insert as probe. This procedure yielded 7 clones which together accounted for approximately 30 kb of continuous sequence. Since the size of porcine uPA hnRNA on electrophoresis was approximately 31S (18)

FIG. 2. Nucleotide and deduced amino acid sequences of pYN15; comparison with human uPA amino acid sequence. The upper register is the nucleotide sequence of the porcine cDNA insert in pYN15, each hundredth nucleotide being numbered. The middle register presents the deduced amino acid sequence of single chain porcine pro-uPA, 40 amino acids to the line. The lower register shows the human pro-uPA sequence: only non-identical residues are indicated and the two non-homologous deleted regions are indicated by (xxx). (▼) Indicates sites at which the sequence is interrupted by introns. (▽) Indicates the cleavage site for proenzyme activation. (✖) Amino acids at the active site of the enzyme. (✱) Glycosylation sites. (O) Amino acid identity with murine EGF.

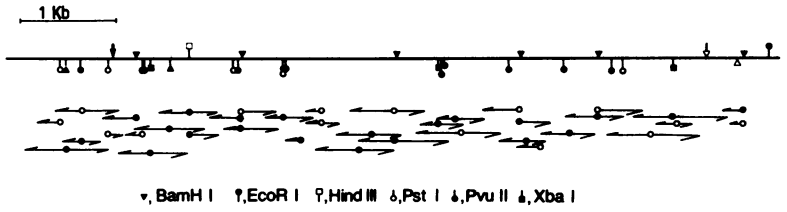


FIG. 3. Restriction map and sequencing strategy for uPA gene. Only a portion of the insert of λYN4 is shown. DNA was mapped with the indicated restriction enzymes by the method of Smith and Birnstiel (49). Open and filled circles represent 3' and 5' ends, respectively, and the arrows indicate the direction and extent of reading. (↓) The putative transcription initiation site, and (⇩) poly(A) addition site.

the corresponding genomic sequence was expected to account for some 7 kb of sequence. With this in mind a particular clone,λYN4, was selected for analysis because the pPK79 probe hybridized to its central region and it therefore seemed likely that it contained the entire uPA sequence.

The restriction map of part of λYN4 and the sequencing strategy are given in Fig. 3 , and the result of sequencing in Fig.4 . The sequence obtained accounts for the entire cDNA present in pYN15 and includes in addition 1 kb of 5'- and 300 bp of 3'-flanking regions. Comparison with the cDNA indicates that the genomic sequence is composed of 11 exons separated by 10 introns; the 5' non-coding region contains only one intron, whereas the 3' non-coding segment consists of a single, uninterrupted stretch of 942 bp.

S1 nuclease mapping

S1 nuclease protection analysis was undertaken to locate the initiation site for transcription. For this purpose a Taq I-Ava II fragment (220 bp) was prepared from the genomic clone and end-labeled at the Taq I site as shown in Fig. 5a; the fragment was then hybridized to size-fractionated poly(A)$^+$RNA, incubated with S1 nuclease as described in Methods, and the products analyzed by electrophoresis. The procedure yielded a population of bands with graded intensity corresponding to a region around 18-25 nucleotides downstream of the TATA box (38) which is closest to the protein coding sequence (Fig. 5b). It is known that S1 nucle-

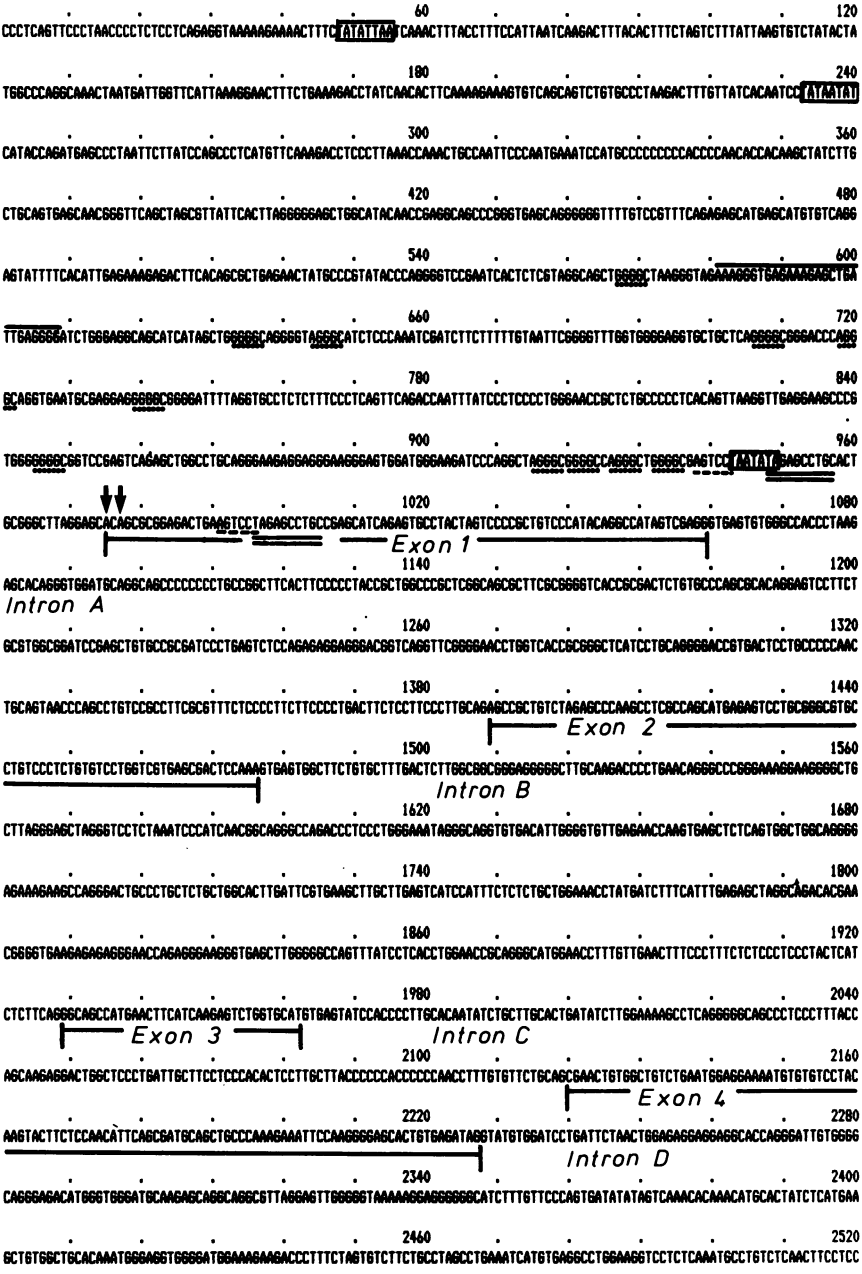
ase-generated fragments migrate differently than those produced by chemical digestion (39), and that the presence of cap structure at the 5' end of mRNA interferes with S1 action (40). Both of these factors may cause the mobility of oligonucleotides to deviate from that expected on the basis of their assumed length. These considerations, and the fact that eukaryotic mRNA are entirely initiated at an adenine residue that is preceded by a cytosine residue (41), point to adenine 975 and/or 977 as the putative initiation site(s).

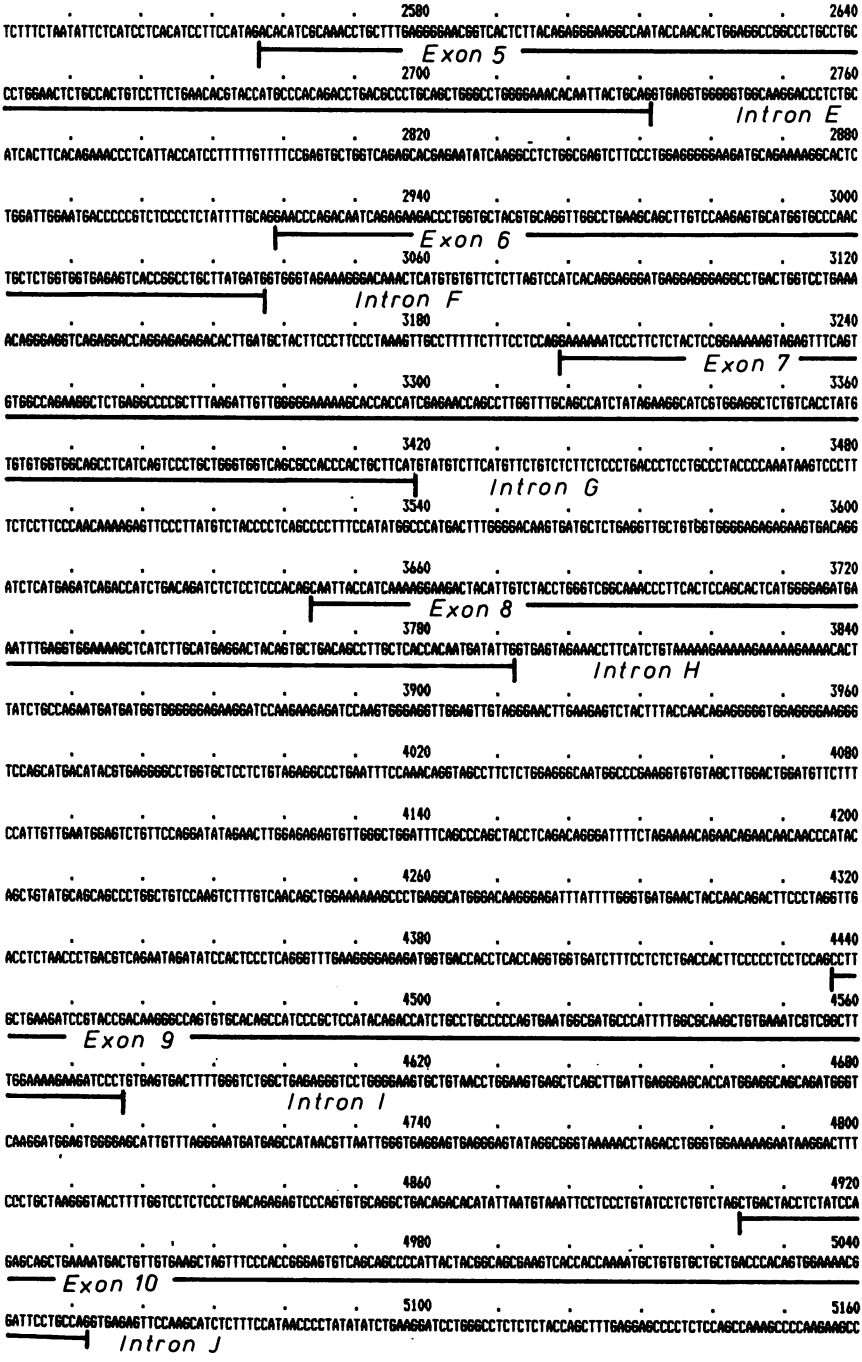
Features of the genomic nucleotide sequence

i) Transcribed region. This portion of the gene comprises 6 kb of nucleotide sequence, of which 2.4 kb are represented in mature mRNA. This fraction - 40% - is rather large, especially in comparison with other kringle-containing serine proteases such as prothrombin (2 kb mRNA/>20 kb transcribed) (Degen, S., personal communication), plasminogen (3 kb/> 21 kb) (42), or tissue plasminogen activator (2.6 kb/ 30 kb) (Degen, S. and Rajput, B., personal communication). When compared with the cDNA sequence the gene is seen to consist of 11 exons separated by 10 introns, but this segmentation pattern does not seem to be correlated with any obvious structural or functional domains in the protein (Fig. 6): for example, the region of EGF homology comprises two exons, the kringle structure is split in mid-sequence, and there is no genetic discontinuity at or near the site where limited proteolysis converts the one-chain proenzyme to the active, two-chain enzyme. There are no apparent homologies between the sequences of different introns, and comparison of these intron sequences both with known repetitive sequences and with the data base of class II genes failed to produce any evidence of homology. An intriguing repeating sequence, $(AAAAG)_3AAAA$, is situated at positions 3747-3768 in intron H, but there are as yet no clues to its function or significance.

ii) Flanking regions. We determined approximately 1 kb of 5' flanking sequence and identified several features of potential interest, especially in the vicinity of the first TATA box. This structure, TAATATA, is 25 nucleotides upstream of the presumptive transcription initiation site and the following elements are located nearby: a) the pentanucleotide sequences GGGGC or

pig-uk gene




2580
2640
Exon 5
2700
2760
Intron E
2820
2880
2940
3000
Exon 6
3060
3120
Intron F
3180
3240
Exon 7
3300
3360
3420
3480
Intron G
3540
3600
3660
3720
Exon 8
3780
3840
Intron H
3900
3960
4020
4080
4140
4200
4260
4320
4380
4440
4500
4560
Exon 9
4620
4680
Intron I
4740
4800
4860
4920
4980
5040
Exon 10
5100
5160
Intron J

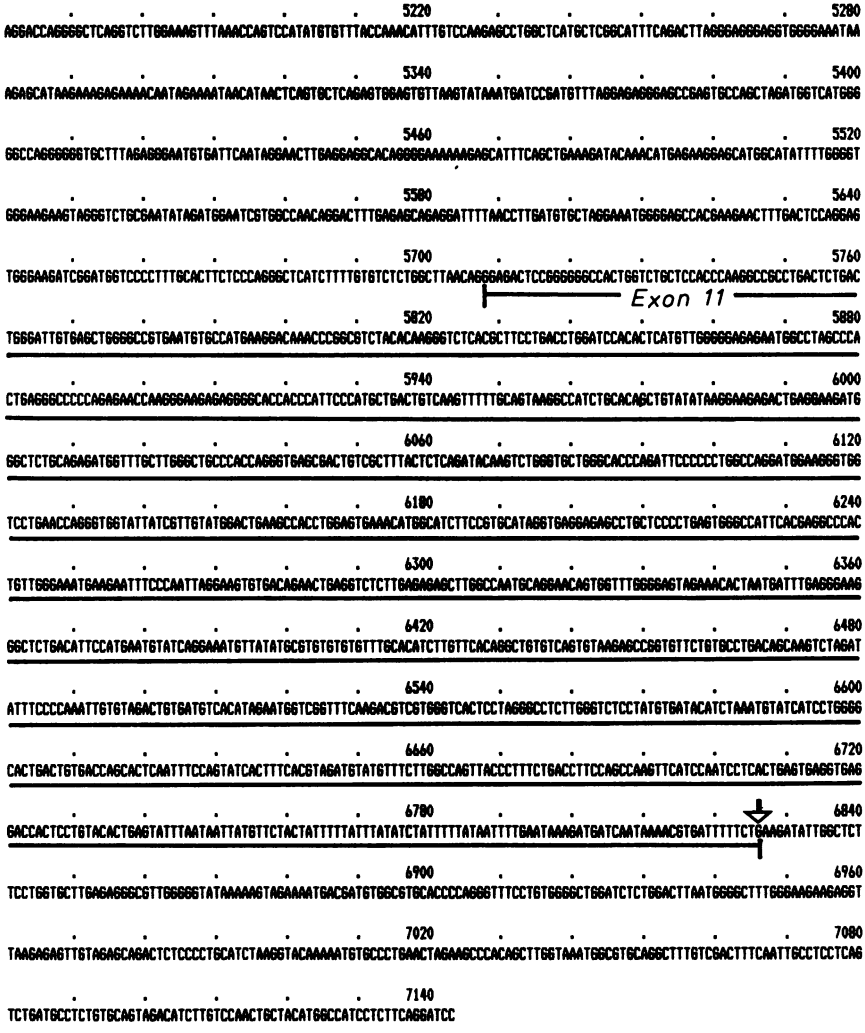
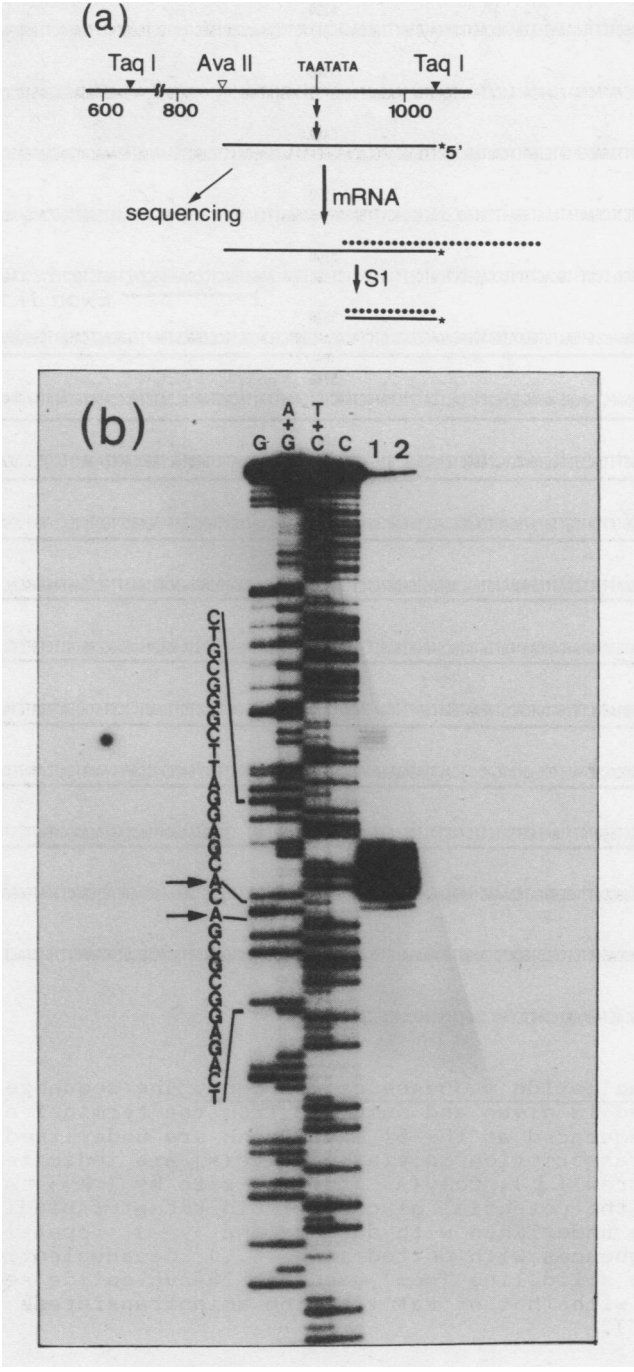


FIG. 4. Nucleotide sequence of uPA gene. The sequence of the sense strand is given and numbered from the terminal nucleotide that was sequenced at the 5' end. Exons are underlined. The putative transcription initiation site(s) are indicated by the vertical arrow (↓): poly(A) addition site by (⇩); TATA boxes are boxed; the potential glucocorticoid receptor binding sequences are underlined with dashed line (---), repeating hexapeptide sequences with dotted line (...), decanucleotide repeats with double solid line (═══), and the 28-nucleotide sequence homologous with that of rat tyrosine aminotransferase is overlined (‾‾‾‾).


(a)
Taq I
Ava II
TAATATA
Taq I
600
800
1000
*5'
sequencing
mRNA
S1
(b)
A + G
T + C
G
C
1
2
CTGCGGGCTTAGGAGCACACAGCGCGGAGACT

AGGGC are repeated 11 times in the interval 7-376 bp upstream of the TATA box; b) a decanucleotide sequence - TAGAGCCTGC - overlaps the 3' end of the TATA box and is repeated once in exon 1, 48 nucleotides downstream; c) the hexanucleotide sequence - AGTCCT - which is reported to act as a core sequence for binding the glucocorticoid receptor (43), overlaps the 5' end of the TATA box and is repeated once in exon 1, 53 nucleotides downstream; d) a 28 nucleotide sequence - AAAGGGTGAGAAAGA-GCTGATTGAGGGG - which begins 335 nucleotides upstream of the TATA box and shows considerable homology (~ 75% identity) with a comparable sequence in the 5' flanking region of the rat tyrosine aminotransferase gene (44); e) two additional TATA-like structures beginning 703 (TATAATAT) and 888 nucleotides (TATA-TTAA), respectively, upstream of the TATA box.

The purine-rich domain which begins 37 nucleotides upstream of the TATA box is only one of many that are scattered throughout the untranslated regions of the gene.

In the 3' flanking region we surveyed 300 bp of sequence downstream of the established poly(A) attachment site but found no indication of any further poly(A)-attachment signal.

DISCUSSION

In this paper we have presented the complete nucleotide sequence of the cDNA and gene for porcine uPA. cDNA sequences have previously been obtained for human (34,35) and mouse uPA (45) but, as far as we are aware, ours is the first report of the uPA genomic structure and sequence. The nucleotide sequence data, and the deduced amino acid sequence, reveal extensive homology as well as a few interesting differences between the porcine and human enzymes, as already indicated in Results. The major differences are accounted for by two insertions consisting of 2 and

FIG. 5. Nuclease S1 digestion analysis of 5' end of uPA mRNA. A Taq I-Ava II fragment (220 bp) which covers the TATA box and 5' end of cDNA was end labeled at the Taq I site, hybridized to uPA mRNA, and sequenced. (a) Experimental design; (b) sequence of the Taq I-Ava II fragment, and S1 digestion of the heteroduplex of fragment and uPA-mRNA. The duplex was digested at 37°C (lane 1) or 45°C (lane 2). Complementary nucleotides are shown beside the sequencing gel, and the arrows indicate the putative initiation site(s).

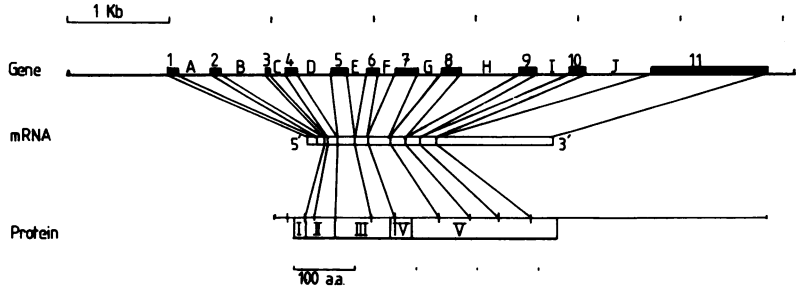


FIG. 6. Relationship between gene, mRNA and protein. The gene and mRNA are drawn to the same scale. The heavy lines of the gene represent exons. The numbering of exons and the alphabetic designation of introns correspond to those in Fig. 4. The protein domains are: (I) leader sequence; (II) "growth factor" domain; (III) kringle structure; (IV) connecting peptide; (V) catalytic domain.

9 amino acids, respectively, that are found in porcine cDNA; and, because we used a slightly unconventional procedure for second strand synthesis during cDNA preparation, it can be asked whether these sequences might somehow be due to a procedural or cloning artifact. Several findings argue against such an interpretation: firstly, the same sequences were found both in genomic and in cDNA clones; and, secondly, the longer of the two, corresponding to the nona-peptide insert, was present also in pPK79 which was prepared by conventional methods and independently isolated in a separate experiment. It is unlikely that the insertions are the result of some rare processing intermediate because their boundaries do not obey the GT.....AG rule (46) and, in addition, they are too short to be removed by splicing (47). Another interesting point of difference concerns the glycosylation sites; their separation by nearly 200 residues in the two enzymes suggests that, whatever its function, glycosylation is probably not important for catalytic activity.

The complex tissue-specific patterns of hormonal responses that affect uPA production might be achieved by modulating the expression either of a single pair of alleles or of a family of closely related genes. We do not yet have definitive evidence on this question, but Southern blot analysis of cellular DNA and the restriction maps of 7 independent genomic clones all are consist-

ent with the assumption that the porcine genome contains only one pair of uPA alleles.

As already noted (v. supra, Results), the 5' flanking region of the uPA gene is rich in sequence elements that could potentially mediate a variety of regulatory mechanisms. For example, the proximity to the TATA box of two possible glucocorticoid receptor binding sequences could account for the repression of uPA production by dexamathasone (11, and Nagamine, Y., unpublished observations); and the presence of additional TATA boxes and short repeating sequences may well be significant for regulatory phenomena. Perhaps of greatest potential interest is the 28-nucleotide upstream sequence previously referred to (Results): this sequence has intriguing homologies both with a comparable stretch in the 5' flanking domain of rat tyrosine amino transferase (44) another enzyme whose expression is cAMP regulated) and with some regulatory sequences in bacteria; these homologies and their implications are the subject of a separate report (48). Whatever the relevance of these and other particular sequences may be for the control of uPA expression, the relatively small size of the gene is convenient for analysing their role, and experiments designed to clarify some of the evident questions are in progress.

ACKNOWLEDGEMENT

We thank Drs. R.D. Estensen, B. Hemmings, S. Shibahara and C. Moroni for critical reading of the manuscript and Dr. Y. Nabeshima for discussion on sequencing method.

*To whom reprint requests should be sent

REFERENCES

1. Beers, W.H., Strickland, S., and Reich, E. (1975) Cell 6, 389-394.
2. Strickland, S., Reich, E., and Sherman, M.I. (1976) Cell 9, 231-240.
3. Ossowski, L., Biegel, D., and Reich, E. (1979) Cell 16, 929-940.
4. Granelli-Piperno, A., Vassalli, J.-D., and Reich, E. (1977) J. Exp. Med. 146, 1693-1706.
5. Lacroix, M., Smith, F.E., and Fritz, I.B. (1977) Mol. Cell. Endocrinol. 7, 227-236.
6. Valinsky, J.E., Reich, E., and Le Douarin, N.M. (1981) Cell 25, 471-476.
7. Markus, G., Camiolo, S.M., Kohga, S., Madeja, J.M., and Mittelman, A. (1983) Cancer Res. 43, 5517-5525.

8. Mira y Lopez, R., Reich, E., and Ossowski, L. (1983) Cancer Res. 43, 5467-5477.
9. Ossowski, L., and Reich, E. (1983) Cell 35, 611-619.
10. Strickland, S., and Beers, W.H. (1976) J. Biol. Chem. 251, 5694-5702.
11. Dayer, J.-M., Vassalli, J.-D., Babbitt, J.L., Hull, R.N., Reich, E., and Krane, S.H. (1981) J. Cell Biol. 91, 195-200.
12. Mak, T.W., Rutledge, G., and Sutherland, J.A. (1976) Cell 7, 223-226.
13. Vassalli, J.-D., Hamilton, J., and Reich, E. (1976) Cell 8, 271-281.
14. Wilson, E.L., and Reich, E. (1978) Cell 15, 385-392.
15. Wigler, M., and Weinstein, I.B. (1976) Nature 259, 232-233.
16. Vassalli, J.-D., Hamilton, J., and Reich, E. (1977) Cell 11, 695-705.
17. Virji, M.A.G., Vassalli, J.-D., Estensen, R.D., and Reich,E. (1980) Proc. Natl. Acad. Sci. USA 77, 875-879.
18. Nagamine, Y., Sudol, M., and Reich, E. (1983) Cell 32, 1181-1190.
19. Okayama, H., and Berg, P. (1982) Mol. Cell. Biol. 2, 161-170.
20. Gubler, U., and Hoffman, B.J. (1983) Gene 25, 263-269.
21. Deng, G., and Wu, R. (1981) Nucleic Acids Res. 9, 4173-4188.
22. Villa-Komanoff, L., Efstratiadis, A., Broome, S., Lomedics, P., Tizard, R., Naker, S.P., Chick, W.L., and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA 75, 3727-3731.
23. Hanahan, D., and Meselson, M. (1980) Gene 10, 63-67.
24. Blin, N., and Stafford, D.W. (1976) Nucleic Acids Res.3, 2303-2308.
25. Karn, J., Brenner, S., Barnett, L., and Cesareni, G. (1980) Proc. Natl. Acad. Sci. USA 77, 5172-5176.
26. Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. (1983) J. Mol. Biol. 170, 827-842.
27. Hohn, B. (1979) Methods Enzymol. 68, 299-309.
28. Benton, W.D., and Davis, R.W. (1977) Science 196, 180-182.
29. Maxam, A., and Gilbert, W. (1980) Methods Enzymol. 65, 499-580.
30. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
31. Queen, C., and Korn, L.J. (1984)Nucleic Acids Res. 12, 581-599.
32. Steffens, G.J., Guenzler, W.A., Oetting, F., Frankus, E., and Flohé, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058.
33. Guenzler, W.A., Steffens, G.J., Oetting, R., Kim, S.-M.A., Frankus, E., and Flohé, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165.
34. Heyneker, H.L., Holmes, W.E., and Vehar, G.A. (1983) European Patent Applic Publ. No. 0092182.
35. Verde, P., Stoppelli, M.P., Galleffi, M., Di Nocera, P., and Blasi, F. (1984) Proc. Natl. Acad. Sci. USA 81, 4727-4731.
36. Zweig, S.E. (1984) Nucleic Acids Res. 12, 767-776.
37. Claeys, H., Sottrup-Jensen, L., Zajdel, M., Petersen, T.E., and Magnusson, S. (1976) FEBS Letters 61, 20-24.
38. Goldberg, M.L. (1979) Ph.D. Thesis Stanford Univ.
39. Tomizawa, J., Ohmori, H., and Bird, R. (1977) Proc. Natl. Acad. Sci. USA 74, 1865-1869.

40. Weaver, R.F., and Weissmann, C. (1979) Nucleic Acids Res. 7, 1175-1193.
41. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349-383.
42. Malinowski, D.P., Sadler, J.E., and Davie, E.W. (1984) Biochemistry 23, 4243-4250.
43. Cato, A.C.B., Geisse, S., Wenz, M., Westphal, H.M., and Beato, M. (1984) EMBO J., in press.
44. Shinomiya, T., Sherer, G., Schmid, W., Zentgraf, H., and Schuetz, G. (1984) Proc. Natl. Acad. Sci. USA 81, 1346-1350.
45. Belin, D., Vassalli, J.-D., Combepine, C., Godeau, F., Nagamine, Y., Reich, E., Kocher, H.P., and Duvoisin, R.(1984) Submitted for publication.
46. Sharp, P. (1981) Cell 23, 643-646.
47. Wieringa, B., Hofer, E., and Weissmann, C. (1984) Cell 37, 915-925.
48. Nagamine, Y., and Reich, E., manuscript in preparation.
49. Smith, H.D., and Birnstiel, M.L. (1976) Nucleic Acids Res. 3, 2387-2398.