Characterization of a Complementary Deoxyribonucleic Acid Coding for Human and Bovine Plasminogen†

Douglas P. Malinowski, J. Evan Sadler, and Earl W. Davie*

ABSTRACT: A cDNA library was constructed in pBR322 from bovine liver mRNA that was enriched for plasminogen mRNA by polysome immunoprecipitation. A ^{32}P-labeled single-stranded cDNA was then prepared from the enriched bovine mRNA and employed as a probe to screen the cDNA library. The screening was carried out by testing for clones that protect the hybridized ^{32}P-labeled cDNA from S1 nuclease digestion. The longest clone that was found was 581 base pairs in length and coded for the C-terminal 107 amino acids of bovine plasminogen, a 3′ noncoding region of 246 nucleotides and a poly(A) tail. The bovine cDNA clone was then used as a probe to screen a human liver cDNA library of 18 000 recombinants. Six isolates were found to contain human plasminogen sequences. The longest clone consisted of 1851 base pairs corresponding to amino acid residues 272–790, followed by a 3′ noncoding region of 227 base pairs and a poly(A) tail. Restriction fragments of the human cDNA were then used as probes to screen a human genomic DNA library present in a Charon 4A λ phage library. Approximately 50 isolates from 10^6 recombinants were identified that hybridized to varying degrees with the cDNA probe. Among these, 10 corresponding to the gene for human plasminogen have been analyzed, and 3 that overlap have been shown to extend from kringle 3 through the 3′ noncoding region of the gene. A 160 base pair exon with flanking splice junctions was then characterized and shown to encode for the first half of plasminogen kringle 4, including amino acid residues 346–399.

Plasminogen is a glycoprotein (M_r ~92 000) that participates in the final stages of fibrinolysis (Robbins, 1982). During these reactions, plasminogen is converted to plasmin, a serine protease that hydrolyzes fibrin clots. Plasminogen is converted to plasmin by a number of activators, including tissue plasminogen activator, urokinase, or a complex of plasminogen and streptokinase. During the conversion of plasminogen to plasmin, a specific internal peptide bond between Arg-560 and Val-561 is cleaved in the single-chain precursor molecule. Plasmin is composed of two polypeptide chains held together by disulfide bonds. The serine protease portion of the enzyme is located in the carboxyl-terminal region of the protein (Groskopf et al., 1969; Robbins et al., 1973). This region shows considerable amino acid sequence homology with the serine proteases involved in blood coagulation, as well as those involved in other physiological processes such as the digestive enzymes of the pancreas (Davie et al., 1979).

Two major forms of plasminogen have been separated on lysine–agarose (Brockway & Castellino, 1972). Form 1 contains two carbohydrate chains linked to Asn-288 and Thr-345, while form 2 contains one carbohydrate chain linked to Thr-345 (Hayes & Castellino, 1979a,b). Other forms of human plasminogen have also been isolated from plasma, including a Glu–plasminogen and a Lys–plasminogen. The latter plasminogen is a partially degraded protein resulting from the cleavage of a peptide of 76 residues from the N-terminal portion of the molecule (Wallén & Wiman, 1972; Sottrup-Jensen et al., 1978a).

The complete amino acid sequence of human plasminogen has been determined by Wiman & Wallén (1975a,b), Wiman (1977), and Sottrup-Jensen et al. (1978a,b). An interesting feature of the molecule is the presence of five tandem repeats, called kringles. These structures, containing about 80 amino acids, are present in the amino-terminal region of the protein (Sottrup-Jensen et al., 1978a). They are also homologous to two kringles present in the amino-terminal region of tissue plasminogen activator (Pennica et al., 1983), two present in the amino-terminal region of prothrombin (Magnusson et al., 1975), and a single kringle present in the amino-terminal region of urokinase (Günzler et al., 1982).

Plasminogen is synthesized in the liver (Raum et al., 1980; Bohmfalk & Fuller, 1980; Saito et al., 1980). Recently, two forms of plasminogen have been shown to be synthesized with different mRNA populations from monkey liver (23S and 18S) by employing an mRNA-dependent reticulocyte lysate (Gonzalez-Gronow & Robbins, 1984). In this report, we describe the enrichment of plasminogen messenger RNA from bovine liver and the construction of a bovine cDNA clone for plasminogen. This clone was then used as a hybridization probe for the isolation of a cDNA clone coding for human plasminogen from a human liver cDNA library. Preliminary reports of this work have been presented (Malinowski & Davie, 1983; Malinowski et al., 1983).

Experimental Procedures

Materials. Sepharose 4B was purchased from Pharmacia Fine Chemicals, and CNBr was purchased from Pierce Chemical Co. Freund's complete and incomplete adjuvants were purchased from Difco Laboratories. Phenol was purchased from J. T. Baker Chemical Co., and prior to use, it was redistilled, saturated with an equal volume of 0.5 M Tris-HCl, pH 7.5, and stored at 4 °C in the dark.

All buffers, glassware, and plasticware were sterilized by autoclaving prior to use. Oligo(dT)–cellulose and deoxynucleotide triphosphates were purchased from P-L Biochemicals. Terminal transferase was kindly provided by Dr. R. L. Ratliff (Los Alamos Scientific Laboratory, Los Alamos, NM); reverse transcriptase was provided by Dr. Joseph Beard (Life Sciences, St. Petersburg, FL); S1 nuclease was provided by Dr. Richard Palmiter of this department. DNA polymerase (*Escherichia coli*) was purchased from Boehringer Mannheim, and T4 DNA kinase, Klenow fragment, and bacterial alkaline

† From the Department of Biochemistry, University of Washington, Seattle, Washington 98195. *Received February 14, 1984.* This work was supported in part by a research grant (HL 16919 to E.W.D.) and a postdoctoral fellowship (HL 05968 to D.P.M.) from the National Institutes of Health. J.E.S. in an Associate of the Howard Hughes Medical Institute. A portion of this work was presented at the Sixth International Congress on Fibrinolysis in Lausanne, Switzerland (Malinowski & Davie, 1983).

0006-2960/84/0423-4243$01.50/0

phosphatase were purchased from Bethesda Research Laboratories. Restriction endonucleases were purchased from either Bethesda Research Laboratories or New England Biolabs. [^{35}S]Methionine, [^{32}P]ATP, and deoxynucleotide [α-^{32}P]triphosphates were purchased from New England Nuclear.

Isolation of Bovine Plasminogen and Preparation of Plasminogen Antibodies. Bovine plasminogen was isolated from freshly collected bovine blood by utilizing a modification (Malinowski & Davie, 1983) of the original procedure of Deutsch & Mertz (1970). The purified protein was homogeneous as evidenced by NaDodSO$_4$[1]–polyacrylamide gel electrophoresis (Weber & Osborn, 1969) and N-terminal sequence analysis (Malinowski & Davie, 1983). This protein preparation was used to immunize New Zealand white rabbits (Harboe & Ingild, 1973). Each rabbit was injected subcutaneously with 100 μg of plasminogen in Freund's complete adjuvant. At weekly intervals, the injection was repeated with incomplete adjuvant. After 6 weeks, blood was collected from the rabbits and allowed to clot at room temperature for 1 h. The clot was allowed to retract overnight at 4 °C and then separated from the serum by centrifugation for 10 min at 9000 rpm on a Sorvall SS34 rotor at 4 °C. The serum supernatant was removed, made 0.001 M in phenylmethanesulfonyl fluoride, and stored at −20 °C.

Plasminogen was coupled to CNBr-activated Sepharose 4B as previously described (Porath & Kristansen, 1975). Bovine serum albumin (BSA) was coupled to CNBr-activated Sepharose 4B and was kindly provided by Dr. Ross MacGillivray. Plasminogen antisera were first passed through a column of BSA–Sepharose that had been equilibrated in 0.01 M sodium phosphate–0.15 M NaCl, pH 7.4 (PBS). This step removed traces of anti-BSA antibodies that were found in the plasminogen antisera. Failure to perform this step resulted in an antisera preparation that specifically immunoprecipitated both plasminogen and BSA in a cell-free translation system when total bovine liver mRNA was employed (D. P. Malinowski, unpublished observations). Presumably, the isolated bovine plasminogen, although apparently homogeneous based upon physicochemical characterization, still contained traces of BSA, which elicited a strong immune reaction in the immunized rabbits. The post-BSA–Sepharose antisera were then passed through plasminogen–Sepharose 4B according to the general procedures of Schechter (1974) and Gough & Adams (1978). The plasminogen–Sepharose was packed into a sterile glass column (0.9 × 17 cm) and washed overnight with 1 L of sterile PBS. The antisera were applied to the column and then washed with sterile PBS until the A_{280} of the effluent was less than 0.05 OD unit. Generally, this required an overnight wash using 2 L of PBS at 23 °C. The bound plasminogen antibodies were eluted from the column by washing with sterile 3.5 M NaSCN–PBS, pH 7.4. The eluted antibodies were dialyzed against several changes of sterile PBS at 4 °C and stored at −20 °C. They were shown to be free of contaminating ribonuclease as previously described (Gough & Adams, 1978).

Isolation of Bovine Liver Polysomes and Poly(A+)-mRNA. Bovine calf liver was obtained from a local abattoir, rinsed in sterile PBS, sliced into small sections, and frozen in liquid nitrogen. The liver slices were homogenized to a powder in a Waring blender that had been prechilled to −20 °C, and the powder was then stored at −70 °C. Polysomes were isolated from the liver powder according to the magnesium precipitation method of Palmiter (1974). Analysis of the sedimentation profile in a 10–40% sucrose gradient revealed that the polysomes were intact (Palmiter, 1974).

Total mRNA was isolated from the polysomes by proteinase K digestion and LiCl precipitation as described (MacGillivray et al., 1978). Poly(A+)-mRNA was isolated by chromatography on oligo(dT)–cellulose (Aviv & Leder, 1972).

Cell-free translation assays were performed in a mRNA-dependent rabbit reticulocyte lysate as described by Pelham & Jackson (1976). Immunoprecipitation of the translation products was performed by using the affinity-purified plasminogen antibodies according to the procedure of Palmiter et al. (1977). The translation products were analyzed by NaDodSO$_4$–polyacrylamide gel electrophoresis (Laemmli, 1970) and fluorography (Laskey & Mills, 1975).

Polysomes were enriched for plasminogen mRNA by using the technique of immunoprecipitation (Gough & Adams, 1978) as described in detail elsewhere (Malinowski & Davie, 1983).

cDNA Synthesis and Recombinant Library Screening. Double-stranded complementary DNA (cDNA) was prepared from the enriched plasminogen mRNA and cloned into the *Pst*I site of pBR322 by the method of GC tailing (Stein et al., 1978). The chimeric plasmids were used to transform *E. coli* strain RR1 according to the NIH Guidelines for Recombinant DNA Research.

^{32}P-Labeled cDNA was prepared from the enriched mRNA and was used as a hybridization probe to screen the enriched cDNA library. Positive colonies were rescreened by a solution hybridization assay (MacGillivray et al., 1980) to identify bovine plasminogen cDNA clones.

A cDNA library prepared from total human liver mRNA was kindly provided by Dr. Savio L. C. Woo (Chandra et al., 1983). Human plasminogen cDNA clones were identified by the use of bovine plasminogen cDNA plasmid 5 (vide infra) as a hybridization probe. The *Pst*I insert of plasmid 5 was cleaved with *Hpa*II and *Sph*I to remove the GC tails of the insert, and the *Hpa*II–*Sph*I fragment was isolated by preparative gel electrophoresis and recovered by electroelution. This fragment was used as a hybridization probe under conditions of reduced stringency to allow for some mismatching between the bovine plasminogen cDNA and the human plasminogen cDNAs. The hybridization conditions employed were 0.05 M citrate–0.9 M NaCl, pH 7.0 (6 × SSC), 2 × Denhardt's solution (Southern, 1975), and 0.5% NaDodSO$_4$–1 mM EDTA at 58 °C. Nick translations were performed according to Maniatis et al. (1975) using α-^{32}P-labeled dATP, dCTP, and dTTP (each at 600 Ci/mmol).

Restriction Mapping and DNA Sequence Analysis. Restriction enzymes, purchased from Bethesda Research Laboratories and New England Biolabs, were used according to the manufacturer's recommendations. DNA fragments were analyzed by agarose (Stein et al., 1978) and polyacrylamide gel (Maniatis et al., 1975b) electrophoresis. DNA fragments were labeled at the 5′ end with [γ-^{32}P]ATP and T4 kinase (Maxam & Gilbert, 1980). Labeling at the 3′ end was accomplished by fill-in synthesis using the Klenow fragment of *E. coli* DNA polymerase I and [α-^{32}P]dNTPs (Smith et al., 1979) or with terminal transferase in the presence of cobalt chloride and ^{32}P-labeled cordycepin (New England Nuclear) (Tu & Cohen, 1980). The labeled DNA was separated from the unincorporated dNTPs by gel filtration chromatography

[1] Abbreviations: BSA, bovine serum albumin; PBS, phosphate-buffered saline (10 mM Na_3PO_4, 0.15 M NaCl, pH 7.4); NaDodSO$_4$, sodium dodecyl sulfate; EDTA, ethylenediaminetetraacetic acid; SSC, standard saline citrate (15 mM citrate, 0.15 M NaCl, pH 7.0); dNTP, deoxynucleotide triphosphate; NaSCN, sodium thiocyanate; cDNA, complementary deoxyribonucleic acid; Denhardt's solution, 0.02% (w/v) BSA, 0.02% (w/v) poly(vinylpyrrolidone), and 0.02% (w/v) Ficoll.

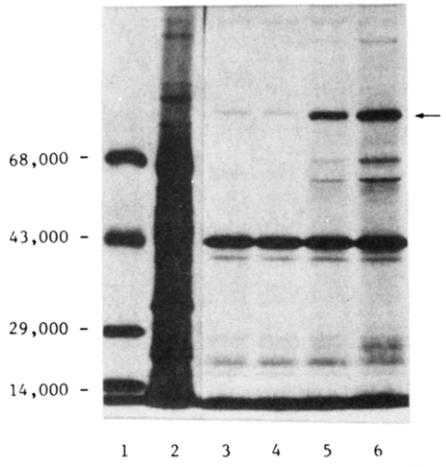

FIGURE 1: $NaDodSO_4$–polyacrylamide gel electrophoresis of the translation products from enriched bovine plasminogen mRNA. Lane 1, molecular weight standards: ^3H-labeled bovine serum albumin (68 000), ovalbumin (43 000), carbonic anhydrase (29 000), and lysozyme (14 000). Lane 2, translation products of total liver mRNA. Lane 3, translation products in the absence of exogenous mRNA. Lanes 4–6, total translation products synthesized from successive fractions of the enriched mRNA eluted from oligo(dT)–cellulose. Nearly all of the plasminogen mRNA was recovered in the fractions represented in lanes 5 and 6. The arrow indicates the position of native plasminogen on the gel.

on AcA54 (LKB) and concentrated by ethanol precipitation. The labeled ends were separated by cleavage with a secondary restriction enzyme, followed by preparative gel electrophoresis (Rixon et al., 1983; Maxam & Gilbert, 1980). Sequence analysis of the labeled DNA fragments was performed according to a modified protocol (Rixon et al., 1983) of the Maxam & Gilbert procedure (1980).

Results

Enrichment and Translation of Bovine Liver mRNA. Polysomes were isolated from calf liver by the magnesium precipitation method of Palmiter (1974) and enriched for plasminogen mRNA by immunoprecipitation (Gough & Adams, 1978; Malinowski & Davie, 1983). The total liver poly(A)-containing RNA and the enriched poly(A)-containing RNA were then translated in a cell-free protein-synthesizing system prepared from rabbit reticulocytes (Pelham & Jackson, 1976), and the translation products were analyzed by $NaDodSO_4$–polyacrylamide gel electrophoresis, followed by fluorography (Figure 1). As expected, a large number of translation products were formed in the presence of total bovine liver mRNA (lane 2). The product precipitated with antibody to plasminogen in these experiments was calculated at approximately 0.5% of the total product synthesized (data not shown). This was estimated by the amount of radiolabeled product precipitated by antibody to plasminogen vs. the amount of radioactivity in the total translation products formed. The products synthesized from enriched bovine plasminogen mRNA are shown in lanes 5 and 6. In the presence of the enriched plasminogen mRNA, a single major protein (M_r 80 000) and several minor proteins (M_r <70 000) were observed. The band with M_r 43 000 was probably met-tRNA, which was also present in the control (lane 3). From these experiments, it was estimated that the plasminogen mRNA was enriched 40–60-fold by the polysome immunoprecipitation method.

Construction and Screening of a Bovine Liver cDNA Library. Double-stranded cDNA was synthesized from enriched bovine mRNA and then made blunt ended by digestion with S1 nuclease. The product was then subjected to fill-in repair

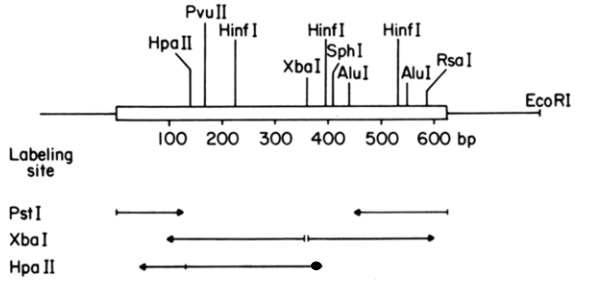

FIGURE 2: Restriction map of the bovine plasminogen cDNA insert in plasmid 5. The cDNAs identified by solution hybridization and S1 nuclease digestion were sequenced and found to encode the carboxyl-terminal region of bovine plasminogen. The longest cDNA identified was found in plasmid 5. The strategy used to determine the sequence of this cDNA is also shown.

```
684                  690                                  700
Trp Gly Glu Thr Gln Gly Thr Phe Gly Glu Gly Leu Leu Lys Glu Ala His Leu Pro Val
TGG GGA GAA ACC CAA GGC ACT TTT GGT GAG GGC CTC CTG AAG GAA GCG CAC CTG CCT GTG
          10                   20                   30                   40

                     710                                  720
Ile Glu Asn Lys Val Cys Asn Arg Asn Glu Tyr Leu Asp Gly Arg Val Lys Pro Arg Glu
ATC GAG AAC AAG GTG TGT AAT CGC AAC GAG TAT CTC GAC GGG AGG GTC AAG CCC AGG GAG
          70                   80                   90                  100

                     730                                  740
Leu Cys Ala Gly His Leu Ile Gly Gly Thr Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro
CTG TGT GCC GGC CAT CTA ATT GGA GGG ACT GAC AGC TGC CAG GGT GAC AGC GGC GGG CCT
          130                  140                  150                  160

                     750                                  760
Leu Val Cys Phe Glu Lys Asp Lys Tyr Ile Leu Gln Gly Val Thr Ser Trp Gly Leu Gly
CTG GTC TGC TTT GAG AAG GAC AAA TAC ATT CTG CAA GGA GTC ACT TCT TGG GGT CTT GGC
          190                  200                  210                  220

                     770                                  780
Cys Ala Arg Pro Asn Lys Pro Gly Val Tyr Val Arg Val Ser Pro Tyr Val Pro Trp Ile
TGT GCA CGC CCC AAT AAG CCT GGT GTG TAT GTT CGT GTT TCC CCC TAT GTT CCT TGG ATT
          250                  260                  270                  280

                     790
Glu Glu Thr Met Arg Arg Asn Stop
GAA GAA ACT ATG AGA AGG AAT TAA TTA GAT GGG AAT TAT TAG ATG GGA GAC AGT GAC CAT
          310                  320                  330                  340

GGT AGG GAA TCT AGC ATG CCA GAA ATA ATT GAC AGT AAT CCA ACC AAG TAA GAC ACT CTC
          370                  380                  390                  400

CTT AGC TGC CAT CTT CTG CCA AAC CTC AAC ATT TTT TTG TGA TGC TGT GAA TAA CAT GTT
          430                  440                  450                  460

CCT GAC TGT GGA CTG CTG GAT TCT GTA ATA CCA AGG TAA TAC AGC TAT GAC ATT TGT TGA
          490                  500                  510                  520

AAA TAA ACT CGG TAC TTC TTT TGA TTT AAAAAAAAAAAAA
          550                  560                  570        581
```

FIGURE 3: Nucleotide sequence of the bovine plasminogen cDNA insert in plasmid 5 and the corresponding amino acid sequence.

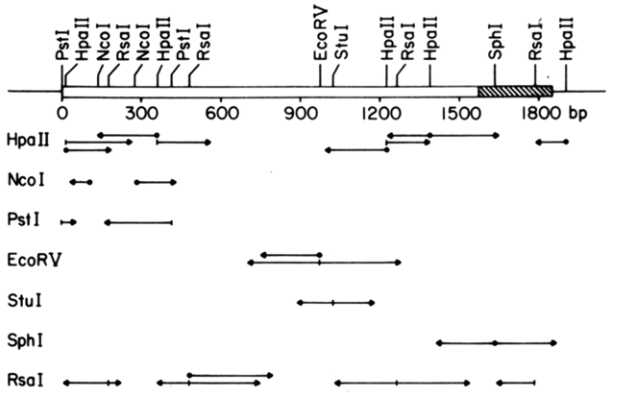

FIGURE 4: Restriction map of a human plasminogen cDNA and the strategy used to determine its nucleotide sequence. The shaded region corresponds to the 3′ noncoding sequence. The site of ^{32}P labeling for DNA sequencing is indicated as a perpendicular bar (–|–) for 5′ labelings and as a solid circle (–•–) for 3′ labelings.

with the Klenow fragment of DNA polymerase, and the double-stranded cDNA was tailed with dCTP in the presence

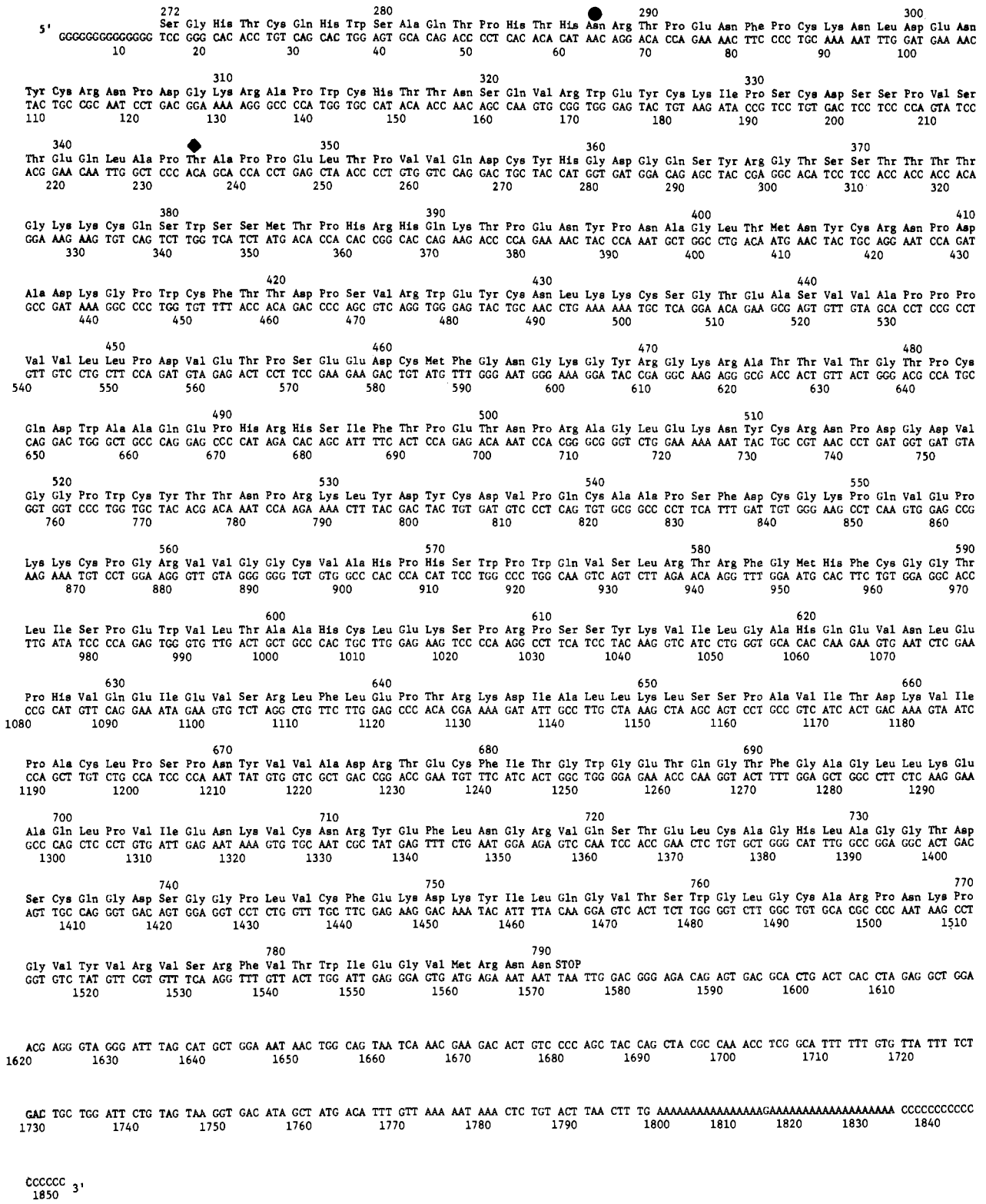

FIGURE 5: Nucleotide sequence of the human plasminogen cDNA described in Figure 4 and the corresponding amino acid sequence. The two carbohydrate binding sites (Hayes & Castellino, 1979a,b) are shown by solid diamonds.

of terminal transferase. The tailed cDNA was then annealed to plasmid pBR322, which had been previously cleaved at the *Pst*I site and tailed with dGTP. The chimeric plasmids were then used to transform *E. coli* K12, strain RR1 (Stein et al., 1978). Approximately 500 colonies were screened with a ^{32}P-labeled cDNA prepared from enriched plasminogen mRNA as described by MacGillivray et al. (1980). Fourteen colonies that were positive were identified, and these plasmids were then subjected to a second screening by a solution hybridization assay with a ^{32}P-labeled cDNA probe. Three plasmids, numbered 1, 4, and 5, provided significant protection of the single-stranded cDNA probe from S1 nuclease digestion. Clone 5, which provided approximately 15% protection compared to 8% for negative controls, was found to have an insert of about 580 base pairs. A restriction map for this clone was prepared (Figure 2), followed by a sequence analysis by the method of Maxam & Gilbert (1980). The sequencing strategy is summarized at the bottom of Figure 2. This cDNA insert was found to code for the carboxyl-terminal 107 amino acids of bovine plasminogen, followed by a stop codon of TAA, a

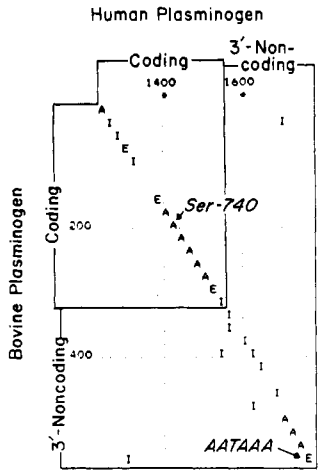

FIGURE 6: Homology matrix comparison of the bovine and human plasminogen cDNA nucleotide sequences. The human plasminogen sequence is residues 1251–1798 (Figure 5), and the bovine sequence is residues 1–567 (Figure 3). Parameters selected for the CINTHOM program of Pustell and Kafatos were range = 6, scale = 0.99, minimum value plotted = 78, and compression = 20. The letters (A, E, I) indicate the approximate percent of conserved residues over a 13 base pair segment according to the rule A = 99–100%, B = 97–98%, C = 95–96%, and so on. The positions of the active site serine (Ser-740) and the polyadenylation or processing signal (AATAAA) are indicated.

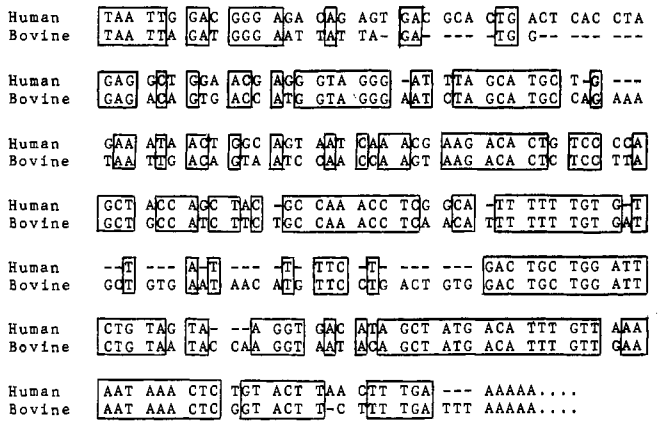

FIGURE 7: Alignment of the 3′ noncoding regions of the human and bovine plasminogen cDNA sequences starting with the stop codon of TAA in each cDNA. Conserved residues between the two sequences are enclosed in boxes. Gaps shown as (---) were inserted for maximal alignment of the two 3′ noncoding regions in each cDNA sequence.

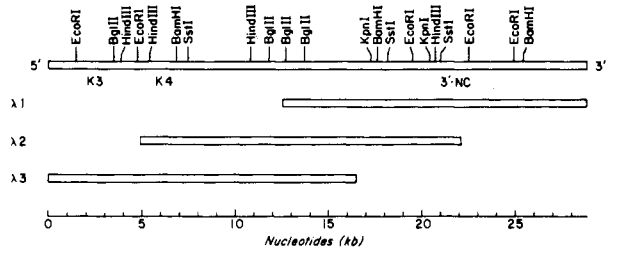

FIGURE 8: Restriction map of the partial human plasminogen gene. The regions of the gene carried in three representative recombinant λ phage isolates are indicated. The approximate location of sequences encoding plasminogen kringle 3, kringle 4, and the 3′ noncoding region of the plasminogen mRNA are shown as K3, K4, and 3′-NC, respectively.

noncoding region of 243 base pairs, and a poly(A) tail of 14 base pairs (Figure 3). The 3′ noncoding region also contained an AATAAA sequence that was 20 base pairs upstream from the poly(A) tail. This sequence is apparently involved in a polyadenylation or processing reaction (Proudfoot & Brownlee, 1976). Clone 4 contained an insert of 420 base pairs and was shown by sequence analysis to code for the carboxyl-terminal 42 amino acids of bovine plasminogen.

Human Plasminogen cDNA. A human liver cDNA library of 18 000 transformants was then screened for cDNAs coding for plasminogen by employing a radiolabeled nick-translated bovine plasminogen cDNA insert as a probe. Six plasmids that were positive with the radioactive bovine cDNA insert for plasminogen were identified and further characterized by restriction mapping and DNA sequencing. The longest cDNA insert contained approximately 1800 base pairs. A partial restriction map of this cDNA is shown in Figure 4, along with the sequencing strategy employing the method of Maxam & Gilbert (1980). The complete sequence for this cDNA is shown in Figure 5, along with the predicted amino acid sequence. Over 80% of the cDNA sequence was determined from both strands, and the remainder was sequenced at least twice on the same strand. This cDNA started with Ser-272 in human plasminogen and was 1851 base pairs in length, including 14 G's and 17 C's on the 5′ end and 3′ end, respectively. The carboxyl-terminal asparagine at position 790 was followed by a stop codon of TAA and 224 base pairs of noncoding nucleotides at the 3′ end. The region contained an AATAAA sequence (Proudfoot & Brownlee, 1976) 17 base pairs upstream from the poly(A) tail of 36 base pairs. The poly(A) tail also contained a single G, which is probably the result of a cloning artifact that occurred during the preparation of the human cDNA library. The amino acid sequence predicted from this cDNA sequence differs in only two amide assignments from the reported protein sequence (Sottrup-Jensen et al., 1978b). This includes residue 341, where the cDNA sequence predicts glutamine rather than glutamic acid, and an aspartic acid at residue 452 rather than asparagine.

Comparison of Bovine and Human cDNA Sequences. The bovine and human plasminogen cDNA sequences are compared in Figure 6 by using the CINTHOM matrix program of Pustell & Kafatos (1982). The coding regions correspond to the C-terminal 107 residues of human plasminogen and include the active site serine. Overall, 82% of the nucleotides are identical between the coding regions of the two partial cDNA sequences. Highly conserved sequences are evident around the active site serine (Residue 740 in the human protein) and especially for a stretch of 53 identical nucleotides corresponding to amino acid residues 755–772. This region probably accounts for the success of the cross-species hybridization employing the bovine cDNA as a probe against the human cDNA library. As noted for other comparisons between bovine and human cDNA sequences (Degen et al., 1983), the 3′ noncoding regions have diverged more than the coding regions, preserving approximately 60% of the residues with the alignment shown in Figure 7.

Human Plasminogen Gene. A human fetal liver genomic DNA library (Lawn et al., 1978; Maniatis et al., 1978) was screened with fragments of the human cDNA clone. The *Nco*I–*Sph*I fragment, composed of 1362 base pairs, was labeled by nick translation and employed to screen 10^6 recombinants of Charon 4A λ phage by the method of Benton & Davis (1977), as modified by Woo (1979). Approximately 50 phage were identified that hybridized to varying degrees with the probe. Each was plaque purified and the DNA isolated. Ten isolates have been partially characterized by restriction mapping and DNA sequence analysis and shown to contain plasminogen sequences. A restriction map of the plasminogen gene is shown in Figure 8. The portions of the genome included in representative λ phage isolates are indicated also. The approximate locations of sequences corre-

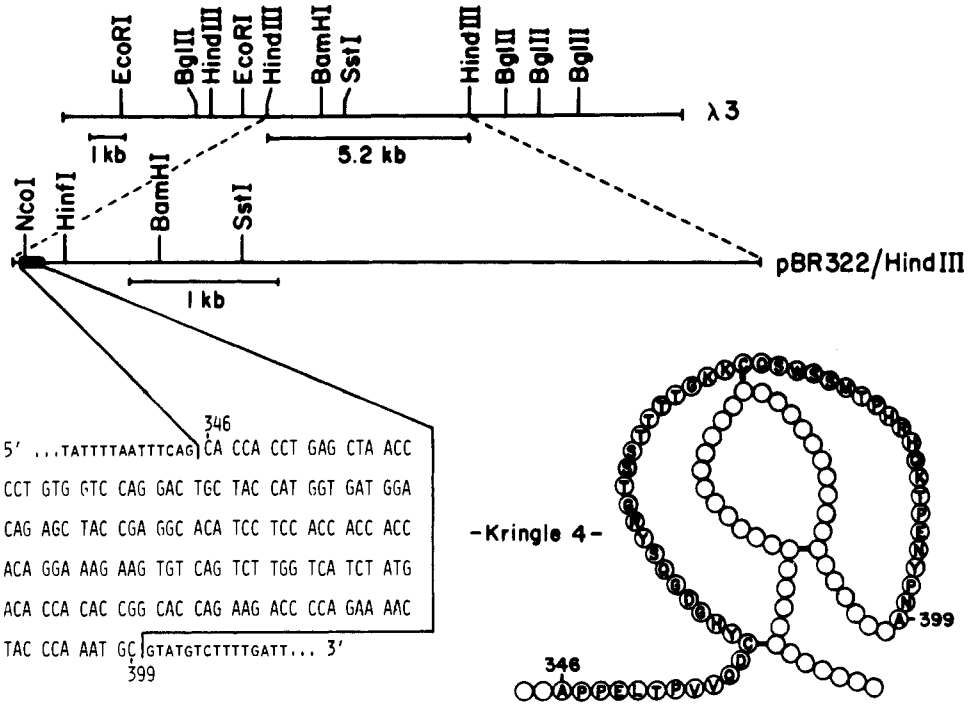

FIGURE 9: Location and DNA sequence of an exon encoding part of human plasminogen kringle 4. The 5.2-kb *Hin*dIII fragment of λ3 was subcloned into pBR322 and partially sequenced. Both 3′ and 5′ labelings at the unique *Nco*I site allowed the determination of the sequence shown from both strands. The exon encodes amino acid residues 346–399 of human plasminogen.

sponding to the third kringle or the 5′ end of the cDNA clone, and the 3′ noncoding region as determined by Southern blotting, are shown (J. E. Sadler and E. W. Davie, unpublished data). The length of this partial gene is over 21 kb. Since the cDNA employed to define the gene is only 1.8 kb in length, over 90% of the genomic DNA must consist of intervening sequences. Among the restriction sites shown in Figure 8, none occurs in the cDNA sequence; thus, all of these sites must lie within introns. Southern blotting studies indicate a minimum of six exons (J. E. Sadler and E. W. Davie, unpublished data).

The 5.2-kb restriction fragment spanning the second and third *Hin*dIII sites of the restriction map (Figure 8) has been subcloned into the *Hin*dIII site of pBR322 by employing purified λ3 DNA. Sequence studies have yielded the sequence of an exon corresponding to the first half of the fourth kringle of human plasminogen (Figure 9). Also shown is a restriction map of the insert from this subclone. The position of the exon is shown as a solid bar, along with the corresponding portion of plasminogen kringle 4 and the genomic DNA sequence. The 160 base pairs of the exon are identical with the plasminogen cDNA residues 238–397 and encode the protein sequence from amino acid 346 to amino acid 399. In each case, the codons for alanine that begin and end the exon are split by the flanking splice junctions.

Discussion

A cDNA clone corresponding to the carboxyl-terminal two-thirds of human plasminogen has been isolated by exploiting the high degree of nucleotide sequence conservation among homologous bovine and human serine proteases. A similar approach has been successful in the isolation of human liver cDNA clones for prothrombin using a bovine cDNA probe (Degen et al., 1983) and for factor IX using a baboon cDNA probe (Kurachi & Davie, 1982). Comparison of the human and bovine plasminogen sequences reveals the expected similarities of nucleotide sequence in regions of highly conserved amino acid sequence. This comparison is also true for the plasminogen cDNA sequence with the same region of other serine proteases, as well as for comparisons among the cDNA sequences of the kringle regions of human plasminogen with those of prothrombin (Degen et al., 1983) and tissue plasminogen activator (Pennica et al., 1983). However, no consistent similarity can be demonstrated between the 3′ noncoding region of human plasminogen and the 3′ noncoding regions of either human prothrombin or human factor IX cDNAs (J. E. Sadler and E. W. Davie, unpublished results).

The two disagreements between the protein sequence predicted from the cDNA sequence and that reported earlier (Sottrup-Jensen et al., 1978b) are possibly the result of incorrect amide assignments made during protein sequencing, although polymorphisms in protein structure or artifacts during cDNA cloning cannot be excluded. Recent NMR studies of the first plasminogen kringle (DeMarco et al., 1982) suggest discrepancies in the number of methionine, phenylalanine, and tyrosine residues when compared to earlier protein sequence reports (Sottrup-Jensen et al., 1978b). Cloning of the 5′ end of a plasminogen cDNA may help resolve these differences, and these experiments are in progress.

Aoki and colleagues have characterized a variant of plasminogen that is catalytically inactive. This variant form contains a single amino acid substitution of Ala-600 by Thr (Miyata et al., 1982; Sakata & Aoki, 1980). As first suggested by these investigators, this variant could result from a G → A transition in the first nucleotide of the Ala-600 codon (nucleotide 999 in Figure 5).

The present data do not clarify the interesting findings of Gonzalez-Gronow & Robbins (1984), who observed the synthesis of plasminogen form 1 by 23S mRNA and plasminogen form 2 by 18S mRNA isolated from monkey liver. Further analysis of a number of near-full-lenth cDNAs and the gene for plasminogen may provide some explanation for these data.

The limited structure of the plasminogen gene presented herein can be compared with the organization of the prothrombin gene (Degen et al., 1983; S. J. F. Degen and E. W. Davie, unpublished results). As in kringle 4 of plasminogen, the coding sequence for kringle 1 of prothrombin is split by

an intron. The prothrombin intron actually occurs between residues 98 and 99, nine amino acid residues prior to the corresponding splice junction in plasminogen kringle 4. In contrast, the entire kringle 2 of prothrombin is encoded by a single exon. Tissue plasminogen activator (Pennica et al., 1983) and urokinase (Gunzler et al., 1982) are now known to contain two and one kringle(s), respectively. As the number of kringle-containing proteins increases, so does the need to define the structural and evolutionary relationships between them. Further characterization of the plasminogen gene will allow comparisons of the genomic organization of the kringle domains for prothrombin and plasminogen, and perhaps of other proteins. Similar comparisons for the protease domains of plasminogen, prothrombin, factor IX, other coagulation factors, and related serine proteases will reveal details of the evolution and possibly the transcriptional regulation of these enzymes.

Registry No. Plasminogen, 9001-91-6.

References

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.

Benton, W. D., & Davis, R. W. (1977) *Science (Washington, D.C.) 196*, 180–182.

Bohmfalk, J. F., & Fuller, G. M. (1980) *Science (Washington, D.C.) 209*, 408–410.

Brockway, W. J., & Castellino, F. J. (1972) *Arch. Biochem. Biophys. 151*, 194–199.

Chandra, T., Stackhouse, R., Kidd, V. J., & Woo, S. L. C. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1845–1848.

Davie, E. W., Fujikawa, K., Kurachi, K., & Kisiel, W. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol. 48*, 277–318.

Degen, S. J. F., MacGillivray, R. T. A., & Davie, E. W. (1983) *Biochemistry 22*, 2087–2097.

DeMarco, A., Hochschwender, S. M., Laursen, R. A., & Llinás, M. (1982) *J. Biol. Chem. 257*, 12716–12721.

Deutsch, D. G., & Mertz, E. T. (1970) *Science (Washington, D.C.) 170*, 1095–1096.

Gonzalez-Gronow, M., & Robbins, K. C. (1984) *Biochemistry 23*, 190–196.

Gough, A., & Adams, J. M. (1978) *Biochemistry 17*, 5560–5566.

Groskopf, W. R., Summaria, L., & Robbins, K. C. (1969) *J. Biol. Chem. 244*, 3590–3597.

Gunzler, W. A., Steffens, G. J., Ötting, F., Kim, S-M. A., Frankus, E., & Flohé, L. (1982) *Hoppe-Seyler's Z. Physiol. Chem. 363*, 1155–1165.

Harboe, N., & Ingild, A. (1973) *Scand. J. Immunol., Suppl. 2*, 161–164.

Hayes, M., & Castellino, F. J. (1979a) *J. Biol. Chem. 254*, 8772–8776.

Hayes, M., & Castellino, F. J. (1979b) *J. Biol. Chem. 254*, 8777–8780.

Kurachi, K., & Davie, E. W. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6461–6464.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

Laskey, R. A., & Mills, A. D. (1975) *Eur. J. Biochem. 56*, 335–341.

Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G., & Maniatis, T. (1978) *Cell (Cambridge, Mass.) 15*, 1157–1174.

MacGillivray, R. T. A., Chung, D. W., & Davie, E. W. (1978) *Eur. J. Biochem. 98*, 477–485.

MacGillivray, R. T. A., Friezner-Degen, S. J., Chandra, T., Woo, S. L. C., & Davie, E. W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5153–5157.

Magnusson, S., Petersen, T. E., Sottrup-Jensen, L., & Claeys, H. (1975) in *Proteases and Biological Control* (Reich, E., Rifkin, D. B., & Shaw, E., Eds.) pp 123–249, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.

Malinowski, D. P., & Davie, E. W. (1983) in *Progress in Fibrinolysis* (Davidson, J. F., Bachmann, F., Bouvier, C. A., & Kruithof, E. K. O., Eds.) Vol. 6, pp 321–326, Churchill Livingstone, Edinburgh, Scotland.

Malinowski, D. P., Sadler, J. E., & Davie, E. W. (1983) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 42*, 1761.

Maniatis, T., Jeffrey, A., & Kleid, D. G. (1975a) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1184–1188.

Maniatis, T., Jeffrey, A., & van de Sande, H. (1975b) *Biochemistry 14*, 3787–3794.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., & Efstratiadis, A. (1978) *Cell (Cambridge, Mass.) 15*, 687–701.

Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 560–580.

Miyata, T., Iwanaga, S., Sakata, Y., & Aoki, N. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6132–6136.

Palmiter, R. D. (1974) *Biochemistry 13*, 3606–3615.

Palmiter, R. D., Gagnon, J., Ericsson, L. H., & Walsh, K. A. (1977) *J. Biol. Chem. 252*, 6386–6393.

Pelham, H. R. B., & Jackson, R. J. (1976) *Eur. J. Biochem. 67*, 247–256.

Pennica, D., Holmes, W. E., Kohr, W. J., Harkins, R. N., Vehar, G. A., Ward, C. A., Bennett, W. F., Yelverton, E., Seeburg, P. H., Heyneker, H. L., Goeddel, D. V., & Collen, D. (1983) *Nature (London) 301*, 214–221.

Porath, J., & Kristansen, T. (1975) *Proteins (3rd Ed.) 1*, 115–121.

Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature (London) 263*, 211–214.

Pustell, J., & Kafatos, F. C. (1982) *Nucleic Acids Res. 10*, 4765–4782.

Raum, D., Marcus, D., Alper, C. A., Levey, R., Taylor, P. D., & Starzl, T. E. (1980) *Science (Washington, D.C.) 208*, 1036–1037.

Rixon, M. W., Chan, W.-Y., Davie, E. W., & Chung, D. W. (1983) *Biochemistry 22*, 3237–3244.

Robbins, K. C. (1982) in *Hemostasis and Thrombosis* (Colman, R. W., Hirsh, J., Marder, V. J., & Salzman, E. W., Eds.) pp 623–639, J. B. Lippincott, Philadelphia.

Robbins, K. C., Bernabe, P., Arzadon, L., & Summaria, L. (1973) *J. Biol. Chem. 248*, 1631–1633.

Saito, H., Hamilton, S. M., Travill, A. S., Louis, L., & Ratnoff, O. D. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 6837–6840.

Sakata, Y., & Aoki, N. (1980) *J. Biol. Chem. 255*, 5442–5447.

Schechter, I. (1974) *Biochemistry 13*, 1875–1885.

Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L., & Hall, B. D. (1979) *Cell (Cambridge, Mass.) 16*, 753–761.

Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T. E., & Magnusson, S. (1978a) in *Fibrinolysis and Thrombolysis* (Davidson, T. F., Rowan, R. M., Samama, M. M., & Desnoyer, P. C., Eds.) Vol. 3, pp 191–209, Raven Press, New York.

Sottrup-Jensen, L., Petersen, T. E., & Magnusson, S. (1978b) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, p 91, National Biomedical Research Foundation, Silver Spring, MD.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503-517.
Stein, J. P., Catterall, J. F., Woo, S. L. C., Means, A. R., & O'Malley, B. W. (1978) *Biochemistry 17*, 5763-5772.
Tu, C. D., & Cohen, S. N. (1980) *Gene 10*, 177-183.
Wallén, P., & Wiman, B. (1972) *Biochim. Biophys. Acta 257*, 122-134.
Weber, K., & Osborn, M. (1969) *J. Biol. Chem. 244*, 4406-4412.
Wiman, B. (1977) *Eur. J. Biochem. 76*, 129-137.
Wiman, B., & Wallén, P. (1975a) *Eur. J. Biochem. 50*, 489-494.
Wiman, B., & Wallén, P. (1975b) *Eur. J. Biochem. 58*, 539-547.
Woo, S. L. C. (1979) *Methods Enzymol. 68*, 389-395.

Differential Inhibition of Histone and Polyamine Acetylases by Multisubstrate Analogues†

Bradley G. Erwin, Lo Persson, and Anthony E. Pegg*

ABSTRACT: Mammalian cells contain a number of enzymes catalyzing the acetylation of polyamines and histones including an inducible spermidine/spermine N^1-acetyltransferase which may play a key role in regulating the interconversion of polyamines [Matsui, I., Wiegand, L., & Pegg, A. E. (1981) *J. Biol. Chem. 256*, 2454-2459]. The present experiments were carried out in order to provide a method to distinguish this enzyme from other polyamine/histone acetylases and to test whether specific inhibitors of its activity could be obtained. Rabbit antiserum to homogeneous rat liver spermidine/spermine N^1-acetyltransferase had no effect on the activity of a crude nuclear extract from rat liver, indicating that its spermidine acetylating capability is not related to the cytosolic spermidine/spermine N^1-acetyltransferase induced by hepatotoxins. Potential multisubstrate analogues were prepared by attaching various polyamines to coenzyme A via an acetic acid linkage and tested as potential inhibitors of the acetylation of spermidine and histones. There was little difference in the potency of these polyamine derivatives as inhibitors of histone or spermidine acetylation by the crude nuclear extracts which appeared to contain at least two such activities, one inhibited completely by 20-30 μM and the other amounting to 50% of the total being unaffected by 100 μM. Spermidine/spermine N^1-acetyltransferase was also inhibited by all the derivatives, but the potency toward this enzyme differed widely. The derivative from *sym*-norspermidine was a very strong inhibitor, giving 50% inhibition at 0.3 μM, and was more than 1 order of magnitude more active than the others. These results are consistent with *N*-[2-(*S*-coenzyme A)acetyl]-*sym*-norspermidine amide acting as a multisubstrate analogue since *sym*-norspermidine is a preferred substrate of spermidine/spermine N^1-acetyltransferase, having a K_m (9 μM) 14 times less than that for spermidine (130 μM). Comparisons of the effects of these inhibitors on cells and nuclear extracts may be valuable in understanding the physiological role of polyamine and histone acetylases.

A number of reports have appeared describing the presence of enzymes acetylating polyamines (Seiler & Al-Therib, 1974; Blankenship & Walle, 1977, 1978; Libby, 1978, 1980; Matsui et al., 1981, 1983; Della Ragione & Pegg, 1982; Cullis et al., 1982). Some but not all of the preparations also acetylate histones, and multiple forms of the enzymes acetylating both polyamines and histones have been described (Libby, 1978, 1980; Sures & Gallwitz, 1980; Garcea & Alberts, 1980). Intact cells and crude cell extracts are, therefore, likely to contain a mixture of proteins capable of catalyzing the formation of acetylated spermidine and spermine using acetyl coenzyme A (acetyl-CoA) as substrate. One of these enzymes, which we have described as spermidine/spermine N^1-acetyltransferase (Matsui et al. 1981; Della Ragione & Pegg, 1982), may play an important role in the regulation of intracellular polyamine concentrations. This enzyme is the rate-limiting step in the interconversion of polyamines which is brought about by the degradation of N^1-acetylspermidine and N^1-acetylspermine by polyamine oxidase, forming 3-acetamidopropanal and either putrescine or spermidine, respectively (Pegg et al., 1981; Seiler et al., 1981). Spermidine/spermine N^1-acetyltransferase is highly inducible in rodent liver in response to treatment with the hepatotoxin carbon tetrachloride (Matsui et al., 1981; Pösö & Pegg, 1982). After such induction, it is the predominant form of polyamine acetylase in cytosolic extracts, but it was not clear to what extent it contributed to the nuclear activity. Also, the importance of the oxidase/acetylase pathway for the interconversion of polyamines in maintaining normal polyamine content is not yet well understood. A specific and potent inhibitor of the spermidine/spermine N^1-acetyltransferase would provide one means by which to investigate these questions.

In the present paper, we describe tests of a number of potential inhibitors based on the synthesis of multisubstrate analogues of CoA with various polyamines. Cullis et al. (1982) have previously reported that such a spermidine derivative was a powerful inhibitor of the histone acetylase A from calf thymus. Since cytosolic spermidine/spermine N^1-acetyltransferase has a much lower K_m for *sym*-norspermidine than for spermidine and the reaction has been shown to occur via an ordered Bi-Bi mechanism (Della Ragione & Pegg, 1982, 1983), we reasoned that *N*-[2-(*S*-coenzyme A)acetyl]-*sym*-norspermidine amide should be a potent inhibitor of this enzyme, and this was found to be the case. This inhibitor was less active against nuclear histone/spermidine acetylase and could be used to distinguish between these enzymes. The

† From the Department of Physiology and Cancer Research Center, The Milton S. Hershey Medical Center, The Pennsylvania State University, Hershey, Pennsylvania 17033. *Received January 27, 1984.* This work was supported in part by Research Grants GM-26290, 1P30 CA 18450, and CA 18138 from the National Institutes of Health and by a travel grant to L.P. from the Swedish Medical Research Council.

* Correspondence should be addressed to this author at the Department of Physiology.

0006-2960/84/0423-4250$01.50/0