Characterization of Bovine Prothrombin mRNA and Its Translation Product†

Ross T. A. MacGillivray and Earl W. Davie*

ABSTRACT: Prothrombin mRNA has been enriched 20–60-fold by using specific immunoadsorption of bovine liver polysomes. The enriched mRNA was translated in a cell-free protein synthesizing system derived from rabbit reticulocytes, and the radiolabeled translation product was isolated by immunoprecipitation and sodium dodecyl sulfate–polyacrylamide gel electrophoresis. The radiolabeled prothrombin synthesized in the cell-free system was then subjected to automated Edman degradation and shown to contain a leader sequence of at least 30 residues that was rich in leucine, phenylalanine, and alanine. In order to fully characterize the leader sequence for prothrombin, a bovine liver cDNA library was constructed containing DNA inserts of over 1000 base pairs. Two cDNA clones coding for bovine prothrombin were isolated from the library and their nucleotide sequences determined. A leader sequence of 43 amino acids was predicted from the sequence of the cDNA, and the first 30 residues were in agreement with the partial sequence obtained by the cell-free protein synthesizing system. From the amino acid sequence of the leader sequence, it is proposed that bovine prothrombin is synthesized with a prepro leader sequence starting with a methionine residue at position −43. The amino acid sequence of the mature prothrombin molecule circulating in plasma was also predicted from the cDNA and shown to be in good agreement with that determined previously by conventional amino acid sequence analysis.

In the final stages of blood coagulation, prothrombin is converted to thrombin by factor Xa in the presence of factor Va, calcium ions, and phospholipid. In this reaction, factor Xa cleaves two internal peptide bonds in bovine prothrombin, producing an activation peptide (residues 1–274) and thrombin (residues 275–323 and 324–582 held together by a disulfide bond). Thrombin, a serine protease, then converts fibrinogen to fibrin by limited proteolysis, releasing fibrinopeptides A and B.

Bovine prothrombin is a glycoprotein consisting of a single polypeptide chain of 582 amino acid residues (Magnusson et al., 1975). It also contains γ-carboxyglutamic acid in the amino-terminal region of the molecule (Stenflo et al., 1974; Nelsestuen et al., 1974; Magnusson et al., 1975). These γ-carboxyglutamic acid residues result from the carboxylation of the first 10 glutamic acid residues in the protein by a membrane-bound vitamin K dependent carboxylase (Friedman & Shia, 1976; Sadowski et al., 1976; Helgeland, 1977). The physiological function of the γ-carboxyglutamic acid residues appears to be the binding of calcium ions to the protein (Stenflo & Suttie, 1977).

The γ-carboxyglutamic acid region of prothrombin is followed by two regions of internal homology called kringle structures (Magnusson et al., 1975). These structures are composed of about 80 amino acids with a typical disulfide bond arrangement. Although kringles have also been found in other plasma proteins such as plasminogen (Sottrup-Jensen et al., 1978) and plasminogen activator (Pennica et al., 1983), their function remains unclear.

The carboxyl-terminal half of prothrombin contains the catalytic region of the proteolytic enzyme. This region is highly homologous with the catalytic region in other serine proteases, including most of the blood coagulation factors (Davie et al., 1979).

The major site of prothrombin synthesis is the liver (Anderson & Barnhart, 1964). Nardacci et al. (1975) translated rat liver poly(A)-containing RNA in a cell-free protein synthesizing system and identified prothrombin among the translation products. MacGillivray et al. (1979) reported similar experiments and showed that prothrombin represented approximately 1% of the translation products of bovine liver poly(A)-containing RNA. Besmond et al. (1981) have also reported similar results using human liver poly(A)-containing RNA. In the latter case, however, prothrombin represented only 0.1% of the in vitro translation products.

Studies on the mechanism of prothrombin biosynthesis have been difficult due to the low level of prothrombin mRNA in liver. In this paper we report the 20–60-fold enrichment of bovine prothrombin mRNA by immunoprecipitation of polysomes. This enriched preparation was then used to characterize the translation product of prothrombin mRNA in a cell-free system. We also report the isolation and characterization of a nearly full-length bovine prothrombin cDNA clone which codes for the complete amino acid sequence of prepro-prothrombin. This clone was much larger than that previously isolated (MacGillivray et al., 1980b) and was similar to a cDNA clone recently characterized for the human molecule (Degen et al., 1983).

Experimental Procedures

Materials

Restriction enzymes were purchased from Bethesda Research Laboratories (BRL) (Gaithersburg, MD) and New England Biolabs (Beverly, MA) and were used according to the manufacturer's recommendations. BRL also supplied T4 DNA ligase and terminal deoxynucleotidyltransferase. *Escherichia coli* DNA polymerase I was purchased from Boehringer Mannheim. *E. coli* DNA polymerase I (Klenow fragment) and T4 polynucleotide kinase were purchased from

† From the Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5 (R.T.A.M.), and the Department of Biochemistry, University of Washington, Seattle, Washington 98195 (E.W.D.). *Received October 3, 1983.* This work was initiated in Seattle where it was supported by Grant HL 16919 from the National Institutes of Health (to E.W.D.) and was completed in Vancouver where it was supported by Grant MA-7716 from the Medical Research Council of Canada and Grant 19(82-1) from the British Columbia Health Care Research Foundation (to R.T.A.M.). A portion of this work was presented at the 8th Steenbock Symposium, University of Wisconsin—Madison, June 10–13, 1979 (MacGillivray et al., 1980a).

0006-2960/84/0423-1626$01.50/0 © 1984 American Chemical Society

New England Nuclear. Avian myeloblastoma virus reverse transcriptase (RNA-dependent DNA nucleotidyltransferase) was generously supplied by Dr. Joseph Beard (Life Sciences, St. Petersburg, FL). Bovine prothrombin and α-thrombin were kindly supplied by Dr. W. Kisiel, University of Washington.

P-L Biochemicals supplied deoxy- and dideoxyribonucleotide triphosphates, M13-specific primer (pentadecanucleotide), and oligo(dT)–cellulose (type 7). Tritium- and ^{35}S-labeled amino acids were purchased from New England Nuclear, Amersham, and Schwarz/Mann and were of the highest specific activity available. All ^{32}P-labeled nucleotides were purchased from New England Nuclear. Human placental ribonuclease inhibitor was purchased from Enzo Biochemicals Inc. (NY).

Reagents for the automatic sequenator were "Sequenal" grade purchased from Pierce Chemical Co. or were sequencer grade from Beckman Instruments. Phenol was distilled prior to use and stored at −20 °C. Precautions taken during the preparation of RNA were as described (MacGillivray et al., 1979).

Methods

Bovine Prothrombin mRNA. Total RNA was isolated from frozen liver powder by using the method of MacGillivray et al. (1979) for the isolation of small amounts or the guanidinium chloride method of Chirgwin et al. (1979) for larger amounts of RNA. Poly(A)-containing RNA was isolated by chromatography with oligo(dT)–cellulose (Aviv & Leder, 1972).

Bovine mRNA was enriched for prothrombin by immunoadsorption of polysomes using rabbit anti-prothrombin antibodies and protein A containing *Staphylococcus aureus* bacteria (Gough & Adams, 1978). The method described was developed in collaboration with Dr. Dominic W. Chung, University of Washington.

Bovine liver polysomes were prepared from frozen bovine liver powder by a modification (MacGillivray et al., 1979) of the magnesium precipitation method (Palmiter, 1974). The polysomes were used immediately or stored at −70 °C for up to 6 months. After thawing, the polysomes were centrifuged at 27000*g* for 10 min to remove aggregates. The immunoadsorption appeared to work equally well with either fresh polysomes or frozen polysomes.

Rabbit anti-prothrombin antibodies were purified by affinity chromatography on prothrombin–cellulose. Prior to the chromatographic step, the antibodies were rendered ribonuclease free by employing the washing step of Gough & Adams (1978). In some experiments, the antibodies were freed of ribonuclease by chromatography on a column of DEAE-cellulose and CM-cellulose (Whatman DE-52 and CM-52) as described by Palacios et al. (1972). Bovine prothrombin was coupled to CNBr-activated Sepharose 4B by using the method of Cuatrecasas (1970).

Staphylococcus aureus cells ("Pansorbin"; Calbiochem) were used as the insoluble matrix for the isolation of specific polysomes. Prior to use, the cells were washed 3 times in buffer A [50 mM Tris-HCl,[1] pH 7.5, 5 mM EDTA, 0.15 M NaCl, and 0.05% Nonidet P-40 (BDH Chemicals, Ltd., Poole, England)], 3 times in buffer B (50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 0.3 M NaCl, 0.5% Nonidet P-40, and 0.5% sodium deoxycholate), and twice in buffer C (50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 2 mM EGTA, and 0.3 M NaCl). Each wash consisted of suspending 1 g of cells in 30 mL of buffer, followed by centrifugation at 13000*g* in a Sorvall HB-4 rotor at 4 °C for 10 min. Cells were stored at 4 °C as a suspension in buffer C and centrifuged and resuspended in fresh buffer C prior to use. These washing steps removed contaminating ribonuclease in the Pansorbin as well as fines.

For a typical enrichment experiment, bovine liver polysomes (3000 A_{260}) were diluted to 25 A_{260}/mL with ice-cold PB and gently stirred at 4 °C. Rabbit anti-prothrombin antibodies were added (3.3 $\mu g/A_{260}$ polysomes), and the mixture was gently stirred at 4 °C for 1 h. Washed Pansorbin cells were added such that a 2-fold excess of Pansorbin over antibodies was present, assuming no losses during the washing steps. The binding capacity of each lot of Pansorbin was supplied by the manufacturer. The suspension was stirred gently at 4 °C for an additional 10 min, and 17-mL aliquots were layered over a step gradient of sucrose (4 mL of 0.5 mM sucrose in buffer C containing 0.5% sodium deoxycholate and 0.5% NP-40 layered over 7 mL of 1.0 M sucrose in the same solution) in a 30-mL Corex tube. The gradients were centrifuged at 13000*g* in a Sorvall HB-4 rotor at 4 °C for 15 min. The solution was aspirated into the 0.5 M sucrose layer, and the sides of the tube were rinsed with water. The solution was then aspirated to the pellet, and the sides of the tube were wiped with a tissue. The pellets were resuspended in buffer C and recentrifuged on the sucrose step gradient. After aspiration of the supernatant, the bound polysomes were released by resuspending the pellets in a total volume of 3 mL of 10 mM Tris-HCl, pH 7.5, 20 mM EDTA, 0.15 M NaCl, and 1.5% $NaDodSO_4$. The suspension was heated at 50 °C for 2 min and centrifuged at 27000*g* in a Sorvall SS-34 rotor at 20 °C for 10 min. The supernatant was removed and the pellet resuspended in 1.5 mL of EDTA–$NaDodSO_4$ solution and centrifuged. The supernatants were pooled and extracted with an equal volume of phenol, followed by three extractions with an equal volume of chloroform. The enriched RNA was precipitated with ethanol in the presence of 0.2 M NaCl and recovered by centrifugation. Poly(A)-containing RNA was isolated by chromatography with oligo(dT)–cellulose.

RNA was translated in a cell-free protein-synthesizing system prepared from rabbit reticulocytes (Pelham & Jackson, 1976). For some experiments, the lysate was gel filtered on a column of Sephadex G-50 to reduce endogenous amino acid levels (Palmiter et al., 1977). Immunoprecipitation of the translation products with specific antisera was carried out as described (MacGillivray et al., 1979). Protein samples were analyzed by polyacrylamide gel electrophoresis in the presence of $NaDodSO_4$ (Laemmli, 1970) and visualized by fluorography after treatment of the gel with ENHANCE (New England Nuclear).

Translation of RNA for Automatic Sequenator Analysis. For the large-scale translation of mRNA, 1 μg of enriched poly(A) RNA was incubated with the Sephadex G-50 treated reticulocyte lysate containing the labeled amino acid (200 μCi/mL) in a final volume of 0.5 mL. After incubation at 26 °C for 1.5 h, the translation reaction was terminated by the addition of Triton X-100 to 0.3% (v/v) and unlabeled amino acid to 1 mM. The radiolabeled product was immunoprecipitated (Palmiter et al., 1977) and dissolved in a minimum volume of 2% $NaDodSO_4$ by warming at 37 °C for a few minutes. The product was then diluted with 10 mM Tris-HCl, pH 7.5, so that the final concentration of $NaDodSO_4$ was 0.02%. Thrombin was added (25% w/w compared to the amount of preprothrombin in the original immunoprecipita-

[1] Abbreviations: Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; EGTA, [ethylenebis(oxyethylenenitrilo)]tetraacetic acid; PB, 10 mM Tris-HCl buffer, pH 7.5, 0.15 M NaCl, 5 mM $MgCl_2$, and 0.1 mg/mL heparin; $NaDodSO_4$, sodium dodecyl sulfate; bp, base pair(s).

tion) and the reaction mixture incubated at 37 °C for 10 min. The reaction mixture was lyophilized, redissolved in sample loading buffer (75 mM Tris-HCl, pH 7.5, 1% $NaDodSO_4$, 5% glycerol, 5% 2-mercaptoethanol, and 0.02% bromophenol blue), and loaded onto a 10% polyacrylamide gel containing 0.1% $NaDodSO_4$ (Laemmli, 1970) with a 3% stacking gel. After electrophoresis, the gel was cut into 1-mm slices, and the slices were shaken overnight at 37 °C in a buffer containing 10 mM sodium phosphate, pH 7.2, 0.1% $NaDodSO_4$, and 0.25 mg/mL bovine serum albumin (Miles). Radioactive polypeptides were located by counting a small aliquot of each tube in a scintillation counter. Fractions containing radioactivity were pooled and prepared for the automatic sequenator by the acetone–HCl precipitation method of Strauss et al. (1977).

Automated sequenator analyses were carried out on a Beckman 890C instrument, employing the dimethylbenzylamine program of Hermodson et al. (1972). For protein samples labeled with a single amino acid, the phenylthiohydantoin derivatives were counted directly (Palmiter et al., 1977). For protein samples labeled with more than one amino acid, the phenylthiohydantoin derivatives were separated by high-performance liquid chromatography as described (Palmiter et al., 1977). Under these conditions, the counting efficiency for ^{35}S was >90%, and the counting efficiency for ^{3}H was 45%.

Construction and Screening of the Bovine Liver cDNA Library. The cDNA library was prepared by employing human placental ribonuclease inhibitor (Martynoff et al., 1980) at a concentration of 100 μg/mL during the first-strand synthesis. Double-stranded cDNA was then synthesized by using the method of Land et al. (1981). The cDNA was tailed with dCMP residues (Deng & Wu, 1981) and subjected to electrophoresis on a 1% low melting point agarose gel (BRL) together with DNA size markers. After autoradiography, DNA fragments greater than one kilobase pair were recovered by the procedure supplied by BRL, annealed to dG-tailed pBR322 (Bolivar et al., 1977), and used to transform *E. coli* strain RR1 (Dagert & Ehrlich, 1979). The dG-tailed pBR322 was a generous gift from Dr. W. R. McMaster, University of British Columbia.

The bovine liver cDNA library was screened by in situ hybridization (Grunstein & Hogness, 1975) using the *Pst*I insert of pBII3 (MacGillivray et al., 1980b) previously labeled by nick translation (Maniatis et al., 1975) as a probe. Colonies were also screened at high colony density by using the method of Hanahan & Meselson (1980). Plasmid DNA was isolated primarily by the methods of Katz et al. (1973, 1977).

DNA Sequence Analysis. DNA sequence analysis was carried out by using both the chemical cleavage method (Maxam & Gilbert, 1980) and the chain termination method (Sanger et al., 1977) using the single-stranded phage M13 (Messing et al., 1981; Messing & Vieira, 1982) as a cloning vector. Restriction endonuclease fragments were ligated individually or in mixtures into the *Bam*HI site of M13mp7 (*Sau*3A fragments) or the *Sma*I sites of M13mp8 and M13mp9 (*Hae*III, *Rsa*I, or *Fnu*DII fragments). This DNA was used to transform *E. coli* strain JM103 (Messing et al., 1981). Recombinants were screened by in situ hybridization (Benton & Davis, 1977), hybridization of phage DNA on dot blots (Kafatos et al., 1979), and "T-tracking" (Sanger et al., 1982). These analyses were necessary because some clear plaques generated by blunt-end ligation into the *Sma*I site of M13 were found to be the result of religation of exonuclease-digested vector DNA. Single-stranded DNA was prepared from single plaques (Winter & Fields, 1980) and used as a template for the chain termination reactions with a synthetic pentadecadeoxyribonucleotide as a primer. The reaction products were analyzed by denaturing polyacrylamide gels. DNA sequence data were stored and analyzed by using the computer programs of Delaney (1982).

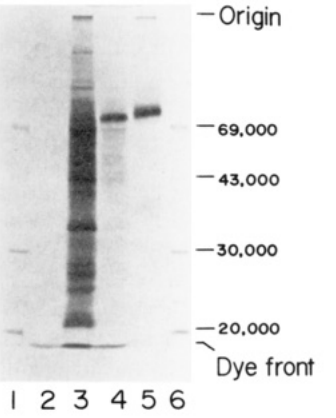

FIGURE 1: Fluorogram of [^{3}H]leucine-labeled translation products of total bovine liver poly(A) RNA and enriched prothrombin poly(A) RNA. Samples were analyzed by $NaDodSO_4$–polyacrylamide gel electrophoresis after reduction with 2-mercaptoethanol. The stained gel was treated with ENHANCE, dried, and exposed to Kodak XRP-1 film. Lanes 1 and 6, tritiated protein standard comprised of bovine albumin (M_r 69 000), ovalbumin (M_r 43 000), carbonic anhydrase (M_r 30 000), and soybean trypsin inhibitor (M_r 20 000); lane 2, translation products in the absence of added RNA; lane 3, translation products of total bovine liver poly(A) RNA; lane 4, translation products of enriched prothrombin poly(A) RNA; lane 5, immunoprecipitate of the sample shown in lane 4 employing antibodies to prothrombin. The positions of the origin and dye front are indicated.

Experiments were performed in compliance with NIH and MRC guidelines for Recombinant DNA Research.

Results

Enrichment of Bovine Prothrombin mRNA. Bovine liver polysomes were incubated with ribonuclease-free anti-prothrombin antibodies followed by incubation with protein A containing *S. aureus* cells. After extensive washing, bound polysomes were released with $NaDodSO_4$, and the poly(A)-containing RNA was isolated. This enriched RNA and total liver poly(A)-containing RNA were translated in a cell-free protein-synthesizing system derived from rabbit reticulotyes, and the translation products were analyzed by $NaDodSO_4$–polyacrylamide gel electrophoresis followed by fluorography (Figure 1). In the absence of added mRNA, the reticulocyte lysate synthesized several short polypeptides that migrated with the dye front (lane 2). The addition of total liver poly(A)-containing RNA to the lysate, however, resulted in the synthesis of a large number of proteins (lane 3). In the presence of enriched prothrombin mRNA, a single major protein (M_r 71 000) together with several minor proteins of M_r <71 000 were synthesized (lane 4). Immunoprecipitation of the translation products employing enriched mRNA also showed a major band corresponding to prothrombin (lane 5). Depending on the particular enriched mRNA preparation employed, 30–60% of the translation products were immunoprecipitated by the antibodies to prothrombin.

Characterization of the Radiolabeled Product from a Translation Assay. The radiolabeled product from a translation assay was further characterized by examining the degradation products generated by thrombin. In these experiments, enriched prothrombin mRNA was translated with the reticulocyte lysate in the presence of [^{3}H]proline, and the newly synthesized prothrombin was isolated by immunoprecipitation. The immunoprecipitate (containing [^{3}H]pro-

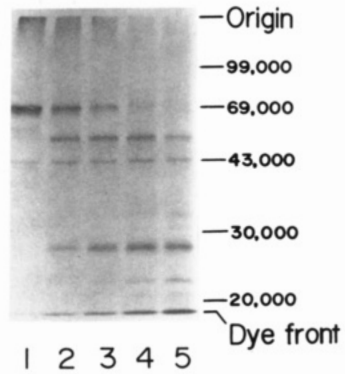

FIGURE 2: Thrombin digestion of [^{3}H]proline-labeled prepro-prothrombin. Enriched prothrombin poly(A) RNA was translated in vitro in the presence of [^{3}H]proline and the prepro-prothrombin isolated by immunoprecipitation with anti-prothrombin antibodies and carrier prothrombin (40 μg). The immunoprecipitate was solubilized with $NaDodSO_4$, diluted with 10 mM Tris buffer, pH 7.5, to a final $NaDodSO_4$ concentration of 0.02%, and divided into aliquots each containing 4 μg of carrier prothrombin. Increasing amounts of thrombin were added, and the reaction mixture was incubated at 37 °C for 10 min before lyophilization. Samples were analyzed by $NaDodSO_4$–polyacrylamide gel electrophoresis as described in the legend to Figure 1. Lane 1, no thrombin added; lane 2, 0.1 μg of thrombin; lane 3, 0.2 μg of thrombin; lane 4, 0.5 μg of thrombin; lane 5, 1 μg of thrombin. Protein standards are as in Figure 1 with the addition of phosphorylase (M_r 99 000). The positions of the origin and dye front are indicated.

thrombin, carrier prothrombin, and anti-prothrombin antibodies) was then incubated with increasing concentrations of thrombin, and the reaction products were analyzed by $NaDodSO_4$–polyacrylamide gel electrophoresis, followed by fluorography. Initially, the [^{3}H]prothrombin was present as a major single band of M_r 70 000 (Figure 2, lane 1). On addition of increasing amounts of thrombin, the prothrombin band disappeared with the concomitant appearance of two smaller bands of M_r 49 000 and 27 000 (lanes 2–5). The M_r 49 000 band was tentatively identified as intermediate I[2] (residues 157–582; M_r 48 143; Owen et al., 1974) and the M_r 27 000 band as fragment I (residues 1–156; M_r 17 973) containing an amino-terminal extension. To test this possibility, automatic sequenator analyses were then performed on fragment I. In these experiments, enriched prothrombin mRNA was translated in the presence of [^{3}H]leucine, and the radiolabeled product was recovered by immunoprecipitation. The immunoprecipitate was then dissolved in $NaDodSO_4$, diluted such that the $NaDodSO_4$ concentration was 0.02%, and digested with thrombin. Fragment I was isolated by preparative $NaDodSO_4$–polyacrylamide gel electrophoresis and subjected to 30 rounds of automatic Edman degradation. The PTH derivatives were identified and quantitated. In this analysis, radioactivity was released at rounds 8, 12, 14, 17, 20, and 28 (Figure 3, top panel). This procedure was repeated for fragment I labeled with [^{3}H]phenylalanine, [^{3}H]alanine, [^{3}H]proline, or [^{3}H]valine (Figure 3, bottom four panels). Sequenator analyses were also carried out with fragment I labeled with [^{3}H]tyrosine, [^{3}H]isoleucine, or [^{35}S]methionine, but these analyses released no radioactive peaks during the first 28 rounds. By use of the criteria of Palmiter et al. (1977), a partial amino-terminal sequence was deduced for the precursor for bovine prothrombin (Figure 4). These data suggest that bovine prothrombin is synthesized as a precursor having an amino-terminal extension of at least 30 residues. Since the partial sequence data did not overlap the amino-terminal sequence of prothrombin present in plasma, the exact size of the extension could not be deduced. It was rich, however, in hydrophobic residues (Leu, Phe, and Val) which is typical of signal sequences (Blobel et al., 1979). Final proof that the fragment I analyzed in these experiments originated from a prepro-prothrombin was established by cloning of the mRNA for bovine prothrombin as described below.

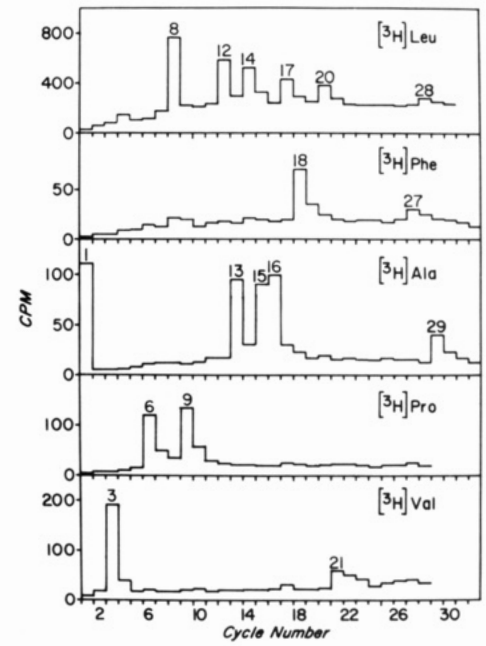

FIGURE 3: Automated Edman degradation of bovine prepro-prothrombin fragment I synthesized in five different reaction mixtures containing radiolabeled Leu, Phe, Ala, Pro, or Val. Numbers above the histograms indicate positive assignments in the sequence of prepro-prothrombin.

	-43			-40					-35				
Sequenator	X	Ala	X	Val	X	X	Pro	X	Leu	Pro	X	X	Leu
Predicted from cDNA	Met	Ala	Arg	Val	Arg	Gly	Pro	Arg	Leu	Pro	Gly	Cys	Leu

	-30					-25					-20		
Sequenator	Ala	Leu	Ala	Ala	Leu	Phe	X	Leu	Val	X	X	X	X
Predicted from cDNA	Ala	Leu	Ala	Ala	Leu	Phe	Ser	Leu	Val	His	Ser	Gln	His

			-15					-10					-5
Sequenator	X	Phe	Leu	Ala	-	-	-	-	-	-	-	-	-
Predicted from cDNA	Val	Phe	Leu	Ala	His	Gln	Gln	Ala	Ser	Ser	Leu	Leu	Gln

				-1	+1				+5			
Sequenator	-	-	-	-	-	-	-	-	-	-	-	-
Predicted from cDNA	Arg	Ala	Arg	Arg	Ala	Asn	Lys	Gly	Phe	Leu	Glu	Glu
Plasma prothrombin					Ala	Asn	Lys	Gly	Phe	Leu	Gla	Gla

FIGURE 4: Amino-terminal sequence of bovine prepro-prothrombin as determined by automatic Edman degradation of prepro-prothrombin fragment I. The sequenator data are from Figure 3. X represents positions where no amino acid could be assigned from the automatic sequenator analysis. The alanine residue at position −42 was released at cycle 1 of the sequenator analysis and is preceded by an initiator methionine (see text for details). No residues could be assigned past cycle 29 of the sequenator analysis (corresponding to the alanine at position −14). The amino acid sequences are numbered backward from the site of cleavage that gives rise to plasma prothrombin.

Construction and Screening of a Bovine Liver cDNA Library. Because of the difficulty of obtaining sufficient amounts of fragment I labeled with high specific activity, and because the sequenator analyses were providing little data past residue 30, it was decided to prepare a full-length complementary DNA of prothrombin mRNA. The complete amino acid sequence for the amino-terminal extension could then be predicted from the DNA sequence, and the sequenator analysis would predict precisely where the leader sequence was initiated.

A bovine liver cDNA library of 90 000 independent transformants was then constructed from 150 ng of sized DNA.

[2] Intermediate I is also referred to as prethrombin I.

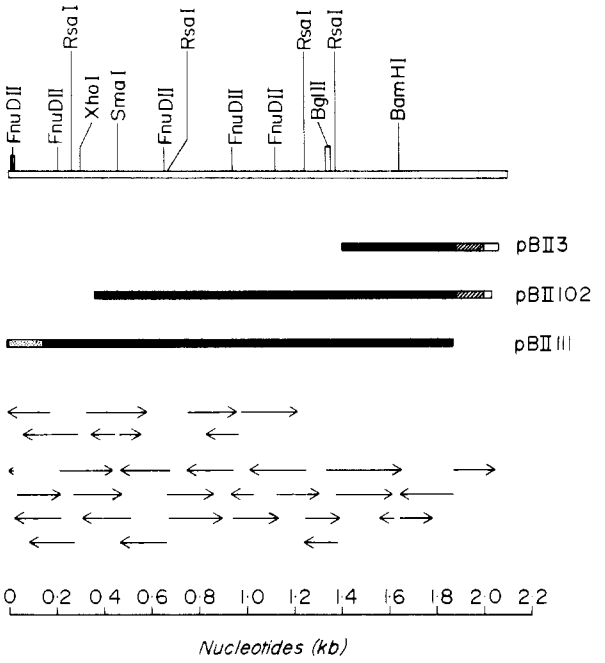

FIGURE 5: Restriction map and sequencing strategy for three bovine prothrombin cDNAs. The bars below the restriction map represent the three clones pBII3 (MacGillivray et al., 1980b), pBII102, and pBII111. The region coding for the leader sequence is represented by a dotted bar, the coding region by a solid bar, the 3′-noncoding region by a slashed bar, and the poly(A) tail by an open bar. Plasmid pBII111 contains 5 bp of 5′-noncoding sequence. The extent of sequencing is indicated by the length of the arrow. DNA sequence determined on the coding strand is indicated by an arrow pointing right; sequence determined on the noncoding strand is indicated by an arrow pointing left. The top two rows of arrows indicate DNA sequence determined by the chemical cleavage method; the bottom four rows indicate sequence determined by the chain-terminator method. The scale shown at the bottom represents nucleotides in kilobases.

In order to maximize the synthesis of full-length cDNAs for this library, human placental ribonuclease inhibitor was included in the reverse transcription reaction. After base hydrolysis of the mRNA, the cDNA was tailed with dCMP, and the second strand synthesis was primed with oligo(dG) (Land et al., 1981). The double-stranded cDNA was again tailed with dCMP and sized by agarose gel electrophoresis prior to insertion into the *Pst*I site of pBR322 by homopolymeric G·C tailing. Approximately 40 000 recombinant colonies from the cDNA library were screened by in situ hybridization using nick-translated *Pst*I insert of pBII3 as a probe (MacGillivray et al., 1980b). About 150 positive colonies were obtained, and seven (pBII106–112) were studied in detail. Restriction endonuclease digestion of plasmids isolated from these colonies showed that they contained bovine prothrombin cDNA inserts of 1100, 640, 1600, 1400, 1300, 1950, and 1000 bp, respectively. Plasmid pBII111 containing the largest cDNA insert was subjected to restriction mapping (Figure 5).

DNA Sequence Analysis of Bovine Prothrombin cDNAs. Plasmids pBII102 and pBII111 were both employed for determining the nucleotide sequence of prothrombin cDNA. The sequencing strategy used is summarized at the bottom of Figure 5. Initially, the DNA sequences around the *Pst*I, *Xho*I, *Sma*I, and *Bgl*II sites were determined by the chemical cleavage method. Extensive sequence data were then determined by the chain terminator method by analyzing small fragments cloned into the single-stranded phage M13. The final overlaps were then determined by sequencing around three *Dde*I sites by the chemical cleavage method. Each nucleotide was determined at least twice, and 73% of the sequence was determined on both strands. Every labeling/cloning site was overlapped, with the exception of the *Bam*HI site (nucleotides 1636–1641). However, the sequence of this region has already been determined in pBII3 (MacGillivray et al., 1980b).

The complete sequence of prothrombin cDNA and the predicted amino acid sequence for the protein are shown in Figure 6. Nucleotides 135–1880 code for residues 1–582 of plasma prothrombin. The coding region is followed by a stop codon TAG (nucleotides 1881–1883) and a 3′-untranslated region of 121 nucleotides (1884–2005). This is two nucleotides longer than the 3′-untranslated region of pBII3 (MacGillivray et al., 1980b). Plasmid pBII102 also contains a poly(A) tract of 30 nucleotides. The cDNA in pBII111 corresponded to nucleotides 1–1877 and terminated at Gly-581. Nucleotides 6–134 code for an amino-terminal extension of 43 amino acid residues. An alignment of the predicted sequence of the extension with the partial sequence of prepro-prothrombin determined by in vitro translation is shown in Figure 4. The partial amino acid sequence agrees well with that predicted from the cDNA sequence. Only the valine residue at −17 was not detected during the sequenator analyses. In addition, amino acids that gave no positive cycles during the sequenator analyses (isoleucine, tyrosine, and methionine) were not present in the first 30 residues of the prepro-prothrombin sequence predicted from the cDNA. As the reticulocyte lysate has a methinone aminopeptidase that removes methionyl residues in the sequence Met-Ala (Thibodeau et al., 1979), it is concluded that the methionine at position −43 is the initiator methionine. Accordingly, pBII111 contains five nucleotides of 5′-untranslated sequence.

Discussion

From the present data, it is clear that bovine prothrombin is synthesized as a precursor having an amino-terminal leader extension of M_r ~4500. Partial amino-terminal sequence analysis indicated that this leader sequence was hydrophobic in nature and corresponded to the signal sequence found in most other secreted proteins (Blobel et al., 1979). The complete amino acid sequence of the precursor was predicted from the cDNA sequence, and this confirmed the partial sequence data obtained from fragment I synthesized in vitro (Figure 4). Conversion of the precursor to plasma prothrombin occurs by cleavage of the peptide bond between Arg^{-1}–Ala^{+1}, where Ala^{+1} represents the amino-terminal residue of plasma prothrombin. This cleavage is not typical of signal peptidase which has an elastase-like specificity (Blobel et al., 1979). This suggests that the newly synthesized bovine prothrombin contains a prepro leader sequence that is cleaved first by the signal peptidase to yield pro-prothrombin, and this molecule is then cleaved by a second protease to yield plasma prothrombin. This biosynthetic pathway would be analogous to several other plasma proteins synthesized in liver, including albumin (Strauss et al., 1977; MacGillivray et al., 1979; Lawn et al., 1981), factor IX (Kurachi & Davie, 1982), and human prothrombin (Degen et al., 1983). It is also consistent with the presence of prothrombin precursors found in rat liver microsomal preparations (Suttie & Jackson, 1977).

The position of the putative signal peptidase cleavage site in prepro-prothrombin and any possible functions of the pro peptide in pro-prothrombin are unknown at present. However, it might be expected that the sequences of functional regions of the precursor peptides may be conserved. Figure 7 shows a comparison of the leader sequence for bovine prepro-prothrombin, human prepro-prothrombin (Degen et al., 1983), and human prepro-factor IX (Kurachi & Davie, 1982; Jaye

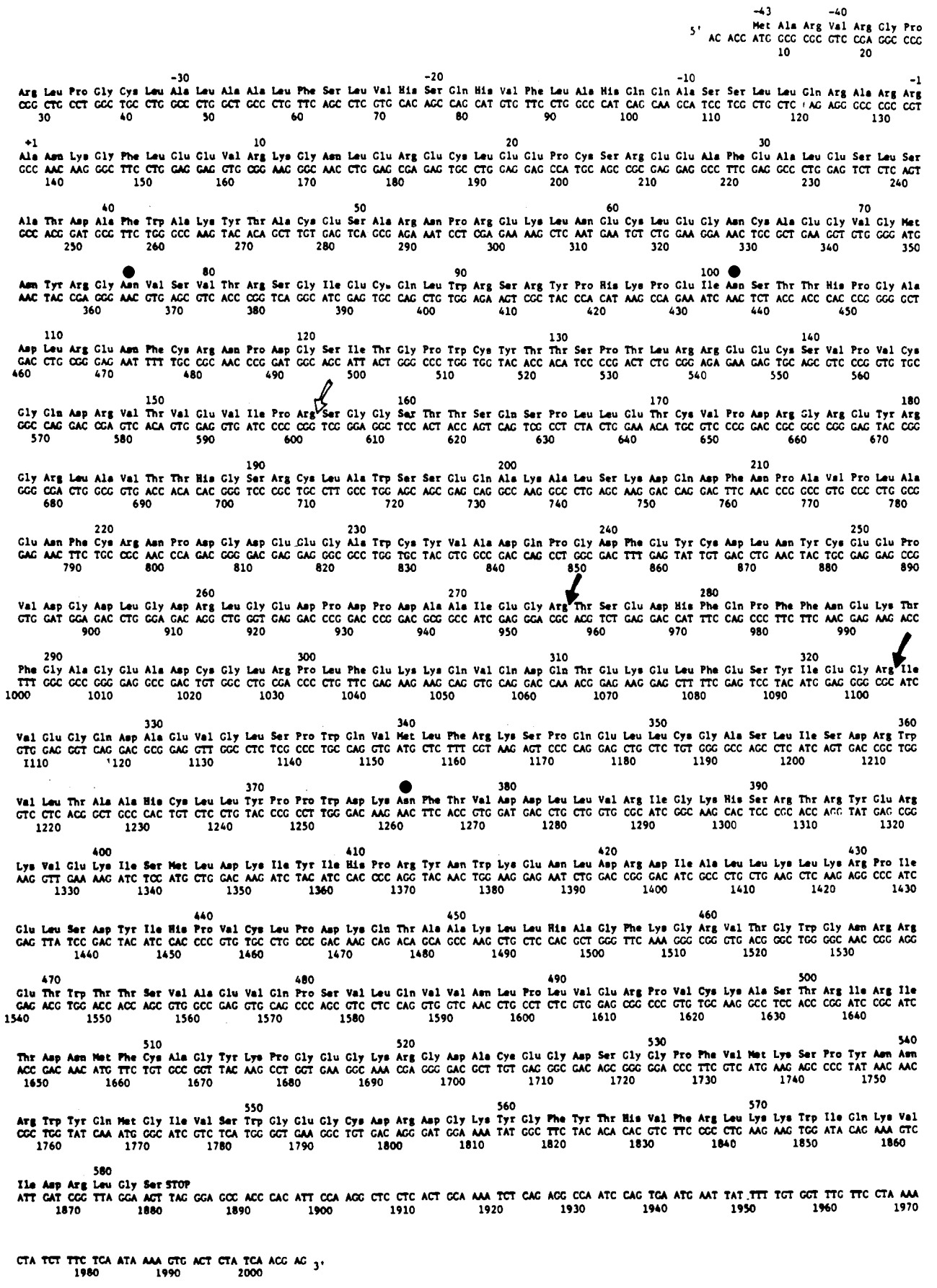

FIGURE 6: Complete nucleotide sequence of bovine prothrombin cDNA. The sequence was determined by analysis of the three overlapping clones shown in Figure 4. The predicted amino acid sequence of bovine prepro-prothrombin is shown above the DNA sequence. Residues −43 to −1 represent the leader sequence for bovine prothrombin, and residues 1–582 represent the sequence of plasma prothrombin. The factor Xa cleavage sites are shown by the solid arrows, and the thrombin cleavage site is shown by the open arrow. The attachment sites for carbohydrate on Asn-77, -101, and -376 are indicated by (♦).

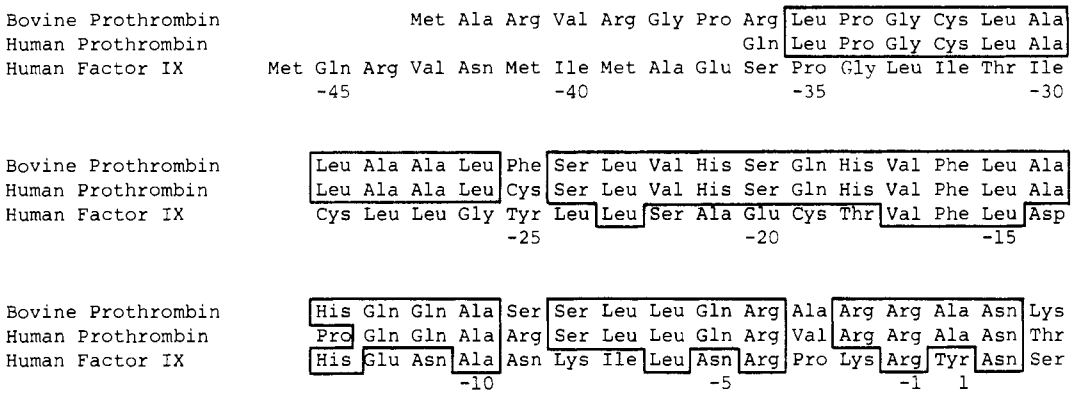

FIGURE 7: Comparison of the leader sequences of bovine prothrombin, human prothrombin (Degen et al., 1983), and human factor IX (Kurachi & Davie, 1982; Jaye et al., 1983; K. Kurachi and E. W. Davie, unpublished results) as predicted from the cDNA sequences. Identical residues in corresponding positions are shown in boxes. The sequences are numbered backward from the site of cleavage that gives rise to the mature proteins found in plasma. The recombinant plasmid encoding human prepro-prothrombin contained only 36 residues of the leader sequence. The exact position of the initiator methionine residue for factor IX is not known.

et al., 1983). Considerable sequence homology is evident between the bovine and human leader sequences for prothrombin where 31 out of 36 residues in corresponding positions are identical. Also, there is some homology between the prothrombin leader sequences and the factor IX leader sequence, particularly at the carboxyl end where 7 of the last 17 residues are identical in all three proteins. However, the amino-terminal region of the leader sequence shows no homology between all three proteins. This is consistent with signal peptides being nonhomologous for proteins synthesized in the same tissue (Compere et al., 1981). However, if the prepro leader sequence of prothrombin and factor IX play a common role in the biosynthetic pathway of these two proteins, the signal peptidase cleavage site may be conserved in all three sequences. Since this enzyme has an elastase-like specificity, the most probable cleavage site would be between alanine at position −10 and the adjacent amino acid at position −9. This would leave a pro peptide of nine residues on each of the three proteins.

There are seven differences between the amino acid sequence of bovine prothrombin predicted from the cDNA of pBII102 and that determined by classical protein sequence analysis employing prothrombin prepared from pooled plasma (Magnusson et al., 1975). These include His-188, Asp-245, Gln-310, Glu-312, Asp-506, Asn-507, and Asp-557 which were reported to be Ser, Asn, Glu, Gln, Asn, Asp, and Asn, respectively. Five of the last six differences very likely reflect errors in amide assignments. The codon for residue 557, however, was found to be AAU coding for Asn in pBII3 (MacGillivray et al., 1980b) and GAU coding for Asp in pBII102. Thus, polymorphism or errors by reverse transcriptase during the cDNA synthesis cannot be excluded. The evidence for Ser in position 188 as determined by amino acid sequence analysis is convincing (Staffan Magnusson, personal communication), also indicating the likelihood of polymorphism or a cloning artifact at this position in prothrombin.

The amino acid composition of bovine prothrombin predicted from the cDNA sequence of pBII102 and pBII111 is Asn_{23}, Asp_{37}, Thr_{29}, Ser_{35}, Glu_{43}, Gla_{10}, Gln_{18}, Pro_{35}, Gly_{48}, Ala_{34}, Cys_{24}, Val_{35}, Met_{6}, Ile_{20}, Leu_{46}, Tyr_{19}, Phe_{20}, Lys_{31}, His_{10}, Arg_{45}, and Trp_{14}. This amino acid composition includes the 10 γ-carboxyglutamic acid residues determined by amino acid sequence analysis (Magnusson et al., 1975). This composition corresponds to a molecular weight of 66 226 in the absence of carbohydrate and 73 585 with the addition of 10% carbohydrate (DiScipio et al., 1977). The carbohydrate attachment sites were established previously as Asn-77, -101, and -376 in bovine prothrombin (Magnusson et al., 1975).

Table I: Codon Usage for Bovine and Human[a] Prothrombin mRNA

		bovine	human			bovine	human
Ala	GCG	6	3	Lys	AAG	27	26
	GCA	1	5		AAA	4	3
	GCT	6	8	Met	ATG	6	8
	GCC	21	20	Phe	TTT	4	5
Arg	AGG	6	9		TTC	16	15
	AGA	3	1	Pro	CCG	10	3
	CGG	14	9		CCA	4	4
	CGA	7	6		CCT	6	12
	CGT	1	1		CCC	15	12
	CGC	14	13	Ser	AGT	6	8
Asn	AAT	4	4		AGC	12	9
	AAC	19	22		TCG	3	3
Asp	GAT	6	12		TCA	3	6
	GAC	31	23		TCT	3	
Cys	TGT	9	12		TCC	8	9
	TGC	15	12	Thr	ACG	6	8
Gln	CAG	16	19		ACA	7	9
	CAA	2	2		ACT	3	6
Glu	GAG	43	39		ACC	13	12
	GAA	10	12	Trp	TGG	14	14
Gly	GGG	14	19	Tyr	TAT	5	4
	GGA	8	6		TAC	14	17
	GGT	6	7	Val	GTG	22	21
	GGC	20	15		GTA		1
His	CAT	2	5		GTT	2	2
	CAC	8	5		GTC	11	9
Ile	ATA	1	2	stop	TGA		
	ATT	2	8		TAG	1	1
	ATC	17	12		TAA		
Leu	TTG		4				
	TTA	2					
	CTG	29	25				
	CTA	1	3				
	CTT	2	3				
	CTC	12	7				

[a] Data taken from Degen et al. (1983).

The base composition of prothrombin mRNA is somewhat G·C rich (21% A, 32.6% G, 29.6% C, and 16.8% T). This is reflective of the coding region where 80% of the codons end in either G or C. Codon usage in bovine prothrombin is nonrandom, as shown in Table I. Two codons are not used (UUG for leucine and GUA for valine), and some are used rarely (17 of 20 isoleucine residues are encoded by AUC). Again, this mainly reflects the use of G and C in the third position. Codon usage for bovine and human prothrombin is very similar (Table I), and all γ-carboxyglutamic acid residues are encoded by GAG in each protein. The γ-carboxyglutamic

acid residues in factor IX, however, are encoded mainly by GAA (Kurachi & Davie, 1982; Jaye et al., 1983). The significance of this difference in codon usage is unclear.

Acknowledgments

We thank our colleagues for helpful suggestions throughout this study, including Drs. Dominic Chung, Richard Palmiter, Walter Kisiel, Douglas Malinowski, Joan McPherson, Mark Zoller, and Rob McMaster. We also thank Bill Kimmerly for confirming the DNA sequence of several M13 clones and Rob McMaster and Joan McPherson for critically reading the manuscript.

Registry No. Prothrombin, 9001-26-7; prepro blood coagulation factor II, 88980-83-0; prepro blood coagulation factor II (ox. protein moiety reduced), 88980-84-1; blood coagulation factor II (ox. protein moiety reduced), 88980-82-9; thrombin (ox. A chain), 88980-85-2; thrombin (ox. B-chain protein moiety reduced), 88980-86-3; deoxyribonucleic acid (ox. thrombin-specifying messenger RNA complementary), 88980-90-9; thrombin, 9002-04-4.

References

Anderson, G. F., & Barnhart, M. I. (1964) *Proc. Soc. Exp. Biol. Med. 116*, 1–6.

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.

Besmond, C., Benarous, R., & Kahn, A. (1981) *Biochem. Biophys. Res. Commun. 103*, 587–594.

Blobel, G., Walter, P., Chang, C. N., Goldman, B. M., Erickson, A. H., & Lingappa, R. (1979) in *Secretory Mechanisms* (Hopkin, C. R., & Duncan, C. J., Eds.) Vol. 33, pp 9–36, Cambridge University Press, London.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H., & Falkow, S. (1977) *Gene 2*, 95–113.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Compere, S. J., Lively, M. O., & MacGillivray, R. T. A. (1981) *Eur. J. Biochem. 116*, 437–440.

Cuatrecasas, P. (1970) *J. Biol. Chem. 245*, 3059–3065.

Dagert, M., & Ehrlich, S. D. (1979) *Gene 6*, 23–28.

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

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

Delaney, A. (1982) *Nucleic Acids Res. 10*, 61–67.

Deng, G., & Wu, R. (1981) *Nucleic Acids Res. 9*, 4173–4188.

DiScipio, R. G., Hermodson, M. A., Yates, S. G., & Davie, E. W. (1977) *Biochemistry 16*, 698–706.

Friedman, P. A., & Shia, M. (1976) *Biochem. Biophys. Res. Commun. 70*, 647–654.

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

Grunstein, M., & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3961–3965.

Hanahan, D., & Meselson, M. (1980) *Gene 10*, 63–67.

Helgeland, L. (1977) *Biochim. Biophys. Acta 499*, 181–193.

Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., & Walsh, K. A. (1972) *Biochemistry 11*, 4493–4502.

Jaye, M., de la Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, P. T., & Lecocq, J.-P. (1983) *Nucleic Acids Res. 11*, 2325–2335.

Kafatos, F. C., Jones, C. W., & Efstratiadis, A. (1979) *Nucleic Acids Res. 7*, 1541–1552.

Katz, L., Kingsbury, D. T., & Helinski, D. R. (1973) *J. Bacteriol. 114*, 577–591.

Katz, L., Williams, P. H., Sato, S., Leavitt, R. W., & Helinski, D. R. (1977) *Biochemistry 16*, 1677–1683.

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.

Land, H., Grez, M., Hauser, H., Lindenmaier, W., & Schultz, G. (1981) *Nucleic Acids Res. 9*, 2251–2266.

Lawn, R. M., Adelman, J., Bock, S. C., Franke, A. E., Houck, C. M., Najarian, R. C., Seeburg, P. H., & Wion, K. L. (1981) *Nucleic Acids Res. 9*, 6103–6114.

MacGillivray, R. T. A., Chung, D. W., & Davie, E. W. (1979) *Eur. J. Biochem. 98*, 477–485.

MacGillivray, R. T. A., Chung, D. W., & Davie, E. W. (1980a) in *Vitamin K Metabolism and Vitamin K-dependent Proteins* (Suttie, J. W., Ed.) pp 546–552, University Park Press, Baltimore.

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

Magnusson, S., Peterson, T. E., Sottrup-Jensen, L., & Claeys, H. (1975) in *Proteases and Biological Control* (Reich, E., Rifkin, D. B., & Shaw, E., Eds.) pp 123–149, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.

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

Martynoff, G., Pays, E., & Vassart, G. (1980) *Biochem. Biophys. Res. Commun. 93*, 645–653.

Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Messing, J., & Vieira, J. (1982) *Gene 19*, 269–276.

Messing, J., Crea, R., & Seeburg, P. H. (1981) *Nucleic Acids Res. 9*, 309–321.

Nardacci, N. J., Jones, J. P., Hall, A. L., & Olsen, R. E. (1975) *Biochem. Biophys. Res. Commun. 64*, 51–58.

Nelsestuen, G., Zytkovicz, T. H., & Howard, J. B. (1974) *J. Biol. Chem. 249*, 6347–6350.

Owen, W. G., Esmon, C. T., & Jackson, C. M. (1974) *J. Biol. Chem. 249*, 594–605.

Palacios, R., Palmiter, R. D., & Schimke, R. T. (1972) *J. Biol. Chem. 247*, 2316–2321.

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

Palmiter, R. D., Gagnon, J., Ericsson, L. H., & Walsh, K. A. (1977) *J. Biol. Chem. 255*, 6386–6393.

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

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

Sadowski, J. A., Esmon, C. T., & Suttie, J. W. (1976) *J. Biol. Chem. 251*, 2770–2775.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F., & Petersen, G. B. (1982) *J. Mol. Biol. 162*, 729–773.

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

Stenflo, J., & Suttie, J. W. (1977) *Annu. Rev. Biochem. 46*, 157–172.

Stenflo, J., Fernlund, P., Egan, W., & Roepstorff, P. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2730-2733.
Strauss, A. W., Bennett, C. D., Donohue, A. M., Rodkey, J. A., & Alberts, A. W. (1977) *J. Biol. Chem. 252*, 6846-6855.
Suttie, J. W., & Jackson, C. M. (1977) *Physiol. Rev. 57*, 1-70.
Thibodeau, S. N., Lee, D. C., & Palmiter, R. D. (1978) *J. Biol. Chem. 253*, 3771-3774.
Winter, G., & Fields, S. (1980) *Nucleic Acids Res. 8*, 1965-1974.

Sequence Specificity of Heat-Labile Sites in DNA Induced by Mitomycin C†

Kazumitsu Ueda, Junji Morita,‡ and Tohru Komano*

ABSTRACT: The sequence specificity of the mitomycin C–DNA interaction was directly determined by using DNA sequencing techniques and by using 3′- or 5′-end-labeled DNA fragments of defined sequence as substrates. Mitomycin C reduced with sodium borohydride induced heat-labile sites in DNA preferentially at specific sequences. The heat-labile sites were induced most preferentially at the dinucleotide sequence G-T (especilly Pu G-T), which was determined by scanning autoradiograms with a microdensitometer after gel electrophoresis. DNA was cleaved at the 3′ side of deoxyguanosines and of some deoxyadenosines by heat treatment. Oligonucleotides produced by heat treatment after reaction with reduced mitomycin C contained phosphoryl groups at the 5′ termini. The 3′ termini seemed not to have simple structures, judging from their electrophoretic mobilities. Oxygen radicals such as singlet oxygen and hydroxyl radical were possibly involved in the induction of heat-labile sites.

Mitomycin C, a potent anticarcinogenic antibiotic, interacts with DNA, resulting in a covalent binding of the drug to DNA, as well as in the formation of cross-links between the complementary strands of DNA (Iyer & Szybalski, 1963, 1964; Matsumoto & Lark, 1963). These DNA modifications, in which the former is predominant 10–20-fold over the latter (Szybalski & Iyer, 1964a), are believed to be essential for the cytotoxicity of mitomycin C (Iyer & Szybalski, 1963; Weiss et al., 1968, Kinoshita et al., 1971; Mercado & Tomasz, 1972). The aziridine and methylurethane moieties are suggested to be involved in the binding to DNA (Schwartz et al., 1963; Iyer & Szybalski, 1964; Hashimoto et al., 1982). The binding sites of mitomycin C in DNA are the O-6 position or the 2-amino group of guanine residues or the 6-amino group of adenine residues (Tomasz et al., 1983; Hashimoto et al., 1982). However, the details of the interaction of mitomycin C with DNA have yet to be elucidated.

Mitomycin C contains a quinone moiety beside aziridine and methylurethane. Reduction of mitomycin C, by chemical or enzymatic methods, followed by exposure to air results in the generation of superoxide anion and hydrogen peroxide (Handa & Sato, 1975; Tomasz, 1976). Oxygen radicals were generated not only by free mitomycin C but also by mitomycin C irreversibly bound to DNA (Tomasz, 1976). Lown et al. (1976) and Ueda et al. (1980, 1981, 1982) reported that chemically reduced mitomycin C induces single-strand scission in single-stranded and double-stranded DNAs. The DNA strand scission is considered to involve the oxygen radicals, such as hydroxyl radical and singlet oxygen, and mitomycin C semiquinone radical (Lown et al., 1976; Ueda et al., 1980, 1981, 1982).

DNA cleavage via a mechanism involving oxygen radicals is reported for some antitumor antibiotics such as bleomycin (Lown & Sim, 1977; Sausville et al., 1976). Strand scission by bleomycin occurs preferentially at specific sequences (Takeshita et al., 1978; D'Andrea & Haseltine, 1978) and at specific sites in DNA (Lloyd et al., 1978), and the sequence specificity of single-strand scission is related to the site-specific double-strand scission by bleomycin (Mirabelli et al., 1982).

We have investigated the interaction of mitomycin C with DNA by using DNA substrates of defined sequence. In this paper, we show that reduced mitomycin C induces heat-labile sites in DNA preferentially at specific sequences and that oxygen radicals are possibly involved in the induction of heat-labile sites.

Materials and Methods

Chemicals and Enzymes. Mitomycin C was kindly supplied by Kyowa Hakko Co. Ltd., Tokyo, Japan. The restriction enzymes *Hae*III, *Taq*I, and *Hin*fI and T4 polynucleotide kinase were obtained from Takara Shuzo Co. Ltd., the Klenow fragment of DNA polymerase I of *Escherichia coli* was from Bethesda Research Laboratories GmbH, and calf intestine alkaline phosphatase was from Boehringer Mannheim GmbH. [α-^{32}P]dTTP, [α-^{32}P]dCTP, and [γ-^{32}P]ATP (specific activity about 3000 Ci/mmol) were purchased from New England Nuclear, Du Pont, and Amersham International.

DNA Substrates. Three DNA fragments of defined sequence were obtained from bacteriophage ϕX174 replicative form DNA. Double-stranded ϕX174 replicative form DNA was prepared as previously described (Ueda et al., 1981) and digested with *Hae*III, and 194 and 234 base pair fragments [Z_7 and Z_8 fragments in the map reported by Sanger (Sanger et al., 1978)] were purified. Fragment Z_7 was digested with *Taq*I and was labeled by extension of the 3′ termini with Klenow polymerase in the presence of [α-^{32}P]dCTP (Maniatis et al., 1982). Fragment Z_8 was digested with *Hin*fI and labeled at the 3′ termini in the presence of [α-^{32}P]dTTP and unlabeled

† From the Laboratory of Biochemistry, Department of Agricultural Chemistry, Kyoto University, Kyoto 606, Japan. *Received June 23, 1983.* This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.
‡ Present address: Laboratory of Food Chemistry, Department of Food Science, Doshisha Women's College of Liberal Art, Kyoto 602, Japan.

0006-2960/84/0423-1634$01.50/0 © 1984 American Chemical Society