Primary Structure of Human Placental Anticoagulant Protein†

Takayuki Funakoshi, Lee E. Hendrickson, Brad A. McMullen, and Kazuo Fujikawa*
Department of Biochemistry, University of Washington, Seattle, Washington 98195
Received July 2, 1987; Revised Manuscript Received August 10, 1987

ABSTRACT: The primary structure of human placental anticoagulant protein was determined by a combination of amino acid and nucleotide sequencing techniques. The carboxymethylated protein was digested with cyanogen bromide, and the resulting peptides were separated by gel filtration and high-performance liquid chromatography. A total of 239 out of 319 amino acid residues were identified from 7 cyanogen bromide fragments. A full-length cDNA clone encoding placental anticoagulant protein was isolated from a human placenta cDNA library. This clone was 1.6 kilobases long and contained a translation initiation site coding for methionine, 957 nucleotides encoding for the mature protein, a stop codon, a poly(A) recognition site, and a poly(A) tail. Analysis of the tryptic-blocked peptide that originated from the NH_2-terminus of the protein showed that the terminal methionine was removed and the adjacent alanine residue was acetylated by posttranslational events. Placental anticoagulant protein is composed of 319 amino acids with acetylalanine as the NH_2-terminus and has a high degree of sequence identity with lipocortins I and II. It contains four internal repeats, each including a sequence corresponding to a putative Ca^{2+}-dependent phospholipid binding site. Placental anticoagulant protein is a member of the lipocortin/calpactin family.

Recently, we isolated an anticoagulant protein, called placental anticoagulant protein (PAP),[1] from human placenta, and we studied its inhibition mechanism and partial amino acid sequence (Funakoshi et al., 1987). PAP is a soluble cellular protein with a molecular weight of 36 500 and has no carbohydrate. It inhibits prothrombinase activity in a dose-dependent manner. It also inhibits the activation of factor IX and factor X by the factor VIIa/tissue factor complex (Kondo et al., 1987). PAP is capable of binding to acidic phospholipid vesicles in the presence of Ca^{2+}. Thus, the inhibitory effect of PAP is thought to occur through competition with vitamin K dependent factors for binding to phospholipid vesicles. Similar proteins have been isolated from human liver (Deykin et al., 1969), placenta (Maki et al., 1984), and umbilical cord (Reutelingsperger et al., 1985).

Partial amino acid sequence analysis of PAP showed that it had a high degree of sequence identity with lipocortins I and II. The lipocortins belong to a group of Ca^{2+}-dependent membrane binding proteins, which are distributed widely in cells and tissues (Kretsinger & Creutz, 1986). Although the physiological functions of these proteins have not been clearly defined, they have been implicated in exocytosis and membrane cytoskeleton interaction (Moore & Dedman, 1982; Creutz et al., 1983; Gerke & Weber, 1984; Sudhof et al., 1984; Geisow et al., 1984; Shadle et al., 1985). Lipocortins I and II bind to phospholipid vesicles in a Ca^{2+}-dependent manner and inhibit hydrolysis of phospholipid by phospholipase A_2 (Schlaepfer & Haigler, 1987; Davidson et al., 1987). This inhibition is thought to suppress prostaglandin synthesis leading to antiinflammatory action (Flower & Blackwell, 1979; Blackwell et al., 1980; Wallner et al., 1986).

Thus far, four of these proteins, human lipocortins I and II (Huang et al., 1986; Wallner et al., 1986), the murine protein-tyrosine kinase substrate p36[2] (Saris et al., 1986), and bovine calpactin I heavy chain (p36) (Kristensen et al., 1986), have been completely sequenced. In this paper, we report the complete primary structure of PAP and discuss the functional and structural relationships of PAP and the human lipocortins.

Experimental Procedures

PAP was purified from fresh human placenta as described previously (Funakoshi et al., 1987), and its concentration was determined by using $E^{1\%}_{280nm} = 6.0$. A λgt11 cDNA library prepared from human placenta poly(A) mRNA was purchased from Clontech. *N*-Formyl-L-leucine, *N*-acetyl-DL-leucine, dansyl chloride, and precoated 100% C18-silanized silica gel plates were purchased from Sigma. Hybridization probe primer was obtained from Pharmacia. Synthetic oligonucleotides were kindly prepared by Dr. Patrick O'Hara, Zymogenetics, Seattle, WA. DNA sequencing kit was obtained from New England Biolabs. Anhydrous hydrazine was prepared by distillation and stored in sealed glass tubes until use.

The protein was S-carboxymethylated by the method of Crestfield et al. (1963) with a minor modification. CM-PAP, 6.7 mg, was digested overnight at room temperature in 0.5 mL of 70% formic acid containing 2% CNBr, and the resulting peptides were first fractionated by gel filtration on a Sephadex G-50 superfine column (0.9 × 50 cm) with 4 M guanidine hydrochloride. Each subfraction was further separated by a Waters HPLC system using an Altex Ultrapore C3 reverse-phase separation column (0.46 × 7.5 cm). Tryptic digestion was performed by incubating peptides at 37 °C in 0.1 M NH_4HCO_3 for 6 h with 1% (w/w) chymotrypsin-free bovine trypsin (Fujikawa & McMullen, 1985), and the resulting peptides were separated by HPLC using a Waters μBondapak C18 column (0.49 × 30 cm). A gradient system composed of 0.1% trifluoroacetic acid (solvent A) and 0.08% trifluoroacetic acid in 80% acetonitrile (solvent B) was used for elution

†This work was supported by Research Grant HL 16919 from the National Institutes of Health.
*Address correspondence to this author.

[1] Abbreviations: CM, carboxymethyl; CNBr, cyanogen bromide; EGF, epidermal growth factor; PAP, placental anticoagulant protein; HPLC, high-performance liquid chromatography; kb, kilobase(s); SDS, sodium dodecyl sulfate; kDa, kilodalton(s).

[2] The murine protein kinase substrate p36 and bovine calpactin I heavy chain are analogous to lipocortin II.

0006-2960/87/0426-8087$01.50/0

of both the C3 and C18 columns at a flow rate of 1.5 mL/min. The effluents were monitored at 214 nm, and peptides were collected manually.

Amino acid sequence analysis of peptides was performed by a Beckman Sequencer Model 890C according to the method of Edman and Begg (1967). Phenylthiohydantoin–amino acids were identified by two complementary HPLC systems (Bridgen et al., 1976; Ericsson et al., 1977). Amino acid composition of peptide hydrolysates (6 N HCl, 110 °C for 24 h) was determined by a Waters Picotag system (Bidlingmeyer et al., 1984).

Detection of the acetyl group was performed by the method of Schmer and Kreil (1969). The blocked peptide, CNBr-1-T1, was heated at 110 °C for 5 h with anhydrous hydrazine in a sealed tube. Excess hydrazine was removed in vacuo over concentrated sulfuric acid. Hydrazides were dissolved in 0.3 mL of 0.1 M citrate, pH 3.0, and treated at 37 °C overnight with 0.3 mL of ethanol containing dansyl chloride at a 5–10-fold molar excess to peptides. Dansyl hydrazides were developed on silanized silica gel plates (5 × 10 cm) by two different solvent systems: 1-butanol/acetic acid/water (4/1/1 v/v/v) and chloroform/1-butanol/acetic acid (6/3/1 v/v/v).

Affinity-purified anti-PAP was prepared as previously described (Funakoshi et al., 1987), and 0.2 mg of the antibody was radiolabeled with 0.5 mCi of $Na^{125}I$ (19.6 mCi/μg of iodine, Amersham) using Iodo-gen (Pierce). The specific activity of the antibody was 1×10^6 cpm/μg. Phage (5×10^5) from the cDNA library were screened by affinity-purified ^{125}I-anti-PAP according to Young and Davis (1983) as modified by Foster and Davie (1984). The same library was also screened by using a recombinant phage M13 template containing a 1.5-kb PAP–cDNA insert and an M13 template containing a 59-base *Pst*I fragment from the 5′ end of the 1.5-kb insert subcloned into plasmid pUC18. These hybridization probes were labeled by employing a synthetic hybridization probe primer by the method of Hu and Messing (1982). Screening was performed by the method of Hu and Messing (1982). Filters were washed for 1 h at 60 °C with 2 × SSC buffer (8.2 g of sodium citrate, pH 7.0, and 17.5 g of NaCl/L) containing 0.5% SDS. Positive clones obtained by screening with antibody and cDNA were amplified by the liquid lysis method (Maniatis et al., 1982), and the phage were purified by CsCl banding. Restriction fragments from isolated inserts were cloned into either M13mp18 or M13mp19 vector for sequence analysis by the dideoxy chain termination method (Sanger et al., 1977). Two clones, 1.6 (λHPAP1.6) and 1.5 kb (λHPAP1.5), were subcloned into plasmid vector pUC18 for completion of the sequence. Some of the sequences were determined by using complementary synthetic 17-base oligonucleotides corresponding to nucleotides of the 1.5-kb insert as sequencing primers.

Results

Separation and Sequence of CNBr Peptides of PAP. As observed previously (Funakoshi et al., 1987), sequence analysis of intact CM-PAP failed to detect any sequence, indicating that the NH_2-group of the NH_2-terminal residue of PAP was blocked. Thus, sequence analysis of PAP was initiated with CNBr fragments. CNBr peptides were initially separated into three fractions by gel filtration on a Sephadex G-50 column (Figure 1). Three fractions, Fr-I, -II, and -III, were then individually subjected to the C3 HPLC column for further separation (Figure 2). Four homogeneous peptides, CNBr-5, -7, -1, and -6, were obtained from Fr-III (left panel). Two major peaks, CNBr-3 and -2, were obtained from Fr-II (center panel). One homogeneous peptide, CNBr-4, was also obtained

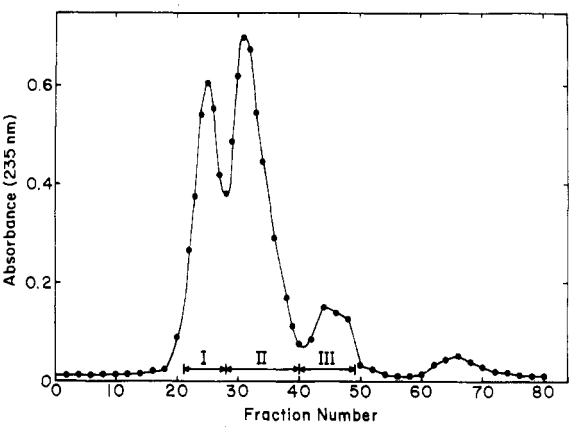

FIGURE 1: Gel filtration of CNBr digest of CM-PAP. The CNBr digest was applied to a column of Sephadex G-50, superfine, and peptides were eluted with 4 M guanidine hydrochloride. A 1.0-mL fraction was collected in each tube. Three fractions, Fr-I, -II, and -III, were pooled and further purified by HPLC (see Figure 2).

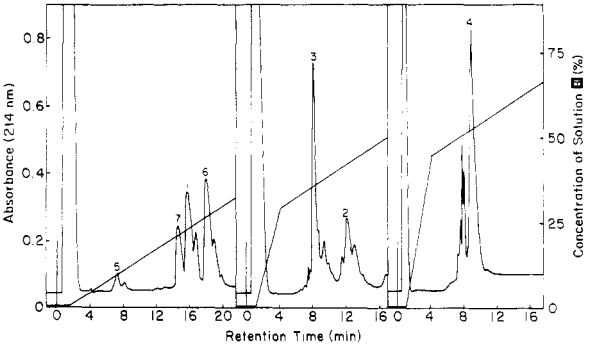

FIGURE 2: Separation of CNBr peptides by HPLC. Three CNBr fractions obtained from the gel filtration column were further separated by the C3 HPLC column. Peptides were eluted as described under Experimental Procedures. Separations of Fr-III (left panel), Fr-II (center panel), and Fr-I (right panel) are shown.

from Fr-I (right panel). The second peak, CNBr-2, from Fr-II was further purified by HPLC using the C18 column before sequence analysis. In all, seven homogeneous peptides were obtained and sequenced.

No sequence was found in CNBr-1, indicating that this peptide originated from the NH_2-terminus. This peptide was then further digested with trypsin, and three tryptic peptides (CNBr-1-T1, CNBr-1-T2, and CNBr-1-T3) were separated. Peptides CNBr-1-T2 and CNBr-1-T3 were sequenced, but the peptide CNBr-1-T1 was found to be blocked. The amino acid composition of this peptide is the same as the NH_2-terminal five residues deduced from the nucleotide sequence of the PAP clone λHPAP1.6 (see below). The sequences of the 239 residues obtained from the seven CNBr peptides and their subpeptides are shown in Table I.

Isolation and Characterization of cDNAs Cloned for PAP. Twenty-two positive clones were isolated by screening a human placenta cDNA library (5×10^5 phage) using affinity-purified antibody. The largest insert isolated, λHPAP1.5, was 1.5 kb long and encoded for PAP from the Leu residue at 38 to the 3′ noncoding region. In order to obtain a full-length cDNA encoding for PAP, a radiolabeled recombinant phage template probe containing the 1.5-kb cDNA was used to rescreen the placenta cDNA library. Eighty-two positive clones were isolated from 5×10^5 phage. To isolate the longest clone from the 82 positives, these clones were then screened by using a recombinant phage template probe containing a 59-base fragment isolated from the 5′ end of the 1.5-kb insert. The clones that produced the 12 strongest signals were selected,

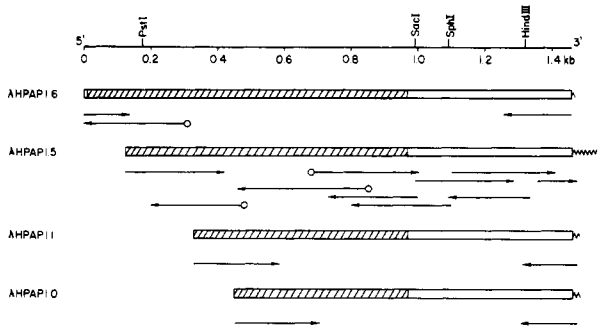

FIGURE 3: Summary of sequence strategy of the cDNA inserts. The arrows under the bars indicate the direction and extent of sequence analysis of the respective clones. Arrows with circles indicate that the sequences were determined by using specific synthetic oligonucleotides as primers. The open bars represent the noncoding regions, and the slashed bars represent the coding region. The poly(A) tail is shown by wavy lines.

Table I: Amino Acid Sequences of the Peptides Derived from the CNBr Digest of PAP[a]

peptides	locations	sequences
CNBr-1-T1	1–5	(AQVLR)
CNBr-1-T2	6–17	GTVTDFPGFDER
CNBr-1-T3	18–25	ADAETLRK
CNBr-2	28–72	KGLGTDEESILTLLTSRSNAQRQEISA-AFKTLFGRDLLDDLKSEL
CNBr-3	85–141	KPSRLYDAYELKHALKGAGTNEKVL-TEIIASRTPEELRAIKQVYEEEYGSS-LEDDVV
CNBr-4	152–211	LVVLLQANRDPDAGIDEAQVEQDAQA-LFQAGELKWGTDEEKFITIFGTRSV-SHLRKVFDK
CNBr-5	259–270	KGAGTDDHTLIR
CNBr-6	273–296	VSRSEIDLFNIRKEFRKNFATSLY
CNBr-7	299–319	IKGDTSGDYKKALLLLCGEDD

[a] Residues in parentheses were not sequenced due to the blockage of the terminal residue.

and the sizes of their inserts were determined. Six clones were found to contain 1.5-kb or longer inserts. The longest insert, λHPAP1.6, was 1.6 kb long and found to contain a full-length cDNA coding for PAP.

Most of the sequence was determined by using the 1.5-kb insert. The sequence coding for the NH_2-terminal portion of PAP was determined with the 1.6-kb insert. By combining the sequences obtained from two additional inserts, λHPAP1.1 and λHPAP1.0, the entire nucleotide sequence of both strands was determined. A summary of the sequence strategy and restriction mapping is shown in Figure 3.

The complete nucleotide sequence and the predicted amino acid sequence deduced from the PAP clones are shown in Figure 4. The 1.6-kb clone consists of a 5′ noncoding region, 957 bases for the mature protein, and a 3′ noncoding region. A continuous open-reading frame contained a stop codon at each end, TGA (1–3) and TAA (973–975). Twelve nucleotides downstream from the TGA stop codon, an ATG codon (13–15) was present coding for Met, which is the probable initiation site for translation of protein. This Met residue was followed by a sequence of Ala-Gln-Val-Leu-Arg, the composition of which agreed with that of the terminally blocked peptide CNBr-1-T1. This peptide had the following amino acid composition: Glx, 1.0; Arg, 1.0; Ala, 0.98; Val, 0.98; and Leu, 1.0. These results indicate that the initiator Met residue was removed by a cotranslational or posttranslational event and the sequence of the mature protein begins with Ala. The absence of a signal peptide in the PAP cDNA sequence sug-

```
TGA GTA GTC GCC ATG GCA CAG GTT CTC AGA GGC ACT GTG ACT GAC TTC CCT GGA TTT GAT GAG CGG GCT GAT GCA GAA ACT CTT CGG AAG      90
stop           (M) A  Q  V  L  R  G  T  V  T  D  F  P  G  F  D  E  R  A  D  A  E  T  L  R  K                            25

GCT ATG AAA GGC TTG GGC ACA GAT GAG GAG AGC ATC CTG ACT CTG TTG ACA TCC CGA AGT AAT GCT CAG CGC CAG GAA ATC TCT GCA GCT     180
 A  M  K  G  L  G  T  D  E  E  S  I  L  T  L  L  T  S  R  S  N  A  Q  R  Q  E  I  S  A  A                               55

TTT AAG ACT CTG TTT GGC AGG GAT CTT CTG GAT GAC CTG AAA TCA GAA CTA ACT GGA AAA TTT GAA AAA TTA ATT GTG GCT CTG ATG AAA     270
 F  K  T  L  F  G  R  D  L  L  D  D  L  K  S  E  L  T  G  K  F  E  K  L  I  V  A  L  M  K                               85

CCC TCT CGG CTT TAT GAT GCT TAT GAA CTG AAA CAT GCC TTG AAG GGA GCT GGA ACA AAT GAA AAA GTA CTG ACA GAA ATT ATT GCT TCA     360
 P  S  R  L  Y  D  A  Y  E  L  K  H  A  L  K  G  A  G  T  N  E  K  V  L  T  E  I  I  A  S                               115

AGG ACA CCT GAA GAA CTG AGA GCC ATC AAA CAA GTT TAT GAA GAA GAA TAT GGC TCA AGC CTG GAA GAT GAC GTG GTG GGG GAC ACT TCA     450
 R  T  P  E  E  L  R  A  I  K  Q  V  Y  E  E  E  Y  G  S  S  L  E  D  D  V  V  G  D  T  S                               145

GGG TAC TAC CAG CGG ATG TTG GTG GTT CTC CTT CAG GCT AAC AGA GAC CCT GAT GCT GGA ATT GAT GAA GCT CAA GTT GAA CAA GAT GCT     540
 G  Y  Y  Q  R  M  L  V  V  L  L  Q  A  N  R  D  P  D  A  G  I  D  E  A  Q  V  E  Q  D  A                               175

CAG GCT TTA TTT CAG GCT GGA GAA CTT AAA TGG GGG ACA GAT GAA GAA AAG TTT ATC ACC ATC TTT GGA ACA CGA AGT GTG TCT CAT TTG     630
 Q  A  L  F  Q  A  G  E  L  K  W  G  T  D  E  E  K  F  I  T  I  F  G  T  R  S  V  S  H  L                               205

AGA AAG GTG TTT GAC AAG TAC ATG ACT ATA TCA GGA TTT CAA ATT GAG GAA ACC ATT GAC CGC GAG ACT TCT GGC AAT TTA GAG CAA CTA     720
 R  K  V  F  D  K  Y  M  T  I  S  G  F  Q  I  E  E  T  I  D  R  E  T  S  G  N  L  E  Q  L                               235

CTC CTT GCT GTT GTG AAA TCT ATT CGA AGT ATA CCT GCC TAC CTT GCA GAG ACC CTC TAT TAT GCT ATG AAG GGA GCT GGG ACA GAT GAT     810
 L  L  A  V  V  K  S  I  R  S  I  P  A  Y  L  A  E  T  L  Y  Y  A  M  K  G  A  G  T  D  D                               265

CAT ACC CTC ATC AGA GTC ATG GTT TCC AGG AGT GAG ATT GAT CTG TTT AAC ATC AGG AAG GAG TTT AGG AAG AAT TTT GCC ACC TCT CTT     900
 H  T  L  I  R  V  M  V  S  R  S  E  I  D  L  F  N  I  R  K  E  F  R  K  N  F  A  T  S  L                               295

TAT TCC ATG ATT AAG GGA GAT ACA TCT GGG GAC TAT AAG AAA GCT CTT CTG CTG CTC TGT GGA GAA GAT GAC TAA CGTGTCACGG GGAAGAGCTC     995
 Y  S  M  I  K  G  D  T  S  G  D  Y  K  K  A  L  L  L  L  C  G  E  D  D  stop                                              319

CCTGCTGTGT GCCTGCACCA CCCCACTGCC TTCCTTCAGC ACCTTTAGC TGCATTTGTAT GCCAGTGCTT AACACATTGC CTTATTCATA CTAGCATGCT CATGACCAAC    1105

ACATACACGT CATAGAAGAA AATAGTGGTG CTTCTTTCTG ATCTCTAGT GGAGATCTCTT TGACTGCTGT AGTACTAAAG TGTACTTAAT GTTACTAAGT TTAATGCCTG    1215

GCCATTTTCC ATTTATATAT ATTTTTTAAG AGGCTAGAGT GCTTTTAGC CTTTTTTAAAA ACTCCATTTA TATTACATTT GTAACCATGA TACTTTAATC AGAAGCTTAG    1325

CCTTGAAATT GTGAACTCTT GGAAATGTTA TTAGTGAAGT TCGCAACTA AACTAAACCTG TAAAATTATG ATGATTGTAT TCAAAAGATT AATGAAAAAT AAACATTTCT    1435

GTCCCCCTGA AAAAAAAAAA AAAA                                                                                                  1459
```

FIGURE 4: Complete nucleotide sequence of the PAP cDNA inserts. The residues that are underlined were also determined from the CNBr fragments (see Table I). The Met residue in parentheses is the initiation site and is removed in a posttranslational event. The nucleotide sequence that is underlined twice is the potential recognition signal for polyadenylation.

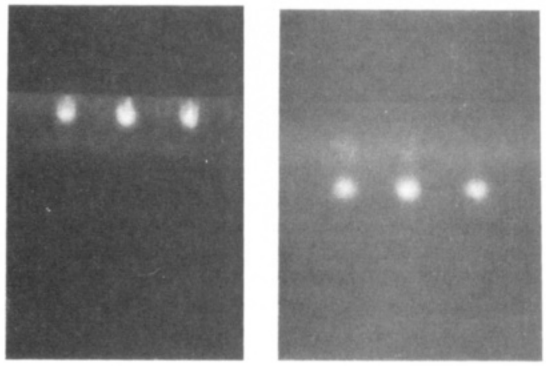

FIGURE 5: Detection of the acetyl group in the NH_2-terminal peptide of PAP on thin-layer chromatography. Acetyldansylhydrazide (A), sample (B), and a mixture of acetyldansylhydrazide and sample (C). Left panel with chloroform/butanol/acetic acid (6/3/1) and right panel with butanol/acetic acid/water (4/1/1).

gests that PAP is not constitutively secreted from cells. The amino acid sequence deduced from the nucleotide sequence of the PAP clones was in complete agreement with the 239 residues determined by Edman degradation of the 7 CNBr peptides.

Detection of Acetyl Group in the NH_2-Terminal Peptide. Identification of the acyl group of PAP was performed by the method of Schmer and Kreil (1965). The NH_2-terminal peptide, CNBr-1-T1, was treated with hydrazine, as was an authentic standard, acetylleucine. Resulting hydrazide was dansylated, and the dansylhydrazide was identified by thin-layer chromatography. The dansyl derivative of the sample migrated to the same position as acetyldansylhydrazide in two different solvent systems (Figure 5). We tentatively identified the blocking group of PAP as an acetyl group. However, a possibility that the blocking group is formyl is not ruled out, since separation of acetyl- and formyldansylhydrazide was not convincing in the solvent systems that we have tried. Acetylation is by far the most common acylation in blocked eukaryotic proteins and occurs predominantly as acetylserine or acetylalanine (Persson et al., 1985). Formylation is usually found as formyl-Met in prokaryotes. Formylation to other amino acids is quite rare (Wold, 1981), and only formyl-Gly has been found in honeybee melitin (Kreil & Kreil-Kiss, 1967). Thus, the blocking group of PAP is most likely acetyl.

The mature protein is composed of 319 amino acids and 1 acetyl group with the following composition: $acetyl_1$, Ala_{26}, Cys_1, Asp_{25}, Glu_{29}, Phe_{13}, Gly_{22}, His_3, Ile_{18}, Lys_{22}, Leu_{38}, Met_7, Asn_6, Pro_5, Gln_{12}, Arg_{19}, Ser_{21}, Thr_{23}, Val_{16}, Trp_1, and Tyr_{12}. The molecular weight was calculated to be 35 847, which is in good agreement with the value of 36 500 previously estimated by SDS–polyacrylamide gel electrophoresis. The presence of one Cys residue in the protein also agreed with the value of 0.8 mol of free sulfhydryl group/protein determined by Ellman reagent. No Asn residues are followed by the consensus signal sequence, X-Ser or X-Thr, for N-glycosylation. This is consistent with our previous result that carbohydrate was not detected in PAP by staining the SDS gel with Schiff base reagent. The isolation of eight CNBr peptides is expected from the presence of seven Met residues in the PAP sequence. However, we could isolate only seven peptides. The failure in isolating an eighth peptide can be explained by the finding of a Met-Thr sequence (residue 213–214), which is known to be resistant to CNBr cleavage (Cunningham et al., 1968).

Discussion

Computer-assisted analysis of the partial amino acid sequence of PAP revealed that it was highly homologous to the lipocortins (Funakoshi et al., 1987). The complete PAP sequence determined in the present study is compared with those of human lipocortins I and II (Figure 6). PAP has 42% and 44% identities with lipocortins I and II, respectively. A 51% identity is obtained when lipocortin I is compared with lipocortin II. Many nonidentical positions are also occupied by homologous amino acids. The major structural difference in these proteins is the variable length of the NH_2-terminal region. PAP is shorter than lipocortin I by 28 residues and 19 residues shorter than lipocortin II. It is unlikely that this region

```
                                                        1         10        20        30
PAP                                                  Ac AQVLRGTVTDFPGFDERADAETLRKAMKGLGT
Lipocortin I   MAMVSEFLKQAWFIENEEQEYVQTVKSSKGGPGSAVSPYPTFNPSSDVAALHKAIMVKGV
Lipocortin II  MSTVHEILCKLSLEGDH----------STPPSAYGSVKAYTNFDAERDALNIETAIKTKGV

                        40        50        60        70        80        90
PAP            DEESILTLLTSRSNAQRQEISAAFKTLFGRDLLDDLKSELTGKFEKLIVALMKPSRLYDA
Lipocortin I   DEATIIDILTKRNNAQRQQIKAAYLQETGKPLDETLKKALTGHLEEVVLALLKTPAQFDA
Lipocortin II  DEVTIVNILTNRSNAQRQDIAFAYQRRTKKELASALKSALSGHLETVILGLLKTPAQYDA

                       100       110       120       130       140       150
PAP            YELKHALKGAGTNEKVLTEIIASRTPEELRAIKQVYEEEYGSSLEDDVVGDTSGYYQRML
Lipocortin I   DELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDLAKDITSDTSGDFRNAL
Lipocortin II  SELKASMKGLGTDEDSLIEIICSRTNQELQEINRVYKEMYKTDLEKDIISDTSGDFRKLM

                       160       170       180       190       200       210
PAP            VVLLQANRDPDAGIDEAQV-EQDAQALFQAGELKWGTDEEKFITIFGTRSVSHLRKVFDK
Lipocortin I   LSLAKGDRSEDFGVNE-DLADSDARALYEAGERRKGTDVNVFNTILTTRSYPQLRRVFQK
Lipocortin II  VALAKGRRAEDGSVIDYELIDQDARDLYDAGVKRKGTDVPKWISIMTERSVPHLQKVFDR

                       220       230       240       250       260       270
PAP            YMTISGFQIEETIDRETSGNLEQLLLAVVKSIRSIPAYLAETLYYAMKGAGTDDHTLIRV
Lipocortin I   YTKYSKHDMNKNLDLELKGDIEKCLTAIVKCATSKPAFFAEKLHQAMKGVGTRHKALIRI
Lipocortin II  YKSYSPYDMLESIRKEVKGDLENAFLNLVQCIQNKPLYFADRLYDSMKGKGTRDKVLIRI

                       280       290       300       310      319
PAP            MVSRSEIDLFNIRKEFRKNFATSLYSMIKGDTSGDYKKALLLLCGEDD
Lipocortin I   MVSRSEIDMNDIKAFYQKMYGISLCQAILDETKGDYEKILVALCGGN
Lipocortin II  MVSRSEVDMLKIRSEFKRKYGKSLYYYIQQDTKGDYQKALLYLCGGDD
```

FIGURE 6: Comparison of the amino acid sequence of PAP with human lipocortins I and II. Residue numbers represent the numbering system of PAP. The positions that have identical residues in three proteins are boxed. Gaps (–) were placed to gain the best alignment. Sequences of lipocortins were taken from Huang et al. (1986).

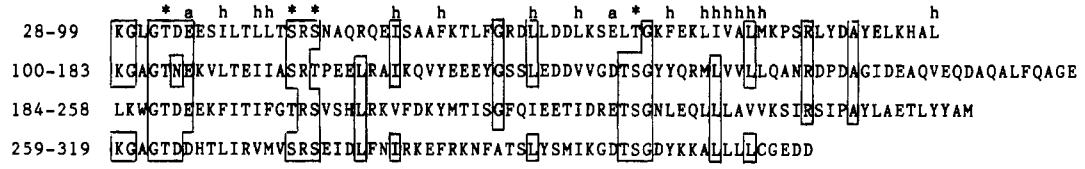

FIGURE 7: Alignment for the four repeat sequences in PAP. The positions that have three or four identical residues are boxed. The positions that have hydrophobic, hydroxy, and acidic amino acids are marked by h, asterisks, and a, respectively.

is involved in functional binding activity, because PAP has a strong binding affinity for phospholipid vesicles and also inhibits phospholipase A_2 activity (J. F. Tait, unpublished results). This concept is also supported by the fact (Davidson et al., 1987) that lipocortins I and II have essentially the same inhibitory activities against phospholipase A_2 despite the absence of homology in the NH_2-terminal regions of these two proteins. Furthermore, it was recently reported (Huang et al., 1987) that a proteolytically produced 33-kDa fragment, which lacks the NH_2-terminal 32 residues of lipocortin I, retains essentially full inhibitory activity against phospholipase A_2.

As observed in lipocortins (Weber & Johnson, 1986; Geisow, 1986; Huang et al., 1986; Saris et al., 1986; Wallner et al., 1986; Kristensen et al., 1986), the PAP sequence contains four internal repeats starting from Lys-28 (Figure 7). These repeats account for the entire sequence from this Lys to the end of the molecule. Without insertion of gaps, two Gly residues, one Thr residue, and one Arg residue can be aligned in all of the four repeats. In addition, hydrophobic amino acids occupy 15 positions (marked by "h" in the figure), and hydroxyamino acids are conserved in 4 positions (marked by asterisks). Acidic amino acids are also conserved in two positions (marked by "a").

Each of the four repeating sequences of PAP contains two regions that are commonly present in phospholipid binding proteins. The first region, the NH_2-terminal 17 residues, conforms to a consensus sequence (Geisow et al., 1986) of Lys-Gly-X-Gly-Thr-Asp-Glu-X-X-h-h-X-h-h-X-Ser-Arg, where h represents hydrophobic amino acids. This type of sequence has been found in the Ca^{2+} regulated membrane binding proteins, such as endonexin and calelectrin (Geisow et al., 1986), and in lipocortins (Huang et al., 1986; Saris et al., 1986; Wallner et al., 1986; Kristensen et al., 1986). Phospholipase A_2 and a viper venom phospholipase A_2 inhibitor also have closely related sequences (Mancheva et al., 1984). The second homologous region in these proteins is a stretch of six residues of hydrophobic amino acids at the C-terminal portion of each repeat. These two regions are thought to be directly involved in binding to phospholipid (Geisow, 1986). It is evident from its structural features that PAP belongs to this group of phospholipid binding proteins. We attribute the strong anticoagulant activity observed in our previous study to the presence of the phospholipid binding regions in the PAP molecule. Although it has not been reported, other members of the lipocortin family are expected to have anticoagulant activity because of their structural homology with PAP.

Tyr at 24 and Ser at 22 are identified as phosphorylation sites in the protein-tyrosine kinase substrate p36 (Glenney & Tack, 1985; Gould et al., 1986). This protein was later identified as an analogue of human lipocortin II (Huang et al., 1986). Tyr-21 was phosphorylated by EGF receptor/kinase, and phosphorylated lipocortin I required a lower Ca^{2+} concentration for maximal binding to phospholipid vesicles (Schlaepfer & Haigler, 1987). Tyr-187 and Ser-183, which are surrounded by acidic residues, are also thought to be potential phosphorylation sites in lipocortin II (Hirata, 1981; Hirata et al., 1984). Four corresponding positions (positions 5, 7, 165, and 169 by the PAP numbering system; see Figure 6) in both lipocortin I and PAP do not have either a Tyr or a Ser residue. In our preliminary study, PAP was not stoichiometrically phosphorylated by the insulin and EGF-receptor kinase (N. K. Tonks, unpublished results) or by protein kinase C (K. E. Meier, unpublished results). Thus, PAP does not seem to be a substrate of these protein kinases.

ACKNOWLEDGMENTS

We are grateful to Dr. Earl W. Davie for his support and encouragement throughout this work. We thank Drs. Dominic W. Chung, Jonathan Tait, and Akitada Ichinose for their helpful advice and discussions.

Registry No. DNA (human blood coagulation factor PAP messenger RNA complementary), 111237-03-7; blood coagulation factor PAP (human precursor), 111237-10-6; blood coagulation factor PAP (human), 111237-09-3.

REFERENCES

Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. C. (1984) *J. Chromatogr. 336*, 93.

Blackwell, G. J., Carnuccio, R., Di Rosa, M., Flower, R. J., Parente, L., & Persico, P. (1980) *Nature (London) 287*, 147–149.

Bridgen, P. J., Cross, G. A. M., & Bridgen, J. (1976) *Nature (London) 263*, 613–614.

Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622–627.

Creutz, C. E., Dowling, L. G., Sando, J. J., Villar-Palasi, C., Whipple, J. H., & Zaks, W. J. (1983) *J. Biol. Chem. 258*, 14664–14674.

Cunningham, B. A., Gottlieb, P. D., Konigsberg, W. H., & Edelman, G. M. (1968) *Biochemistry 7*, 1983–1994.

Davidson, F. F., Dennis, E. A., Powell, M., & Glenney, J. R., Jr. (1987) *J. Biol. Chem. 262*, 1698–1705.

Deykin, D., Cochios, F., & Mosher, D. (1969) *Biochem. Biophys. Res. Commun. 34*, 245–251.

Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1*, 80–91.

Ericsson, L. H., Wade, R. D., Gagnon, J., McDonald, R. M., Granberg, R. R., & Walsh, K. A. (1977) in *Solid Phase Methods in Protein Sequence Analysis* (Previero, A., & Coletti-Previero, M. A., Eds.) pp 137–142, Elsevier, Amsterdam.

Flower, R. J., & Blackwell, G. J. (1979) *Nature (London) 278*, 456–459.

Foster, D. C., & Davie, E. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4766–4770.

Fujikawa, K., & McMullen, B. A. (1983) *J. Biol. Chem. 258*, 10924–10933.

Funakoshi, T., Heimark, R. L., Hendrickson, L. E., McMullen, B. A., & Fujikawa, K. (1987) *Biochemistry 26*, 5572–5578.

Geisow, M. J. (1986) *FEBS Lett. 203*, 99–103.

Geisow, M., Childs, J., Dash, B., Harris, A., Panayotou, G., Sudhof, T., & Walker, J. H. (1984) *EMBO J. 3*, 2969–2974.

Geisow, M. J., Fritsche, U., Hexham, J. M., Dash, B., & Johnson, T. (1986) *Nature (London) 320*, 636–638.

Gerke, V., & Weber, K. (1984) *EMBO J. 3*, 227-233.
Glenney, J. R., Jr., & Tack, B. F. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7884-7888.
Gould, K. L., Woodgett, J. R., Isacke, C. M., & Hunter, T. (1986) *Mol. Cell. Biol. 6*, 2738-2744.
Hirata, F. (1981) *J. Biol. Chem. 256*, 7730-7733.
Hirata, F., Matsuda, K., Notsu, Y., Hattori, T., & DelCarmine, R. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4717-4721.
Hu, N., & Messing, J. (1982) *Gene 17*, 271-277.
Huang, K.-S., Wallner, B. P., Mattaliano, R. J., Tizard, R., Burne, C., Frey, A., Hession, C., McGray, P., Sinclair, L. K., Chow, E. P., Browning, J. L., Ramachandran, K. L., Tang, J., Smart, J. E., & Pepinsky, R. B. (1986) *Cell (Cambridge, Mass.) 46*, 191-199.
Huang, K.-S., McGray, P., Mattaliano, R. J., Burne, C., Chow, E. P., Sinclair, L. K., & Pepinsky, R. B. (1987) *J. Biol. Chem. 262*, 7639-7645.
Kondo, S., Noguchi, M., Funakoshi, T., Fujikawa, K., & Kisiel, W. (1987) *Thromb. Res.* (in press).
Kreil, G., & Kreil-Kiss, G. (1967) *Biochem. Biophys. Res. Commun. 27*, 275-280.
Kretsinger, R. H., & Creutz, C. E. (1986) *Nature (London) 320*, 573.
Kristensen, T., Saris, C. J., Hunter, T., Hicks, L. J., Noonan, D. J., Glenney, J. R., & Tack, B. F. (1986) *Biochemistry 25*, 4497-4503.
Maki, M., Murata, M., & Shidara, Y. (1984) *Eur. J. Obstet., Gynecol. Reprod. Biol. 17*, 149-154.
Mancheva, I., Kleinschmidt, T., Aleksiev, B., & Braunitzer, G. (1984) *Hoppe-Seyler's Z. Physiol. Chem. 365*, 885-894.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, pp 76-85, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Moore, P. B., & Dedman, J. R. (1982) *J. Biol. Chem. 257*, 9663-9667.
Persson, B., Flinta, C., von Heijne, G., & Jornvall, H. (1985) *Eur. J. Biochem. 152*, 523-527.
Reutelingsperger, C. P. M., Hornstra, G., & Hemker, H. C. (1985) *Eur. J. Biochem. 151*, 625-629.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Saris, C. J. M., Tack, B. F., Kristensen, T., Glenney, J. R., Jr., & Hunter, T. (1986) *Cell (Cambridge, Mass.) 46*, 201-212.
Schlaepfer, D. D., & Haigler, H. T. (1987) *J. Biol. Chem. 262*, 6931-6937.
Schmer, G., & Kreil, G. (1969) *Anal. Biochem. 29*, 186-192.
Shadle, P. J., Gerke, V., & Weber, K. (1985) *J. Biol. Chem. 260*, 16354-16360.
Sudhof, T. C., Ebbecke, M., Walker, J. H., Fritsche, U., & Boustead, C. (1984) *Biochemistry 23*, 1103-1109.
Wallner, B. P., Mattaliano, R. J., Hession, C., Cate, R. L., Tizard, R., Sinclair, L. K., Foeller, C., Chow, E. P., Browning, J. L., Ramachandran, K. L., & Pepinsky, R. B. (1986) *Nature (London) 320*, 77-81.
Weber, K., & Johnsson, N. (1986) *FEBS Lett. 203*, 95-98.
Wold, F. (1981) *Annu. Rev. Biochem. 50*, 783-814.
Young, R. A., & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194-1198.

Conformational Properties of the Main Intrinsic Polypeptide (MIP26) Isolated from Lens Plasma Membranes†

Joseph Horwitz* and Dean Bok

Jules Stein Eye Institute, University of California, Los Angeles, School of Medicine, Los Angeles, California 90024-1771

Received May 6, 1987; Revised Manuscript Received July 28, 1987

ABSTRACT: The conformational properties of the main intrinsic polypeptide (MIP26) isolated from lens plasma membranes were studied by using near- and far-ultraviolet circular dichroism. The far-ultraviolet spectrum of MIP26 solubilized with octyl β-D-glucopyranoside indicates an α-helical content of ~50% and a β-structure content of ~20%. A detergent-free membrane suspension of MIP26 produced a typically distorted far-ultraviolet spectrum which was caused by differential light scattering and absorption flattening. However, decreasing the size of the membrane fragments by sonication produced a far-ultraviolet spectrum free of distortion, and with a rotatory strength profile similar to that obtained for MIP26 solubilized with octyl β-D-glucopyranoside. This implies similar secondary structure properties for the protein in both the suspension and the sugar detergent. The cleavage of MIP26 with *Staphylococcus aureus* protease, which results in removal of a 5-kilodalton peptide and which mimics the age-dependent posttranslational changes that take place in the lens, did not significantly affect the conformation of the core protein as judged by the near-ultraviolet circular dichroism spectra.

The vertebrate eye lens is an avascular tissue with a highly ordered architecture (Maisel et al., 1981; Benedetti et al., 1981). The bulk of the lens is composed of a multitude of elongated, nondividing cells which are called lens fibers. These lens fibers are coupled to one another by many intercellular junctions situated along their plasma membranes (Rae, 1979; Goodenough et al., 1980; Benedetti et al., 1981). The most abundant protein in the lens fiber plasma membrane is an intrinsic membrane protein known as MIP26 or MP26 (Broekhuyse et al., 1976; Benedetti et al., 1981). This protein,

† This work was supported in part by USPHS Grants EY 00331, EY 00444, and EY 03897. D.B. is the Dolly Green Professor of Ophthalmology.

0006-2960/87/0426-8092$01.50/0