Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 3708–3712, June 1988
Biochemistry

Characterization of cDNA encoding human placental anticoagulant protein (PP4): Homology with the lipocortin family

(DNA sequence analysis/amino acid sequence identity/thromboplastin)

ULRICH GRUNDMANN*†, KARL-JOSEF ABEL*, HANS BOHN*, H. LÖBERMANN‡, FRIEDRICH LOTTSPEICH§, AND HANS KÜPPER*

*Department of Molecular Biology, Research Institutes, Behringwerke AG, Postfach 1140, 3550 Marburg, Federal Republic of Germany; ‡Department of Biotechnology-Production, Sandoz AG, 4002 Basel, Switzerland; and §Max-Planck-Institut, 8033 Martinsried, Federal Republic of Germany

Communicated by Frank W. Putnam, January 19, 1988 (received for review November 18, 1987)

ABSTRACT **A cDNA library prepared from human placenta was screened for sequences encoding the placental protein 4 (PP4). PP4 is an anticoagulant protein that acts as an indirect inhibitor of the thromboplastin-specific complex, which is involved in the blood coagulation cascade. Partial amino acid sequence information from PP4-derived cyanogen bromide fragments was used to design three oligonucleotide probes for screening the library. From 10^6 independent recombinants, 18 clones were identified that hybridized to all three probes. These 18 recombinants contained cDNA inserts encoding a protein of 320 amino acid residues. In addition to the PP4 cDNA we identified 9 other recombinants encoding a protein with considerable similarity (74%) to PP4, which was termed PP4-X. PP4 and PP4-X belong to the lipocortin family, as judged by their homology to lipocortin I and calpactin I.**

The protein PP4 was first described by Bohn *et al.* (1) as a placental protein (M_r, 35,000), which is mainly found in the placental tissue (50 mg per placenta) with very little secretion into the maternal bloodstream (2 μg/liter). Subsequently, the protein PP4 was shown to be an anticoagulant protein that appeared to be a potent inhibitor of thromboplastin and immunologically identical or related to the thromboplastin inhibitor described by Shidara (2). It binds specifically to phospholipid membranes and inhibits the blood coagulation cascade by preventing formation of the extrinsic pathway activator, which consists of phospholipid, thromboplastin (FIII), and FVII in the presence of Ca^{2+} (3). A thromboplastin inhibitor prevents blood coagulation without proteolytic inactivation and prolongs the thromboplastin-induced clotting time, but it does not influence the thrombin-dependent fibrin formation. Nearly the same anticoagulatory mechanism was described for a protein of M_r 32,000 isolated from arteries of human umbilical cord (4). An increased serum level of PP4 has been reported for patients with cardiac infarction as well as in tumor tissues such as serous and mucinous cystadenocarcinomas (5, 6). These findings suggest that PP4 could serve as a diagnostic marker. Partial amino acid sequences of PP4 showed a striking similarity to proteins of the phospholipase A2 inhibitor family—e.g., lipocortin I and calpactin I. These proteins are thought to block production of the mediators of inflammation (prostaglandins and leukotrienes) by inhibiting the release of arachidonic acid from membranes by phospholipase A2 (7). Lipocortin I and calpactin I (or lipocortin II) are the first identified members of the family, but calelectrin, endonexin, and some other proteins of apparent M_r 32,000–38,000 also belong to this class (8–12). Human lipocortin I is defined as an anti-inflammatory protein that is inducible by glucocorticoids. It can be phosphorylated by tyrosine and serine-threonine kinases (13). Recently, it was shown that lipocortin I has the same inhibitory profile as the glucocorticoids but with a more rapid onset of action (14). Lipocortin I exists as a monomeric protein. On the other hand, calpactin I was isolated as a monomer as well as a heterologous tetramer of two large subunits (M_r, 36,000) and two small subunits (M_r, 10,000) (15). It acts on the cytoskeletal proteins spectrin and actin, binds phospholipid in a Ca^{2+}-dependent manner, and is a major substrate for tyrosine-protein kinases (16). Both proteins, lipocortin I and calpactin I from humans, occur in placenta at ≈0.2% of the total protein (7). Analysis of the amino acid sequence of both proteins revealed four internal repeating units of 70 amino acids each (7, 9, 17).

Here we report the isolation and determination of the coding sequence for PP4, a placental anticoagulant protein that is an indirect thromboplastin inhibitor.¶ We also show the similarity between this protein and the above mentioned proteins of the phospholipase A2 inhibitor family. Recently, Funakoshi *et al.* (18) isolated and characterized a human placental anticoagulant protein (PAP), which seems to be the same protein as PP4 described by Bohn *et al.* (1) The partial amino acid sequence reported for PAP is in good agreement with our amino acid sequence deduced from the cDNA. PAP also shows similarities with other proteins of the phospholipase A2 inhibitor family.

MATERIALS AND METHODS

Reagents. Restriction endonucleases were purchased from New England Biolabs or Boehringer Mannheim. Reverse transcriptase was from P. H. Stehelin (Basel). Phage λgt10 DNA and the *in vitro* packaging system were obtained from Vector Cloning Systems (San Diego, CA).

Isolation of RNA and Preparation of cDNA Libraries. RNA was prepared from a mature human placenta by the standard guanidinium thiocyanate extraction procedure (19). Poly(A)$^+$ RNA was selected by oligo(dT)-cellulose chromatography (20) and 4 μg was copied into cDNA with reverse transcriptase; double-stranded cDNA was prepared by the method of Gubler and Hoffman (21). After *Eco*RI linker addition, the cDNA was ligated into the λgt10 vector as described (22) and packaged *in vitro*. Phages were plated on *Escherichia coli* C600HFL. About 1×10^6 independent recombinant plaques were obtained.

Screening of a Human Placenta cDNA Library. The library was screened on blotted duplicate filters with two ^{32}P-labeled oligonucleotide probes (multiprime labeling system) derived on the basis of amino acid sequences of two PP4 cyanogen

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†To whom reprint requests should be addressed.

¶This sequence is being deposited in the EMBL/GenBank data base (Intelligenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03264).

bromide fragments. Sequence analysis was performed on an Applied Biosystems 470A gas-phase sequencer. We chose the oligonucleotide sequence 5′ ATGAAGGGCC TGGGCACAGA TGAGGAGAGC ATCCT 3′ (probe I) for the peptide sequence Met-Lys-Gly-Leu-Gly-Thr-Asp-Glu-Glu-Ser-Ile-Leu, 5′ CAGGAGATCT CTGCTGCCTT CAAGACCCTG TTTGGC 3′ (probe II) for Gln-Glu-Ile-Ser-Ala-Ala-Phe-Lys-Thr-Leu-Phe-Gly, and 5′ GACCCTGATG CTGGCATTGA TGAGGCCCAG GTGGAGCAGG ATGCCCAG 3′ (probe III) for Asp-Pro-Asp-Ala-Gly-Ile-Asp-Glu-Ala-Gln-Val-Glu-Gln-Asp-Ala-Gln according to the rules of Lathe (23). For hybridization the blotted filters were hybridized and washed as described at the calculated temperature (22, 23).

Characterization of cDNA Inserts. Recombinant phages were plaque-purified and grown in liquid culture, and phage DNA was extracted as described (22). cDNA inserts were characterized by restriction enzyme analysis and the sequence was determined by the dideoxy sequencing strategy of Sanger *et al.* (24) after subcloning into the Bluescript M13 vector (Stratagen Cloning Systems, La Jolla, CA). Oligonucleotides were synthesized on a Biosearch 8750 DNA-Synthesizer to provide sequence primers from either end of the sequence. Sequence data were analyzed with the UWGCG sequence analysis software as implemented on the MicroVax II.

RESULTS

Isolation of Human PP4 cDNA. Hybridization of the complete placental cDNA library with PP4-derived oligonucleotide probe I yielded 46 positive recombinant plaques on duplicate filters (25). By sequence analysis, we found that 9 had inserts that represent coding sequences for a similar, but not identical, protein to PP4, which we named PP4-X. Therefore, we rescreened the 46 positive phages separately with two newly designed oligonucleotide probes derived from the PP4 amino acid sequence (probes II and III). Both probes hybridized to 18 of the positives, which were different from PP4-X cDNA-containing phages. Phage DNA from three phages was isolated and digested with *Eco*RI, and the size of the cDNA inserts was analyzed by agarose gel electrophoresis. Each of the phages had one single *Eco*RI fragment ranging from 1000 to 1660 base pairs (Fig. 1). The cDNA insert of phage PP4-14 was used to rescreen the 46 recombinant phages. Again, all 18 phages that had hybridized to probes II and III showed positive signals. cDNA of PP4-20 was sequenced and a 5′-terminal synthetic oligonucleotide probe derived from PP4-20 was used to identify inserts with maximal 5′ sequence information. Two phages were identified (PP4-26 and -48). Both cDNA-specific inserts had identical 5′ termini, but with a different length of the poly(A) tail, and therefore are probably derived from different cDNA molecules. In total, of the 18 isolated positives, 4 cDNA inserts were identified that carry the complete protein coding sequence of PP4 including the poly(A) sequence, of which PP4-26 has the longest sequence.

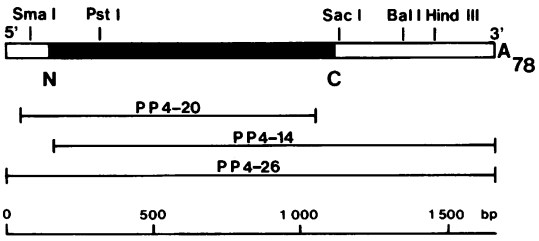

FIG. 1. Restriction map of human cDNA encoding PP4 and the location of the coding sequence (solid bar). N and C, amino and carboxyl termini of the PP4 protein, respectively. The overlapping cDNA inserts of clones PP4-14, -20, and -26 are shown, whereby PP4-26 covers the complete cDNA sequence. bp, Base pairs.

Identification of PP4-Specific mRNA. The size of the PP4 mRNA from different tissues was determined by RNA blot analysis. Poly(A)$^+$ RNA from human liver, kidney, and placenta was separated in an agarose gel containing formaldehyde, blotted onto nitrocellulose filters, and hybridized to the ^{32}P-labeled cDNA insert of PP4-20 (Fig. 2). Placental poly(A)$^+$ RNA showed the strongest signal while 1/5th to 1/10th the mRNA was detected in liver and kidney cells. The size of the single RNA band (≈1700 nucleotides) is in good agreement with the size of the cDNA insert of PP4-26 (1660 nucleotides).

Analysis of cDNA and Deduced Amino Acid Sequence. After subcloning into the Bluescript M13 vector (Stratagene), the complete nucleotide sequence of PP4 cDNA was determined. The nucleotide sequence as determined from clones PP4-14, -20, -26, and -48 is displayed in Fig. 3 including the deduced amino acid sequence. All four clones showed identical sequences; however, one of the clones (PP4-20) contained cytosine at nucleotide position 214, which would result in an amino acid change from glutamic acid to glutamine at amino acid position 22. The cDNA contains an open reading frame encoding a protein of 320 amino acids. Assuming the first AUG is the initiator codon the predicted molecular weight of PP4 is 35,935 (including initiator methionine).

Similar to PP4-X and lipocortin I, the first two amino acids are methionine and alanine. The protein does not contain N-linked glycosylation sites and has only one cysteine residue at the C-terminal end (position 316). Upstream of the initiator codon, the cDNA carries another 150 nucleotides. The 3′ untranslated region contains 550 nucleotides, including the stop codon at position 1111, a poly(A) tail of 78 residues, and the poly(A) addition site AATAAA 17 nucleotides from the start of the poly(A) tail.

The deduced amino acid sequence reveals the presence of four internal repeat units of ≈70 amino acids in a similar fashion as demonstrated for lipocortin I and calpactin I (11, 17, 26). Computer analysis of the amino acid sequence for hydrophobicity according to Kyte and Doolittle (27) shows that the repeat units carry stretches of hydrophobic amino

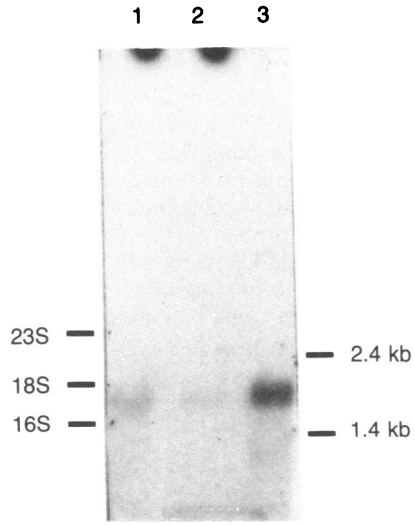

FIG. 2. Blot hybridization analysis of mRNA. Poly(A)$^+$ RNA (5 μg) of human liver (lane 1), kidney (lane 2), and placenta (lane 3) were separated in a denaturing agarose gel containing formaldehyde and transferred to a nitrocellulose filter. The filter was hybridized with a ^{32}P-labeled cDNA insert of PP4-20 at 65°C, washed as described (22), and exposed to Kodak XAR-5 film. The positions of 16S, 18S, and 23S RNA of human and bacterial (*E. coli*) ribosomal RNA as RNA standards are indicated. kb, Kilobases.

```
        10                  30                  50                  70                  90                  110
CTAGGTGCAGCTGCCAGATCCTTCAGCGTCTGCATCTCGGCGTCGCCCCGCGTACCGTCGCCCGGCTCTCCGCCGCTCTCCCGGGGGTTCGGGGCACTTGGGTCCCACAGTCTGGTCCTG

        130                 150                 170                 190                 210                 230
CTTCACCTTCCCCTGACCTGAGTAGTCGCTATGGCACAGGTTCTCAGAGGCACTGTGACTGACTTCCCTGGATTTGATGAGCGGGCTGATGCAGAAACTCTTCGGAAGGCTATGAAAGGC
                                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  A  M  K  G

        250                 270                 290                 310                 330                 350
TTGGGCACAGATGAGGAGAGCATCCTGACTCTGTTGACATCCCGAAGTAATGCTCAGCGCCAGGAAATCTCTGCAGCTTTTAAGACTCTGTTTGGCAGGGATCTTCTGGATGACCTGAAA
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  F  K  T  L  F  G  R  D  L  L  D  D  L  K

        370                 390                 410                 430                 450                 470
TCAGAACTAACTGGAAAATTTGAAAAATTAATTGTGGCTCTGATGAAACCCTCTCGGCTTTATGATGCTTATGAACTGAAACATGCCTTGAAGGGAGCTGGAACAAATGAAAAAGTACTG
S  E  L  T  G  K  F  E  K  L  I  V  A  L  M  K  P  S  R  L  Y  D  A  Y  E  L  K  H  A  L  K  G  A  G  T  N  E  K  V  L

        490                 510                 530                 550                 570                 590
ACAGAAATTATTGCTTCAAGGACACCTGAAGAACTGAGAGCCATCAAACAAGTTTATGAAGAAGAATATGGCTCAAGCCTGGAAGATGACGTGGTGGGGGACACTTCAGGGTACTACCAG
T  E  I  I  A  S  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  G  Y  Y  Q

        610                 630                 650                 670                 690                 710
CGGATGTTGGTGGTTCTCCTTCAGGCTAACAGAGACCCTGATGCTGGAATTGATGAAGCTCAAGTTGAACAAGATGCTCAGGCTTTATTTCAGGCTGGAGAACTTAAATGGGGGACAGAT
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  Q  A  L  F  Q  A  G  E  L  K  W  G  T  D

        730                 750                 770                 790                 810                 830
GAAGAAAAGTTTATCACCATCTTTGGAACACGAAGTGTGTCTCATTTGAGAAAGGTGTTTGACAAGTACATGACTATATCAGGATTTCAAATTGAGGAAACCATTGACCGCGAGACTTCT
E  E  K  F  I  T  I  F  G  T  R  S  V  S  H  L  R  K  V  F  D  K  Y  M  T  I  S  G  F  Q  I  E  E  T  I  D  R  E  T  S

        850                 870                 890                 910                 930                 950
GGCAATTTAGAGCAACTACTCCTTGCTGTTGTGAAATCTATTCGAAGTATACCTGCCTACCTTGCAGAGACCCTCTATTATGCTATGAAGGGAGCTGGGACAGATGATCATACCCTCATC
G  N  L  E  Q  L  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  H  T  L  I

        970                 990                 1010                1030                1050                1070
AGAGTCATGGTTTCCAGGAGTGAGATTGATCTGTTTAACATCAGGAAGGAGTTTAGGAAGAATTTTGCCACCTCTCTTTATTCCATGATTAAGGGAGATACATCTGGGGACTATAAGAAA
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  Y  S  M  I  K  G  D  T  S  G  D  Y  K  K

        1090                1110                1130                1150                1170                1190
GCTCTTCTGCTGCTCTGTGGAGAAGATGACTAACGTGTCACGGGGAAGAGCTCCCTGCTGTGTGCCTGCACCACCCCACTGCCTTCCTTCAGCACCTTTAGCTGCATTTGTATGCCAGTG
A  L  L  L  C  G  E  D  D

        1210                1230                1250                1270                1290                1310
CTTAACACATTGCCTTATTCATACTAGCATGCTCATGACCAACACATACACGTCATAGAAGAAAATAGTGGTGCTTCTTTCTGATCTCTAGTGGAGATCTCTTTGACTGCTGTAGTACTA

        1330                1350                1370                1390                1410                1430
AAGTGTACTTAATGTTACTAAGTTTAATGCCTGGCCATTTTCCATTTATATATATTTTTTAAGAGGCTAGAGTGCTTTTAGCCTTTTTTAAAAACTCCATTTATATTACATTTGTAACCA

        1450                1470                1490                1510                1530                1550
TGATACTTTAATCAGAAGCTTAGCCTTGAAATTGTGAACTCTTGGAAATGGTATTAGTGAAGTTCGCAACTAAACTAAACCTGTAAAATTATGATGATTGTATTCAAAAGATTAATGAAA

●       1570                1590                1610                1630                1650
AATAAACATTTCTGTCCCCCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 3. Nucleotide sequence of human placental cDNA encoding PP4. The nucleotide sequence was determined by analysis of the overlapping inserts shown in Fig. 1. The deduced amino acid sequence (single-letter code) is displayed below the DNA sequence. The poly(A) addition site 17 base pairs upstream of the poly(A) sequence is marked by a solid circle. The positions of oligonucleotide probes II (positions 301–336) and III (positions 634–681) are underlined. Probe I lies within a sequence of high homology among the repeat units and therefore cannot be placed unequivocally.

acids in two conserved regions. One is at the end of each repeat unit and one is in the middle (Fig. 4).

DISCUSSION

We have isolated the cDNA for the placental anticoagulant protein PP4 (formerly called thromboplastin inhibitor), which is a member of the phospholipase A2 inhibitor family. The earlier name thromboplastin inhibitor is not correct in the sense that PP4 does not inhibit thromboplastin directly by interacting with the protein in a competitive, steric, or irreversible way but rather inhibits the reaction caused by thromboplastin through binding of the necessary phospholipids. This mechanism could be confirmed by prolongation of the coagulation time, which is dependent on the concentration of phospholipids (unpublished results). This is in accordance with results published by Davidson *et al.* (10), which suggest that the inhibition of pancreatic phospholipase A2 by calpactin I is created by phospholipid depletion.

The amino acid composition of PP4 described by Bohn *et al.* (1) is in good agreement with the composition deduced from the cDNA sequence. Although it was reported that PP4 contained 2.4% carbohydrates, we could not detect any N-linked glycosylation site in the deduced protein sequence of PP4, whereas one glycosylation site each is found in PP4-X, lipocortin I, and calpactin I. Since both proteins PP4 and PP4-X are very similar in size and amino acid composition, it is possible that the glycosylation content of PP4

```
                    20                  40                  60                  80
MAQVLRGTVTDFPGFDERADAETLRKAMKGLGTDEESILTLLTSRSNAQRQEISAAFKTLFGRDLLDDLKSELTGKFEKLIVALM
                    100                 120                 140
          KPSRLYDAYELKHALKGAGTNEKVLTEIIASRTPEELRAIKQVYEEEYGSSLEDDVVGDTSGYYQRMLVVLLQ
160                 180                 200                 220                 240
  ANRDPDAGIDEAQVEQDAQALFQAGELKWGTDEEKFITIFGTRSVSHLRKVFDKYMTISGFQIEETIDRETSGNLEQLLAVV
                    260                 280                 300                 320
          KSIRSIPAYLAETLYYAMKGAGTDDHTLIRVMVSRSEIDLFNIRKEFRKNFATSLYSMIKGDTSGDYKKALLLLCGEDD

PP4 Consensus sequence:   A....GTDE.....I..SRS......I...F....A..I...I..E.SG.....IL.L
                          G     ND      L  T T      V   Y    G  L   L  D T      LV V
                                        V                            V
```

FIG. 4. Amino acid consensus sequence for the repeat structure of PP4. The amino acid sequences (single-letter code) of all four repeating units of PP4 were aligned. The consensus sequence was derived from the identical or chemically related residues present in all four repeats. Stretches of three and more hydrophobic amino acids [according to Kyte and Doolittle (27)] are underlined.

```
          1                                                  51                                                 100
PP4                                maqv lrgtVtdfpg FderaDaetl rkAmkglGtD EesIltlLts RsnaQRQeIs aAFktlfgrd LlddLKseLt

Lipo  amvseflkqa wfieneeqey vqtvksskgg pgsaVspypt FnpssDvaal hkAimvkGvD EatIidiLtk RnnaQRQqIk aAYlqetgkp LdetLKkaLt
Cal           ms tvheilckls legdhstpps aygsVkaytn FdaerDalni etAiktkGvD EvtIvniLtn RsneQRQdIa fAYqrrtkke LasaLKsaLs
PP4-X                        mamat kggtVkaasg FnameDaqtl rkAmkglGtD EdaIisvLay RntaQRQeIr tAYkstigrd LiddLKseLs

Consensus  ---------- ---------- ---------- ----V----- F----D---- --A----G-D E--I---L-- R---QRQ-I- -AY------- L---LK--L-


          101                                                151                                                200
PP4   GkFEklIvaL MkpsrlYDay ELkhaLKGaG TnEkvLtEIi aSRTpeElra IkqvYeeeyg ssLedDVvgD TSgyYqrmLv vLlqanRdpd agidEaqveq

Lipo  GhLEevVlaL LktpaqFDad ELraaMKGlG TdEdtLiEIl aSRTnkEird InrvYreelk rdLakDItsD TSgdFrnaLl sLakgdRsed fgvnE.dlad
Cal   GhLEtvIlgL LktpaqYDas ELkasMKGlG TdEdsLiEIi cSRTnqElqe InrvYkemyk tdLekDIvsD TSgdFrklMv aLakgrRaed gsviDyelid
PP4-X GnFEqvIvgM MtptvlYDvq ELqraMKGaG TdEgcLiEIl aSRTpeEirr IsqtYqqqyg rsLedDIrsD TSfmFqrvLv sLsaggRdeg nyldD.alvr

Consensus  G--E--I--L ------YD-- EL---MKG-G T-E--L-EI- -SRT--E--- I---Y----- --L--DI--D TS--F---L- -L----R--- ----------


          201                                                251                                                300
PP4   .DAqaLFqAG elkwGTDeek FitIFgtRsv shLrkVFdkY mtiSgfqiee tidrEtsGnl EqlLlavVks irsipaYLAE tLyyaMKGaG TddhtLIRVM

Lipo  sDAraLYeAG errkGTDvnv FntILttRsy pgLrrVFqkY tkySkhdmnk vldlElkGdi EkcLtaiVkc atskpaFFAE kLhqaMKGvG TrhkaLIRIM
Cal   qDArdLYdAG vkrkGTDvpk WisIMteRsv chLqkVFerY ksySpydmle sikkEvkGdl EnaFlnlVqc iqnkplYFAD rLydsMKGkG TrdkvLIRIM
PP4-X qDAqdLYeAG ekkwGTDevk FltVLcsRnr nhLlhVFdeY kriSqkdieq siksEtsGsf EdaLlaiVkc mrnksaYFAE kLyksMKGlG TddntLIRVM

Consensus  -DA--LY-AG ----GTD--- F--IL--R-- --L--VF--Y ---S------ ----E--G-- E--L---V-- ------YFAE -L---MKG-G T----LIR-M


          301
PP4   VSRsEIDLfn IrkeFrknFa tSLYsmIkgD TsGDYkKaLl lLCGedd

Lipo  VSRsEIDMnd IkafYqkmYg iSLCqaIldE TkGDYeKiLv aLCGgn.
Cal   VSRsEVDMlk IrseFkkkYg kSLYyyIqqD TkGDYqKaLl yLCGgdd
PP4-X VSRaEIDMld IrahFkrlYg kSLYsfIkgD TsGDYrKvLl vLCGgdd

Consensus  VSR-EIDM-- I---F---Y- -SLY--I--D T-GDY-K-L- -LCG---
```

FIG. 5. Similarity between the human PP4, lipocortin I, calpactin I, and PP4-X sequences. Amino acid sequences (single-letter code) of all four proteins were aligned and similarities are shown on the consensus line. Capital letters denote amino acids that are identical or chemically of the same class in all four sequences. Numbering starts with the alanine residue of lipocortin I.

described by Bohn *et al.* (1) was caused by a contamination or comigration with PP4-X.

The PP4 cDNA contains only one large open reading frame with a stop codon four codons upstream of the initiator codon. Protein sequence analysis of PP4 showed that its amino terminus is blocked. Since an alanine residue follows the initiator methionine, one might assume that the first methionine of PP4 is cleaved off and the following alanine residue is modified by N-acetylation. Similar situations have been observed for lipocortin I, calpactin I, PP4-X, and factor XIIIa (8, 9, 22, 25, 26). As in the case with these proteins, the cDNA for PP4 does not encode a signal peptide. All of these proteins, however, are secreted through the cellular membrane by a mechanism yet to be established.

Comparison of the PP4 protein sequence reveals a high degree of similarity to members of the phospholipase A2 inhibitory proteins including the presence of four internal repeat units (Fig. 5). The sequence similarity of PP4 to lipocortin I and calpactin I is 62–63% and to PP4-X it is 74%. The C-terminal part is more conserved among the proteins and most variability is seen at the very N-terminal end. PP4 is most similar to PP4-X and has nearly the same number of amino acid residues. It also crossreacts with antiserum raised against PP4, as shown after expression of PP4-X cDNA in *E. coli* (25). The biological function of the protein is not yet known. Although the similarity is only weakly retained in the amino-terminal region, there exists a potential protein kinase C phosphorylation site in PP4 at amino acid position 23 (threonine) as well as in PP4-X at position 24 and lipocortin I (26).

Common to all members of the phospholipase A2 inhibitor family is a 4-fold repeat structure of ≈70 amino acid residues (28). A consensus sequence for the repeat structures was first described by Geisow *et al.* (11) for calelectrin. Within the repeat structures of PP4, there are 20 residues identical or chemically related (Fig. 4). Interestingly, the repeat units are separated by stretches of hydrophobic amino acids and another stretch of at least seven hydrophobic residues exists approximately in the middle of each repeat unit. Minor hydrophobic stretches are scattered over the whole protein sequence. A similar pattern is seen with the other proteins of the phospholipase A2 inhibitor family. The somewhat regular arrangement of hydrophobic stretches might very well have to do with the specific association of phospholipids with these proteins. Reutelingsperger *et al.* (4) proposed that their anticoagulatory protein isolated from human umbilical cord, which leads in a similar fashion to a prolongation of the prothrombin time, interferes with the lipid binding of factor X and/or prothrombin.

By RNA blot analysis of poly(A)$^+$ RNA from tissues of human origin, we could show that the PP4 mRNA has a size of ≈1700 nucleotides. The analysis indicates that the concentration of PP4 mRNA is 5–10 times higher in placenta than in the other two organs. Bohn *et al.* (1) have shown by immunological methods that PP4 is present not only in placenta but also in different adult tissues from humans (e.g., stomach, kidney, bladder, spleen, adrenal gland, and liver). The clones PP4-26 and -48 have identical 5′ ends. Although one could take this as an indication that this specifies the 5′ end of the mRNA, we could not identify any sequences representing homology to the 3′ end of ribosomal RNA as described by Baralle and Brownlee (29). Within the phospholipase A2 inhibitor family the PP4 mRNA contains the most extended 3′ noncoding region found so far. It is interesting that within the family the coding sequences show similarities of 50–60%, while the 5′ and 3′ noncoding regions show similarities of <40%.

Note Added in Proof. After submission of the manuscript, we learned about the publications by Funakoshi *et al.* (30) and Iwasaki *et al.* (31). Both papers report on the isolation of a cDNA encoding a placental anticoagulant protein that is identical to the one described in our paper.

We thank C. Nerlich and T. Rein for technical assistance in isolation of cDNA clones and DNA sequencing, M. Rühl for her preparation of oligonucleotides, and H. Karges for critical review of the manuscript.

1. Bohn, H., Kraus, W. & Winkler, W. (1985) *Arch. Gynaekol.* **236,** 225–233.
2. Shidara, Y. (1984) *Acta Obstet. Gynaecol. Jpn.* (Engl. Ed.) **36,** 2583–2592.
3. Osterud, B. (1984) *Scand. J. Haematol.* **23,** 337–345.
4. Reutelingsperger, C., Hornstra, G. & Hemker, C. (1985) *Eur. J. Biochem.* **151,** 625–629.
5. Inaba, N., Satto, N., Ijichi, M., Fukazawa, I., Nito, A., Takamizawa, H., Lüben, G. & Bohn, H. (1984) *Tumor Biol.* **5,** 75–85.
6. Bohn, H. (1985) *Behring Inst. Mitt.* **78,** 70–82.
7. Huang, K. S., Wallner, B. P., Mattaliano, R. J., Tizard, R., Burne, C., Fray, A., Hession, C., McGray, P., Sinclair, K. L., Chow, E. P., Browning, J. L., Ramachandran, K. L., Tang, J., Smart, J. & Pepinsky, R. B. (1986) *Cell* **46,** 191–199.
8. 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. & Pepinskey, R. B. (1986) *Nature (London)* **320,** 77–81.
9. Kristensen, T., Saris, C. J. M., Hunter, T., Hichs, L. J., Noonan, D. J., Glenney, J. R. & Tack, B. F. (1986) *Biochemistry* **25,** 4497–4503.
10. Davidson, F. F., Dennis, E. A., Powell, M. & Glenney, J. R. (1987) *J. Biol. Chem.* **262,** 1698–1705.
11. Geisow, M. J., Fritsche, U., Hexham, J. M., Dash, B. & Johnson, T. (1986) *Nature (London)* **320,** 636–638.
12. Haigler, H. T., Schlaepfer, D. D. & Burgess, W. H. (1987) *J. Biol. Chem.* **262,** 6921–6930.
13. Schlaepfer, D. D. & Haigler, H. T. (1987) *J. Biol. Chem.* **262,** 6931–6937.
14. Cirino, G., Flower, R. J., Browning, J. L., Sinclair, L. K. & Pepinsky, R. B. (1987) *Nature (London)* **328,** 270–272.
15. Glenney, J. R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4258–4262.
16. Glenney, J. R. & Tack, B. F. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7884–7888.
17. Geisow, M. J. & Walker, J. H. (1986) *Trends Biol. Sci.* **11,** 420–423.
18. Funakoshi, T., Heimark, R. L., Hendrickson, L. E., McMullen, B. A. & Fujikawa, K. (1987) *Biochemistry* **26,** 5572–5578.
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
20. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
21. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
22. Grundmann, U., Amann, E., Zettlmeissl, G. & Küpper, H. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8024–8028.
23. Lathe, R. (1985) *J. Mol. Biol.* **183,** 1–12.
24. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
25. Grundmann, U., Abel, K. J., Amann, E. & Küpper, H. A. (1988) *Behring Inst. Mitt.* **82,** 59–67.
26. Saris, J. M., Tack, B. F., Kristensen, T., Glenney, J. R. & Hunter, T. (1986) *Cell* **46,** 201–212.
27. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
28. Kretsinger, R. H. & Creutz, C. E. (1986) *Nature (London)* **320,** 573.
29. Baralle, F. E. & Brownlee, G. G. (1978) *Nature (London)* **274,** 84–87.
30. Funakoshi, T., Hendrickson, L. E., McMullen, B. A. & Fujikawa, K. (1987) *Biochemistry* **26,** 8087–8092.
31. Iwasaki, A., Suda, M., Nakao, H., Nagoya, T., Saino, Y., Arai, K., Mizoguchi, T., Sato, F., Yoshizaki, H., Hirata, M., Miyata, T., Shidara, Y., Murata, M. & Maki, M. (1987) *J. Biochem.* **102,** 1261–1273.