

Molecular Cloning of the cDNA for a Human Amyloid Precursor Protein Homolog: Evidence for a Multigene Family

Cindy A. Sprecher,[‡] Francis J. Grant,[‡] Gloria Grimm,[‡] Patrick J. O'Hara,[‡] Fanny Norris,[§] Kjeld Norris,[§] and Donald C. Foster^{*,‡}

ZymoGenetics, Inc., 4225 Roosevelt Way, NE, Seattle, Washington 98105, and Novo-Nordisk A/S, Novo Alle DK-2880, Bagsvaerd, Denmark

Received January 15, 1993; Revised Manuscript Received March 3, 1993

ABSTRACT: Alzheimer's disease is a degenerative neurological disorder characterized by neural loss and brain lesions associated with plaques containing large amounts of the β /A4 amyloid peptide. Molecular cloning of the cDNA for this peptide from human brain has shown it to be derived by proteolysis from a much larger precursor called the amyloid precursor protein (APP). The biological role of the precursor is unknown, but it has been shown to be transcribed in many human tissues in addition to brain. In the present report, we describe the molecular cloning from a human placental library of a full-length cDNA for a molecule closely related to APP. This novel molecule, which we have called amyloid precursor protein homolog (APPH), shares overall domain organization with APP. It is 763 amino acids in length and appears to encode a signal peptide, a large apparent extracellular domain including a Kunitz inhibitor domain, a transmembrane region, and a short cytoplasmic domain. Northern analysis indicates that it occurs in at least two molecular forms and is transcribed in human brain, heart, lung, liver, and kidney, in addition to placenta. On the basis of its extensive sequence similarity and conservation of domain structure, APPH is the nearest relative of APP yet identified in an emerging multigene family.

A characteristic feature of Alzheimer's disease is the accumulation of an insoluble peptide called the β -amyloid or A4 peptide (β /A4) in pathologic brain lesions or senile plaques. This 42–43 amino acid peptide has been shown by molecular cloning in several groups to be derived from a much larger protein precursor of 751 amino acids called the amyloid precursor protein (APP).¹ The domain structure of APP appears to resemble a cell-surface receptor comprised of a signal peptide, a large "extracellular" domain including a Kunitz protease inhibitor domain, a transmembrane region, and a short presumed "intracellular" domain (Kang et al., 1987; Ponte et al., 1988). The β /A4 peptide is encoded within the precursor on the N-terminal side of the transmembrane region and is partly derived from the membrane-spanning region. The mechanism by which the β /A4 peptide is liberated from the precursor is unknown.

The biosynthesis of APP is complex. Several forms of mRNA encoding the precursor have been reported. The first cDNA cloned, APP₆₉₅, encoded 695 amino acids and appeared to represent a full-length form of the precursor (Kang et al., 1987). Subsequently, two additional forms, APP₇₅₁ (Ponte et al., 1988; Tanzi et al., 1988) and APP₇₇₀ (Kitaguchi et al., 1988), were identified in which a Kunitz inhibitor domain of 57 amino acids was inserted by alternative mRNA splicing into the middle of the extracellular domain. A truncated form, APP₅₆₃ (Sauvage & Octave, 1989), has been cloned in which the C-terminal 208 amino acids have been replaced with a 20 amino acid sequence similar to a human Alu repeat. APP₅₆₃ probably occurs due to alternative splicing and appears to encode a truncated, soluble form of the APP molecule lacking the cell-anchoring transmembrane domain.

In addition, a soluble form of the protein has been identified by three laboratories as protease nexin II (Van Nostrand et al., 1989; Oltersdorf et al., 1989) and as coagulation factor XIa inhibitor (Smith et al., 1990).

The presence of a serine protease inhibitor domain within some molecular forms of the precursor structure has prompted speculation that this domain serves to inhibit protease(s) which can otherwise cleave the precursor to liberate the A4 peptide and lead to disease pathology. This concept is supported by the observations that the precursor can be proteolytically processed to the peptide by normal cells in culture (Haass et al., 1992; Seubert et al., 1992) and also that the Kunitz inhibitor domain is apparently functional toward a wide spectrum of serine proteases, even within the context of the larger, unprocessed soluble precursor structure (Oltersdorf et al., 1989; Van Nostrand et al., 1989, 1990a; Smith, et al., 1990). Putative proteases which may be involved in such cleavage in Alzheimer's disease have not yet been identified.

The normal biological role of the APP molecule is currently not understood. It is apparently expressed in a diverse spectrum of human tissues, including the brain, spleen, liver, kidney, lung, and heart (Tanzi et al., 1988). The soluble form of the protein has been detected in human plasma, platelet releasates (Van Nostrand et al., 1990b), and cell lines derived from human liver (Smith et al., 1990) and fibroblasts (Van Nostrand et al., 1990a). The only current links with biological roles are the pathologic effects of the β /A4 peptide, the inhibitory effects of the Kunitz domain toward coagulation proteases (Smith et al., 1990; Van Nostrand et al., 1990a), and observations that some portion of the secreted soluble form has growth factor activity toward fibroblasts (Saito et al., 1989). The rest of the structure of APP does not suggest a relationship with other molecules whose biology is known.

In an effort to further explore the biological significance of the APP molecule, we have examined mRNA from several human tissues to determine whether APP is part of a multigene family. Such structural relationships have frequently facil-

* To whom correspondence should be addressed.

[‡] ZymoGenetics, Inc.

[§] Novo-Nordisk A/S.

¹ Abbreviations: APP, amyloid precursor protein; APPH, amyloid precursor protein homolog; APLP, amyloid precursor-like protein; RSMF, rat sperm membrane protein.

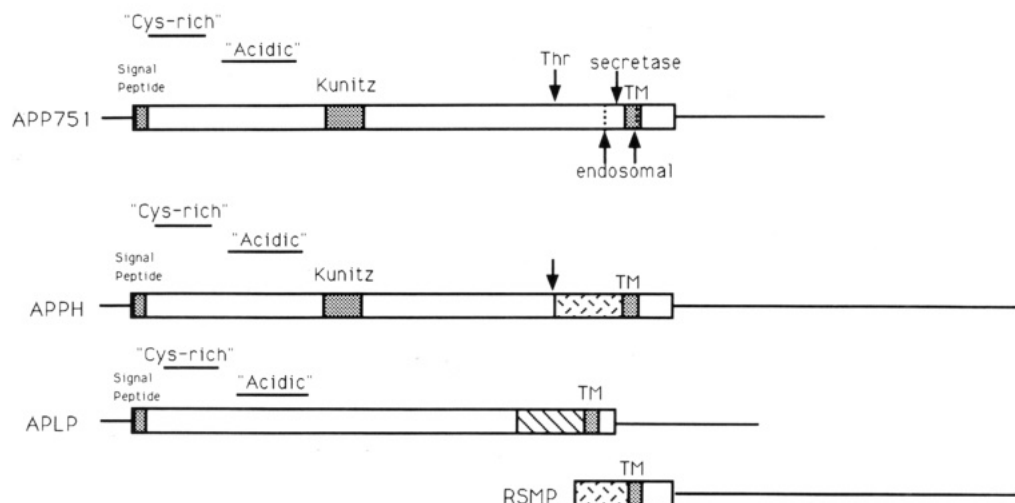


FIGURE 1: Diagram of the domain structures for APP, APPH, APLP, and RSMP. The regions corresponding to the signal peptide, Kunitz domain, and transmembrane (TM) domain are labeled and stippled. The region of highest sequence divergence on the N-terminal side of the TM domain in these molecules is crosshatched. Arrows indicated probable proteolytic cleavage sites by thrombin (Igarashi et al., 1992) or by proteases in the secretase vs endosomal pathways (Haass et al., 1992; Seubert et al., 1992). The cysteine-rich and acidic domains are indicated by solid bars above each diagram.

itated structure/function correlations with other macromolecules, as has been the case for the Kunitz domain of APP itself. When mRNAs from human tissues were hybridized with oligonucleotides derived from regions of APP under conditions of reduced stringency, hybridizing bands were observed in placenta which suggested the existence of homologous precursor structures. In the present study, we report the cloning and full-length sequence of a cDNA (APPH) with similar domain organization and considerable amino acid sequence homology with human APP₇₅₁. This new cDNA, like APP, is transcribed in a broad spectrum of human tissues.

MATERIALS AND METHODS

cDNA Isolation and Sequencing. A human placenta cDNA library was screened using an antisense 30-mer oligo, 5'-GT-TGTTGCTGTTGCCTCCGCAGCCTCCGTA3'. This is a λ GT11 library constructed at ZymoGenetics and kindly provided by Dr. Fred Hagen. A total of 1×10^6 plaques were screened at low hybridization stringency. Nylon membranes were hybridized at 55 °C in 5 \times SSPE (1 \times SSPE = 0.15 M NaCl, 0.01 M NaH₂PO₄, 0.001 M EDTA, pH 7.4), 5 \times Denhardt's solution, 0.5% SDS, and 100 μ g/mL salmon sperm DNA. Filters were washed at 50 °C in 2 \times SSC and 0.1% SDS. Positive plaques were purified, and their *Eco*RI inserts were subcloned into pUC19. These were sequenced by the dideoxy chain termination method (Sanger et al., 1977).

Northern Analysis. A blot of poly(A)⁺ mRNA from human tissues (CLONTECH Laboratories, multiple tissue Northern blot) was screened using a random primed 1-kb fragment from APPH. Hybridization was at 42 °C in 5 \times SSPE, 10 \times Denhardt's solution, 2% SDS, 100 μ g/mL salmon sperm DNA, and 50% formamide. The blot was washed at 65 °C in 0.2 \times SSC and 0.1% SDS.

RESULTS

A human placental cDNA library was screened with an oligonucleotide probe which could encode a protein homologous to the Kunitz inhibitor domain of APP and which had shown positive signals on a Northern blot of human placental mRNA under conditions of reduced hybridization stringency (data not shown). Since the sizes of the placental mRNAs did not correspond to the known transcript sizes of APP or any other

previously identified proteins containing Kunitz inhibitor domains, these data implied the existence in placenta of additional protein(s) with structures possibly related to APP. During the library screening, plaques were identified which showed positive hybridization with the oligonucleotide probe at reduced hybridization stringency (50 °C) but which were negative for hybridization at more stringent temperatures (65 °C). Eleven of the plaques were purified, ten of which are related to each other by DNA sequence analysis. One of these clones (APPH13) contains an insert of approximately 3.8 kb and appears to be a full-length transcript of an open reading frame which encodes a 763 amino acid protein. The cDNA contains a 5' noncoding region of 71 nucleotides, an open reading frame of 2283 nucleotides encoding 763 amino acids, and a 3' noncoding region of 1361 nucleotides.

The open reading frame encodes a protein with considerable sequence and overall domain organization similarity with the human amyloid precursor protein (APP) (Figure 1). Like APP, APPH appears to encode a receptor-like protein with a classic signal peptide which directs cotranslational insertion into the ER, a large extracellular domain of 662 amino acids, a transmembrane domain of 23 amino acids, and a short intracellular domain of 47 amino acids (Figure 2). Following the signal peptide, the N-terminal 180 amino acids of the extracellular domain are rich in cysteines, all twelve of which are conserved in APP with precisely the same spacing profile. Thus, this region is likely similarly folded and stabilized by disulfide bridges in both molecules. The next 72 residues are extremely rich in acidic amino acids, both in APP and in APPH. This region (61% acidic amino acids) includes stretches of up to 18 consecutive acidic residues and has been proposed (Kang et al., 1987) to be involved in binding of cations on the surface of electrically active cells in neuronal tissues.

Following the acidic region is a domain (residues 309–364) with sequence homology with the Kunitz family of protease inhibitors. This domain of 57 amino acids is either present or absent in differently spliced forms of the APP mRNAs (Ponte et al., 1988; Tanzi et al., 1988; Kitaguchi et al., 1988) which differ in size by 167 nucleotides and also differ in tissue distribution. It is present in all of the APPH clones isolated from the human placental library. Comparison of the amino acid sequence of this domain from APPH with the known

GTCGCGGTGTCTAAGCGAGGAGTCCGAGTGTGTGAGCTTGAGAGCCGCGCTAGAGCGACCCGGCGAGGATGGCGGCCACCGGACCGCGGCCCGCAGCCAGGGCAGGCTCCTGCTTCT 125
 M A A T G T A A A A A T G R L L L L L 18
 GCTGCTGGTGGGGCTCACGGCGCTGCCCTGGCGCTGCGCGGTACATCGAGGCTCTTGAGCCAATGCCGGAACAGGATTGTGCTGTGCTGAGCCTCAAATCGCAATGTTTGTGGGAAGTTAA 250
 L L V G L T A P A L A L A G Y I E A L A A N A G T G F A V A E P Q I A M F C G K L 59
 ATATGCATGTGAACATTGAGACTGGGAAATGGGAACCTGATCCAACAGGCACCAAGAGCTGCTTTGAAACAAAAGGAAGTTCTTCAGTACTGTGAGGAGATGTATCCAGAGCTACAGATCACA 375
 N M H V N I Q T G K W E P D P T G T K S C F E T K E E V L Q Y C Q E M Y P E L Q I T 101
 AATGTGATGGAGGCAACACAGCGGTTAGTATTGACAACCTGGTGCCGAGGGACAAAAGCAATGCAAGAGTCGCTTTGTACACCTTTCAAGTGTCTCGTGGGTGAATTTGTAAGTGTGCT 500
 N V M E A N Q R V S I D N W C R R D K K Q C K S R F V T P F K C L V G E F V S D V L 143
 GCTAGTTCAGAAAAGTCCAGTTTTTCCACAAAGAGCGGATGGAGGTGTGTGAGAATCACCAGCACTGGCACACGGTAGTCAAAGAGGCATGTCTGACTCAGGGAATGACCTTATATAGCTACG 625
 L V P E K C Q F F H K E R M E V C E N H Q H W H T V V K E A C L T Q G M T L Y S Y 184
 GCATGCTGCTCCCATGTGGGTAGACAGTTCATGGCACTGAATATGTGTGCTGCCCTCAGACAAAGATTATTGGATCTGTGTCAAAGAAGAGGAAGAGGAGATGAAGAGGAAGAGGAAGAG 750
 G M L L P C G V D Q F H G T E Y V C C P Q T K I I G S V S K E E E E E D E E E E E E 226
 GAAGATGAAGAGGAAGACTATGATGTTTATAAAAGTGAATTTCTACTGAAGCAGATCTGGAAGACTTCACAGAAGCAGCTGTGGATGAGGATGATGAGGATGAGGAAGAGGGGAGGAGTGGT 875
 E D E E E D Y D V Y K S E F P T E A D L E D F T E A A V D E D D E D E E E G E E V V 268
 GGAGGACCGAGATTACTACTATGACACCTTCAAAGGAGATGACTACAATGAGGAGAATCCTACTGAACCCGGCAGCGAGCCACCATGTGAGACAAGGAATTTACTCATGATGTCAAAGCTGTCT 1000
 E D R D Y Y Y D T F K G D D Y N E E N P T E P G S D G T M S D K E I T H D V K A V 309
GCTCCAGGAGGCGATGACGGGGCCCTGCCGGGCCGTGATGCCTCGTTGGTACTTCGACCTCTCCAAGGAAAGTGCCTGCCGCTTTATATATGGTGGCTGCCGGGGCAACAGGAACAATTTTGAG 1125
C S Q E A M T G P C R A V M P R W Y F D L S K G K C V R F I Y G G C G G N R N N F E 351
TCTGAGGATTATTGTATGGCTGTGTGAAGCGATGATTCCTCAACTCCTCTGCCAACCAATGATGTTGATGTGATTTTCGAGACCTCTGCAGATGATAATGAGCATGCTCGCTCCAGAAGGC 1250
S E D Y C M A V C K A M I P P T P L P T N D V D V Y F E T S A D D N E H A R F Q K A 393
 TAAGGAGCAGCTGGAGATTCCGACCCGCAACCGAATGGACAGGGTAAGAAGGAATGGGAAGGCGAGAGCTTCAAGCTAAGAACCTCCCAAAGCAGAGAGGCGAGCTCTGATTGAGCACTTCC 1375
 K E Q L E I R H R N R M D R V K K E W E E A E L Q A K N L P K A E R Q T L I Q H F 434
 AAGCCATGGTTAAAGCTTTAGAGAAGGAAGCAGCCAGTGAGAAGCAGCAGCTGGTGGAGCCACCTGCCCGGAGTGAAGCTATGCTGAATGACCGCGCTGGATGGCTGGAGAACTACCTG 1500
 Q A M V K A L E K E A A S E K Q L V E T H L A R Y V A E M L N D R R M A L E N Y L 476
 GCTGCCCTGCAGTCTGACCCGCCACGGCCTCATCGCATTCTCCAGGCCCTACGGCGTTATGTCGCTGCTGAGAACAAGATCGCTTACATACCATCCGTCATTACAGCATGTGTTGGCTGTGA 1625
 A A L Q S D P P R P H R I L Q A L R R Y V R A E N K D R L H T I R H Y Q H V L A V D 518
 CCCAGAAAAGGGGGCCAGATGAAATCCAGGTGATGACACATCTCCACGTGATTGAAGAAAGGAGGAACCAAGCCTCTCTGCTCTACAAAGTACCTTATGTAGCCCAAGAAATTCAGAGG 1750
 P E K A A Q M K S Q V M T H L H V I E E R R N Q S L S L L Y K V P Y V A Q E I Q E 559
 AAATGTAGAGCTCCTTCAGGAGCAGCGTGCAGATATGGACAGTTCAGTCCCTCAATCTCAGAGACCCCTGTGGACGTCCGGGTGAGCTCTGAGGAGAGTGAAGAGATCCACCGTTCACCC 1875
 E I D E L L Q E Q R A D M D Q F T A S I S E T P V D V R V S S E E S E E I P P F H P 601
 TCCACCCCTTCCAGCCCTACCTGAGAAGCAAGACTCAGCCGAGTGTGACCACTGAAAAAGGATCTGGAGTGGGAGCAGGATGGGGGACTGATCGGTGCCGAAGAGAAAGTGA 2000
 F H P F P A L P E N E D T Q P E L Y H P M K K G S G V G E Q D G G L I G A E E K V I 643
 TAACAGTAAGAATAAAGTGGATGAAAACATGGTCATTGACGAGACTCTGGATGTTAAGGAATGATTTTCAATGCCGAGAGAGTGGAGGCCTCGAGGAAGAGCGGAATCCGTGGGCCACTGC 2125
 N S K N K V D E N M V I D E T L D V K E M I F N A E R V G G L E E E R E S V G P L 684
 GGGAGGACTTCAGTCTGAGTAGCAGTGTCTCATTGGCCTGCTGGTCATCGCAGTGGCCATTGCCACGGTCATCGTCATCAGCCTGGTGTGCTGAGGAAGAGGAGATGGCACCATCAGCCAC 2250
 R E D F S L S S S A L I G L L V I A V A I A T V I V I S L V M L R K R Q Y G T I S H 726
 GGGATCGTGGAGTTGATCCAATGCTCACCCAGAAGAGCGTCACCTGAACAAGTGCAGAACCATTGGCTATGAGAACCCACCTACAAATACCTGGAGCAGATGAGATTTAGGTGGCAGGGAG 2375
 G I V E V D P M L T P E E R H L N K M Q N H G Y E N P T Y K Y L E Q M Q I . 763
 CGCGGCAGCCCTGGCGAGGGATGCAGGTGGCCGGAAGATCCACGATTCGATCGACTGCCAAGCAGCAGCCGCTGCCAGGGGCTGCGTCTGACATCCTGACCTCTGGACTGTAGGACTATA 2500
 TAAAGTACTACTGTAGAAGTGAATTTCCATTCTTTTAAATGGGTGAAAAATGGTAATATAACAATATATGATATATAAACCTTAAATGAAAAAATGATCTATTGCAGATATTGTAGTATT 2625
 TCTTTTTTAAATTAATCAGAAACCCCACTTCCATTGTATTGTCTGACACATGCTCTCAATATATAATAATGGGAAATGTCGATTTTCAATAATAGACTTATATGAGGCTGTGCTTCCGGTTAT 2750
 GTTGTGAAGTCAACTCTTCAGCCTCATTCACTGTCTGGCTTTTATTTAAAGAAAAAAGGAGTATTCCTTTTTAAATGAGCTTTTCAGGAAGTGTCTGAGAAATGGGGTGAATAGGGAAC 2875
 TGTAATGGCCACTGAAGCAGCTGAGAGACCCCTCGCAAAATGATGTGAAGGACCAGTTTCTTGAAGTCCAGTGTTCACGGCTGGATACCTGTGTCTCCATAAAGTCTGTCCACCAAGGAC 3000
 GTTAAAGGCATTTTATCCAGCGCTCTTAGAGAGCTTAGTGTATACAGATGAGGGGTGCTGCTGCTTCCCTCGGAATCAGTGTCTCCACAGAGATTAGCCTGTAGCTTATATTGACAT 3125
 TCTTCACTGTCTGTTTACCTACCGTAGCTTTTACCGTTCACCTCCCTTCCAACATATGTCCAGATGTGACGGCTCCTCTCTGGACTTCTCCAAAGGCACTGACCTCGGCCCTCACT 3250
 TTGTCCCTCACCTCCACCCCTCCTGTACCGGCCTTGTGACATTCACCTCAGAGAAGCACACCAAGGAGGGGCGCGGCTGGCCAGGAGAGAACACGGGGAGGTTGTTGTGTGAAAGGA 3375
 AAGTAGTCCAGGCTGTCCCTGAAACTGAGTCTGTGGCACTGTGGAAGCTTTGAACAATGTGTTTTCGTCACAGGAGTCTTTGTAATGCTGTGACAGTTGATGTGATGCTCACTGCTTCTGC 3500
 TTTTCTTTCTTTTATTTTAAAAAATCTGAAGGTTCTGGTAACCTGTGGTGTATTTTATTTTCTGTGACTGTTTTGTTTTGTTTTTCTTTTCTCCCTTTAGCCCTATTATGTCT 3625
 CTACCCACTATGCACAGATTAACCTCACCTACAACTCCTTAATATGATCTGTGGAGAATGTACACAGTTTAAACACATCAATAAATACTTTAACTTCCA(poly A) 3726

FIGURE 2: Nucleotide sequence of cDNA clone APPH13. Potential signal peptidase cleavage sites are shown by asterisks, the Kunitz domain is boxed, the N-linked glycosylation signal is indicated by a broken underline, the transmembrane region is indicated by a solid underline, and the apparently alternatively spliced section is indicated by a dashed overline.

Kunitz domains of human origin (Figure 3) indicates the highest degree of similarity with the domain from APP.

Following the Kunitz domain is a region of about 200 amino acids (residues 365–566) with continued high overall sequence similarity with APP. The similarity then breaks down

abruptly, with no significant similarity apparent from residues 567–694, until the beginning of the transmembrane domain in both molecules. Notably absent are sequence identities corresponding to Asp₆₅₃–Lys₆₈₀ in the APP sequence which correspond to the apparent pretransmembrane portion of the

APPH	VKAVCSQEAMTGPCRAVMPRWYFDLSK GKCVRFIYGGCGGNRRNFESEDYCMVCKA
	* * * * *
APP	VREVCSEQAETGPCRAMISRWFYFDVTEGKCAPFFYGGCGGNRRNFDTEEYCMVCGS
	* * * * *
a3 (VI)	ETDICKLPKDEGTCRDFILKWYYDPNTKSCARFWYGGCGGNENKFGSQKECEKVCAP
	* * * * *
ITI-1	KEDSCQLGYSAGPCMGMTSRFYNGTSMACETFQYGGCMGNGNMFVTEKECLQTCRT
	* * * * *
ITI-2	TVAACNLPIVRGPCRAFIQLWAFDAVGKCVLFYGGCQNGNKFYSEKECREYCGV
	* * * * *
TFPI-1	MHSFCAFKADDGPCKAIMKRFFFNIFTRQCEEFIYGGCEGNQNRFSLEECKKMCRT
	* * * * *
TFPI-2	KPDFCFLEEDPGICRGYITRYFYNNQTKQCFERFKYGGCLGNMNNFETLEECKNICED
	* * * * *
TFPI-3	GPSWCLTPADRGLCRANENRFYNSVIGKCRPFKYSGCCGNENNFTSKQECLRACKK

FIGURE 3: Alignment of the Kunitz domain sequence of APPH with known human Kunitz domains. Identities with APPH are indicated for each sequence with asterisks. Abbreviations: a3(VI), α_3 -collagen, type VI; ITI, inter- α trypsin inhibitor; TFPI, tissue factor pathway inhibitor.

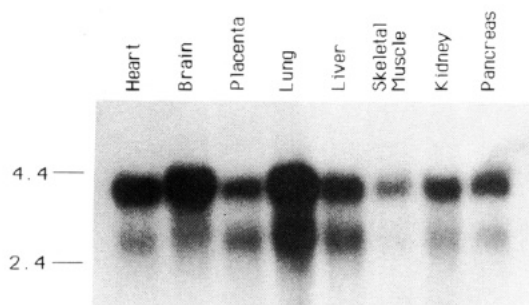


FIGURE 4: Northern blot of mRNA from human tissues probed with APPH cDNA. Mobilities for 4.4- and 2.2-kb DNA markers are indicated. Equivalent amounts of mRNA from the indicated tissues were run and probed as described under Materials and Methods.

pathologic β /A4 peptide. Thus, it is not clear that the APPH molecule bears any structural counterpart to the β /A4 peptide region. One of our cDNA clones (APPH20) contained a deletion of 12 amino acids (613–624) in this region while maintaining the same reading frame. It is likely that the APPH20 form arose by an alternative splicing event.

Following the conserved hydrophobic transmembrane domain is a cytoplasmic region of 47 amino acids which is very similar (32/42 identical) to the same region in APP and also shows even stronger sequence homology with a partial cDNA corresponding to a rat sperm membrane surface protein of unknown function (Yan et al., 1990). The transmembrane sequence is immediately followed by three basic amino acids which presumably serve a stop-transfer function, anchoring the transmembrane region in the plasma membrane of the cell. Possible functions of the cytoplasmic domain remain obscure.

The cDNA sequence encodes a single potential N-linked glycosylation site, Asn-Gln-Ser, at positions 541–543 in the extracellular domain of the molecule. Two other sequences of Asn-Pro-Thr are present but presumably are not glycosylated due to the presence of proline in the structure (Bause, 1983).

Northern blot analysis of human tissues with APPH (Figure 4) reveals that this gene is transcribed in a wide variety of tissues. The transcript is most abundant in the lung and brain but is also obvious in heart, placenta, liver, skeletal muscle, kidney, and pancreas. Two major transcripts are apparent in all tissues observed, a major mRNA of 3.8 kb and a less abundant mRNA of 3.0 kb. From the size of our full-length clones, we are apparently reporting the sequence of the larger mRNA. The size of the smaller mRNA does not correspond to that for APP, and the hybridization of the blot was of

sufficient stringency to preclude hybridization to inexact matches. The molecular identity of the smaller mRNA is unknown at this time but could represent a truncated molecule which could encode a soluble form of APPH, similar to those observed for APP.

DISCUSSION

The striking degree of amino acid sequence similarity and the conservation of overall domain structure between the APP molecule and APPH indicate that these two genes have probably arisen from gene duplication. In this respect, it will be of interest to determine the chromosomal localization of APPH, since the APP gene has been localized to a region of chromosome 21 which has been shown to undergo duplication in association with Down's syndrome (Kang et al., 1988). Like APP, APPH appears to have a domain organization resembling a cell-surface receptor (Figure 1) with a very large extracellular domain, a single transmembrane domain, and a short intracellular domain. The highly conserved regions of sequence within the "Cys-rich" and Kunitz inhibitor regions of the extracellular domain and in the cytoplasmic domains especially suggest similarity of function for both molecules within their relative biological contexts. The normal functions of both of these molecules are presently unknown.

Some aspects of possible function can be inferred from the molecular structures. The inhibitory properties of the Kunitz inhibitor domain of APP have been studied both within the context of the truncated, secreted form of the native molecule (Oltersdorf et al., 1989; Van Nostrand et al., 1989, 1990b; Smith et al., 1990) and also as an individual domain expressed in bacterial systems (Sinha et al., 1990) and yeast systems (Wagner et al., 1992). In both cases, the inhibitor profile toward trypsin-like serine proteases was similar. The Kunitz domain was a potent inhibitor of trypsin, plasmin, chymotrypsin, and notably factor XIa, all with inhibition constants of $<1 \times 10^{-9}$ M (Oltersdorf et al., 1989; Van Nostrand et al., 1989, 1990b; Smith et al., 1990). In fact, the soluble form of APP was purified from conditioned media from HepG2 cells as a factor XIa inhibitor activity and also was found in platelet releasates (Smith et al., 1990). Its potency toward factor XIa, its activation by heparin, and synthesis in platelets and liver-derived cells have prompted several laboratories to speculate that this form of APP may play a role in regulation of hemostasis. The Kunitz domain of APPH is highly similar (37/57 identical amino acids) to the Kunitz domain of APP, including strong homology surrounding the active site P1 residue and protease contact regions. It may, therefore, be expected to have similar serine protease inhibitor properties.

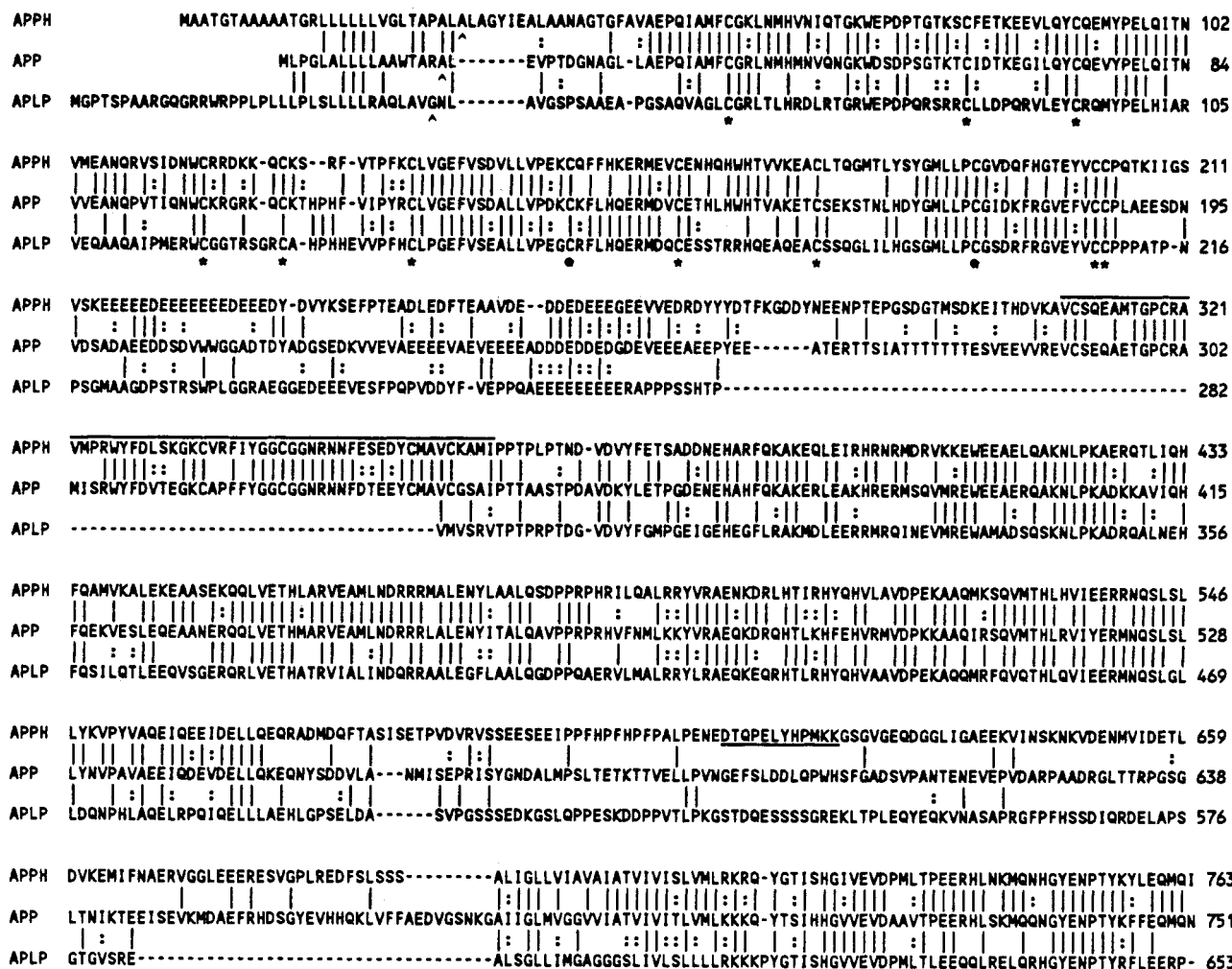


FIGURE 5: Amino acid sequence alignment of APPH, APP, and APLP. The last amino acid of each potential signal peptide is marked by a ^ symbol. Cysteines in the Cys-rich region are marked with asterisks. The Kunitz domain is indicated by a solid underline. The transmembrane domain is indicated by a wavy underline. Sequence identities with APP are indicated by vertical lines. Conservative amino acid substitutions are indicated by colons.

A conceivable role in regulation of hemostasis by the inhibitor domain of APPH is supported by our observation (data not shown) that this molecule is transcribed in human endothelial cells. Studies are underway in this laboratory to express the Kunitz domain of APPH in yeast and examine its inhibition profile.

Two additional comparisons with APP are noteworthy. The strikingly high similarity between APP and APPH breaks down specifically in one region, between amino acids 567 and 694 of APPH, just on the N-terminal side of the transmembrane domain. This is the region which encodes the pathogenic β /A4 peptide in APP, and no corresponding peptide sequence is found in APPH. It is unlikely, therefore, that APPH is similarly processed into such pathogenic products as the amyloid peptide. However, the point at which the two sequences begin to diverge occurs very near the sequence where APP has been shown to be cleavable by thrombin (Igarashi et al., 1992; Smith & Broze, 1992), which could potentially release a soluble form of the protein. In the APPH sequence a similar site is encoded three amino acids away at Asp-Val-Arg₅₈₇, which is also a potential processing site for thrombin (Le Bonniec et al., 1992). This suggests that APPH may be susceptible to thrombin generation of a similar soluble form of the protein released from the cell surface (Figure 5).

The transcript processing of APP proceeds via multiple pathways to produce different forms with or without the Kunitz

inhibitor domain (Ponte et al., 1988; Tanzi et al., 1988; Kitaguchi et al., 1988) and also to produce a truncated form which apparently encodes a soluble, secreted form of the protein (Sauvage & Octave, 1989). Several laboratories have recently shown that the full-length APP cDNA can be processed by mammalian cells in culture to produce both the amyloid peptide and also a truncated, secreted form of the extracellular protein domain (Haass et al., 1992; Seubert et al., 1992). In our studies of ten APPH clones from the placental library, we have not seen forms without the encoded Kunitz domain; however, these clones were isolated with a molecular probe corresponding to that region and may represent a selected population. Northern blot analysis of APPH mRNA in many human tissues did indicate the presence of a shorter form which could encode a soluble truncated protein.

Northern blot results (Figure 4) have indicated that APP is transcribed in a wide variety of tissues in addition to brain, including nearly every tissue examined (Tanzi et al., 1988). Similarly, we have shown that APPH is transcribed, possibly in two molecular forms, in many tissues, including lung, brain, heart, placenta, liver, skeletal muscle, kidney, and pancreas, with the highest levels in lung and brain tissue. In addition, we have evidence of its presence in a cDNA library from human endothelial cells (data not shown).

Several additional molecules have been identified in other laboratories with significant structural relationship with APP.

A partial-length cDNA called rat sperm membrane protein (RSMP) has been isolated from a rat sperm library which shows sequence homology with APP in the transmembrane and cytoplasmic domain (Yan et al., 1990) and which is dissimilar to APP in the amyloid peptide region. The remainder of the coding sequence for RSMP is currently unavailable for further comparison. Careful sequence comparison suggests that APPH may represent a full-length human version of RSMP, since the two molecules are highly similar for the extent of sequence available for RSMP, including the 3' noncoding region. If APPH is the human version of RSMP, the present, more complete characterization of the human cDNA would suggest that any specific correlation with sperm/egg adherence, as suggested by the authors, or other sperm biology would have to be extended to a more generalized role in cell adhesion, given the wide tissue distribution of the APPH molecule. APP itself has been shown to be present in testis (Shoji et al., 1990) and proposed to be involved in spermatogenesis. Secondly, a *Drosophila* protein cDNA corresponding to the *vnd* (ventral nervous system development) genomic locus has been isolated and bears a significant, but more distant relationship to APP (Rosen et al., 1989). The *vnd* protein bears sequence similarity with two regions of the extracellular domain and with the cytoplasmic domain of APP but differs in overall domain organization, having regions highly dissimilar to APP and also lacking both the Kunitz domain and the region corresponding to the β /A4 amyloid peptide. Finally, a cDNA (called APLP) recently isolated from mouse brain (Wasco et al., 1992) bears overall structural similarity and 64% sequence similarity to APP but notably lacks sequences which correspond to either the Kunitz inhibitor domain or the β /A4 peptide. Taken together, these observations suggest that APP and APPH belong to a multigene family whose members encode proteins highly similar in structure and which may share common functions. APPH bears the closest structural relationship (and presumably, therefore, functional relationship) to APP of any presently known member of this emerging family.

Given the inhibitory profile of the soluble secreted form of APP toward coagulation factors and its potential role in regulation of hemostasis, and given the nearly ubiquitous expression of APPH in highly vascularized tissues and in endothelial cells, it will be of interest to examine whether APPH itself can be processed to a soluble circulating form and whether it has inhibitory properties toward molecules involved in maintenance of hemostatic balance. To that end, studies are underway to express the cDNA for APPH in mammalian cell culture in order to examine both its biosynthesis and processing and also its protease inhibitory profile.

NOTE ADDED IN PROOF

Following the acceptance of this paper, another laboratory (Nishimoto et al., 1993) has published evidence that APP is a functional neuronal receptor which couples to intracellular signaling pathways through the GTP-binding protein G_o . Since APPH is identical to APP in 16 positions of the 20 amino acid intracellular peptide through which APP putatively interacts

with G_o , it is possible that APPH also interacts with cellular G-protein signaling pathways.

REFERENCES

- Bause, E. (1983) *Biochem. J.* 209, 331–336.
- Haass, C., Schlossmacher, M. G., Hung, A. Y., Vigo-Pelfrey, C., Mellon, A., Ostaszewski, B. L., Lieberburg, I., Koo, E. H., Schenk, D., Teplow, D. B., & Selkoe, D. J. (1992) *Nature* 359, 322–325.
- Igarashi, K., Murai, H., & Asaka, J. (1992) *Biochem. Biophys. Res. Commun.* 185, 1000–1004.
- Kang, J., Lemaire, J. G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K. H., Multhaup, G., Beyreuther, K., & Muller-Hill, B. (1987) *Nature*, 733–736.
- Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S., & Ito, H. (1988) *Nature* 331, 530–532.
- Le Bonniec, B. F., Guinto, E. R., & Esmon, C. T. (1992) *J. Biol. Chem.* 267, 6970–6976.
- Nishimoto, I., Okamoto, T., Matsuura, Y., Takahashi, S., Okamoto, T., Murayama, Y., & Ogata, E. (1993) *Nature* 362, 75–79.
- Oltersdorf, T., Fritz, L. C., Schenk, D. B., Lieberburg, I., Johnson-Wood, K. L., Beattie, E. C., Ward, P. J., Blacher, R. W., Dovey, H. F., & Sinha, S. (1989) *Nature* 341, 144–147.
- Ponte, P., Gonzalez-DeWhitt, P., Schilling, J., Miller, J., Hsu, D., Greenberg, B., Davis, K., Wallace, W., Lieberburg, I., Fuller, F., & Cordell, B. (1988) *Nature* 331, 525–527.
- Rosen, D., Martin-Morris, L., Luo, L., & White, K. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 2478–2482.
- Saito, T., Sundsmo, M., Rock, J.-M., Kimura, N., Schubert, D., Oltersdorf, T., & Schenk, D. B. (1989) *Cell* 58, 612–622.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5467.
- Sauvage, F., & Octave, J.-N. (1989) *Science* 245, 651–653.
- Seubert, P., Vigo-Pelfrey, C., Esch, F., Lee, M., Dovey, H., Davis, D., Sinha, S., Schlossmacher, M., McCormack, R., Wolfert, R., Selkoe, D., Lieberburg, I., & Schenk, D. (1992) *Nature* 359, 325–327.
- Shoji, M., Kawarabayashi, T., Harigaya, Y., Yamaguchi, H., Shunsaku, H., Kamimura, T., & Sugiyama, T. (1990) *Am. J. Pathol.* 137, 1027–1032.
- Sinha, S., Dovey, H. F., Seubert, P., Ward, P. J., Blacher, R. W., Blaber, M., Bradshaw, R. A., Arici, M., Mobley, W. C., & Lieberburg, I. (1990) *J. Biol. Chem.* 265, 8983–8985.
- Smith, R. P., & Broze, G. J. (1992) *Blood* 80, 2252–2260.
- Smith, R. P., Higuchi, D. A., & Broze, G. J. (1990) *Science* 248, 1126–1128.
- Tanzi, R. E., McClatchey, A. I., Lamperti, E. D., Villa-Komaroff, L., Gusella, J. F., & Neve, R. L. (1988) *Nature* 331, 528–530.
- Van Nostrand, W. E., Wagner, S. L., Suzuki, M., Choi, B. H., Farrow, J. S., Geddes, J. W., Cotman, C. W., & Cunningham, D. D. (1989) *Nature* 341, 546–549.
- Van Nostrand, W. E., Wagner, S. L., Farrow, J. S., & Cunningham, D. D. (1990a) *J. Biol. Chem.* 265, 9591–9594.
- Van Nostrand, W. E., Schmaier, A. H., Farrow, J. S., & Cunningham, D. D. (1990b) *Science* 248, 745–748.
- Wagner, S. L., Siegel, R. S., Vedvick, T. S., Raschke, W. C., & Van Nostrand, W. E. (1992) *Biochem. Biophys. Res. Commun.* 186, 1138–1145.
- Wasco, W., Bupp, K., Magendantz, M., Gusella, J. F., Tanzi, R. E., & Solomon, F. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 10758–10762.
- Yan, Y. C., Bai, Y., Wang, L., Miao, S., & Koide, S. S. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 2405–2408.