BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **236,** 178–183 (1997)
ARTICLE NO. RC976927

cDNA Cloning of a Novel Amphiphysin Isoform and Tissue-Specific Expression of Its Multiple Splice Variants

Ken Tsutsui,*[,1] Yukihide Maeda,† Kimiko Tsutsui,† Shuji Seki,* and Akira Tokunaga†
**Department of Molecular Biology, Institute of Cellular and Molecular Biology, and †Third Department of Anatomy, Okayama University Medical School, 2-5-1 Shikata-cho, Okayama 700, Japan*

Received May 28, 1997

Amphiphysin, an SH3-domain containing protein concentrated in nerve terminals, is believed to be involved in the synaptic vesicle recycling. We have cloned cDNAs of a novel isoform of amphiphysin (amphiphysin II) by exploiting sequence information for homologous ESTs deposited in databases. At least 9 different subtypes of the isoform with 50–60% amino acid identity to the human amphiphysin were identified by a conventional library screening and PCR amplification of cDNA libraries. Each subtype probably represents a splice variant of a single gene transcript. Analysis of mRNA expression in various tissues by RT-PCR showed that the isoform is ubiquitously distributed. The expression spectrum of the isoform subtypes, however, is significantly different in several tissues examined, suggesting that they are involved in the regulation of endocytic processes that are unique to each cell type.

Amphiphysin was discovered in 1992 as a protein partially associated with synaptic vesicles and its cDNA was first cloned from a chicken forebrain expression library (1). The human homolog of amphiphysin was then identified as the autoantigen which is recognized by sera from patients with breast cancer and Stiff-Man Syndrome, a rare neurological disorder with progressive muscle rigidity accompanied by painful spasms (2, 3). Amphiphysin has an SH3 domain which binds specifically to a proline-rich region of dynamin (4), a GTPase directly involved in internalization of clathrin-coated vesicles (5). A number of proteins that have been assigned to signal transduction pathways possess the SH3 domain (6). In addition to the SH3-mediated interaction with dynamin, amphiphysin has been shown to bind AP2, the clathrin-adaptor complex, via an unidentified region in the molecule (4, 7). Thus, amphiphysin recruits dynamin in proximity of clathrin coats, and the elevated local concentration of dynamin may facilitate the self assembly of dynamin rings at the base of budding endocytic vesicles (8). It has also been reported that the SH3 domain of amphiphysin interacts specifically with synaptojanin, a presynaptic protein with inositol-5-phosphatase activity, suggesting a link between phosphoinositide metabolism and synaptic vesicle recycling (9).

We have found an autoantigen (p130) which is reactive to the antibody produced in a patient with ductal mammary carcinoma. The patient showed neurological symptoms, mainly sensory disturbance, that are clearly distinct from the Stiff-Man Syndrome. Recently, however, p130 was identified also as amphiphysin (manuscript in preparation). Since the antibody cross-reacts with other proteins of different molecular size on immunoblots and a number of expressed sequence tags (ESTs) with moderate sequence homology to amphiphysin have been accumulated for some time in the database, we suspected that there exist an isoform of amphiphysin which might be cross-reactive to the patient's antibody. We report here the molecular structure of the isoform (designated amphiphysin II) and the presence of its multiple subtypes or splice variants that showed a tissue-specific expression.

MATERIALS AND METHODS

Database search. DNA sequences homologous to the human amphiphysin cDNA (accession number: U07616) were searched in the GenBank and dbEST databases at NCBI using the BLAST algorithm. Out of many EST hits, only ESTs showing 50–70% identity were selected. Being guided by homology, 22 ESTs were aligned along with amphiphysin, and a contig sequence (AP II1) was deduced by merging the overlapping ESTs. Conserved features in protein sequences were detected in the BLOCKS database using Block

[1] To whom correspondence should be addressed. Fax: 81-086-224-5150. E-mail: tsukken@cc.okayama-u.ac.jp.

The nucleotide sequence for the amphiphysin II (subtype 2) has been submitted to the GenBank under the accession No. AF001383. Sequences of other subtypes will also be submitted.

0006-291X/97 $25.00

Searcher (10). Potential coiled-coil regions in proteins were predicted by the COILS program (11).

cDNA cloning. The cDNA libraries used in this study were human placenta cDNA in Uni-Zap (Stratagene), human fetal brain cDNA in pCMV-SPORT2 (Gibco BRL), and rat brain stem/spinal cord cDNA in Lambda Zap II (Stratagene). About 10^7 pfu of phage particles or 5 ng of purified plasmid DNA, in case of the fetal brain library, were used as templates for PCR amplification. The Taq DNA polymerase (AmpliTaq Gold, Perkin–Elmer) was used mostly for PCR under the standard conditions recommended by the supplier. PCR products were isolated by electrophoresis in 1% agarose gels and cloned into the pGEM-T vector (Promega). Purified plasmids were subjected to sequencing reactions with a dye terminator/cycle sequencing kit (Perkin–Elmer) and sequenced on a DNA sequencer (373A, Applied Biosystems).

Original clones of the amphiphysin isoform were obtained by PCR amplification of the placenta library using a primer pair, 5′-GGGCAGTAAAGGGGTGA-3′ and 5′-TGGGACCCTCTCAGTGAA-3′, designed to span the whole coding region of AP II1. To further acquire the full length cDNA, the fetal brain library was probed with a ^{32}P-labeled *Pst*I fragment (708 bp) excised from the AP II2 insert, one of the PCR clones obtained. The colony hybridization was performed according to a standard protocol. In the first screening, 32 colonies out of 3×10^5 recombinants gave a positive signal. After the second screening, only one clone was selected for further analysis. Nucleotide sequencing was done as stated above with appropriate oligonucleotide primers and this particular clone was identified as AP II2.

Analysis of subtype expression in various tissues. The PolyATtract System 1000 and Access RT-PCR System (both from Promega) were used for mRNA purification and RT-PCR, respectively. Tissue samples (ca. 40 mg) were obtained from a male Wistar rat (7 week old) and homogenized immediately in 400 μl of the GTC extraction buffer containing 4 M guanidine thiocyanate and 2% 2-mercaptoethanol in Eppendorf tubes. Following procedures were performed according to the manufacturer's protocol. The RT-PCR mixture (50 μl) contained the reaction buffer, 5 units each of AMV reverse transcriptase/Tfl DNA polymerase, 0.2 mM dNTPs, 1 μM primers, and 10 ng of mRNA isolated from various tissues. Primers used here were AP8 (5′-AAGAAGGATGAAGCCAAAATTGC-3′), R5 (5′-CTCATGGTTCACTCTGATCTC-3′), AP10 (5′-AGATCAGAGTGAACCATGAGCC-3′), and AP9 (5′-GCCAGCTTTGAGTTGCAGCT-3′). Reactions for reverse transcription and PCR were carried out successively in a thermal cycler (GeneAmp PCR System 2400, Perkin–Elmer) using a standard program recommended by the supplier. Resulting reaction products were separated in 1.5% agarose gels and transferred with 0.4 M NaOH onto the Hybond-N+ membrane (Amersham) in a vacuum blotter. Specific amplification products were then detected by a Southern hybridization procedure using the cDNA insert of the rat AP II4 clone as a ^{32}P-labeled probe.

RESULTS

cDNA cloning of an amphiphysin isoform. Many EST entries were found to be 50–70% homologous to the amphiphysin sequence when GenBank and dbEST databases were BLAST-searched against human amphiphysin cDNA. These ESTs most likely to represent partial cDNA segments of an unknown isoform of amphiphysin. In this report we designate the novel isoform as "amphiphysin II" (AP II) and the canonical one as "amphiphysin I" (AP I). The homologous ESTs were so abundant in the database that it was possible to construct a contig, a hypothetical cDNA (AP II1), simply by splicing the overlapping ESTs. The contig contained a long open reading frame encoding a sequence of 607 amino acids with 49% identity to AP I.

To actually clone AP II cDNA we first synthesized a pair of oligonucleotides encompassing the putative coding region of AP II1 (1.8 kb) and used them as PCR primers to amplify a cDNA library from human placenta. The resulting PCR products (1.45 and 1.3 kb), although significantly shorter than the expected size, were cloned into a plasmid vector and sequenced. Analysis of the sequence for these clones, termed AP II2 and AP II3, showed that both had an uninterrupted coding frame and were nearly identical to AP II1 except for deletions in 3 positions (aligned in Fig. 2).

We then isolated a full length cDNA of AP II2 by screening a human fetal brain cDNA library using a probe excised from the AP II2 PCR clone. The cDNA (2.1 kb) contained 5′ and 3′ noncoding regions in addition to the coding sequence which was identical to that of the PCR clone isolated from the placenta library. The deduced amino acid sequence of AP II2 was compared with that of AP I and AP II1, together with other AP II2-related sequences retrieved from the database (Fig. 1). Homology between AP I and AP II is significantly high in the N-terminal 1/3, in the central portion between "e" and "f", and in the SH3 domain at the C terminus. In other regions they are largely dissimilar. Interestingly, amphiphysin II2 is identical to BIN1, a nuclear protein reported to be interacting with the MYC oncoprotein (12), except that i) the 43 amino acids between "e" and "f" are deleted in BIN1, and ii) BIN1 has an insertion of a nuclear localization signal (NLS) composed of 15 amino acids between "c" and "d". AP II2 and BIN1 are also highly homologous to the SH3P9, a mouse protein containing an SH3 domain that was recently identified by a cDNA cloning strategy involving selection of cDNA expression libraries with synthetic polypeptide ligands (13). SH3P9 is probably a mouse homolog of the human AP II3 since the sequences can be aligned throughout with only short gaps and their amino acid identity is as high as 95% (alignment not shown).

Amphiphysin II has multiple splice variants. These results strongly suggest the existence of additional subtypes in the novel amphiphysin isoform. To isolate them, therefore, cDNA libraries from human fetal brain and from adult rat brain stem/spinal cord were amplified by PCR with various primers shown in Fig. 2. More than 10 products of different sizes were isolated, cloned, and sequenced. The predicted amino acid sequences revealed that most of them belong to the AP II family. Seven clones of these are shown schematically in Fig. 2. This comparison brings about a remarkable implication that all the cDNA clones are derived from a primary transcript of amphiphysin II gene

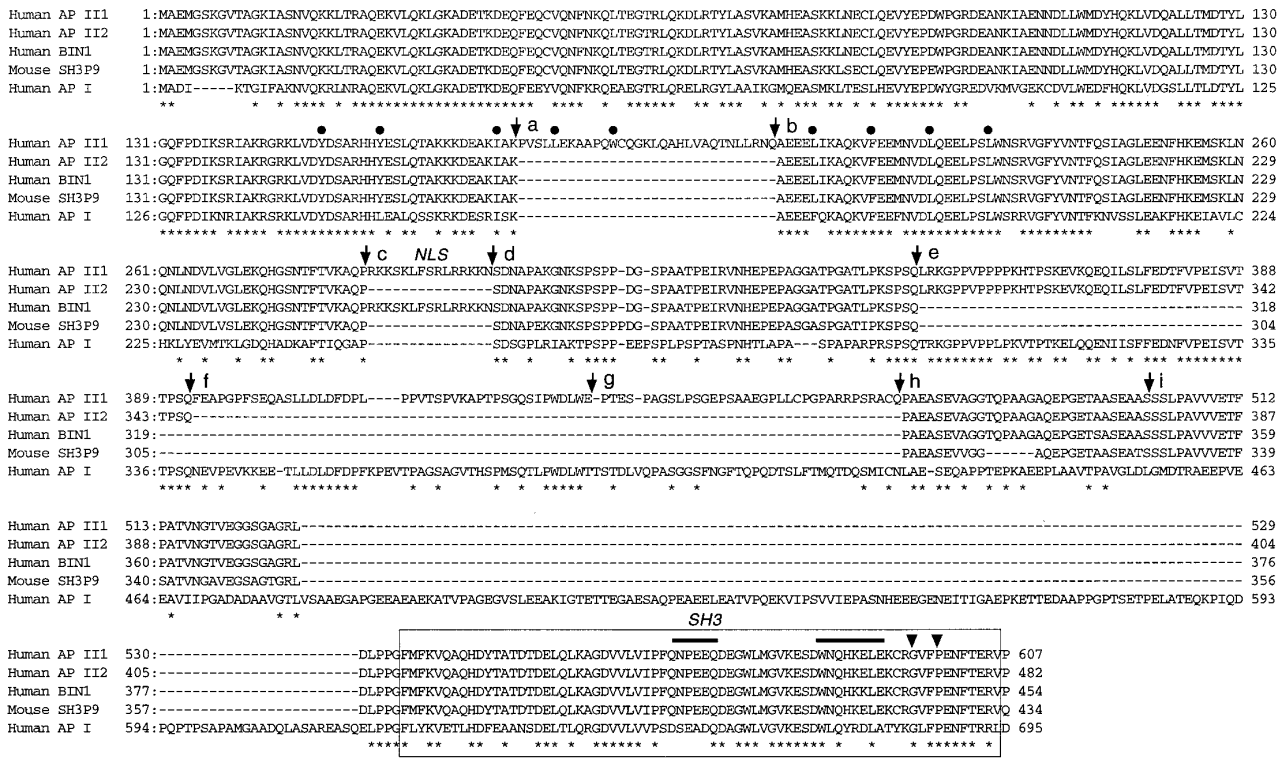

FIG. 1. Amino acid sequence alignment of amphiphysin isoforms. Note that AP II1 is the EST contig. Hyphenated positions are the gaps introduced for alignment, and identical residues in AP I and AP II1 are marked by asterisks. Sequences of other AP II2-related proteins, BIN1 (U68485) and SH3P9 (U60884), are also shown here. The nuclear localization signal (NLS) in BIN1 and the C-terminal SH3 domain (SH3) are indicated. The glycine and proline residues strictly conserved in all SH3 domain-containing proteins and the insertions characteristic of the amphiphysin SH3 are indicated by arrowheads and bars, respectively. Arrows denote putative exon–exon junctions as in Fig. 2. Heptad repeats of hydrophobic amino acids within the coiled coil region are marked by dots.

through an alternative splicing mechanism (see Discussion for evidence). Nucleotide sequences for corresponding regions of these clones and AP II1 were the same except for a few conflicts probably due to PCR errors and, in some cases, polymorphism. The insertion sequence found in F5/R4-b differed from the g-h segment of other clones and AP II1, suggesting that more than 2 alternative exons can be inserted at this position. We have not detected a clone corresponding to AP II1 although the subtype S1/R3-a may well represent a partial segment of AP II1 which is devoid of the NLS exon. Since all the patterns of putative exon organization are unique to each clone, we conclude that AP II consists of more than 10 variants or subtypes.

Tissue-specific expression of AP II subtypes. Ubiquitous expression of AP II mRNA in various tissues has already been shown by a Northern blot analysis with BIN1 cDNA probe which should detect most of the messages for AP II subtypes (see Fig. 5a of Ref. 12). To investigate the expression of individual AP II subtypes, poly $(A)^+$ RNA was isolated from rat tissues and subjected to RT-PCR analysis (Fig. 3). The variants with different exon organization between the primer sites are expected to give products of different length. Two pairs of primers were selected to amplify different cDNA regions. With primers AP8/R5 that are targeted to a central portion, a dominant product of the same size (365 bp) was produced in all the tissues tested, except for retina and skeletal muscle (marked by arrow in Fig. 3A). In all the nerve tissues (lanes 1–5), an additional band of 460 bp was present (double arrowhead). These major products are likely to be generated from the subtypes homologous to the human counterparts without and with the exon a-b, respectively, since their sizes coincide well each other. A band with intermediate size (435 bp) was detected only in skeletal muscle (oblique arrow). In some tissues larger products of 1 kb< were seen, implying that insertional events of unknown exon(s) may occur between these primer sites. The largest band in nervous tissues (arrowhead) is slightly larger than that found in other tissues (open arrowhead). The primers targeted for a C-terminal portion (AP10/AP9) gave a dominant product of 248 bp which is probably derived from a subtype with the same exon organization as in the rat AP II4 (arrow in Fig. 3B). The 690 bp band detected in the cerebral cortex

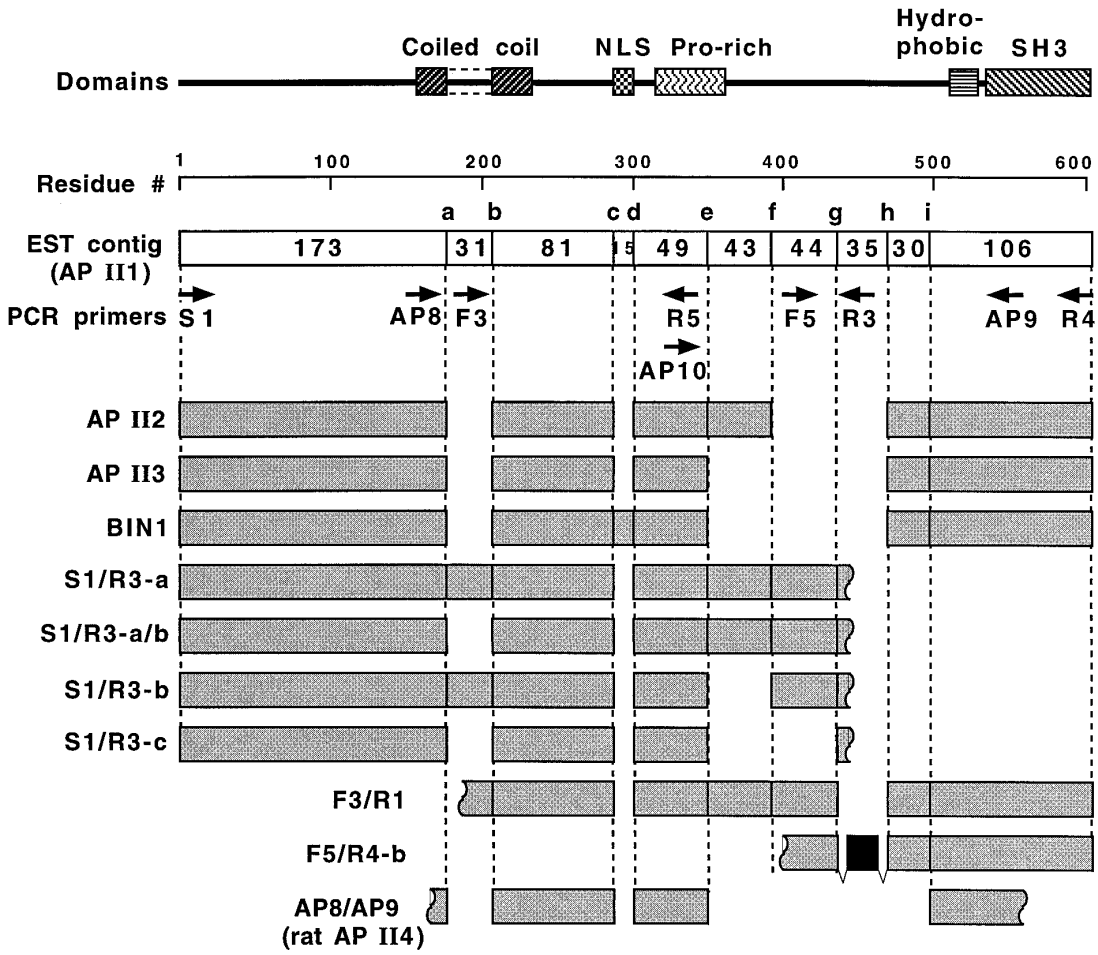

FIG. 2. Summary of amphiphysin II subtypes. Shown here is a schematic representation of amino acid sequences. The uppermost sequence shown by white bars is a hypothetical isoform II (AP II1) reconstituted from overlapping ESTs. Putative exon–exon junctions are labeled with roman letters and numbers inside the bar designate length of exons. Note that the bars separated by gaps actually constitute a consecutive sequence. Wavy ends illustrate that these are the clones truncated at the position of PCR primers. Except for AP II2 and BIN1, shaded sequences were obtained by PCR amplification of cDNA libraries either from human placenta (AP II3), rat brain stem/spinal cord (AP8/AP9), or from human fetal brain (all the others). PCR clones are named after primers used for amplification. Location of the primer "R1" is not shown because it resides in the 3′ noncoding region. The filled bar in F5/R4-b depicts the insertion of an alternative exon which is absent from other subtypes. Structural domains predicted from the AP II1 sequence are shown on the top. The potential coiled coil region shown here adopts this conformation with a probability higher than 0.9.

and the pons (arrowhead) is comparable in size to a product expected from the hypothetical subtype AP II1. These results are consistent with the notion that amphiphysin II is expressed ubiquitously in various tissues although its subtype composition shows a significant tissue specificity.

DISCUSSION

We have demonstrated through cDNA cloning that messages for a novel family of proteins homologous to amphiphysin, thus termed amphiphysin II (AP II), is expressed in a wide variety of tissues. Overall sequence identity between AP II2, one of the subtypes identified in this study, and amphiphysin I was about 59% at the amino acid level. As shown in Figs. 1 and 2, the colinearity between AP I and AP II is disrupted by 2 alternative insertions into isoform II (between the labels "a/b" and "c/d") and by the lack of homology and large deletion in AP II within the region between "g" and the SH3 domain, where a sequence divergence has also been noted between human and chicken amphiphysin I (3). The SH3 domain identified in AP I is longer than that of other proteins because of insertions in 2 positions (3,13). These insertions are also present in AP II and the amino acid sequence in the SH3 domain, including the invariable glycine and proline residues, is relatively conserved between the isoforms (Fig. 1). Whether the SH3 domain of AP II also interacts with dynamin/synaptojanin or it has different target

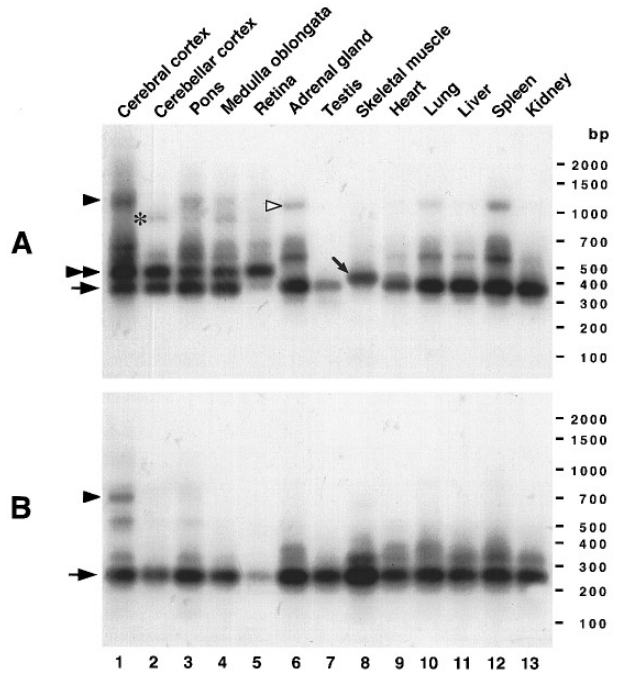

FIG. 3. Tissue-specific expression of amphiphysin II subtypes. Poly $(A)^+$ RNAs isolated from various rat tissues (indicated on the top) were subjected to RT-PCR analysis with primer pairs AP8/R5 (panel A) or AP10/AP9 (panel B). After electrophoresis in agarose gels, specific amplification products were detected by Southern blotting with a probe prepared from the rat AP II4 clone. Autoradiograms are shown here. Migration of a ladder marker fragments are designated on the right. Other symbols in this figure are explained in the text.

proteins remains to be clarified. Other structural features of AP I are also present in AP II (Fig. 2, top). The region with clustered prolines might be involved in the binding with SH3 domains of other proteins. Although AP II does not have the homologous exon in which the hydrophobic stretch conserved in human and chicken AP I has been located, AP II does possess a stretch of similar hydrophobicity right next to the SH3 domain. Both amphiphysin isoforms contain a putative coiled coil domain (ca. 50 amino acids) which can be involved in protein-protein interaction. A distinct heptad repeat of hydrophobic amino acids is discernible within this region (Fig. 1). In some of the AP II subtypes, the region is divided into 2 portions by the insertion of 31 amino acids which is detrimental to the coiled coil conformation (Fig. 2).

The most striking feature of AP II is the presence of multiple subtypes that are likely to represent splice variants. The following discussion may support this notion. Even in the absence of knowledge on the genomic sequence, every junction between putative exons could be deduced by aligning nucleotide sequences for the subtypes listed in Fig. 2. In 8 out of 9 exon boundaries, the last 2 nucleotides at the 3′ ends of exons were AG (data not shown), being consistent with the consensus sequence formulated from many genes (14). There is a general preference for intron insertion at a codon boundary, as opposed to the middle of a codon, that has been ascribed to a consequence of gene evolution through the exon shuffling mechanism (14, 15). In 6 junctional sites of AP II, the point of intron insertion occurs at the codon boundary. The frequency (6/9=66%) is significantly higher than the statistical expectation for mammalian genes (43%, Ref. 14). This trait might be somehow related to the extensive splicing events involved in the expression of AP II gene. The splicing pattern is probably conserved at least between man and rodents.

Recent inspection of sequence databases revealed cDNA sequences for a human AP II variant (U87558, Ref. 16) and a putative mouse homolog with 94% amino acid identity (U86405). Interestingly, the human variant is nearly identical to AP II1 without NLS, except for the latter half of the "exon g-h" which differs completely (not shown). Including the unique sequence found in the clone F5/R4-b, therefore, this region appears to accommodate at least 3 alternative exons. How many more subtypes will be identified eventually? Our results suggest that the AP II gene contains more than 10 exons. Assuming that every exon is dispensable and the splicing event is independent each other, a combinational estimate for a total number of subtypes easily exceeds 10^3. Since some exons must be essential, this is clearly an overestimation. However, considering the fact that over 10 variants have been identified without any systematic searches, we could certainly expect additional splice variants to be found. Neurexins, a family of cell surface proteins specific to brain, is an extreme example of this kind since several hundreds of variants have been suggested to be created by extensive alternative splicing (17).

The subtype expression spectrum in each tissue is clearly different. The RT-PCR analysis unequivocally detected subtypes that are specific to nervous tissues and to skeletal muscle (Fig. 3A). Furthermore, the product pattern was also different among nervous tissues: the 365 bp band was negligible in retina and the large band noted in the cerebral cortex was absent from the cerebellar cortex which gave, instead, a smaller product (asterisk in Fig. 3A). We speculate that different AP II subtypes are expressed in each cell type through a regulated splicing mechanism. Therefore, a subset of neurons may well express an unique form of AP II. In order to clarify this situation, experiments with combinations of antipeptide antibodies that are specific to individual exons will be required.

Although we have no information on proteins interacting with AP II, it would be reasonable to assume that AP II is also involved in endocytic pathways by

interacting with dynamin in non-neural tissues where AP I is not expressed. However, AP II may have additional partners that are most likely to be subtype specific. In fact, a splice site-specific partner for neurexins has been identified (18). Presence of AP II subtypes with different organizations in the putative protein-recognition domains, such as coiled coils, is consistent with this idea. Thus AP II, as a family of proteins, can serve as an adaptor for a variety of protein partners, facilitating interaction between the partner and the protein bound to the SH3 domain. Physiological roles of AP II will be established by elucidating its partners which may possibly include the proteins directly involved in signal transduction pathways. Whether AP II can also be an autoantigen in some paraneoplastic conditions remains to be demonstrated.

ACKNOWLEDGMENTS

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan. We thank M. Ishimaru for technical assistance.

REFERENCES

1. Lichte, B., Veh, R. W., Meyer, H. E., and Kilimann, M. W. (1992) *EMBO J.* **11,** 2521–2530.
2. De Camilli, P., Thomas, A., Cofiell, R., Folli, F., Lichte, B., Piccolo, G., Meinck, H-M., Austoni, M., Fassetta, G., Bottazzo, G., Bates, D., Cartlidge, N., Solimena, M., and Kilimann, M. W. (1993) *J. Exp. Med.* **178,** 2219–2223.
3. David, C., Solimena, M., and De Camilli, P. (1994) *FEBS Lett.* **351,** 73–79.
4. David, C., McPherson, P. S., Mundigl, O., and De Camilli, P. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 331–335.
5. Takei, K., McPherson, P. S., Schmid, S. L., and De Camilli, P. (1995) *Nature* **374,** 186–190.
6. Pawson, T. (1995) *Nature* **373,** 573–580.
7. Wang, L. H., Südhof, T. C., and Anderson, R. G. W. (1995) *J. Biol. Chem.* **270,** 10079–10083.
8. Bauerfeind, R., Galli, T., and De Camilli, P. (1996) *J. Neurocytol.* **25,** 701–715.
9. McPherson, P. S., Garcia, E. P., Slepnev, V. I., David, C., Zhang, X., Grabs, D., Sossin, W. S., Bauerfeind, R., Nemoto, Y., and De Camilli, P. (1996) *Nature* **379,** 353–357.
10. Henikoff, S., and Henikoff, J. G. (1994) *Genomics* **19,** 97–107.
11. Lupas, A. (1996) *Meth. Enzymol.* **266,** 513–525.
12. Sakamuro, D., Elliott, K. J., Wechsler-Reya, R., and Prendergast, G. C. (1996) *Nature Genet.* **14,** 69–77.
13. Sparks, A. B., Hoffman, N. G., McConnell, S. J., Fowlkes, D. M., and Kay, B. K. (1996) *Nature Biotechnol.* **14,** 741–744.
14. Tomita, M., Shimizu, N., and Brutlag, D. L. (1996) *Mol. Biol. Evol.* **13,** 1219–1223.
15. Long, M., Rosenberg, C., and Gilbert, W. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 12495–12499.
16. Butler, M. H., David, C., Ochoa, G. C., Freyberg, Z., Daniell, L., Grabs, D., Cremona, O., and De Camilli, P. (1997) *J. Cell Biol.*, in press.
17. Ullrich, B., Ushkaryov, Y. A., and Südhof, T. C. (1995) *Neuron* **14,** 497–507.
18. Ichtchenko, K., Hata, Y., Nguyen, T., Ullrich, B., Missler, M., Moomaw, C., and Südhof, T. C. (1995) *Cell* **81,** 435–443.