Characterization of a Novel Synapse-specific Protein. II. cDNA Cloning and Sequence Analysis of the F1-20 Protein

Shibin Zhou, Rui Sousa, Nancy Hrinya Tannery, and Eileen M. Lafer

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

The F1-20 protein is a novel neuronal-specific, synapse-associated protein that is expressed nonuniformly in mouse brain. Expression of the F1-20 protein is developmentally regulated in a pattern coincident with active synaptogenesis and synaptic maturation. Here we report the cloning of the cDNA sequence for the F1-20 protein. We found two distinct isoforms of F1-20 cDNA that differed by the presence of 15 additional nucleotides, which does not interrupt the open reading frame. RNase protection analysis and PCR amplification of mouse brain RNA revealed that both isoforms are present in cellular RNA. It is likely that the two F1-20 mRNA isoforms are derived from RNA splicing events utilizing alternative 3' acceptor sites. Analysis of the deduced amino acid sequence for the complete open reading frame revealed that the predominant F1-20 mRNA encodes an 896 amino acid polypeptide with a molecular weight of 91,319 Da. The deduced amino acid sequence does not contain a signal sequence, or any extensive hydrophobic regions. The deduced amino acid sequence does contain a number of consensus sequences for protein kinases. Searches of the protein and nucleic acid sequence data bases revealed that the F1-20 protein has not been previously characterized at the primary structure level, although a weak similarity was found between rabbit calpastatin and the C-terminal portion of the F1-20 protein. We then determined biochemically that the F1-20 protein is a substrate for Ca²⁺-dependent proteolysis, which is specifically inhibited by calpain inhibitors in vitro. This indicates that the F1-20 protein is a substrate for neuronal calpain. We observed that treatment of a synaptosomal lysate with alkaline phosphatase led to an increase in the electrophoretic mobility of the F1-20 protein, as well as to an increase in the sharpness of the electrophoretic band. This indicates that the F1-20 protein is phosphorylated in vivo.

The primary structures of many neurotransmitters and their receptors have been determined. However, the molecular events involved in synaptic vesicle synthesis, storage, release, and reuptake, as well as in activity-dependent changes in synaptic efficacy, are less clearly understood (reviewed in Kelly, 1988; Sudhof and Jahn, 1991; Trimble et al., 1991). In order to reach an understanding of these molecular events, a number of groups have been biochemically dissecting the nerve terminal. The isolation and sequencing of cDNA clones have allowed the determination of the primary structures of synapsin I a,b/protein I a,b (Sudhof et al., 1989b), synapsin II a,b/protein III a,b (Sudhof et al., 1989b), synaptophysin/p38 (Buckley et al., 1987; Leube et al., 1987; Sudhof et al., 1987), VAMP-1,2/synaptobrevin (Trimble et al., 1988, 1990; Elferink et al., 1989; Sudhof et al., 1989a), synuclein (Maroteaux et al., 1988), VAT-1 (Linial et al., 1989), synaptoporin (Knaus et al., 1990), p65/synaptotagmin (Perin et al., 1991a,b), and SNAP-25 (Oyler et al., 1989).

In the accompanying article (Sousa et al., 1992), we reported the identification of the novel synaptic molecule, the F1-20 protein. We identified this protein by immunohistochemically screening a monoclonal antibody (MAb) library for synapse-associated proteins. We demonstrated that the F1-20 protein is synapse-associated by both immunohistochemical and biochemical methods. We found that the F1-20 protein and mRNA are neuronal specific and developmentally regulated in a pattern coincident with a period of active synaptogenesis and synaptic maturation. We characterized the neuroanatomical distribution of the F1-20 protein and mRNA. The expression patterns were nonuniform and complex. The highest levels of F1-20 protein were found in the cortically organized regions of the brain. The highest levels of F1-20 mRNA were found in long-distance projection neurons.

In this article, we utilized the F1-20 MAb as a probe to screen a mouse brain cDNA expression library. We went on to isolate a series of overlapping cDNA clones to allow the construction of a composite cDNA sequence containing the full-length open reading frame. We present an analysis of the F1-20 cDNA sequence, and of the deduced amino acid sequence of the F1-20 protein. This analysis led to some predictions about the biochemical properties of the F1-20 protein. We present experiments designed to test these predictions. The availability of the F1-20 cDNA sequence, and the characterization presented here, will make possible future investigations into the role of the F1-20 protein in synaptic function.

Materials and Methods

cDNA cloning, sequencing, and analysis

Expression cloning. A λgt11 cDNA clone expressing a fusion protein reactive with the F1-20 MAb was isolated by screening a mouse brain cDNA library (library ML1002, made from Swiss–Webster adult brain by Clontech Laboratories, Inc., Palo Alto, CA) with the F1-20 MAb according to the procedure described in Huynh et al. (1985), except that the filters were screened using the protocol described under "Western blot analyses" in Sousa et al., 1992. The immunoreactive clone was plaque purified to homogeneity, and subcloned into the Bluescript vector pBS-m⁻ (Stratagene Cloning Systems, San Diego, CA). Double-stranded DNA (ds DNA) sequencing was carried out by the dideoxy method of Sanger et al. (1977) using modified T7 DNA polymerase (Tabor and Richardson, 1987).

Isolation and sequence determination of overlapping cDNA clones. Mouse brain cDNA libraries [oligo(dT)-primed λgt11 (Clontech library ML 1002), random-primed λgt10 (Clontech ML1024a), and λgt11-P10 (J. Merlie)] were screened for multiple rounds with nucleic acid probes (Ausubel et al., 1987). The filter hybridization using nitrocellulose (Schleicher & Schuell) was performed following the manufacturer's instructions with the high-stringency wash at 65°C in 0.1× SSPE/0.1% SDS. The positive cDNA clones were then subcloned into Bluescript phagemid (Stratagene) or M13 phage (New England Biolabs) for sequence analyses. cDNA clones RP1-3, RP1-4, and RP1-18 were obtained using an RNA-PCR technique developed by Frohman et al. (1988). Total cellular RNA (10–30 μg) isolated from mouse brain was reverse transcribed using an oligo-primer (5' CGCCGTATGAACGT-AGACAT 3') complementary to a sequence near the putative translation start codon and avian myoblastosis virus (AMV) reverse transcriptase (Promega). An oligo(dA)-tail was then added to the first-strand cDNA by terminal deoxynucleotidyl transferase [Bethesda Research Laboratories (BRL)]. The oligo(dA)-tailed cDNA was amplified by PCR using an oligo-primer (5' ATCATAACCGTGGGATCCAC 3') located immediately upstream from the primer described above and an oligo(dT)-primer with some restriction sites at its 5'-terminus. The PCR product was then cloned into the M13 vector. Double-stranded or single-stranded DNAs (ssDNAs) were sequenced according to the dideoxy chain-termination method (Sanger et al., 1977) with Sequenase (United States Biochemical Corp.).

Sequence analysis. The sequence was analyzed using the MacFASTA package of Pearson and Lipman (1988), the FASTA-mail service of BIO-NET, and the FASTA program running at the University of Pittsburgh supercomputer.

Polymerase chain reactions

To obtain the cDNA encoding the entire F1-20 open reading frame (ORF), total cellular RNA isolated from mouse brain was reverse transcribed with AMV reverse transcriptase (Promega) and oligo-primer SZ401 (5' CTCGGATCCCGGGTTTGGTACTGGTGG 3'). The first-strand cDNA was then used as the template for PCR. The primers were SZ327 (5' ATGGCACTGATGTCGGGCCAAACGCTC 3') and SZ316 (5' CTCGCGGCCGCGAGCTCTTATCTGAAGTTTCC 3'). PCR was performed with Taq DNA polymerase (Promega) for 35 cycles each consisting of 1 min at 94°C, 1 min at 50°C, and 4 min at 72°C, which was then followed by a final 7 min incubation at 72°C. The PCR product was then resolved in a 1% agarose gel. For demonstration of an alternative splicing event, the first-strand cDNA synthesis and PCR were carried out as described above with the following exceptions. Primers used for PCR were SZ58 (5' GTCCTCGGGTGTCATA 3') and SZ401 (see above). The reaction was done for 35 cycles with 40 sec at 94°C, 20 sec at 45°C, and 1 min at 72°C, which was then followed by an additional 5 min at 72°C. After extraction with phenol/chloroform and precipitation with ethanol, the PCR product was cut with restriction endonuclease StuI (BRL) and resolved by agarose gel electrophoresis.

Antibodies

The preparation of MAb F1-20 was described previously (Sousa et al., 1990), and further characterized in the accompanying article (Sousa et al., 1992). The heterologous antibody control was a MAb raised against a 100 kDa *Escherichia coli* Z-DNA binding protein (Lafer et al., 1988) that does not cross-react with murine brain tissue. Polyclonal antisera were raised against the β-gal–F1-20 fusion protein as follows. The λgt11 clone carrying the 1.2 kilobase (kb) F1-20 cDNA was grown lysogeni-

cally in *E. coli*. The lysogen was induced to make the β-gal–F1-20 fusion protein according to the protocol of Huynh et al. (1985). The cells were lysed and the soluble fraction was electrophoresed on an SDS poly-acrylamide gel. The band containing the fusion protein was excised from the gel and used as an immunogen to generate polyclonal sera in rabbits. The anti-cDNA-encoded fusion protein antiserum was purified by binding to, and subsequent elution from, nitrocellulose filters prepared from isopropyl-1-thio-β-D-galactopyranoside (IPTG)-induced lysogens of the λgt11 clone carrying the 1.2 kb F1-20 cDNA (Huynh et al., 1985).

RNase protection analysis

RNase protection analyses were carried out as described by Chen et al. (1989). Using one of the cDNA clones that includes the mini-exon as a DNA template (R4-6), ³²P-labeled antisense RNA probe was transcribed *in vitro* by T3 RNA polymerase. After overnight hybridization with mouse brain cellular RNA at 50–55°C, T2 RNase digestion was performed at 35°C for 1.5 hr. The reaction mixture was then separated on a 6% sequencing gel containing 8 M urea. The gel was dried and exposed to Kodak XAR-5 film. The M13 DNA sequencing ladder that served as a marker was synthesized as described by Sanger et al. (1977).

Southern blot analysis

DNA isolation, Southern transfer, and Southern hybridization (Southern, 1975) were performed as described in Maniatis (1982). Genomic DNA was isolated from the liver of a 3-month-old Balb/c female mouse. The DNA was digested with either SstI, PstI, HindIII, BamHI, or EcoRI. The restriction fragments were separated on a 0.8% agarose gel, transferred to nitrocellulose, and hybridized to ³²P-R3-11 at 42°C in 5× SSPE, 5× Denhardt's, 0.1% SDS, 100 μg/ml ssDNA, followed by extensive washing. The final wash was in 0.1× SSPE at 65–68°C. Restriction fragments were sized against λHindIII and λBstEII molecular weight markers (BRL).

Proteolysis by neuronal calpain

The P2' fraction ("well-washed crude synaptosomes") was generated by the procedure of Huttner et al. (1983), and lysed in the standard buffer (+0.17% 2-mercaptoethanol) also containing either 0.1 mM EGTA, 0.1 mM EGTA + 1.1 mM CaCl₂, 0.1 mM EGTA + 1.1 mM CaCl₂ + calpain-specific inhibitor N-acetyl-leucyl-leucyl-norleucinal (calpain inhibitor I), 0.1 mM EGTA + 1.1 mM CaCl₂ + the calpain-specific inhibitor N-acetyl-leucyl-leucyl-methioninal (calpain inhibitor II), 0.1 mM EGTA + 1.1 mM CaCl₂ + the thiol protease inhibitor E64, 0.1 mM EGTA + 1.1 mM CaCl₂ + the thiol/serine protease inhibitor leupeptin, 0.1 mM EGTA + 1.1 mM CaCl₂ + the serine protease inhibitor phenolmethyl-sulfonyl fluoride (PMSF), or 0.1 mM EGTA + 1.1 mM CaCl₂ + the acid protease inhibitor pepstatin, incubated 60 min on ice, followed by centrifugation for 20 min at 25,000 × g to generate LS1. The supernatants were all concentrated by a factor of 25, analyzed by SDS-PAGE on 7.5% Phast-gels (Pharmacia), followed by Western blot analyses [all of these methods are detailed in the accompanying article (Sousa et al., 1992)].

Dephosphorylation of the F1-20 protein

The P2' lysate was prepared according to Huttner et al. (1983) and adjusted to 10 mM Tris, pH 8.5, 1 mM dithiothreitol (DTT), 130 μM E64, 130 μM leupeptin, 13 μM pepstatin, 130 μM PMSF, 130 μM calpain inhibitor I, and 130 μM calpain inhibitor II. This lysate was incubated on ice for 30 min and processed for the preparation of LS2. Fractions were then incubated in 10 μl reactions containing either buffer P (10 mM Tris, pH 8.5, 1 mM DTT, 130 μM E64, 130 μM leupeptin, 13 μM pepstatin, 130 μM PMSF, 130 μM calpain inhibitor I, and 130 μM calpain inhibitor II) alone, 20 U of calf intestine alkaline phosphatase (Boehringer) in buffer P, buffer P plus 10 mM EDTA and 10 mM NaH₂PO₄, or 20 U of alkaline phosphatase in buffer P plus 10 mM EDTA and 10 mM NaH₂PO₄. Incubations were carried out for 1 hr at 37°C, and the reactions analyzed by SDS-PAGE on 7.5% Phast-gels (Pharmacia), followed by Western blot analyses.

pI determination

The pI of the F1-20 protein was determined by Western analysis of two-dimensional "O'Farrell" gels (O'Farrell, 1975). The apparatus used was the Protean II slab gel system from Bio-Rad Laboratories and the tube gel apparatus supplied with this apparatus. Two-dimensional elec-
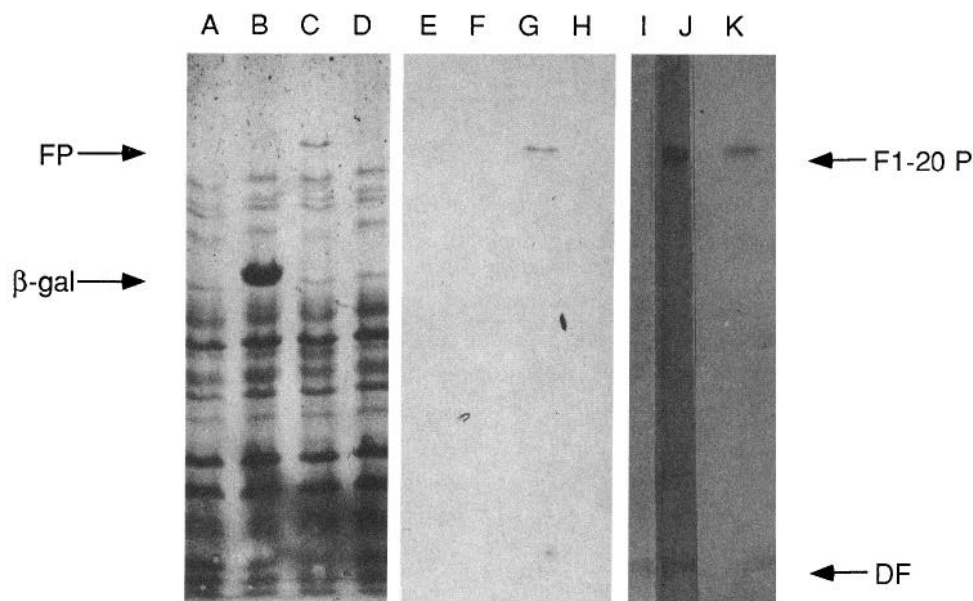
Figure 1. Isolation and expression of a partial cDNA clone encoding the F1-20 protein. Loaded on *lanes A–H* are total protein from uninduced *E. coli* carrying a nonrecombinant λgt11 lysogen (*lanes A* and *E*), total protein from *E. coli* carrying a nonrecombinant λgt11 lysogen IPTG induced to express β-gal (*lanes B* and *F*), protein from *E. coli* carrying the F1-20–immunoreactive λgt11 lysogen IPTG induced to express the fusion protein (*FP*) (*lanes C* and *G*), protein from uninduced *E. coli* carrying the F1-20–immunoreactive λgt11 lysogen (*lanes D* and *H*). Proteins in *lanes A–H* were separated by SDS-PAGE and blotted onto nitrocellulose. In *lanes A–D*, total protein was visualized by staining with india ink. The blots of *lanes E–H* were probed with the F1-20 MAb to visualize F1-20–immunoreactive proteins. *Lanes I–K* show strips cut from a single Western blot from an SDS gel on which total mouse brain protein had been run. *Lane I* was probed with preimmune serum, *lane J* was probed with the anti-fusion protein antiserum, and *lane K* was probed with the F1-20 MAb. The strips are aligned along their dye fronts (*DF*).

trophoresis followed the manufacturer's instructions and the procedure described in O'Farrell (1975) with the following minor modifications. For the sample preparation, the sample run on the first-dimension iso-electric focusing (IEF) gel was the LS2 supernatant prepared as described in the accompanying article (Sousa et al., 1992) except that it was concentrated 100× (rather than 25× as described) by centrifugation in Amicon Centricon concentrators. Fifty microliters were loaded onto 1.5 mm (i.d.) × 18 cm tube gels. The first-dimension gel was composed of 4.86 gm of urea; 2.88 gm of H_2O; 1.18 ml of 30% acrylamide, 0.8% bis; 2.03 ml of 10% Triton X-100; and 0.5 ml of ampholytes. The ampholytes (Pharmacia) were a mixture of 6 parts pH range 2.5–4, 3 parts pH range 3.5–5, and 1 part pH range 4–6. The protein sample was first saturated with solid urea, and then 1 µl of IEF sample buffer [1.7% SDS (w/v), 33% ampholytes (v/v), 17% 2-mercaptoethanol (v/v), 34% Triton X-100] was added to every 5 µl of sample. The IEF gel was then run for 12 hr at 400 V. An important modification to the recommended procedure at this point was the use of a solution of 0.15% H_2SO_4 as the anode buffer rather than the 0.06% phosphoric acid solution recommended by the manufacturer (cathode buffer was 0.1 M NaOH as recommended by Bio-Rad). We found this essential to obtain a linear pH gradient in the very low pH range in which the F1-20 protein was resolved. After electrofocusing, the tube gels were extruded and run on 1.5 mm × 20 cm × 20 cm second-dimension Laemmli SDS gels following the protocol of O'Farrell (1975). Gels were then electroblotted onto nitrocellulose and probed by Western analysis as described in the accompanying article (Sousa et al., 1992).

Results

cDNA cloning of the F1-20 protein

The F1-20 MAb detected a single band with an apparent molecular weight of 190,000 Da in a mouse brain homogenate by Western blot analysis, implying that the F1-20 MAb reacts with a single protein and is negligibly cross-reactive with other proteins in these extracts (Sousa et al., 1992). This suggested that it would be a suitable probe for the isolation of a cDNA clone encoding the F1-20 protein from an appropriate expression li-

brary. To characterize this protein further, we screened a mouse brain λgt11 cDNA expression library with the F1-20 MAb. A clone showing strong reactivity with the F1-20 MAb was identified and used to construct a lysogen for the preparation of fusion protein. Figure 1 shows a Western blot analysis of this expressed fusion protein. The lysogen containing this clone overexpresses a protein larger than the β-gal protein expressed from the λgt11 parent (compare lanes C and B). Expression of this larger fusion protein is dependent on IPTG induction (compare lanes C and D). When proteins from induced (lanes F, G) and uninduced (lanes E, H) cells of a lysogen carrying a λgt11 clone lacking an insert (lanes E, F) or carrying the putative clone for the F1-20 protein (lanes G, H) are probed with the F1-20 MAb, a single immunoreactive band is seen corresponding to the overexpressed fusion protein from this putative clone. This suggests that we have in fact isolated a cDNA encoding the F1-20 protein.

In order to confirm that the clone we had isolated did indeed encode the F1-20 protein that we had been characterizing, the putative F1-20 fusion protein was isolated from SDS gels and was used as an immunogen to raise antisera in rabbits. When this anti-cDNA–encoded fusion protein antiserum was used to probe Western blots of total mouse brain protein (Fig. 1, lane J), it reacted with a band indistinguishable in molecular weight from the band recognized by the F1-20 MAb (Fig. 1, lane K). In other experiments the anti-cDNA–encoded fusion protein antiserum displayed patterns of immunoreactivity identical to the F1-20 MAb. For example, the immunohistochemical staining pattern in the brain of this antisera is identical to the pattern of the F1-20 MAb (Sousa et al., 1992) (Fig. 2). In the cerebellar cortex (Fig. 2A), there is no immunoreactivity in the white matter, or in the granule cell and Purkinje cell bodies. There is

strong punctate staining of synaptic glomeruli in the granule cell layer, and dense staining of the molecular layer. The puncta in the granule cell layer are well illustrated in the higher-magnification view (Fig. 2*B*). In the hippocampus (Fig. 2*C*), there is no immunoreactivity in the white matter, or in the granule cell bodies of the granule cell layer of the dentate gyrus. There is strong staining of the synapse-rich dentate gyrus molecular layer and hillus. We also examined the tissue specificity and developmental specificity of the anti-cDNA–encoded fusion protein antiserum by Western blot analysis (data not shown). A single immunoreactive band with an apparent molecular weight of 190,000 Da was seen in brain extracts, but no immunoreactivity was seen in extracts prepared from liver, spleen, kidney, heart, or lung. This is identical to the tissue specificity observed with the F1-20 MAb (Fig. 4*C* of Sousa et al., 1992). When Western blots of brain extracts prepared from mice at postnatal days 3–16 were stained with the anti-cDNA–encoded fusion protein antiserum, a sharp increase in expression to apparently adult levels was seen on day 7, as was observed with the F1-20 MAb (Fig. 6*C* of Sousa et al., 1992). Because the fusion protein is both immunoreactive with the F1-20 MAb and generates an

antisera with a specificity that is indistinguishable from the F1-20 MAb, we conclude that this clone encodes the F1-20 protein.

In order to obtain the complete cDNA sequence of F1-20, mouse brain cDNA libraries were repeatedly screened with the original 1.2 kb F1-20 cDNA and subsequently isolated cDNA clones as the probes. The "RACE" technique developed by Frohman et al. (1988) was also employed to isolate the clones that cover the beginning of the coding sequence. A partial restriction map of the full-length cDNA sequence, the overlapping cDNA clones most important for the construction of the full-length cDNA sequence, and the sequencing strategy are presented in Figure 3. The entire ORF was sequenced at least twice using independent cDNA clones to ensure the correctness of the sequence information obtained.

Sequence and structural features of F1-20 cDNA

A composite nucleotide sequence (including the mini-exon) was generated by aligning the sequences of the overlapping clones indicated in Figure 4. The predicted amino acid sequence contains a putative translation initiation methionine (ATG) at position 1 following an in-frame stop codon (TAG) 6 codons up-

Figure 3. Composite cDNA map and sequencing strategy for the F1-20 cDNA clones. The 5'- and 3'-UTRs are indicated by *horizontal lines.* The coding region is represented by an *open bar.* The restriction sites utilized for cDNA cloning or sequencin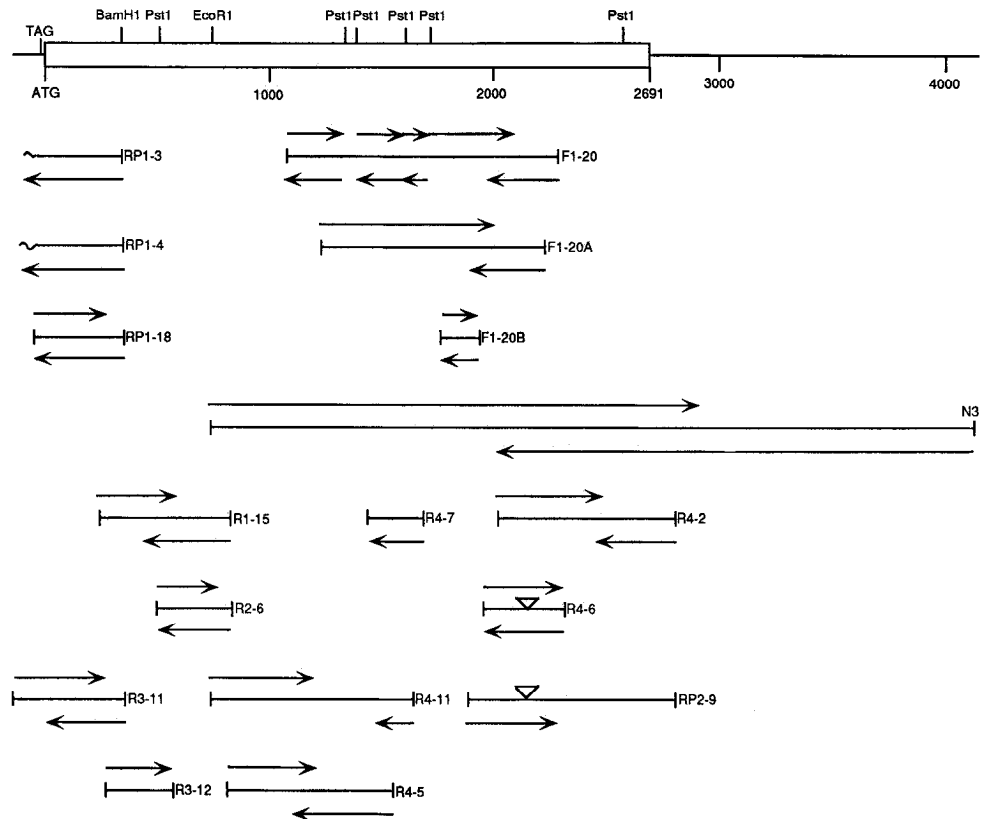g are shown. Also indicated is the upstream in-frame stop codon (TAG). The locations of the cDNA clones used for sequencing are shown. The directionality and extent of sequencing are represented by *arrows.* The position of the alternatively spliced mini-exon in some of the cDNA clones is also indicated (*triangles*). The *wavy lines* represent the intron sequences present in clone *RP1-3* and *RP1-4,* which were constructed using total RNA as the template for reverse transcription.

stream from it. The purine A is located 3 nucleotides upstream from the ATG codon, which is consistent with it serving as an initiation codon for translation (Kozak, 1986). Kozak found that 95% of the mRNAs she examined have a purine at that position and that 75% of them have an A. The authenticity of the upstream in-frame TAG has been confirmed by sequencing both strands of four independently isolated cDNA clones (R3-11, RP1-18, RP1-3, and RP1-4) that extend past this TAG. A termination codon TAA is located at position 2689, which is followed by multiple stop codons in all three ORFs, including two additional in-frame terminators 3 and 11 codons downstream, clearly establishing an end to the long ORF. The long ORF thus predicts a protein of 896 amino acids. The 5'-untranslated region (5'-UTR) of 144 nucleotides is highly GC rich (76%) although the complete 5'-UTR remains to be sequenced. The high GC content may account for the difficulty we have encountered in isolating cDNA clones that cover sequences farther upstream. The size [1410 nucleotides (nt) without polyA tail] of the 3'-untranslated region (3'-UTR) agrees well with the average size of 3'-UTRs of brain-specific mRNAs but is significantly larger than that of non-brain-specific mRNAs (Sutcliffe, 1988). The 3'-UTR terminates with a polyA tail preceded by an AAUAAA polyadenylation signal. A group of AU-rich sequences containing one or more AUUUA motifs in the 3'-UTR of many relatively unstable mRNAs has been shown to be responsible for their instability (Brawerman, 1989). We have identified (underlined in Fig. 4) four ATTTA sequences in the 3'-UTR of the F1-20 cDNA.

To ensure that there were no deletions in the ORF, we independently confirmed its size. PCR reactions were carried out with two primers (SZ327 and SZ316) flanking the long ORF using F1-20 cDNA made from mouse brain RNA as the tem-

plate. The PCR product had a size of 2.8 kb (data not shown), in agreement with the predicted size (2785 nt) of the composite cDNA sequence.

Identification of an alternatively spliced mini-exon

The sequence analysis of two independently isolated cDNA clones, R4-6 and RP2-9, revealed the presence of 15 additional nucleotides at position 2143–2144 that can potentially encode five amino acids without interrupting the ORF. The sequence of these 15 nucleotides and their flanking sequences are shown in Figure 5A. Identities with the consensus splice sites (Mount, 1982) are indicated. The perfect match between the sequence at the 3' end of the mini-exon (CTTTAG/G) and the consensus splice acceptor site supports the possibility of RNA splicing events utilizing alternative 3' acceptor sites. RNase protection analysis (Fig. 5B) and PCR amplification of mouse brain RNA (Fig. 5C) demonstrate that the presence of these 15 additional nucleotides represents an alternative splicing event rather than a cloning artifact. We prepared a ^{32}P-labeled antisense RNA probe (369 nt) that includes the mini-exon, and hybridized it with total cellular RNA. We then treated it with T2 RNase, which digests ssRNA but not dsRNA, and electrophoresed the products on a 6% polyacrylamide, 7 M urea gel. If the probe hybridizes with a mini-exon containing mRNA, we expected to see a protected band of 369 nt. If the probe hybridizes with a mini-exon-less mRNA, we expect that the 15 nt corresponding to the mini-exon will be single stranded, and therefore digested by T2 RNase. This would result in the probe being cleaved into two protected fragments of 171 nt and 183 nt. The autoradiograph of the gel is shown in Figure 5B. The band at 369 nt indicates that a mini-exon–containing mRNA was present in the cellular RNA. The bands at 171 nt and 183 nt indicate that

Figure 4. Nucleotide sequence and predicted amino acid sequence of the 4.3 kb F1-20 cDNA. A composite nucleotide sequence (including the mini-exon) was generated by aligning the sequences of the overlapping clones indicated in Figure 3. The predicted amino acid sequence starts with the first methionine codon in the ORF. Underlined are an upstream in-frame stop codon TAG (*star*); consensus for cAMP-dependent protein kinase R-R-X-S [X = any amino acid (aa)] (Kemp et al., 1977; Czernik et al., 1987) (*open square*); consensus for Ca²⁺/calmodulin-dependent protein kinase R-X-X-S/T (X = any aa) (Pearson et al., 1985; Czernik et al., 1987) (*open diamonds*); consensus for protein kinase C S/T-X-R/K (X is an uncharged aa) (Woodgett et al., 1986) (*solid squares*); consensus for casein kinase II S-X-X-D/E (X = any aa) (Kuenzel et al., 1987) (*open circles*), the alternatively spliced mini-exon (*solid triangle*), and the polyadenylation signal (*cross*). The amino acid translation of the mini-exon is *boxed*. Also *underlined*, *with no symbols*, are ATTTA sequences (implicated in labile RNAs; Brawerman, 1989) in the 3'-UTR.

Figure 5. F1-20 mRNA exists in two alternatively spliced forms. *A* shows the cDNA sequence containing the mini-exon with its flanking regions. The mini-exon sequence is *boxed.* Above are written the consensus splice sites (Mount, 1982). Identities with the consensus are shown in *boldface.* *B* shows the results of RNA protection analyses. Using one of the cDNA clones that includes the mini-exon as a DNA template (R4-6), ^{32}P-labeled antisense RNA probe was transcribed *in vitro* by T3 RNA polymerase. After overnight hybridization with cellular RNA at 50–55°C, T2 RNase digestion was performed at 35°C for 1.5 hr. The reaction mixture was then separated on a 6% sequencing gel containing 8 M urea. The results are shown in the lane labeled *RNA.* The lanes labeled *A, G, C,* and *T* refer to a sequencing ladder that was generated using the Sanger dideoxy method, which served as molecular weight markers. *C* shows the results of an experiment in which we utilized PCR to demonstrate the presence of mini-exon–containing and mini-exon-less mRNA in total mouse brain cellular RNA. cDNA was made from cellular RNA primed with primer SZ401. We then utilized two primers flanking either side of the mini-exon (SZ401 and SZ58) in PCR reactions with the cDNA. The PCR products were cleaved with StuI, which generates fragments of 354 nt representing the mini-exon–containing species, and 339 nt representing the mini-exon-less species. *Lane A,* ΦX174 HaeIII markers; *lane B,* StuI cut PCR product of cDNA/SZ58-SZ401; *lane C,* uncut PCR product of cDNA/SZ58-SZ401; *lane D,* λBstEII markers.

A

$$^C_AAG/GT^A_GAGT \qquad ^{TT}_{CC}N^C_TAG/G$$

ACCCAT**CAG** | TGTTTGATG**CTTTTAG** | **G**CGATCTTC
TGGGTAGTC | AGAAACTACCAAATC | CGCTAGAAG

B

A G C T   RNA

← 369

← 183
← 171

C·

A B C D

← 920+905

← 566

← 354
← 339
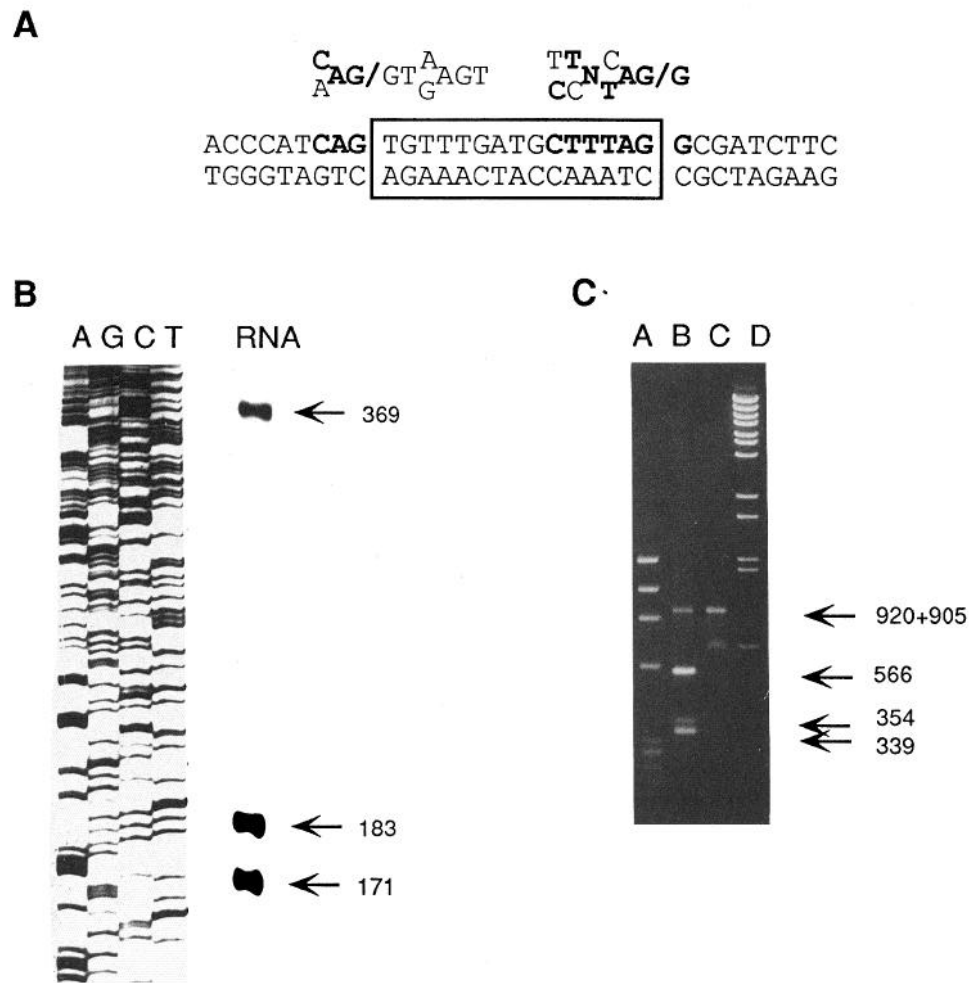
a mini-exon-less mRNA was also present in the cellular RNA. The relative intensities of the bands at 171 nt and 183 nt are greater than the band at 369 nt, indicating that the mini-exon-less mRNA is more abundant than the mini-exon–containing mRNA. We also utilized PCR to demonstrate the presence of mini-exon–containing and mini-exon-less mRNA in total mouse brain cellular RNA (Fig. 5C). cDNA was made from cellular RNA primed with primer SZ401. We utilized two primers flanking either side of the mini-exon (SZ401 and SZ58) in PCR reactions. The expected PCR product from the mini-exon–containing cellular RNA species is 920 nt, and the mini-exon-less cellular RNA species is 905 nt. Because the 15 nt difference between these fragments is too small to be resolved by gel electrophoresis, we cleaved the PCR products with StuI, which generates fragments of 354 nt representing the mini-exon–containing species, and 339 nt representing the mini-exon-less species. Both StuI cut fragments were found, indicating that both RNA species were present in the cellular RNA used to construct the cDNA. The intensity of the band at 339 nt is greater than the band at 354 nt, also indicating that the mini-exon-less mRNA is more abundant than the mini-exon–containing mRNA.

Southern analysis of the F1-20 gene

Southern blot analyses of mouse genomic DNA are consistent with a single-copy F1-20 gene (Fig. 6). Digestion of genomic DNA with five different restriction enzymes, followed by Southern blotting and hybridization with a probe generated from cDNA clone R3-11, yielded a relatively simple pattern (typically approximately two fragments were detected), which statistically is consistent with a single- or low-copy number gene. The probe used is located at the 5′ extreme of the cDNA and spans an intron sequence (Fig. 3).

Analysis of the deduced amino acid sequence of the F1-20 protein

Analysis of the deduced amino acid sequence of the F1-20 protein did not reveal a signal sequence, or any regions of extensive hydrophobicity. A total of 15 possible phosphorylation sites were found and are indicated on Figure 4. Interestingly, one of these potential phosphorylation sites coincides with the mini-exon that is present in a fraction of the F1-20 mRNA as a consequence of an alternative splicing event. The deduced molecular weight of the primary translation product of the mini-exon-less mRNA is 91,319, and the predicted isoelectric point is 4.58. Analysis of the deduced amino acid composition revealed an unusually large number of proline residues (99 prolines/896 residues).

In order to ascertain whether the DNA or amino acid sequence of the F1-20 cDNA or protein had been previously determined and entered in the available data bases, the NBRF-PIR, SWISS-

PROT, GENBANK, and EMBL data bases were searched in their entirety with the F1-20 cDNA and amino acid sequences using the FASTA program (Pearson and Lipman, 1988). This search revealed that sequences corresponding to F1-20 cDNA or protein had not been deposited in these data bases and implies that the F1-20 gene or protein has not been previously characterized at the primary structure level. The similarity with the highest FASTA optimized score that stood up to statistical significance programs (RDF2 analysis) was to rabbit calpastatin (calpain protease inhibitor). A stretch of 92 amino acids in rabbit calpastatin was found that shared 24% amino acid identity with the C-terminal portion of the F1-20 protein sequence. Analysis of this similarity with the RDF2 program revealed that the FASTA similarity score between the unshuffled sequences was 7–8 SDs above the mean scores obtained for shuffled sequences (both uniform and window shuffling using a variety of window sizes was carried out, variation between 7 and 8 SDs reflects variation resulting from use of different shuffling methods). The sequences of both pig and rabbit calpastatin have been deduced from cDNA clones. Calpastatin is composed of four similar ~140 amino acid domains that are believed to form the basis for the ability of one calpastatin molecule to bind and inhibit four to eight calpain molecules (Takano et al., 1986, 1988; Emori et al., 1987; Imajoh et al., 1987; Maki et al., 1987). Alignment of corresponding regions from the four rabbit calpastatin domains and the four similar domains from pig calpastatin and an 85 residue stretch from the C-terminus of the deduced F1-20 sequence (Fig. 7) reveals that residues that are well conserved in the different calpastatin domains are more frequently conserved in the F1-20 sequence than poorly conserved residues. When a consensus is constructed from the alignment shown in Figure 7, it is found to display 42% identity to the F1-20 sequence, allowing for structurally conservative amino acid substitutions.

The F1-20 protein is a substrate for Ca²⁺-dependent proteolysis that is specifically inhibited by calpain inhibitors

Calpastatin is a competitive inhibitor of calpain. Because competitive inhibitors often resemble substrates, we decided to ask whether the F1-20 protein is a substrate for neuronal calpain (Fig. 8). The P2′ fraction (well-washed crude synaptosomes) was generated by the procedure of Huttner et al. (1983) and lysed in the standard buffer (+0.17% 2ME) also containing either 0.1 mM EGTA (lane A), 0.1 mM EGTA + 1.1 mM CaCl₂ (lane B), 0.1 mM EGTA + 1.1 mM CaCl₂ + calpain-specific inhibitor *N*-acetyl-leucyl-leucyl-norleucinal (calpain inhibitor I) (lane C), 0.1 mM EGTA + 1.1 mM CaCl₂ + the calpain-specific inhibitor *N*-acetyl-leucyl-leucyl-methioninal (calpain inhibitor II) (lane D), 0.1 mM EGTA + 1.1 mM CaCl₂ + the thiol protease inhibitor E64 (lane E), 0.1 mM EGTA + 1.1 mM CaCl₂ + the thiol/serine protease inhibitor leupeptin (lane F), 0.1 mM EGTA + 1.1 mM CaCl₂ + the serine protease inhibitor PMSF (lane G), or 0.1 mM EGTA + 1.1 mM CaCl₂ + the acid protease inhibitor pepstatin (lane H), and incubated 60 min on ice. Samples were processed as described in Materials and Methods. It is clear that the intensity of the F1-20–immunoreactive band is dramatically reduced in the Ca²⁺-containing buffer (compare lanes A and B). This is inhibited by the calpain-specific inhibitors I (lane C) and II (lane D), the thiol protease inhibitor E64 (lane E), and the thiol/serine protease inhibitor leupeptin (lane F), all of which are calpain inhibitors. Proteolysis is still observed in the presence of the serine protease inhibitor PMSF (lane G) and the acid protease inhibitor pepstatin (lane H).
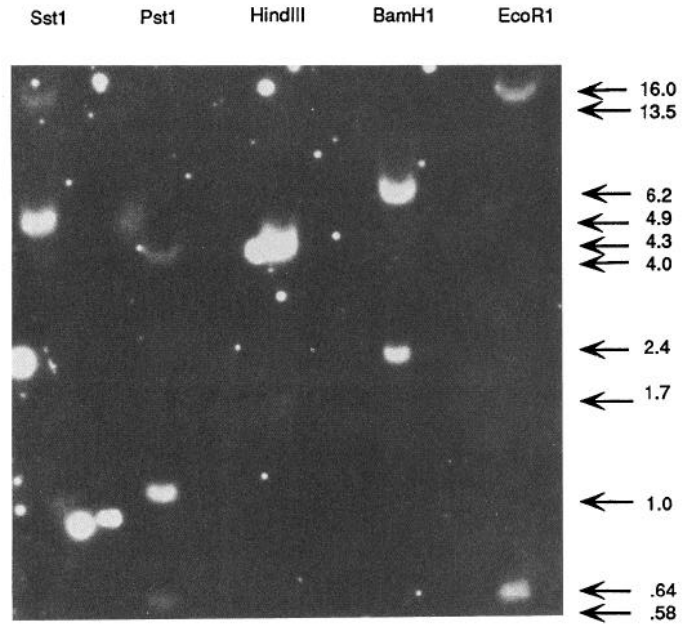
Figure 6. Genomic Southern blot analysis of the F1-20 gene. Genomic DNA was digested with the indicated restriction enzymes. The restriction fragments were separated on a 0.8% agarose gel, transferred to nitrocellulose, and hybridized to ³²P-R3-11 at 42°C in 6× SSPE, 0.01 M EDTA, 5× Denhardt's, 0.5% SDS, and 100 μg/ml ssDNA, followed by extensive washing in 1× SSPE at 65–68°C (Maniatis et al., 1982). Restriction fragments were sized by comparison to λHindIII and λBstEII markers.

Examination of the total protein pattern revealed that while a number of other protein bands are also disappearing in a Ca²⁺-dependent manner, the majority of proteins are unaffected.

The F1-20 protein is likely to be phosphorylated in vivo

We noted some potential phosphorylation sites in the deduced amino acid sequence of the F1-20 protein (Fig. 4). In order to ask whether F1-20 protein extracted directly from mouse brain contains phosphate, we utilized alkaline phosphatase treatment. Alkaline phosphatase catalyzes the hydrolysis of phosphate groups from phosphoproteins, increasing their electrophoretic mobility on SDS-polyacrylamide gels. The P2′ fraction (well-washed synaptosomes) was generated by the procedure of Huttner et al. (1983) and lysed to generate LS1. Aliquots of the synaptosomal lysate were treated with alkaline phosphatase, run on SDS-polyacrylamide gels, and Western blotted. The blots were stained with the F1-20 MAb. Alkaline phosphatase treatment led to a distinct increase in the electrophoretic mobility of the F1-20–immunoreactive band, as well as to an increase in its sharpness (Fig. 9, lane B). This increase in electrophoretic mobility was not seen when alkaline phosphate inhibitors were included in the reaction mix (lane D). Marker proteins revealed that this increase in electrophoretic mobility corresponded to a decrease in the apparent molecular weight of the F1-20 protein from 190,000 Da to 180,000 Da.

The F1-20 protein is extremely acidic

The pI of the F1-20 protein was measured by two-dimensional gel electrophoresis followed by Western blot analysis. In five independent measurements, the pI was found to be 3.85 ± 0.15. A representative blot is shown in Figure 10.

```
                                                  A                                                    E
                              E                    E                        E  SS T         V      P      K
        CONSENSUS: PEPETPVKP-LSEDDLIDA-LSGDFDTCPSTASSSKFTE-KDL--SAAVTAEVVKAVPAPVSMATDSAAPPKETSKQVKSDTMLDDALEALIDSLG
MO. F1-20  688-772: PST-TPVTP-AQNN-LLQPSFEAAFGTTPSTSSSSSFDPSGDL---LMPTMAPSGQ-PAPVSMVPPS--PAMAASKGLGSD--LDSSLASLVGNLG
RA. Calp.  640-718: KKSKEIKKP-AGDQDPIDA-LSGDLDSCPPAAETSQATE-KDK--SKTTTASSSKAA-KHGDKAKDSAQTTEETSKP-KANEKNAS
PI. Calp.  636-713: KKPEAPKKP-EAAQDPIDA-LSGDFDRCPSTTETSENTT-KDK--DKK-TASKSKA-PKNGGKAKDSTKAKEETSKQ-KSDGKSTS
RA. Calp.  500-583: PSEPTAQLPALSEDLLLDA-LSEDFSGPSSSASSLKFDDAM-L--SAAVS-EVVSQSPA--SI-TRATAPPPDTRPSNKE---LDDALDKLSDSLG
PI. Calp.  498-579: KDPKEPVLP-LSEDFVLDA-LSQDFAGPP-AASSL-FEDAK-L--SAAVS-EVVSQTSAP--T--THSAGPPPDTVSDDKK---LDDALDQLSDSLG
RA. Calp.  357-446: PEPAEKPKP-RSESELIDE-LSKDFSQAKSNEKQPKPTG-KTEES KAAVPAPVAEAVPRT-SMCSIQPVPPKPASLQ-KS-TVPDDAVEALAGSLG
PI. Calp.  357-443: PEAEEKPKP-LSESELIDE-LSEDFDQSKRKEKQSKPTE-KTKES QATAPTPVGEAVSRT-SLCCVQSAPPKPAT-----GMVPDDAVEALAGSLG
RA. Calp.  218-302: PPPDS-VTP-LGPDDAIDA-LSSDF-TCSSPVASGKEAG-KEA---AKSAGEVLEAESAKV-M--RAAAPPQEKKRKVEEDAMSDQALEALSASLG
PI. Calp.  219-302: PPPDT-SKP-LGPDDAIDA-LSLDL-TCSSPTADGKKTE-KE-----KSTGEVLKAQSVGV-IKS-AAAPPHEKKRRVEEDTMSDQALEALSASLG
RA. Calp.  150-165:                                                                              KSG--MDAALDDLIDTLG
PI. Calp.  151-166:                                                                              KSD--MDAALDDLIDTLG
```

Figure 7. Alignment and consensus sequence between 92 amino acids of the F1-20 protein and the corresponding similar regions from the multiple pig and rabbit calpastatin domains. Amino acids that are identical or substituted by structurally conservative residues at five or more members of the alignment are shown in *boldface*. Multiple residues are presented for ambiguous positions in the consensus.

Discussion

Because of the apparent discrepancy between the molecular weight of the F1-20 protein as indicated by its migration on SDS-PAGE (180,000 Da for the dephosphorylated form) and its deduced molecular weight from the cDNA clone (91,319 Da), it is important that we have unambiguous confirmation that we have isolated the *correct* and *complete* cDNA. The following pieces of evidence support our conclusion that we have, in fact, isolated the correct cDNA. (1) First, and most importantly, *antibodies raised against the cDNA-encoded fusion protein recognize the same protein recognized by the original F1-20 MAb.* That is, the anti-cDNA–encoded fusion protein antiserum and the original F1-20 MAb recognize a protein of identical apparent molecular weight on SDS-PAGE (190 kDa; Fig. 1) with an identical pattern of developmental regulation and tissue specificity (data not shown). Except for this single

190 kDa band, no other bands are visible on Western blots probed with the anti-cDNA–encoded fusion protein antiserum. Immunohistochemistry with the original F1-20 MAb and the anti-cDNA–encoded fusion protein antiserum also gives identical results (Fig. 2). (2) In high-stringency Northern blots, the cDNA hybridizes to a single band that, like the F1-20 protein, is brain specific and developmentally regulated (Sousa et al., 1992, Figs. 4, 6). It is theoretically possible that the isolated cDNA encodes some other protein (i.e., protein X) that shares an epitope with F1-20. *However, the antiserum against the cDNA-encoded fusion protein detects only the F1-20 protein.* Barring the unlikely possibility that the isolated cDNA encodes a polypeptide that elicits antibodies that recognize the F1-20 protein but do not recognize any other protein (including the protein X, which might be encoded by this cDNA), we conclude that we have isolated the cDNA that encodes the F1-20 protein.

Points 1 and 2 above do not, however, establish that we have

Figure 8. The F1-20 protein is a substrate for calcium-dependent proteolysis *in vitro*, which is inhibited specifically by calpain inhibitors. The P2' fraction (well-washed crude synaptosomes) was generated by the procedure of Huttner et al. (1983) and lysed in the standard buffer (+0.17% 2ME) also containing either 0.1 mM EGTA (*lane A*), 0.1 mM EGTA + 1.1 mM CaCl₂ (*lane B*), 0.1 mM EGTA + 1.1 mM CaCl₂ + calpain-specific inhibitor *N*-acetyl-leucyl-leucyl-norleucinal (calpain inhibitor I) (*lane C*), 0.1 mM EGTA + 1.1 mM CaCl₂ + the calpain-specific inhibitor *N*-acetyl-leucyl-leucyl-methioninal (calpain inhibitor II) (*lane D*), 0.1 mM EGTA + 1.1 mM CaCl₂ + the thiol protease inhibitor E64 (*lane E*), 0.1 mM EGTA + 1.1 mM CaCl₂ + the thiol/serine protease inhibitor leupeptin (*lane F*), 0.1 mM EGTA + 1.1 mM CaCl₂ + the serine protease inhibitor PMSF (*lane G*), or 0.1 mM EGTA + 1.1 mM CaCl₂ + the acid protease inhibitor pepstatin (*lane H*), incubated 60 min on ice, and processed as described in Materials and Methods. All inhibitors were used at 130 μM. A Western blot of LS1 is shown probed with the F1-20 MAb (*A*) or stained for total protein (*B*).
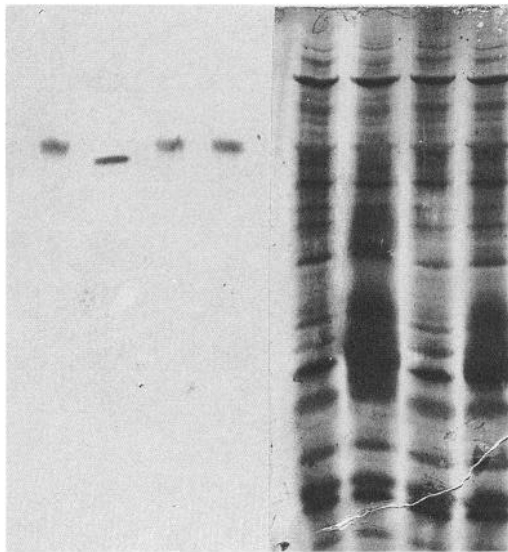
Figure 9. The F1-20 protein is likely to be phosphorylated *in vivo*. The P2′ fraction (well-washed synaptosomes) was generated by the procedure of Huttner et al. (1983) and lysed to generate LS1. Aliquots (10 μl) of LS1 in alkaline phosphatase buffer (10 mM Tris, pH 8.5, 1 mM DTT, 130 μM each of the protease inhibitors E64, leupeptin, PMSF, and pepstatin) were incubated either with no additions (*lanes A, E*); 1 U of calf intestinal alkaline phosphatase (*lanes B, F*); the alkaline phosphatase inhibitors 10 mM EDTA and 10 mM NaH$_2$PO$_4$, (*lanes C, G*); or the alkaline phosphatase inhibitors 10 mM EDTA and 10 mM NaH$_2$PO$_4$, + 1 U of calf intestinal alkaline phosphatase (*lanes D, H*). Reactions were run for 60 min at 37°C, terminated by boiling in SDS sample buffer, and run on 8% SDS-PAGE. Western blots were prepared and stained either with the F1-20 MAb (*lanes A–D*), or for total protein with india ink (*lanes E–H*).

Figure 10. The F1-20 protein is extremely acidic. A representative two-dimensional Western blot of the LS2 fraction of mouse brain synaptosomal proteins probed with the F1-20 MAb reveals a single immunoreactive spot at 190 kDa and a pI of ~3.85 (indicated). The experimentally measured pH gradient is given. Five independent measurements have confirmed this extremely acidic pI of the F1-20 protein (pI range from these measurements was 3.7–4.0).

isolated the *complete* cDNA sequence that encodes the F1-20 protein. Because of cloning artifacts in the construction and isolation of cDNAs, it is always possible that a reconstructed cDNA sequence is missing significant segments. As detailed in the Results, we isolated multiple clones of all segments of the cDNA corresponding to the long ORF and the immediate upstream and downstream regions. Also detailed in the Results is the evidence that unambiguously supports the identification of the initiation AUG and the translation termination codon. In support of the completeness of the given cDNA sequence, we also note the following. (1) The size of the F1-20 mRNA, 4.8 kb, is itself inconsistent with a 180 kDa protein, even if unusually short 5′- and 3′-UTRs are present. (In fact, the 5′-UTR is at least 144 nucleotides long and the 3′-UTR is at least 1410 nucleotides in length.) (2) In order to confirm that the portion of the cDNA that encodes the ORF is complete, we used primers flanking the long ORF to PCR amplify cDNA made from mouse brain mRNA: if the deduced ORF is complete, then the amplified DNA fragment should be 2.785 kb in length. The amplified DNA was found to be 2.8 kb in length, confirming that the deduced ORF is complete.

The molecular weight discrepancy could arise from either additional posttranslational modifications of the primary translation product beyond phosphorylation, anomalous migration on SDS-PAGE, or both. There is precedence for acidic proteins, as well as for proline-rich proteins, to migrate anomalously on
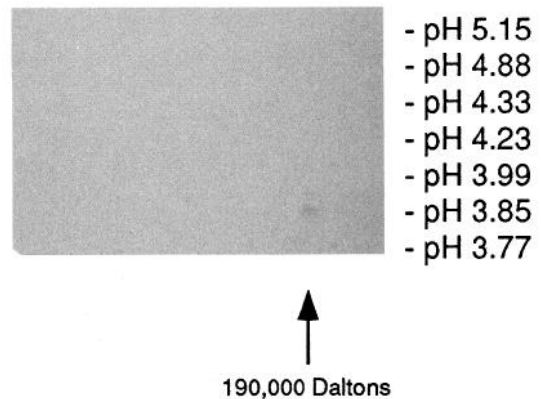
SDS-PAGE, with differences between true and apparent molecular weights as great as a factor of 2 (Julien and Mushynski, 1982; Kleinschmidt et al., 1986; Myers et al., 1987; Takano et al., 1988; Chin and Liem, 1990; See and Jackowski, 1990). In the case of the highly acidic proteins, it is believed that they migrate anomalously due to an inability to bind SDS appropriately, leading to a drastic overestimation of their apparent molecular weight. The F1-20 protein is extremely acidic (pI = 3.85; Fig. 10). The proline content of F1-20 is also quite high (99 prolines/896 residues). These features are consistent with a protein that migrates anomalously on SDS-PAGE. Further characterization of the F1-20 protein will be required to determine the nature and extent of any posttranslational modifications.

We have also identified two distinct isoforms of F1-20 cDNA, differing by the presence of 15 additional nucleotides, which do not interrupt the ORF. We noted a perfect match between the sequence at the 3′ end of the mini-exon and the consensus splice acceptor site, which supports the possibility of RNA splicing events utilizing alternative 3′ acceptor sites. We demonstrated by two independent methods that both mini-exon–containing and mini-exon-less F1-20 RNA molecules can be found in cellular RNA. RNase protection experiments revealed both species of F1-20 RNA in cellular RNA, and PCR amplification of cDNA made from cellular RNA primed with an F1-20–specific primer revealed two distinct PCR products with the expected sizes of the two isoforms. Both experiments indicated that the mini-exon-less mRNA is more abundant than the mini-exon–containing mRNA. The size of the mini-exon is quite small. However, there is precedent in the literature for mini-exons even smaller than that observed in the F1-20 mRNA having functional consequences (Haber et al., 1990; Rauscher et al., 1990; Madden et al., 1991). The possible functional significance of this alternative splicing event is heightened by the observation that it results in the removal of a potential phosphorylation site. cDNAs for most of the well-characterized synaptic proteins have been found to be present in multiple isoforms (Elferink et al., 1989; Sudhof et al., 1989b; Archer et al., 1990; Knaus et al., 1990; Geppert et al., 1991). In many of these cases, the isoforms have been shown to have differential distributions in brain. Experiments are in progress to determine if the two F1-20 mRNA isoforms vary in their neuroanatomical distributions.

Analysis of the deduced amino acid sequence of the F1-20 protein led to some predictions about the biochemical properties of the F1-20 protein. We have presented experiments designed to test some of these predictions. We found that the F1-20 protein is a substrate for Ca^{2+}-dependent proteolysis that is specifically inhibited by calpain inhibitors *in vitro*. This indicates that the F1-20 protein is a substrate for neuronal calpain. The proposed involvement of calpain in such processes as activity-dependent synaptic vesicle exocytosis (Kelly, 1988; reviewed in Trimble et al., 1991), excitatory amino acid–induced structural protein breakdown and neurodegeneration (Siman and Noszek, 1988; Siman et al., 1989), and long-term potentiation (Lynch and Baudry, 1984; reviewed in Madison et al., 1991) may imply a role for the F1-20 protein in these processes. We also found that treatment of a synaptosomal lysate with alkaline phosphatase led to an increase in the electrophoretic mobility of the F1-20 protein, as well as to an increase in the sharpness of the electrophoretic band. Since the effect was blocked by EDTA and NaH_2PO_4, both inhibitors of alkaline phophatase, we conclude that it was due to dephosphorylation and not to a contaminant protease (Sheridan et al., 1989). This indicates that the F1-20 protein is phosphorylated *in vivo*. Protein phosphorylation is involved in signal transduction in many systems including the nervous system and is affected by changes in calcium concentration either directly, through the action of calcium-dependent kinases, or indirectly through complex signal transduction cascades. The phosphorylation and calpain sensitivity of F1-20 may therefore provide mechanisms through which its activity is regulated by changing intracellular calcium concentrations.

References

Archer BT III, Ozcelik T, Jahn R, Francke U, Sudhof TC (1990) Structures and chromosomal localizations of two human genes encoding synaptobrevins 1 and 2. J Biol Chem 265:17267–17273.

Ausubel FM, Brent RB, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) (updated quarterly thereafter) Current protocols in molecular biology. New York: Wiley Interscience.

Brawerman G (1989) mRNA decay: finding the right targets. Cell 57:9–10.

Buckley KM, Floor E, Kelly RB (1987) Cloning and sequence analysis of cDNA encoding p38, a major synaptic vesicle protein. J Cell Biol 105:2447–2456.

Chen J, Neilson K, Van Dyke T (1989) Lymphotropic papovavirus early region is specifically regulated in transgenic mice and efficiently induces neoplasia. J Virol 63:2204–2214.

Chin SSM, Liem RKH (1990) Transfected rat high-molecular-weight neurofilament (NF-H) coassembles with vimentin in a predominantly nonphosphorylated form. J Neurosci 10:3714–3726.

Czernik AJ, Pang DT, Greengard P (1987) Amino acid sequences surrounding the cAMP-dependent and calcium/calmodulin-dependent phosphorylation sites in rat and bovine synapsin I. Proc Natl Acad Sci USA 84:7518–7522.

Elferink LA, Trimble WS, Scheller RH (1989) Two vesicle-associated membrane protein genes are differentially expressed in the rat central nervous system. J Biol Chem 262:11061–11064.

Emori Y, Kawasaki H, Imajoh S, Imahori K, Suzuki K (1987) Endogenous inhibitor for calcium-dependent cysteine protease contains four internal repeats that could be responsible for its multiple reactive sites. Proc Natl Acad Sci USA 84:3590–3594.

Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85:8998–9002.

Geppert M, Archer B III, Sudhof TC (1991) Synaptotagmin II. A novel differentially distributed form of synaptotagmin. J Biol Chem 266:13548–13552.

Haber DA, Buckler AJ, Glaser T, Call KM, Pelletier J, Sohn RL, Douglass EC, Housman DE (1990) An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor. Cell 61:1257–1269.

Huttner WB, Schiebler W, Greengard P, De Camilli P (1983) Synapsin I (protein I), a nerve terminal-specific phosphoprotein. III. Its association with synaptic vesicles studied in a highly purified synaptic vesicle preparation. J Cell Biol 96:1374–1388.

Huynh TV, Young RA, Davis RW (1985) Constructing and screening cDNA libraries in λgt10 and λgt11. In: DNA cloning: a practical approach (Glover DM, ed), pp 49–78. Washington, DC: IRL.

Imajoh S, Kawasaki H, Emori Y, Ishiura S, Minami Y, Sugita H, Imahori K, Suzuki K (1987) A fragment of an endogenous inhibitor produced in Escherichia coli for calcium-activated neutral protease (CANP) retains an inhibitory activity. FEBS Lett 215:274–278.

Julien J-P, Mushynski WE (1982) Multiple phosphorylation sites in mammalian neurofilament polypeptides. J Biol Chem 257:10467–10470.

Kelly RB (1988) The cell biology of the nerve terminal. Neuron 1:431–438.

Kemp BE, Graves DJ, Benjamini E, Krebs EG (1977) Role of multiple basic residues in determining the substrate specificity of cyclic AMP-dependent protein kinase. J Biol Chem 252:4888–4894.

Kleinschmidt JA, Digwall C, Maier G, Franke WW (1986) Molecular characterization of a karyophilic, histone-binding protein: cDNA cloning, amino acid sequence and expression of nuclear protein N1/N2 of Xenopus laevis. EMBO J 5:3547–3552.

Knaus P, Marqueze-Pouey B, Scherer H, Betz H (1990) Synaptoporin, a novel putative channel protein of synaptic vesicles. Neuron 5:453–462.

Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283–292.

Kuenzel EA, Mulligan JA, Sommercorn J, Krebs EG (1987) Substrate specificity determinants for casein kinase II as deduced from studies with synthetic peptides. J Biol Chem 262:9136–9140.

Lafer EM, Sousa RJ, Rich A (1988) Z-DNA-binding proteins in Escherichia coli: purification, generation of monoclonal antibodies and gene isolation. J Mol Biol 203:511–516.

Leube RE, Kaiser P, Seiter A, Zimbelmann R, Franke WW, Rehm H, Knause P, Prior P, Betz H, Reinke H, Beyreuther K, Wiedenmann B (1987) Synaptophysin: molecular organization and mRNA expression as determined from cloned cDNA. EMBO J 6:3261–3268.

Linial M, Miller K, Scheller RH (1989) VAT-1: an abundant membrane protein from torpedo cholinergic synaptic vesicles. Neuron 2:1265–1273.

Lynch G, Baudry M (1984) The biochemistry of memory: a new and specific hypothesis. Science 224:1057–1063.

Madden SL, Cook DM, Morris JF, Gashler A, Sukhatme VP, Rauscher FJ III (1991) Transcriptional repression mediated by the WT1 Wilms tumor gene product. Science 253:1550–1553.

Madison DV, Malenka RC, Nicoll RA (1991) Mechanisms underlying long-term potentiation of synaptic transmission. Annu Rev Neurosci 14:379–397.

Maki M, Takano E, Osawa T, Ooi T, Murachi T, Hatanaka M (1987) All four internally repetitive domains of pig calpastatin possess inhibitory activities against calpains I and II. FEBS Lett 223:174–180.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mount SM (1982) A catalogue of splice junction sequences. Nucleic Acids Res 10:459–472.

Myers MW, Lazzarini RA, Lee VM-Y, Schlaepfer WW, Nelson DL (1987) The human mid-size neurofilament subunit: a repeated protein sequence and the relationship of its gene to the intermediate filament gene family. EMBO J 6:1617–1626.

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021.

Oyler GA, Higgins GA, Hart HA, Battenberg E, Billingsley M, Bloom FE, Wilson MC (1989) The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal subpopulations. J Cell Biol 109:3039–3052.

Pearson RB, Woodgett JR, Cohen P, Kemp BE (1985) Substrate specificity of a multifunctional calmodulin-dependent protein kinase. J Biol Chem 260:14471–14476.

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444–2448.

Perin MS, Johnston PA, Ozcelik T, Jahn R, Francke U, Sudhof TC (1991a) Structural and functional conservation of synaptotagmin (p65) in *Drosophila* and humans. J Biol Chem 266:615–622.

Perin MS, Brose N, Jahn R, Sudhof TC (1991b) Domain structure of synaptotagmin (p65). J Biol Chem 266:623–629.

Rauscher FJ III, Morris JF, Tournay OE, Cook DM, Curran T (1990) Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. Science 250:1259–1262.

Sanger F, Nicklen S, Coulson A (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467.

See YP, Jackowski G (1990) Estimating molecular weights of polypeptides by SDS gel electrophoresis. In: Protein structure (Creighton TE, ed), pp 1–21. Oxford: IRL.

Sheridan PL, Francis MD, Horwitz KB (1989) Synthesis of human progesterone receptors in T47D cells, nascent A- and B-receptors are active without a phosphorylation-dependent post-translational maturation step. J Biol Chem 264:7054–7058.

Siman R, Noszek JC (1988) Excitatory amino acids activate calpain I and induce structural protein breakdown *in vivo*. Neuron 1:279–287.

Siman R, Noszek JC, Kegerise C (1989) Calpain I activation is specifically related to excitatory amino acid induction of hippocampal damage. J Neurosci 9:1579–1590.

Sousa RJ, Hrinya Tannery N, Lafer EM (1990) Two new monoclonal antibodies provide immunohistochemical evidence for the unique biochemical similarity of the mouse globus pallidus, entopeduncular nucleus, and substantia nigra pars reticulata. Neuroscience 43:403–410.

Sousa RJ, Hrinya Tannery N, Zhou S, Lafer EM (1992) Characterization of a novel synapse-specific protein. I. Developmental expression and cellular localization of the F1-20 protein and mRNA. J Neurosci 12:2130–2143.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517.

Sudhof TC, Jahn R (1991) Proteins of synaptic vesicles involved in exocytosis and membrane recycling. Neuron 6:665–677.

Sudhof TC, Lottspeich F, Greengard P, Mehl E, Jahn R (1987) A synaptic vesicle protein with a novel cytoplasmic domain and four transmembrane regions. Science 238:1142–1144.

Sudhof TC, Baumert M, Perin MS, Jahn R (1989a) A synaptic vesicle membrane protein is conserved from mammals to *Drosophila*. Neuron 2:1475–1481.

Sudhof TC, Czernik AJ, Kao H-T, Takei K, Johnston PA, Horiuchi A, Kanazir SD, Wagner MA, Perin MS, De Camilli P, Greengard P (1989b) Synapsins: mosaics of shared and individual domains in a family of synaptic vesicle phosphoproteins. Science 245:1474–1480.

Sutcliffe JG (1988) mRNA in the mammalian central nervous system. Annu Rev Neurosci 11:157–198.

Tabor S, Richardson CC (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc Natl Acad Sci USA 84:4767–4771.

Takano E, Maki M, Hatanaka M, Mori H, Zenita K, Sakihama T, Kannagi R, Marti T, Titani K, Murachi T (1986) Evidence for the repetitive domain structure of pig calpastatin as demonstrated by cloning of complementary cDNA. FEBS Lett 208:199–202.

Takano E, Maki M, Mori H, Hatanaka M, Marti T, Titani K, Sakihama T, Kannagi R, Ooi T, Murachi T (1988) Pig heart calpastatin: identification of repetitive domain structure and anomalous behavior in polyacrylamide gel electrophoresis. Biochemistry 27:1964–1972.

Trimble WS, Cowan DM, Scheller RH (1988) VAMP-1: a synaptic vesicle-associated integral membrane protein. Proc Natl Acad Sci USA 85:4538–4542.

Trimble WS, Gray TS, Elferink LA, Wilson MC, Scheller RH (1990) Distinct patterns of expression of two VAMP genes within the rat brain. J Neurosci 10:1380–1387.

Trimble WS, Linial M, Scheller RH (1991) Cellular and molecular biology of the presynaptic nerve terminal. Annu Rev Neurosci 14:93–122.

Woodgett JR, Gould KL, Hunter T (1986) Substrate specificity of protein kinase C. Eur J Biochem 161:177–184.