# Molecular Cloning of Fyn-associated Molecules in the Mouse Central Nervous System

Nobuyuki Kai

DOCTOR OF PHILOSOPHY

Department of Physiological Sciences
School of life science
The Graduate University for Advanced Studies

1996

Contents

## Introduction

Intracellular signal transduction pathways are implicated in cell differentiation, migration, axonal projection and synaptic connection of neural cells. In the genetic analysis of *Drosophila melanogaster* mutants, several genes which, when altered, affect behavioral responses are implicated in intracellular signal transduction pathways of neural machineries; *dunce* and *rutabaga* gene products are directly involved in cAMP metabolism and their mutations result in defects in associative learning (Han et al., 1992; Levin et al., 1992; Nighorn et al., 1991), *scalloped (sd)* gene product encodes a transcription factor and its mutants show defects in response to a number of taste stimuli (Inamdar et al., 1993). In mouse, a gene-targeting technique shows that inactivation of a certain intracellular signal molecule causes neurological defects or behavioral abnormalities. For example, null mutation of c-*fos* gene causes defects in a series of male sexual behaviors (Baum et al., 1994). Reduced hippocampal long-term potentiation (LTP) and impaired learning are also produced by disruption of each of the following signal-transduction molecules: metabotropic glutamate receptor 1 (mGluR1) (Aiba et al., 1994), alpha-calcium-calmodulin kinase II ($\alpha$-CaMK2) (Silva et al., 1992), PKC gamma (Abeliovich et al., 1993), cyclic AMP responsible element binding protein (CREB) (Bourtchuladze et al., 1994), and Fyn (Grant et al., 1992). These results indicate that the intracellular signaling pathway in the nervous system plays significant roles in the production of behavioral responses.

Fyn is a member of the Src-family kinases, a non receptor-type tyrosine kinase. In addition to the deficiencies of spatial learning in Fyn-lacking mice, gene disruption of Fyn causes several other behavioral abnormalities (Yagi et al., 1993; Yagi et al., 1993; Miyakawa et al., 1994; Miyakawa et al., 1995). The homozygous Fyn mutants are not lethal but the homozygous neonates from homozygous parents die because of a suckling problem. They also exhibited stronger light aversion in the light-dark choice test, higher fear-response scores in the novelty preference and passive avoidance tests, and were

1

hyper responsive to fear-inducing environments and to acoustically primed audiogenic seizures. Anatomical defects in the neural cell layer of the hippocampal formation and in the glomeruli of the olfactory bulb were also observed in the mutants. The Fyn signaling pathway thus should offer insights into the morphological, electrophysiological and behavioral events in the mammalian nervous system.

Biochemical studies have indicated that Fyn phosphorylates N-methyl-D-aspartate (NMDA) receptor subunits NR2A (epsilon 1) and NR2B (epsilon 2) in rat (Suzuki and Okumura-Noji, 1995), and nicotinic acetylcholine receptor (AChR) in Torpedo electric organ (Swope et al., 1995). Furthermore, the Fyn-mediated signaling pathway regulates the NCAM-dependent neurite outgrowth of cerebellar and dorsal root ganglion neurons *in vitro* (Beggs et al., 1994), as well as myelin formation together with a large myelin associated glycoprotein (L-MAG) (Umemori et al., 1994). These findings demonstrate that Fyn is involved in a variety of signal transduction pathways in the nervous system. However, it is likely that many unknown targets of Fyn interaction in this system remain, and a complete understanding of Fyn function requires their identification.

To gain further insight into Fyn signaling pathways in the brain, I performed an *in vivo* analysis to isolate molecules which directly associate with Fyn non-catalytic domains. The yeast two-hybrid system provides an assay of protein-protein interaction in a physiological condition. Here, I identified five clones from the expression cDNA library of the mouse neonatal brain. Expression pattern of each of these genes varied in various tissues and during brain development, respectively. Two of them in particular encode novel proteins which contain the motifs, presumably to associate with Src homology 2 or 3 (SH2 or SH3).

## Materials and Methods

### Strains and Media.

Escherichia coli JM109 [recA1, endA1, gyrA96, thi, supE44, relA1, Δ(lacproAB) F'[traD36, proAB, lacI⁹ZΔM15], mcrA-, mcrB+] was used for plasmid construction. HB101 [supE44, ara-14, hsdS20( rB⁻mB⁻), recA13, ara14, proA2, leuB6, lacY1, galK2, rpsL20, xyl-5mtl-1] was used to recover cDNA carrying plasmids from yeast. The yeast strains and the shuttle vectors (pGAD424, pGBT9 and their derivatives) were provided from CLONTECH Laboratories, Inc.(Palo Alto, CA). The genotype of the Saccharomyces cerevisiae reporter strain HF7c is MATa, ura3-52, his3-200, lys2-801, ade2101, trp1901, leu2-3, 112, gal4-542, gal80538, LYS2:: GAL1HIS3, URA3:: (GAL417mers)3-CYC1-lacZ). The genotype of SFY526 is MATa, ura3-52, his3-200, lys2801, ade2-101, trp1901, leu2-3, 112, gal4-542, gal80-538, URA3::GAL1-lacZ . For drug selections of E. coli transformants, Luria broth (LB) was supplemented with ampicillin (50 μg/ml). Minimal media (M9) plates (47.8 mM Na$_2$HPO$_4$, 22 mM KH$_2$PO$_4$, 9 mM NaCl, 18.6 mM NH$_4$Cl, 0.4% glucose) containing 50 μg/ml ampicillin, 40 μg/ml proline and 1 mM thiamine-HCl were used to isolate cDNA clones. Yeast strains were grown in YPD (2% peptone, 1% yeast extract, 2% glucose) or SD synthetic minimal medium [6.7 g/l yeast nitrogen base without amino acids (Difco), 2% glucose with appropriate supplements according to the manufacturer's instructions].

### Construction of Fyn-GAL4 Binding Domain Plasmid

pBD-Fyn was created as follows: Human fyn cDNA coding sequence lacking kinase domain was obtained by the polymerase chain reaction (PCR) using pSN-MluI (Semba et al., 1986) as a template followed by restriction endonuclease digestion. The primer 5'-TGGATCCGTCGACTAATGGGCTGTGTGCAATGT contains a SalI site 5' to the initiating ATG codon, and the primer 5'-GTTCGAATGGCTACTTTTGTGTT-TCCATT was complementary to nucleotide 1450-1472 of fyn cDNA (Genbank

3

accession M14333). The amplified DNA was digested with *Sal*I and *Xho*I and cloned into the *Sal*I site of pGBT9. The orientation and the junction of reading frames between the Gal4 DNA-binding domain and the *fyn* cDNA was confirmed by sequencing analysis. The expression of the fusion protein in yeast was verified by immunoblotting. Yeast cells were lysed by sonication with acid-washed glass beads (Sigma) in a buffer containing 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris-HCl (pH 8.0), 1 µg/ml leupeptin, and 0.5% aprotinin. Supernatants of cell lysates were separated by SDS/PAGE and electroblotted on a PVDF filter (Millipore). Blots were incubated with anti-human Fyn monoclonal antibody γC3 (Yasuda et al, in preparation). The primary antibody was detected with biotinylated rabbit anti-rat IgG and streptavidin alkaline phosphatase conjugate (Amersham). The alkaline phosphatase activity was visualized with 5-bromo-4-chloro-3-indolylphosphate toluidinium and nitro blue tetrazolium (BICIP and NBT, Sigma).

## cDNA-Activation Domain Fusion Library

To construct the cDNA library, I used a plasmid expression vector (pGAD424) with the Time Saver cDNA Synthesis kit (Pharmacia). Total RNA was isolated from newborn mouse (ICR, Charles River) brains by the method of Chomczynski and Sacchi (Chomczynski, 1987), and poly A$^+$ RNA was isolated by absorption to oligo-dT cellulose (GIBCO-BRL). For first strand synthesis, 5 µg of poly A$^+$ RNA as a template was primed with 0.74 µg of random hexamers, according to the manufacturer's instructions; this resulted in double-stranded cDNA with an *Eco*RI-*Not*I adapter on each end with an average size of ~1.0 kb. About 80 ng of cDNA was ligated to 800 ng of *Eco*RI digested and calf intestinal alkaline phosphatase (CIP)-treated vector pGAD424. Following ligation, the library was desalted using micro concentrators (Microcon-100, GRACE Japan), and electroporated into *E. coli* DH10B (GIBCO-BRL). The library contained ~1.7X10$^6$ primary transformants. Library DNA was obtained by plating primary transformants on LB with ampicillin plates. Colonies were scraped into LB

4

media, and DNA was prepared by alkali lysis and CsCl method. The yield of DNA was ~ 1.5 mg.

## Two-Hybrid Screens

The bait plasmid (pBD-Fyn, *TRP1* marker) and the cDNA library (*LEU2* marker) were introduced sequentially into HF7c. Double-transformants were plated on SD media lacking leucine, tryptophan, and histidine but containing 5 mM 3-amino-1,2,4,triazole (3-AT), a chemical inhibitor of His3 which restores histidine auxotrophy. Approximately $3 \times 10^6$ transformants were analyzed from the titration of a small portion of the transformants on the SD media lacking Leu and Trp but containing His, without 3-AT. The plates were incubated for 2 months at 30°C. Colonies grown on the His- plates were successively picked up and transferred to other SD(Leu⁻, Trp⁻, His⁺) plates (stock plates) and incubated to determine β-galactosidase activity. Patches of the yeast colonies on the stock plates were directly lifted from the plates onto Whatman #1 papers, quickly frozen in liquid nitrogen (~10 sec), and immediately overlaid onto Whatman filters that had been soaked in β-galactosidase buffer [0.1 M phosphate buffer pH7.0, 10 mM KCl, 1 mM MgSO4, 40 mM β-mercaptoethanol, 0.33% Bluo-gal (GIBCO-BRL)]. Filter sandwiches were placed in petri dishes and incubated for 2 weeks at 30°C. Positive colonies which showed a blue color were selected from the stock plates and were grown in SD liquid media without Leu. Plasmids were recovered from the yeast cultures, and were used to transform into *E. coli* HB101 by electroporation. The cDNA plasmids were cloned from bacterial colonies grown on M9 minimal media due to complementation of *leuB* mutation by leu2 gene in cDNA plasmids .

Cloned cDNA plasmids were introduced into *S.cecerevisiae* strain SFY526 and tested for β-galactosidase activity. Colonies were incubated for 6 days on SD (Leu⁻, Trp⁻) plates at 30°C, and scored for the β-galactosidase activity for 24 hours. Verifications of target specificity were performed with pVA3 (containing murine p53 oligomerization domain in pGBT9) and pLAM5' [human lamin C(66-230) in pGBT9] using two yeast strains, HF7c and SFY526.

5

## Sequence analysis and screening of the Full-length cDNA clones

The cDNA inserts were excised from pGAD424 and cloned into pBluescript SK(-) (STRATAGENE) plasmids. The sequences of both strands were determined by an automated sequencer (LI-COR). Homology search of the five cDNA clones by known nucleotide and protein sequences was performed with the BLAST program on the NCBI blast network server. The full-length cDNAs of the clone 34 (termed Fap-1) and 82 (Fap-2) were obtained by screening a λZAPII (STRATAGENE)-based mouse newborn brain cDNA library using each cloned cDNAs as probes under standard hybridization conditions. Sequencing ware performed, mainly by the subcloning of small restriction fragments into pBluescript vectors, and partial primer walking.

## Northern Blot Analysis

Total RNA extraction and poly $A^+$ RNA purification have been described above. Equal amounts of total RNA (20 µg per lane) or polyadenylated RNA (4 µg per lane) from mouse tissues was electrophoresed through a 1% agarose gel containing 2.2M formaldehyde and transferred to nylon membranes (Gene Screen, Du Pont), using a downward capillary blotting system (Chomczynski and Mackey, 1994). Cloned cDNAs and Fyn cDNA fragment from translation initiation site to internal SpII site were labeled by $[^{32}P]$d-CTP(Amersham) with a Random Priming kit (TaKaRa) and used as probes. Membranes were hybridized with each probe at 42°C for 18 hr in 5X SSPE, 5X Denhardt's solution, 1% SDS, 100 µg/ml sperm DNA, and 50% formamide, then washed in 2X SSPE, 2% SDS at 62°C. Filters were exposed to Fuji RX X-ray film for autoradiography.

## RT-PCR assay of Fap-1

Total RNA isolation from each tissue was performed as above, and 5 µg was used in first strand synthesis reactions with 1µg of oligo-d(T) as a primer for the reverse transcriptase (Superscript, GIBCO) reactions. One tenth volume of each product was

used as template for PCR amplifications using 25 pmol of the following primers in a reaction volume of 50 μl: 5'-CAGGGATGGATTTAGAT-3' and 5'-TTACAGAAAG-CCACCAG-3'. These oligonucleotides correspond to the 3' non-coding region of Fap-1 cDNA (nucleotide positions 4744-4760 and 5523-5539). The reaction was cycled with a "hot start" reaction of 96°C (10 min.), 75°C (10 min.), 50°C (5 min.), and 72°C (15 min.), followed by 35 rounds of 95°C (45 sec), 53°C (25 sec), and 72°C (3 min.). Ten μl of the reaction product was electrophoresed in a 1% agarose gel.

# Results

## Cloning of Fyn-associated proteins by the yeast two-hybrid system

I constructed an expression vector (pBD-Fyn) which produces a Gal4 DNA-binding domain-Fyn fusion protein containing Src homology motifs (SH2 and SH3) and the amino terminal region unique to Fyn, but lacking its kinase domain (amino acid 1-269). The fusion protein is described as 'tagged Fyn' in Figure 1A. To prevent the possibility of growth inhibitory effects of yeast by its kinase activity (Florio et al., 1994), I eliminated the kinase domain of Fyn. When the pBD-Fyn plasmid DNA was introduced into a yeast reporter strain HF7c, the fusion protein of expected molecular size was observed by immunoblotting with anti-Fyn antibody. I also confirmed that the transformants could not colonize on His-minus plate and did not express β-galactosidase activity. To screen cDNA-encoding proteins capable of interacting with the fusion protein, I introduced a second expression plasmid which contains sequences for the Gal4 activation domain fused to a cDNA library from mouse neonatal brain into HF7c strain expressing the 'tagged Fyn'. The general scheme is outlined in Figure 1B. If the second fusion protein interacts with the first fusion protein, the binding complex will regain the ability to activate transcription of the reporter genes; *HIS3* (imidazole glycerol phosphatase dehydratase) and *E.coli lacZ* (β-galactosidase). Each reporter gene is controlled by a different promoter but commonly contains Gal4 binding sites [*GAL1* upstream activating sequence (*GAL1* UAS)] (Bartel, 1993).

A total of approximately $3 \times 10^6$ transformants of HF7c were placed on His minus plates, 367 colonies appeared within two months under selection and 169 of them displayed blue colored β-galactosidase activity. The time needed for colonies producing β-galactosidase to turn blue, however, varied from several hours to one week. The plasmid cDNAs were recovered from these 169 His[+] LacZ[+] colonies. Partial sequencing analysis of all of the plasmid cDNAs revealed that at least 15 colonies had a duplicated plasmid cDNA. To eliminate putative false positives, I reintroduced the remaining 154 plasmid cDNAs into another pBD-Fyn transformant of yeast from SFY526 strain,

8

whose β-galactosidase reporter gene expression is under control of different promoter sequences from that of HF7c except for the presence of *GAL1* UAS. Most of the transformants also showed enzymatic activity which turned colonies blue. The time needed to develop color varied among colonies and some showed only faint levels. Therefore, to select those showing reliable enzymatic activity, I performed a β-galactosidase assay in a short time period (within 24 hours) using three independent transformations for each plasmid. Five cDNA clones, clones 27, 34, 40, 48 and 82, showed strong and reproducible β-galactosidase activity in the presence of pBD-Fyn as shown in Fig. 1C, though all five failed to interact with several other GAL4 binding domain-fusions, including murine p53 oligomerization domain (pVA3 of Figure 1C).

**Sequence analysis of obtained clones**

The cDNA inserts of clones 27, 34, 40, 48 and 82 were subjected to sequencing and homology search analyses to estimate functions of the encoded proteins (Fig. 2).

The predicted amino acid sequence of clone 27 cDNA contained a remarkable number of short tandem repeats (Fig. 2). A repeat composed of ten amino acids(-LASN-TMDSQM-) was especially highly conserved in the proximal 190 amino acids to the Gal4 activator sequence. Another type of repeat containing leucine every 7 residues followed but was less conserved (dotted leucine residues in Fig. 2). These stretches of leucine in every 6 residues form a coiled-coil protein dimerization motif, the leucine zipper. The most distal region of the deduced sequence was rich in tyrosine and proline residues, and two possible sites for the direct association with Fyn were found. One was composed of three repeats of -SYT(D/E)-, which was similar to tyrosine phosphorylation and the SH2 association site for Src family kinases observed in the mouse PDGF receptor β subunit (Mori et al., 1993). The other was the -PPLPPEEPP-, which displayed good agreement with the consensus sequence of SH3 binding motif (-PXXP-) (Yu et al., 1994).

The sequence of clone 34 contained clustered polar and hydrophilic residues in the region proximal to the Gal4 activation domain showing similarity to the protease-

9

sensitive domain called PEST sequence (Rogers et al., 1986); this sequence was enriched in serine / threonine, glutamate / aspartate, and proline residues, with flanking positively charged amino acids. A consensus PXXP motif capable of binding to SH3 domain also existed in this region (-PPPEEP-).

Clone 40 cDNA encoded a small (121 amino acids) protein which had no motifs predictable of nuclear localization, membrane spanning, or binding to other proteins.

The polypeptide sequence which was encoded in clone 48 had a similar motif to the SH2 recognition site (Marengere et al., 1994; Zhou et al., 1995) and two poly-proline stretches capable of association with SH3 domain.

Clone 82 cDNA encoded polypeptides 163 amino acids long and had an internal stop codon. The sequence was rich in hydrophobic residues, especially alanine and glycine. Two consensus motifs could possibly interact with Fyn: one was the -PXXP-motif for SH3 binding (Yu et al., 1994), and the other was the binding motif for SH2 domain predicted using an oriented peptide library .

Presence of SH2 or 3 binding motifs in 4 of 5 positive clones indicated that our two-hybrid system was operating well to isolate the molecule binding to the non-catalytic domain of Fyn. As will be mentioned later, the peptide encoded in clone 48 cDNA had in fact been reported to associate with SH3 domain of Fyn.

**Homology to obtained clones**

The homology search analysis revealed that the predicted amino acid sequence of clone 27 cDNA showed a high degree of similarity to the large transcript of human SON DNA binding protein coding sequences (Bliskovskii et al., 1992) (Fig. 3). Therefore, this clone was thought to be a mouse homologue or a protein closely related to the human protein. In the distal end of our cloned region, the sequence homology to the human gene had a gap of nearly 100 amino acids long.

Clone 40 was identical to the predicted product of the tctex-1 gene that maps to the sterile locus of the mouse t-complex (Lader et al., 1989) (Fig. 4). This gene had been isolated from differential screening of testicular cell RNA obtained from wild type and

sterile *t/t* males, and showed 8-fold overexpression in testis of *t/t* homozygous mice relative to wild type.

Homology search analysis revealed that clone 48 encoded the central and carboxyl-terminal portion of a mouse pre-messenger RNA binding protein, heteronuclear ribonucleoprotein K (hnRNP K) (Matunis et al., 1992) (Fig. 5). Fyn SH3 domain binds strongly to the protein via its proline-rich stretches under solution-binding conditions (Weng et al., 1994). The carboxyl terminus of clone 48 was alternatively spliced, which generated the difference in 4 amino acid residues to the previously reported sequences of mouse (Fig. 5B). The same pattern of splicing is known in human (Dejgaard et al., 1994).

Clones 34 and 82 encoded novel sequences. In clone 34 no similar protein or DNA sequence was found in the available data base; in clone 82, there was a weak homology to an extracellular matrix protein, elastin, by virtue of these hydrophobic residues.

## Tissue-distribution of the five clones at adulthood

I performed Northern blot analysis using obtained sequences as probes to examine the expression of these genes. At mRNA level, all five cDNAs were detected in neonatal brain from which the cDNA library was generated (P0 of Fig. 7). To predict the functional roles of the five genes, I first examined their expression patterns and relative amounts in various adult mouse tissues and compared them with those of Fyn. Radiolabeled probes for each of these clones were hybridized to a set of Northern blots containing RNA isolated from forebrain, cerebellum, skeletal muscle, heart, thymus, spleen, liver, and kidney. Three of the 5 cloned genes (clones 27, 40, and 48) were widely expressed among these tissues (Fig. 6A to 6E). Clone 27, a putative homologue of human SON gene, strongly hybridized with a band at ~7 kb in forebrain, cerebellum, and thymus RNA  (Fig. 6A). Relatively modest signals were also detected in muscle, heart, kidney, and liver. The expression of the gene was very faint in spleen. Several lower bands (~2.8 kb) were also detected. These might be due to cross-hybridization or

11

signals from some related products. Clones 40 and 48 were expressed almost ubiquitously in each tissue. The single hybridization signal at 0.9kb of clone 40, almost identical to mouse tctex-1 gene, was found in each tissue (Fig. 6C). Size of the transcript matched well with that of the tctex-1 gene. Clone 48, identical to hnRNP K, also exhibited an ubiquitous hybridization signal as two bands (1.8 kb and 2.8 kb, Fig. 6D) except for spleen. Two forms of hnRNP K, the size of each transcript being 1975 and 2719 bp, were generated by alternative splicing with different sizes of 3'UTR (Ostrowski et al., 1994).

Two putative novel genes (clones 34 and 82), in contrast, showed remarkable features in distribution and relative amount of mRNA in each tissue. Clone 34 mRNAs, ~6.2 kb in size, were faintly detectable in forebrain and cerebellum (Fig. 6B). Using poly(A)$^+$ RNA derived from brain, two transcripts around the size of 6.2 kb were clearly detected (Fig. 8A). To analyze the expression pattern of clone 34 in more detail, I assayed its mRNA level by RT-PCR method with a set of primers corresponding to the 3' portion of the gene. Figure 8B shows the products amplified in various adult tissues. Although one of the highest levels was again found in brain, clone 34 mRNAs were also detected in heart, thymus, lung, ovary, and testis. Very low levels were found in liver, spleen, intestine, kidney, and muscle. This suggested that clone 34 was expressed fairly widely in adult tissues, albeit at low levels.

Clone 82 mRNA, ~6.2kb in size, was predominantly expressed in forebrain and cerebellum, and at lower levels in muscle and heart (Fig. 6E). No signal was detected in thymus, spleen, liver, or kidney. Very faint signals of smaller size (~3 kb) were detected in forebrain, cerebellum and heart. The expression of Fyn was high in forebrain and cerebellum, and was detectable in thymus and heart (Fig. 6F). This pattern of expression was similar to that of clone 82 in adult tissues. These results indicated that each gene had a different expression pattern in various tissues. In brain, where Fyn was extensively expressed, mRNA of all the clones were clearly detected.

## Developmental expression of five clones in mouse brain

To study the developmental profile of the genes, I probed a set of Northern blots containing RNA isolated from the brains of mice ranging in age from embryonic day 15 (E15) to adult (P60). For each clone except clone 34, expression was observed from E15 to P60 in the developing and mature central nervous systems (Fig. 7A to 7E). Expression of clones 27 and 48 was relatively constant throughout neural development similar to that of Fyn (Fig. 7A, D and F). In contrast, clones 40 and 82 showed a pronounced decrease in expression during maturation (Fig. 7C and 7E). The product of clone 40 held at maximum level between E15 and P10, rapidly decreased between P10 and P30, and held at low level until P60. The mRNA of clone 82 was most abundantly expressed at E15, gradually fell until P30 and was relatively constant thereafter.

The expression of clone 34 was faintly detected between E15 to P60 using total RNA blots (Fig. 7B). The RT-PCR analyses described above were performed using seven different stages of whole embryonic bodies, heads or brains (Fig. 8C). The products were undetectable in whole embryo at E13.5, levels were weak in E17.5 head, P0 and P4 brains, and most abundant in P7 brain. Thereafter, their level remained steady or slightly lower. This alteration of each product divided into three types: the first was constant throughout life, the second was down-regulated during post-neonatal development and the third was up-regulated from middle-embryonic to postnatal stages. The roles of these types in brain are discussed below. Since these analysis of expression revealed that the two novel cloned genes, clone 34 and 82, were surely expressed in mice, I termed these two genes Fap-1(clone 34) for Fyn associated protein 1, and Fap-2(clone 82), and performed screening of full-length cDNAs.

## cDNA sequences of Fap-1 and Fap-2

To isolate full-length cDNAs, a neonatal mouse brain cDNA library was screened by hybridization using Fap-1 and Fap-2 inserts as probes. First, Fap-1 was isolated from 4 different library plasmids that contained cDNAs overlapping each other. Sequence analysis revealed that the longest cDNA had two adjacent ATG initiation codon near the

5' end of it (Fig 9A a.a 1 and 3) in a good match to the canonical eukaryotic translation initiation consensus (Kozak, 1987), and encompassed complete coding sequence of Fap-1 encoded by 4296 nucleotides. Multiple stop codons in all three reading frames lay upstream of the distal ATG, suggesting that this may be the start site for translation of the normal mRNA. The 3'-untranslated region of the cDNA contains two polyadenylation signals they are trailed by poly (A). Thus calculated size of Fap-1 transcripts were 5625 and 6887 nucleotides and it matched well to a result obtained from Northern blot analysis. The sequence predicts the protein of 1432 amino acid residues, with a molecular mass of 156,604 daltons. The predicted amino acid sequence of Fap-1 is shown in Figure 9A.

The sequence analysis of Fap-1 cDNA revealed that the Fyn-binding region which was primary isolated in two-hybrid system was placed in the carboxyl terminal portion of the protein (Fig. 9A). Hydrophobicity analysis did not detect typical signal sequence at the N-terminal or any stretches of residues of sufficient length and hydrophobicity to constitute a membrane-spanning domain. Two striking stretches of hydrophilic, charged residues ware found. One consisted of the glutamate repeating 18 times (755-772 a.a), and the other was the lysine / arginine / histidine rich domain which spanned over 50 residues (577-639 a.a). This extremely positive charged domain (calculated pI= 11.46) included multiple nuclear localization signal (NLS) consensus sequences which appears to be sufficient for nuclear translocation of proteins. In the amino-terminal region of Fap-1, a zinc finger motif and a potential tyrosine phosphorylation site were found.

To obtain the full-length cDNA of Fap-2 , a neonatal mouse brain cDNA library was screened with clone 82 insert as a probe. I isolated two clones, which initiated from the same position but differed in size. Both the longer (6408 bp) and the shorter (1582 bp) cDNA clone had a potential initiator ATG codon in the 5' end of the inserts which obeys Kozak's rule, and contained an open reading frame encoding a 394 amino acid residues, with a molecular mass of 41,970 daltons (Fig.10A). However, the frame was open all the way to the 5'end of the clones, and thus I could not rule out the possibility that additional coding sequences lay farther upstream. The 3'-UTR of the longer cDNA

contained an ideal polyadenylation signal at the 3' end of the clone. The shorter cDNA had no typical polyadenylation signal, but the corresponding region of the longer cDNA to the 3' end of the shorter clone had an internal poly-A stretch, long enough to hybridize to oligo dT primer.

The Fyn binding region, which was initially isolated as clone 82, was the carboxyl-terminal portion of the protein and occupied roughly one third of the entire sequence (157 a.a). No hydrophobic stretch or nuclear localization signals was found throughout of the protein.

Comparison of the deduced amino acid sequence databanks demonstrated that Fap-2 belonged to a family of RNA-binding proteins RBPs (Burd and Dreyfuss, 1994). The most significant similarity (33% identity at the amino acid level) was shown with a RBP identified previously in *Caenorhabditis elegans*, termed fox-1 (Hodgkin et al., 1994). Many RBPs contain a conserved sequence of 80-90 amino acids, termed an RNA-binding domain (RBD), that include two short highly conserved motifs called RNP-1 and RNP-2. FAP-2 contained one RBD, in which RNP-1 and RNP-2 sequences were well conserved (Fig.10B). Interestingly, the similarity between Fap-2 and fox-1 was prominent at the middle portion of Fap-2 protein which include RBD. Nevertheless the amino- and carboxyl-portion of the protein shared no homology to fox-1 or other RBPs.

## Discussion

### Isolation of two novel molecules

I identified five cDNAs that interact directly with N-terminal sequence of Fyn tyrosine kinase using the two-hybrid system from mouse brain. Here, I isolated two novel cDNAs (Fap-1 and Fap-2) for positive binding to Fyn. Northern hybridization analysis showed that the mRNA of Fap-1 and Fap-2 were more extensively expressed in brain than other clones isolated. The mRNA of Fap-2 was predominantly expressed in brain and weakly in muscle and heart. Fap-1 mRNA was undetectable except in brain using Northern hybridization analysis (Figs. 6B and 8A), but by RT-PCR analysis it existed in all adult tissues examined. These findings imply the possibility of conserved functions of these genes in nonneural tissues and brains. Notably, these two molecules showed different expression patterns during brain development; Fap-2 had a down-regulated expression while Fap-1 was up-regulated. Although I hardly detected developmental alteration of the fyn mRNA level (Fig. 7F), subcellular localization of Fyn protein is shown to alter during brain development: throughout late fetal brain it is localized in the developing axonal tracts, but in adult brain it was not observed in most axon-rich regions except olfactory and vomeronasal systems, and was restricted to cell bodies and nuclei of neuronal and glial cell subpopulations (Bare et al., 1993). Fyn-deficient mice had two-phase abnormalities in brain: morphological abnormalities and reduction of axonal growth during the developmental stage and decrease of long-term potentiation at the mature stage. These observations suggest that these two novel molecules were independently implicated in Fyn-signalling pathways at early and mature stages.

The Fyn-associated region of Fap-1 was rich in proline residues (15% in all residues of which primary isolated as clone 34) and polar residues that constitute PEST regions, and potential binding sites for SH3 domain-containing proteins. Abi-2, an Abl tyrosine kinase interacting protein (Dai and Pendergast, 1995), partly shows similar

16

compositions of amino acid residues to those of Fap-1 but their functions in Abi-2 protein is not clear. The PEST region was originally found in each of ten short-lived proteins examined (half-lives less than 2h), and is detectable in various proteins including nuclear factors, enzymes, cytoskeletal/ structural and other proteins. This region is considered to correlate better with rapid degradation than with specific compartmentation or functional activity. To date, PEST regions have not been found in any previously reported molecules which can associate with Fyn. The PEST region is also known to exist in a subset of calmodulin-binding proteins and calpain substrates, and is suspected to bind calcium ion ($Ca^{2+}$) and/or mediate $Ca^{2+}$ dependent proteolysis (Wang et al., 1989). Fap-1 may also bind $Ca^{2+}$ to exert its function.

Many molecules which can bind to the non receptor-type tyrosine kinases (NTKs) are known to date, and some of them are relatively well characterized about their molecular nature for signaling functions. They are the receptors (Timson Gauen et al., 1992), linker proteins between the NTKs and other signalling molecules (McPherson et al., 1994), and cytoskeltal proteins (Cobb et al., 1994). These binding molecules to the NTKs are shown to be implicated in common and well distributed signalling pathways for cell proliferation and differentiation. In recent years, another type of the binding molecules to NTKs, like p130Cas (Sakai et al., 1994), HS-1 (Kitamura et al., 1995), Efs (Ishino et al., 1995), Abi-1 (Shi et al., 1995) and Abi-2 (Dai and Pendergast, 1995), are detected and their functions are argued. Fap-1 protein shows some similarities of protein structures to them, although it scarcely shows homology to any of these proteins in amino acid sequence. They are rich in hydrophilic amino acids and commonly have binding motifs to SH2 and/or SH3 domains. They are all thought to be cytoplasmic or nuclear prteins and some of them have motifs for transcription factors or nuclear localization signals. Among them, HS-1 and Efs are expressed in restricted tissues and others are widely expressed. Because the Fap-1 protein does not have any domains for enzymatic activity known to date, it strongly suggests a new signalling cascade of Fyn. The structural similarities of Fap-1 and other NTKs-binding proteins described above may imply that they have overlapped functions to some extent.

17

Data base analysis of clone 82 revealed weak homology to human and bovine elastin at the amino acids level. However, nucleotide sequences for clone 82 cDNA were not conserved in elastin cDNAs. This implies clone 82 is evolutionarly unrelated to elastin. Full-length cDNA of clone 82 (Fap-2) has a motif for RNA binding domain and is suspected to be a member of RNA-binding proteins (RBPs). The neural-enriched putative RBPs have been described in both vertebrates and invertebrates. In *Drosophila*, some RBPs express spesifically in neuronal nuclei. Musashi, one of the fly RBPs, is expressed in sensory organ precursor cells and P element insertion in *msi* locus brings aberrant differentiation of mechanosensory bristles (Nakamura et al., 1994). Another RBP of *Drosophila*, couch potato(*cpo*), is highly expressed in neuronal nuclei in the embryo, and *cpo* mutants exhibit several altered adult behaviors (Bellen et al., 1992). In addition to homology with other RBDs, both msi and cpo proteins have a long stretch of alanine/glutamine-rich regions. The functional significance of this domain is unclear, however, Fap-2 protein was also riche in alanine and glutamine, especially in the region shown to interact with Fyn. Although these RBPs of fly has no nuclear localization signal, they are localized to the nucleus and suggests that they may have roles in RNA processing rather than mRNA stability or nuclear-cytoplasmic transport. Two vertebrate neuronal RBPs have been reported to date. The FuD gene was isolated by screening an expression library with sera from patients with antibody-associated paraneoplasmic encephalomyelitis (Sekido et al., 1994). The other vertebrate RBP, NRP-1, was isolated from *Xenopus laevis*. nervous system (Richter et al., 1990). The identification of these neuronal RBPs suggests that the regulation of gene expression at the posttranscriptional level may play a significant role in the development and function of both verttebrate and invertebrate nervous systems. Details of expression profiles and subcellular localization of Fap-2 products will provide new insights into their cellular and neural function.

## Other genes isolated

Three of the five cDNAs isolated here encoded previously reported sequences, SON, tctex-1 and hnRNP K (clones 27, 40 and 48, respectively). Two of them ( SON

and hnRNP K ) are the nucleic acid binding proteins. Unexpectedly, an immunocytochemistry report stated that Fyn is present in nucleus of neurons and glial cells at significant levels (Bare et al., 1993). I also confirmed nuclear staining in the pyramidal neurons of the hippocampal formation by the monoclonal antibody to Fyn which I developed. This enabled us to speculate that these nucleic proteins are colocalized with Fyn in the nucleus or implicated in nuclear transporting of Fyn.

Human SON gene is originally cloned by hybridization to GC-rich genomic DNA sequences (Bliskovskii et al., 1992), or HLA-DRA promoter pyrimidine rich sequences (Mattioni et al., 1992). Transcripts of the gene have at least two alternative forms which hold a putative nuclear localization signal, a basic region extending over 200 amino acids, and a proline-rich acidic domain in common. Immunofluorescence analysis confirmed its localization in the nucleus (Mattioni et al., 1992). Functions of the SON protein are not clear but overexpression of the recombinant protein can alter cell morphology. This protein has homologous regions to the DNA binding site of the high mobility group Y (HMG-Y) protein, a family of relatively low molecular weight non-histone components in chromatin in its carboxyl terminus, and represents DNA binding activity. This suggests that SON protein participates in morphological dynamics of structures in nucleus.

hnRNP K is originally isolated by association of cysteine rich RNA sequences. Since it is a member of hnRNP particle, this molecule is regarded to function in the processing and transport of nuclear pre-mRNA (Matunis et al., 1992). In addition to binding of hnRNA, hnRNP K also binds tenaciously to the pyrimidine-rich stretch of DNA sequences like kappa B enhancer (Gaillard et al., 1994), and binding to the CT-rich element in c-*myc* promoter region represents transcription factor activity (Tomonaga and Levens, 1995). Recently, it was reported that Fyn shows transcriptional activity via induction of NF kappa B-like DNA-binding proteins (Hohashi et al., 1995), and this suggests Fyn and kappa B-binding proteins may work synergistically in transcriptional regulations.

Nucleotide sequence of clone 40 was homologous to the transcript of the tctex-1 gene family which is highly expressed in homozygous mice of t-haplotype. Since it lacks

signal sequences and hydrophobic stretches in the predicted amino acid sequences, tctex-1 gene product appears to be cytoplasmic proteins but its function has been elusive. Lader et al. (1989) showed that the transcript level in adult testis is 200-fold higher than other tissues (liver, brain, heart, spleen, and kidney) in mice. I did not examine the mRNA level in testis, but it showed ubiquitous expression in various tissues (Fig. 2C). This study is the first to show that tctex-1 is down-regulated during brain development at the mRNA level, perhaps indicating that tctex-1 functions in premature cells of neuron or glia during brain formation.

All three cDNAs described above were broadly expressed in various adult tissues including brain. This suggests the existence of conserved functions of Fyn in nervous systems in common with nonneural tissues.

## Putative binding sequence of isolated clones to Fyn

In the five cDNAs isolated here, all but clone 40 contained Pro-Xaa-Xaa-Pro (PXXP) motif which is considered to be the SH3 domain recognition site. It is probable that isolation of these four molecules is based on interactions with Fyn SH3 domain. The minimal sequence requirement for SH3 domain recognition contains a conserved PXXP motif, and flanking residues which determine binding conformation and specificity for SH3 domain targets (Yu et al., 1994; Rickles et al., 1995). Among the clones isolated here, the PXXP motifs were observed at the Pro330 in clone 27, Pro68 and Pro95 in clone 34, Pro87 and Pro107 in clone 48, Pro34 and Pro73 in clone 82, respectively (Figure. 9).

SH3 domains are widely found in proteins of eukaryotic cells which mediate various cellular functions by interactions with their ligands. For example, SH3 domain of Ash/ Grb-2 adapter molecule binds Sos, a guanine nucleotide exchange factor, and cause Ras activation(Buday and Downward, 1993; Li et al., 1993; Simon et al., 1993). GTPase dynamin, a possible mediator of neurotransmitter endocytosis, binds to and is activated by a subset of SH3 domains (Gout et al., 1993). In addition to enzymatic activation, some transcription factor-like proteins contain potential SH3 domain binding sites (Buyse

20

et al., 1995; Subramaniam et al., 1995). In the case of alpha epithelial amiloride sensitive Na+ channel, interaction with α-spectrin SH3 domain mediates localization of the channel to apical membranes of Na+ transporting epithelia, which is necessary for proper channel function (Rotin et al., 1994). Thus the profiles of expression and cellular localization among SH3 domain binding proteins imply their functional diversity. In our study, the four molecules containing PXXP motif showed distinct expression patterns. These various complexes with Fyn might be generated by alteration of physiological condition during the developmental stages and in different brain regions, and might have different molecular functions.

Possible recognition sites for SH2 domain existed in clones 27, 48, and 82. SH2 domain are conserved among a series of membrane and cytoplasmic signaling molecules which associate with phosphotyrosine in a specific sequence context (Pawson and Gish, 1992; Zhou et al., 1995). These clones may have a potential for phosphorylation and association with SH2 domain *in vivo*.

I isolated a protein (hnRNP K), which was previously reported to associate with Fyn in this study; other molecules known to associate with Fyn, however, were not isolated. There are three possible reasons for this. First, some reported molecules might be indirectly associating with Fyn, because they were detected as an immune complex including Fyn, for example, Thy-1, a glycosylphosphatidylinositol (GPI)-anchored plasma membranous protein, associates with Fyn and other tyrosine kinases via 100kDa transmembrane protein (p100) (Lehuen et al., 1995). Second, there may be some requirements for posttranslational modifications for binding to Fyn which do not occur in yeast. Finally, the library I used might lack the Fyn-binding molecules as fusion protein with the Gal4 activator. Nevertheless, isolation of hnRNP K, known to be a Fyn-binding protein *in vitro*, confirmed that our detection system for protein-protein interaction in yeast operated accurately to screen Fyn-binding molecules.

## Conclusion

This study suggested that Fyn interacted directly with multiple proteins: nuclear, RNA-binding, cytoplasmic and novel proteins in mammalian brain. These indicated that Fyn mediated multiple signaling pathways from the cytoplasm to the nucleus. Fyn-deficiency causes broad alteration in suckling, emotional, and spatial learning behaviors. One or more of five isolated molecules might mediate the signaling pathway to regulate these mammalian behaviors.

## Acknowledgments

# References

Abeliovich A, Paylor R, Chen C, Kim JJ, Wehner JM, Tonegawa S (1993): PKC gamma mutant mice exhibit mild deficits in spatial and contextual learning. Cell 75: 1263-1271.

Aiba A, Chen C, Herrup K, Rosenmund C, Stevens CF, Tonegawa S (1994): Reduced hippocampal long-term potentiation and context-specific deficit in associative learning in mGluR1 mutant mice. Cell 79: 365-375.

Bare DJ, Lauder JM, Wilkie MB, Maness PF (1993): p59fyn in rat brain is localized in developing axonal tracts and subpopulations of adult neurons and glia. Oncogene 8: 1429-1436.

Bartel P, L.,Sternglanz,R.,Fields,S. (1993): Using the two-hybrid system to detect protein-protein interactions. In D Hartley, A (ed): Cellular Interactions in Development. New York: Oxford University Press Inc, pp. 153-180.

Baum MJ, Brown JJ, Kica E, Rubin BS, Johnson RS, Papaioannou VE (1994): Effect of a null mutation of the c-fos proto-oncogene on sexual behavior of male mice. Biol Reprod 50: 1040-1048.

Beggs HE, Soriano P, Maness PF (1994): NCAM-dependent neurite outgrowth is inhibited in neurons from Fyn-minus mice. J-Cell-Biol 127: 825-833.

Bellen HJ, Kooyer S, D'Evelyn D, Pearlman J (1992): The Drosophila couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. Genes Dev 6: 2125-2136.

Bliskovskii VV, Kirillov AV, Zakhar'ev VM, Chumankov IM (1992): The human son gene: the large and small transcripts contains various 5'-terminal sequences. Mol-Biol-Mosk 26: 807-812.

Bourtchuladze R, Frenguelli B, Blendy J, Cioffi D, Schutz G, Silva AJ (1994): Deficient long-term memory in mice with a targeted mutation of the cAMP- responsive element-binding protein. Cell 79: 59-68.

Buday L, Downward J (1993): Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. Cell 73: 611-620.

Burd CG, Dreyfuss G (1994): Conserved structures and diversity of functions of RNA-binding proteins. Science 265: 615-621.

Buyse IM, Shao G, Huang S (1995): The retinoblastoma protein binds to RIZ, a zinc-finger protein that shares an epitope with the adenovirus E1A protein. Proc Natl Acad Sci U S A 92: 4467-4471.

Chomczynski P, Sacchi, N, (1987): Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal.Biochem 162: 156-159.

Chomczynski P, Mackey K (1994): One-hour downward capillary blotting of RNA at neutral pH. Anal Biochem 221: 303-305.

Cicchetti P, Mayer BJ, Thiel G, Baltimore D (1992): Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. Science 257: 803-806.

Cobb BS, Schaller MD, Leu TH, Parsons JT (1994): Stable association of pp60src and pp59fyn with the focal adhesion-associated protein tyrosine kinase, pp125FAK. Mol-Cell-Biol 14: 147-155.

Dai Z, Pendergast AM (1995): Abi-2, a novel SH3-containing protein interacts with the c-Abl tyrosine kinase and modulates c-Abl transforming activity. Genes Dev 9: 2569-2582.

Dejgaard K, Leffers H, Rasmussen HH, Madsen P, Kruse TA, Gesser B, Nielsen H, Celis JE (1994): Identification, molecular cloning, expression and chromosome mapping of a family of transformation upregulated hnRNP-K proteins derived by alternative splicing. J Mol Biol 236: 33-48.

Florio M, Wilson LK, Trager JB, Thorner J, Martin GS (1994): Aberrant protein phosphorylation at tyrosine is responsible for the growth-inhibitory action of pp60v-src expressed in the yeast Saccharomyces cerevisiae. Mol Biol Cell 5: 283-296.

Gaillard C, Cabannes E, Strauss F (1994): Identity of the RNA-binding protein K of hnRNP particles with protein H16, a sequence-specific single strand DNA-binding protein. Nucleic-Acids-Res 22: 4183-4186.

Gout I, Dhand R, Hiles ID, Fry MJ, Panayotou G, Das P, Truong O, Totty NF, Hsuan J, Booker GW, et al. (1993): The GTPase dynamin binds to and is activated by a subset of SH3 domains. Cell 75: 25-36.

Grant SG, O'Dell TJ, Karl KA, Stein PL, Soriano P, Kandel ER (1992): Impaired long-term potentiation, spatial learning, and hippocampal development in fyn mutant mice. Science 258: 1903-1910.

Han PL, Levin LR, Reed RR, Davis RL (1992): Preferential expression of the Drosophila rutabaga gene in mushroom bodies, neural centers for learning in insects. Neuron 9: 619-627.

Hodgkin J, Zellan JD, Albertson DG (1994): Identification of a candidate primary sex determination locus, fox-1, on the X chromosome of Caenorhabditis elegans. Development 120: 3681-3689.

Hohashi N, Hayashi T, Fusaki N, Takeuchi M, Higurashi M, Okamoto T, Semba K, Yamamoto T (1995): The protein tyrosine kinase Fyn activates transcription from the HIV promoter via activation of NF kappa B-like DNA-binding proteins. Int. Immunol. 7: 1851-1859.

Inamdar M, Vijayraghavan K, Rodrigues V (1993): The Drosophila homolog of the human transcription factor TEF-1, scalloped, is essential for normal taste behavior. J Neurogenet 9: 123-139.

Ishino M, Ohba T, Sasaki H, Sasaki T (1995): Molecular cloning of a cDNA encoding a phosphoprotein, Efs, which contains a Src homology 3 domain and associates with Fyn. Oncogene 11: 2331-2338.

Kitamura D, Kaneko H, Taniuchi I, Akagi K, Yamamura K, Watanabe T (1995): Molecular cloning and characterization of mouse HS1. Biochem Biophys Res Commun 208: 1137-1146.

Kozak M (1987): At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J Mol Biol 196: 947-950.

Lader E, Ha HS, O'Neill M, Artzt K, Bennett D (1989): tctex-1: a candidate gene family for a mouse t complex sterility locus. Cell 58: 969-979.

Lehuen A, Beaudoin L, Bernard M, Kearney JF, Bach JF, Monteiro RC (1995): T cell activation through Thy-1 is associated with the expression of a surface protein (p100) on a subset of CD4 cells. Int Immunol 7: 607-616.

Levin LR, Han PL, Hwang PM, Feinstein PG, Davis RL, Reed RR (1992): The Drosophila learning and memory gene rutabaga encodes a Ca2+/Calmodulin-responsive adenylyl cyclase. Cell 68: 479-489.

Li N, Batzer A, Daly R, Yajnik V, Skolnik E, Chardin P, Bar-Sagi D, Margolis B, Schlessinger J (1993): Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature 363: 85-88.

Marengere LE, Songyang Z, Gish GD, Schaller MD, Parsons JT, Stern MJ, Cantley LC, Pawson T (1994): SH2 domain specificity and activity modified by a single residue. Nature 369: 502-505.

Mattioni T, Hume CR, Konigorski S, Hayes P, Osterweil Z, Lee JS (1992): A cDNA clone for a novel nuclear protein with DNA binding activity. Chromosoma 101: 618-624.

Matunis MJ, Michael WM, Dreyfuss G (1992): Characterization and primary structure of the poly(C)-binding heterogeneous nuclear ribonucleoprotein complex K protein. Mol-Cell-Biol 12: 164-171.

27

McPherson PS, Czernik AJ, Chilcote TJ, Onofri F, Benfenati F, Greengard P, Schlessinger J, De Camilli P (1994): Interaction of Grb2 via its Src homology 3 domains with synaptic proteins including synapsin I. Proc Natl Acad Sci U S A 91: 6486-6490.

Miyakawa T, Yagi T, Taniguchi M, Matsuura H, Tateishi K, Niki H (1995): Enhanced susceptibility of audiogenic seizures in Fyn-kinase deficient mice. Brain-Res-Mol-Brain-Res 28: 349-352.

Miyakawa T, Yagi T, Watanabe S, Niki H (1994): Increased fearfulness of Fyn tyrosine kinase deficient mice. Brain Res Mol Brain Res 27: 179-182.

Mori S, Ronnstrand L, Yokote K, Engstrom A, Courtneidge SA, Claesson-Welsh L, Heldin CH (1993): Identification of two juxtamembrane autophosphorylation sites in the PDGF beta-receptor; involvement in the interaction with Src family tyrosine kinases. Embo J 12: 2257-2264.

Nakamura M, Okano H, Blendy JA, Montell C (1994): Musashi, a neural RNA-binding protein required for Drosophila adult external sensory organ development. Neuron 13: 67-81.

Nighorn A, Healy MJ, Davis RL (1991): The cyclic AMP phosphodiesterase encoded by the Drosophila dunce gene is concentrated in the mushroom body neuropil. Neuron 6: 455-467.

Ostrowski J, Van Seuningen I, Seger R, Rauch CT, Sleath PR, McMullen BA, Bomsztyk K (1994): Purification, cloning, and expression of a murine phosphoprotein that binds the kappa B motif in vitro identifies it as the homolog of the human heterogeneous nuclear ribonucleoprotein K protein. Description of a novel DNA-dependent phosphorylation process. J-Biol-Chem 269: 17626-17634.

Pawson T, Gish GD (1992): SH2 and SH3 domains: from structure to function. Cell 71: 359-362.

Richter K, Good PJ, Dawid IB (1990): A developmentally regulated, nervous system-specific gene in Xenopus encodes a putative RNA-binding protein. New Biol 2: 556-565.

Rickles RJ, Botfield MC, Zhou XM, Henry PA, Brugge JS, Zoller MJ (1995): Phage display selection of ligand residues important for Src homology 3 domain binding specificity. Proc Natl Acad Sci U S A 92: 10909-10913.

Rogers S, Wells R, Rechsteiner M (1986): Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science 234: 364-368.

Rotin D, Bar-Sagi D, O'Brodovich H, Merilainen J, Lehto VP, Canessa CM, Rossier BC, Downey GP (1994): An SH3 binding region in the epithelial Na+ channel (alpha rENaC) mediates its localization at the apical membrane. Embo J 13: 4440-4450.

Sakai R, Iwamatsu A, Hirano N, Ogawa S, Tanaka T, Mano H, Yazaki Y, Hirai H (1994): A novel signaling molecule, p130, forms stable complexes in vivo with v-Crk and v-Src in a tyrosine phosphorylation-dependent manner. Embo J 13: 3748-3756.

Sekido Y, Bader SA, Carbone DP, Johnson BE, Minna JD (1994): Molecular analysis of the HuD gene encoding a paraneoplastic encephalomyelitis antigen in human lung cancer cell lines. Cancer Res 54: 4988-4992.

Semba K, Nishizawa M, Miyajima N, Yoshida MC, Sukegawa J, Yamanashi Y, Sasaki M, Yamamoto T, Toyoshima K (1986): yes-related protooncogene, syn, belongs to the protein-tyrosine kinase family. Proc Natl Acad Sci U S A 83: 5459-5463.

Shi Y, Alin K, Goff SP (1995): Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal portion of the Abl protein, suppresses v-abl transforming activity. Genes Dev 9: 2583-2597.

Silva AJ, Paylor R, Wehner JM, Tonegawa S (1992): Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. Science 257: 206-211.

Simon MA, Dodson GS, Rubin GM (1993): An SH3-SH2-SH3 protein is required for p21Ras1 activation and binds to sevenless and Sos proteins in vitro. Cell 73: 169-177.

Subramaniam M, Harris SA, Oursler MJ, Rasmussen K, Riggs BL, Spelsberg TC (1995): Identification of a novel TGF-beta-regulated gene encoding a putative zinc finger protein in human osteoblasts. Nucleic Acids Res 23: 4907-4912.

Suzuki T, Okumura-Noji K (1995): NMDA receptor subunits epsilon 1 (NR2A) and epsilon 2 (NR2B) are substrates for Fyn in the postsynaptic density fraction isolated from the rat brain. Biochem Biophys Res Commun 216: 582-588.

Swope SL, Qu Z, Huganir RL (1995): Phosphorylation of the nicotinic acetylcholine receptor by protein tyrosine kinases. Ann N Y Acad Sci 757: 197-214.

Timson Gauen LK, Kong AN, Samelson LE, Shaw AS (1992): p59fyn tyrosine kinase associates with multiple T-cell receptor subunits through its unique amino-terminal domain. Mol-Cell-Biol 12: 5438-5446.

30

Tomonaga T, Levens D (1995): Heterogeneous nuclear ribonucleoprotein K is a DNA-binding transactivator. J-Biol-Chem 270: 4875-4881.

Umemori H, Sato S, Yagi T, Aizawa S, Yamamoto T (1994): Initial events of myelination involve Fyn tyrosine kinase signalling. Nature 367: 572-576.

Wang KK, Villalobo A, Roufogalis BD (1989): Calmodulin-binding proteins as calpain substrates. Biochem J 262: 693-706.

Weng Z, Thomas SM, Rickles RJ, Taylor JA, Brauer AW, Seidel Dugan C, Michael WM, Dreyfuss G, Brugge JS (1994): Identification of Src, Fyn, and Lyn SH3-binding proteins: implications for a function of SH3 domains. Mol-Cell-Biol 14: 4509-4521.

Yagi T. Aizawa S, Tokunaga T, Shigetani Y, Takeda N, Ikawa Y (1993): A role for Fyn tyrosine kinase in the suckling behaviour of neonatal mice. Nature 366: 742-745.

Yagi T, Shigetani Y, Okado N, Tokunaga T, Ikawa Y, Aizawa S (1993): Regional localization of Fyn in adult brain; studies with mice in which fyn gene was replaced by lacZ. Oncogene 8: 3343-3351.

Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW, Schreiber SL (1994): Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76: 933-945.

Zhou S, Carraway KLr, Eck MJ, Harrison SC, Feldman RA, Mohammadi M, Schlessinger J, Hubbard SR, Smith DP, Eng C, et al. (1995): Catalytic specificity of protein-tyrosine kinases is critical for selective signalling [see comments]. Nature 373: 536-539.

**A**

Fyn

Unique region    SH2 & SH3    Kinase domain

tagged Fyn

Gal4 DNA binding domain

**B**

tagged Fyn only

OFF

UAS    HIS3    His⁺

OFF

UAS    LacZ    White

Transform with
Activation Domain tagged
cDNA Library

Select His⁺ blue
transformants

active complex

ON

UAS    HIS3    His⁺

ON

UAS    LacZ    Blue

**C**

clone. No    27    34    40    48    82

**pBD-Fyn**

**pVA3**

**(p53 oligomerization domain)**

# Figure 1

32
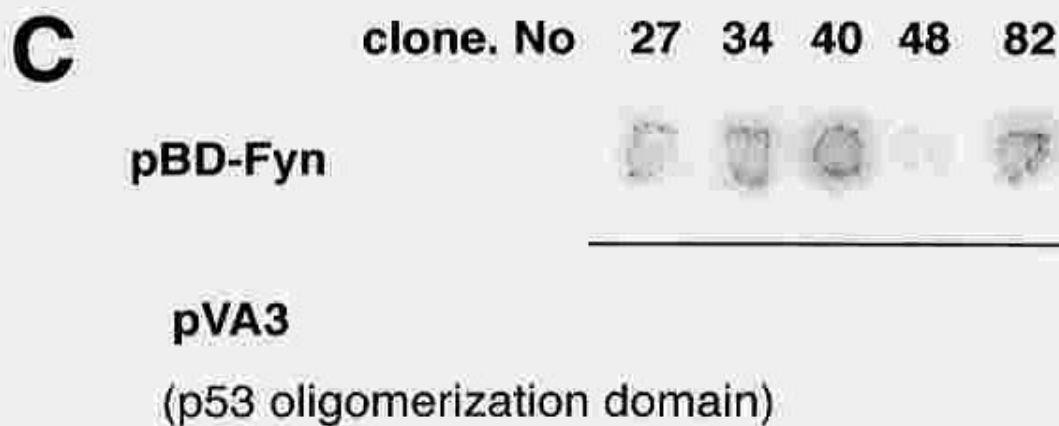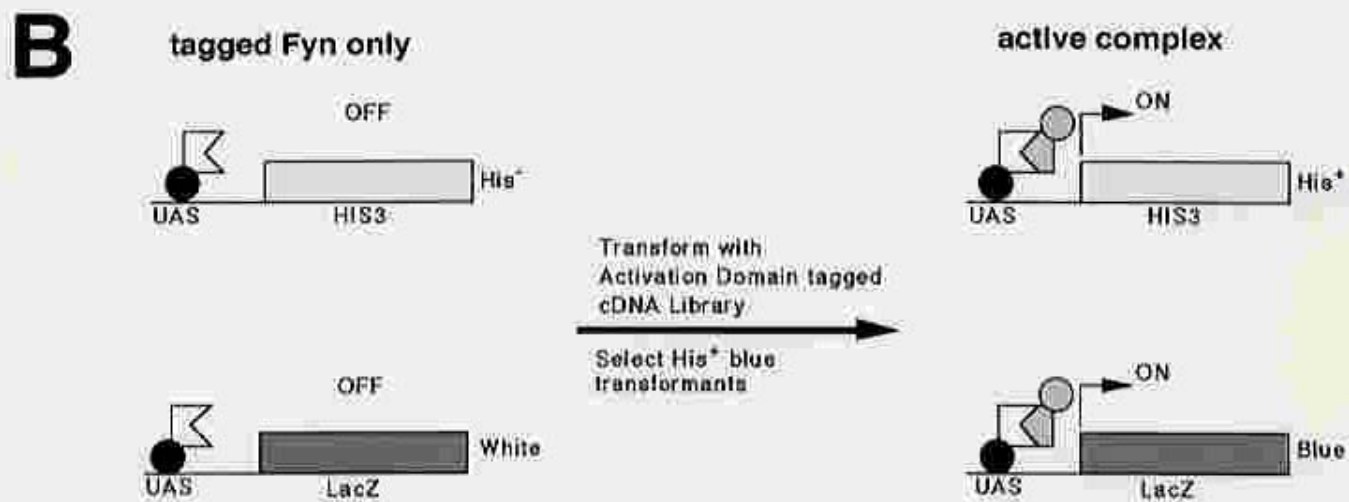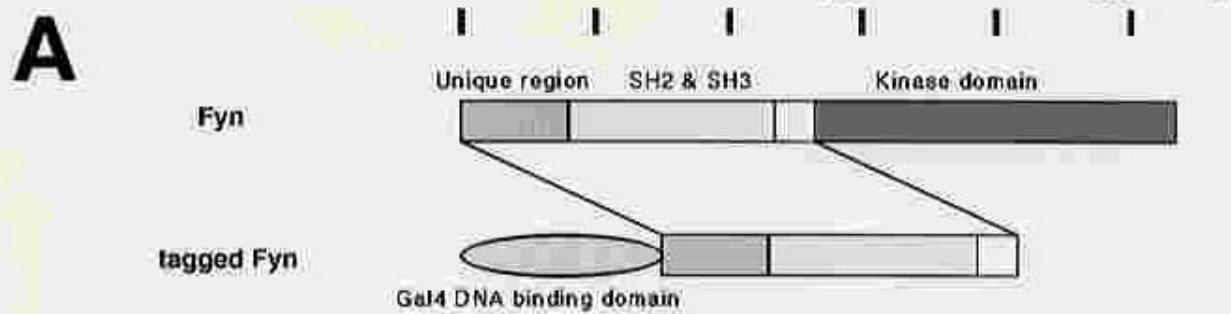
**Figure 1.** Strategy for isolating cDNAs encoding Fyn-associated proteins. **A:** Schematic domain structure of Fyn and tagged Fyn. Fyn protein contains the N-terminal unique region, and SH2, SH3 and tyrosine kinase domains characteristic of Src family members (upper). A fusion protein (tagged Fyn) was generated by connecting Fyn sequence except the kinase domain with Gal4 DNA-binding domain. In molecular cloning of the Fyn-association proteins, I used the tagged Fyn. **B:** The yeast cells (HF7c and SFY526) are initially transformed with a tagged-Fyn expression vector. These transformants are His⁻ (HF7c) and white (HF7c and SFY526) owing to lack of *trans*-activating sequence in tagged Fyn. Subsequent transformation with an activation-domain tagged cDNA of the expression library allows formation of an active transcription complex, when its expressed protein can interact with the tagged Fyn. This active complex induces expression of the reporter genes resulting in His⁺ and blue. Clones displaying these phenotypes are picked up and further analyzed. **C:** Demonstration of the blue selection on β-galactosidase activity from isolated plasmids encoding interacting tagged-Fyn. Each transformant of SFY526 by activator containing plasmids of clones 27, 34, 40, 48 and 82 was transformed again with pBDFyn (upper) or pVA3 (lower) containing the Gal4 DNA-binding domain. All turned blue only with pBDFyn, but not with pVA3. This indicated that the five cloned sequences could specifically interact with the sequence of tagged Fyn. Yeast strains were cultured for 6 days at 30°C and the replicated filters were incubated for 20 hr with Bluo-gal staining filter.

33

```
       Gal4 and Adapter        10         20         30         40         50         60         70         80         90        100
clone27  AKKEIEFAAA   LASNTMETHM LASNTMDSQM LASNTMDSQM LASNTMDSQM LASSTMDSQM LASSTMDSQM LATSTMDSQM LATSSMDSQM LATSSMDSQM LATSSMDSQM
                      LATSSMDSQM LATSSMDSQM LATSSMDSQM LATSSMDSQM LATSSMDSQM LASGAMDSQM LASGTMDAQM LASGTMDAQM LASSTQDSAM MGSKSPDPYR
                      LAQDPYRLAQ DPYRLGHDPY RLGHDAYRLG QDPYRLGHDP YRLTPDPYRV SPRPYRIAPR SYRIAPRPYR LAPRPLMLAS RRSMMMSYAA ERSMMSSYTD
                      RSMMSMAADS YTDSYTISYT EAYMVPPLPP EEPP

clone34  AKKEIEFAAA   LGNKPVLPMI GKLPATRKSN KKCEESGLER GEEQEHSEEE EGSPRSSDAP FGHQFSEEAA GPLSDPPPEE PKSEEATADH SVAPLGTPAH TDCYPGDPAI
                      SHNYLPDPSD GDTLESLDSG SQPGPVESSL LPIAPDLEHF PNYAPPSGEP SIESTDGTED ASLAPLESQP ITFTPEEMEK YSKLQQAAQQ HIQQQLLAKQ
                      VKAFPASTAL APATPAIQPI HIQQPATASA TSITTVQHAI LQHIAAAAAA AIGIHPHPHP QPLAQVHHIP QPHLTPISLS HLTHSIIPGH PATFLASHPI
                      HIIPASAIHP GPFTFHFVPH AALYPTLLAP

clone40  AKKEIEFAAA   ALKMEDFQAS EETAFVVDEV SSIVREAIES AIGGNAYQHS KVNQWTTNVL EQTLSQLTKL GRPFKYIVTC VIMQKNGAGL HSASSCFWDS STDGSCTVRW
                      ENKTMYCIVS TFGLSI*

clone48  AKKEIEFAAA   ECIKIILDLI SESPIKGRAQ PYDPNFYDET YDYGGFTMMF DDRRGRPVGF PMRGRGGFDR MPPGRGGRPM PPSRRDYDDM SPRRGPPPPP PGRGGRGGSR
                      ARNLPLPPPP PPRGGDIMAY DRRGRPGDRY DGMVGFSADE TWDSAIDTWS PSEWQMAYEP QGGSGYDYSY AGGRGSYGDL GGPIITTQVT IPKDLAGSII
                      GKGGQRIKQI RHESGASIKI DEPLEGSEDR IITITGTQDQ IQNAQYLLQN SVKQYSGKFF*

clone82  AKKEIEFAAA   AASSFQADVS LGNEAAVPLS GRGGINTYIP LIIPGFPYPT AATTAAAFRG AHLRGRGRTV YGAVRAVPPT AIPAYPGVVY QDGFYGADLY GGYAAYRYAQ
                      PATATAATAA AAAAAAYSDG YGRVYTADPY HALAPAASYG VGAVASLYRG GYSRFAPY*
```

# Figure 2

**Figure 2.** Deduced amino acid sequences of isolated tagged Fyn-binding cDNAs. The amino acid sequences between Gal4-activation domain and each 5' end of cDNAs containing an adapter are displayed in *italics* at the top left of each clone. Asterisks show the internal stop codons. The single-underlined sequences represent putative SH2-binding motifs including tyrosine residues that can be phosphorylated. Potential SH3-binding sequences are double-underlined. The two PEST-like structures of clone 34 are shaded. The leucine zipper-like repeats of clone 27 are pointed out.

35

**A**

```
          1   142                                    1035        1354 1460
human son ▬▬▬|▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬|▬▬▬▬▬▬▬|▬▬|

               9                                 903        904  1060
clone 27       |▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬|        |▬▬|
                                                   \      /
                                                    \    /
                                                     \  /
                                                      \/
```
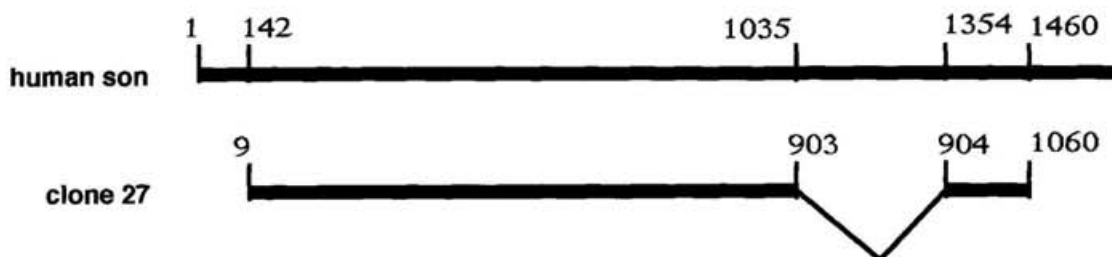
**B**

```
human son  142 :          T TAGCGTCTAG CACCATGGAC TCCCAGATGT TAGCAACTAG CTCCATGGAC
clone27      1 : GCGGCCGCTC ****T****A *********G ****T**** *****T*C*A TA********

           193 : TCCCAAATGC TAGCGTCCAA CACCATGGAC TCCCAGATGC TAGCATCCAA CACCATGGAC
            61 : *****G**** ****A***** ***T****** **T******* ********** T*********

           253 : TCCCAGATGT TAGCGTCTAG CACCATGGAC TCCCAGATGT TAGCAACTAG CTCCATGGAC
           121 : ********** ****C***** ********** ********** ****CT**** *A********

           313 : TCCCAGATGT TAGCAACTAG CTCCATGGAC TCCCAGATGT TAGCAACTAG CACTATGGAC
           181 : ********** ********** *A******** ********** ********** *T********

           373 : TCCCAGATGT TAGCAACCAG TTCCATGGAC TCCCAGATGT TAGCAACCAG CTCCATGGAC
           241 : ********** *******T** T**T****** ********** ********** *AGT******

           433 : TCCCAGATGT TAGCAACCAG CTCCATGGAC TCCCAGATGT TAGCAACCAG CTCCATGGAC
           301 : ********** ********** *AGT****** ********** ********** **********

           493 : TCCCAGATGT TAGCAACCAG CACCATGGAT TCTCAGATGT TAGCAACCAG CACCATGGAC
           361 : ********** ********** **GT****C **C******* ********** **GT******

           553 : TCCCAGATGT TAGCAACTAG CTCAATGGAT TCCCAGATGT TAGCATCTGG CACTATGGAC
           421 : ********** *******C** T**C**** ********** ********** TG*******T

           613 : TCTCAAATGT TAGCTTCTGG CACCATGGAT GCTCAGATGT TAGCGTCTGG TACCATGGAT
           481 : *****G**** ********** ***T****** ********** ****A**** ***T******

           673 : GCCCAGATGT TAGCGTCTAG TACCCAAGAT TCTGCTATGT TGGGTTCAAA ATCTCCTGAT
           541 : *****A**** ****A***** *****G*** *********A ********** G*********

           733 : CCCTATAGGT TAGCTCAGGA TCCTTACAGG TTAGCTCAGG ATCCCTATAG GTTGGGCCAT
           601 : ********** ********** *********A ********** ********** ******T***

           793 : GACCCCTATA GATTAGGTCA TGATGCTTAC AGGTTAGGAC AAGACCCTTA TAGATTAGGC
           661 : ********C* *G****C** ********** ********** ********** **********

           853 : CATGATCCCT ACAGACTAAC TCCTGATCCC TATAGGATGT CACCTAGACC CTACAGGATA
           721 : ********** ********** ***G****** *****G*A* ******G** ***T**A***

           913 : GCACCCAGGT CCTATAGAAT AGCACCCAGG CCATATAGGT TAGCACCTAG ACCCCTGATG
           781 : ********** *******G** ******T*** ****C**** ********** ****T*****

           973 : TTAGCATCTA GACGTTCTAT GATGATGTCC TATGCTGCAG AACGTTCCAT GATGTCATCT
           841 : ********** ********** ********** ********** ********** **********

          1033 : TACGAACGCT CTATGATGTC TTATGAGCGG TCTATGATGT CCCCTATGGC TGAACGCTCT
           901 : ***------- ---------- ---------- ---------- ---------- ----------

          1093 : ATGATGTCAG CCTACGAGCG CTCTATGATG TCAGCCTACG AGCGCTCTAT GATGTCCCCT
                  ---------- ---------- ---------- ---------- ---------- ----------

          1153 : ATGGCTGAGC GCTCTATGAT GTCAGCTTAT GAACGCTCCA TGATGTCAGC TTATGAACGC
                  ---------- ---------- ---------- ---------- ---------- ----------

          1213 : TCCATGATGT CCCCAATGGC TGATCGATCT ATGATGTCCA TGGGTGCTGA CCGGTCTATG
                  ---------- ---------- ---------- ---------- ---------- ----------

          1273 : ATGTCGTCAT ACTCTGCTGC TGACCGGTCT ATGATGTCAT CGTACTCTGC AGCTGACCGA
                  ---------- ---------- ---------- ---------- ---------- ----------

          1333 : TCTATGATGT CATCTTATAC TGCTGATCGT TCAATGATGT CTATGGCTGC TGATTCTTAC
                  ---------- ---------- -A******A ********** ********** **********

          1393 : ACCGATTCTT ACACTGACAC ATATACAGAG GCATATATGG TGCCACCTTT GCCTCCTGAA
           943 : **T******* *T*****TT* C****G*** ********** ********** **********

          1453 : GAGCCCCC
          1003 : *******AG CGGCCGC
```

**Figure 3.** Sequence comparison between clone 27 cDNA and human SON gene. **A:** A schematic diagram represents homologous and gapped region of clone 27 to SON gene. Numbers are positions of nucleotides. **B:** An alignment of nucleic acid sequences between SON (upper) and clone 27 (lower). Adapter sequences of clone 27 cDNA are underlined. Asterisks show identical nucleotides of clone 27 to SON gene. There are 928 nucleotides (92.6%) identical to each other except for a gapped region of clone 27 (dashed line).

```
mouse tctex-1   4 :          C GAGTAGAG[ATG]GAAGACTTC CAGGCCTCCG AGGAGACTGC ATTTGTTGTT
clone40         1 : GCGGCCGCTG CGT**A**** ********** ********** ********** **********

               55 : GATGAAGTGA GCAGCATTGT AAAGGAGGCT ATAGAAAGCG CCATCGGTGG TAATGCCTAC
               61 : ********** ********** ********** ********** ********** **********

              115 : CAGCACAGCA AAGTCAACCA GTGGACCACT AATGTCCTAG AACAGACTTT GAGCCAACTC
              121 : ********** ********** ********** ********** ********** **********

              175 : ACCAAACTGG GGAGACCATT TAAATACATT GTGACCTGTG TGATCATGCA GAAGAACGGT
              181 : ********** ********** ********** ********** ********** **********

              235 : GCTGGGTTAC ACTCCGCAAG TTCCTGCTTC TGGGACAGCT CCACAGACGG AAGCTGCACA
              241 : ********** *** ****** ********** ********** ********** **********

              295 : GTCCGATGGG AGAACAAGAC CATGTACTGC ATCGTCACTA CCTTCGGACT GTCCATC[TGA]
              301 : ********** ********** ********** ********** *** ****** **********

              355 : CCACCTTGCC AGCCTCAGCC TCGTGGTCCC AGCATTTCCA GTCCATCTCA ACCACCAGCT
              361 : ********** ********** ********** ********** ********** **********

              415 : ATGTTGGGCA AACCCCTTCC TCCTAAGTTG TTCTGTGGCA CTCTCAAAAA ACCAAATGAC
              421 : ********** ********** ********** ********** *C******** **********

              475 : TCCCCAGTCA CAGGAACCAC ATGACTTCAG CCAGATTAGC
              481 : ********** ********** ********** ********** AGCGGCCGC
```

# Figure 4

**Figure 4.** Sequence homology of tctex-1 gene to clone 40. Underlined sequences of clone 40 are adapters. Putative initiation ATG codon and stop codon (TGA) of tctex-1 gene are shaded. Asterisks of clone 40 indicate identical nucleotides to tctex-1.

## A

```
L29769   730 :          G AATGCATCAA GATCATCCTT GACCTTATAT CTGAGTCTCC CATCAAAGGA
L31961   700 :        * .......... .......... .......... .......... ..........
clone48    1 : GCGGCCGCT* .......... .......... .......... .......... ..........

         781 : CGTGCACAAC CTTATGATCC CAACTTTTAT GATGAGACCT ATGATTATGG TGGTTTTACA
         751 : .......... .......... .......... .......... .......... ..........
          61 : .......... .......... .......... .......... .......... ..........

         841 : ATGATGTTTG ATGACCGCCG AGGACGACCT GTGGGATTCC CCATGAGGGG AAGAGGTGGT
         811 : .......... .......... .......... .......... .......... ..........
         121 : .......... .......... .......... .......... .......... ..........

         901 : TTTGACAGAA TGCCTCCTGG TCGGGGTGGG CGTCCCATGC CTCCTTCTAG AAGAGATTAT
         871 : .......... .......... .......... .......... .......... ..........
         181 : .......... .......... .......... .......... .......... ..........

         961 : GATGATATGA GCCCTCGTCG AGGACCTCCA CCACCACCAC CTGGTCGAGG TGGCCGGGGT
         931 : .......... .......... .......... .......... .......... ..........
         241 : .......... .......... .......... .......... .......... ..........

        1021 : GGCAGCAGAG CCCGGAATCT GCCTCTTCCT CCTCCACCAC CACCCAGAGG GGAGGATCTA
         991 : .......... .......... .......... .......... .......... **GA*****
         301 : .......... .......... .......... .......... .......... **GA*****

        1081 : ATGGCTTATG ACAGAAGAGG AAGGCCTGGA GACCGCTATG ATGGCATGGT TGGGTTCAGT
        1051 : .......... .......... .......... .......**A G**....... ***C*****
         361 : .......... .......... .......... .......*G A**....... ***G*****

        1141 : GCTGATGAAA CTTGGGATTC TGCAATTGAC ACATGGAGCC CATCAGAATG GCAAATGGCT
        1111 : .......... .......... ********G .......... .......... ..........
         421 : .......... .......... ********C .......... .......... ..........

        1201 : TATGAACCAC AGGGTGTTTC TGGATATGAC TATTCTTATG CAGGGGGCCG TGGCTCATAT
        1171 : .......... ******G** .......... .......... .......... ..........
         481 : .......... ******G** .......... .......... .......... ..........

        1261 : GGTGATCTTG GCGGACCTAT TATCACTACA CAAGTAACTA TTCCCAAAGA TTTGGCTGGA
        1231 : .......... .......... .......... .......... .......... ..........
         541 : .......... .......... .......... .......... .......... ..........

        1321 : TCTATTATTG GCAAAGGTGG TCAGCGGATT AAACAAATTC GTCATGAA*C TGGAGCATCA
        1291 : .......... .......... .......... .......... .......... ..........
         601 : .......... .......... .......... .......... .......... ..........

        1381 : ATCAAAATTG ATGAACCTTT AGAAGGATCT GAAGATCGGA TCATTACCAT TACAGGAACA
        1351 : .......... .......... .......... .......... .......... ..........
         661 : .......... .......... .......... .......... .......... ..........

        1441 : CAGGACCAGA TACAGAACGC ACAGTATTTG CTGCAGAACA GTGTGAAGCA GTATGCAGAT
        1411 : .......... .......... .......... .......... _____ _____
         721 : .......... .......... .......... .......... _____ _____

        1501 : GTTGAAGGAT TC[TAA]TGCAA GATATTTTTT CTTTTTATA GTGTGAAGCA GTATTCTGGA
        1471 : .......... .......... .......... .......... .......... _____
         760 : _____ _____ _____ _____ .......... ._.......

        1561 : AAGTTTTTTC[TGAA]GACTAGTG AAGAACTGAA GGAGTCCTGC ATCTTTTTTT TTTTTATCTG
        1530 : .......... .......... .......... .......... ...._..... ..........
         781 : .......... .......... .......... .......... ...._..... ..........

        1621 : CTTCTGTTTA AAAAGCCAAC ATTCCTCTGC TTCATAGGTG TTCTGCATTT CAGGTGTAGT
         588 : .......... .......... .......... .......... .......... ..........
         810 : .......... .......... .......... .......... .......... ..........

        1681 : GAAATCTTTG CTGTTCACCA GATGTAATGT TTTAGTTCCT TACAAACAGG GTTGGGC
        1648 : .......... .......... .......... .......... ..._...... .......
         900 : .......... .......... .......... .......... ***A***** ****AGCGGCCGCG
```

## B

```
L29769   ATCAAAATTGATGAACCTTTAGAAGGATCTGAAGATCGGATCATTACCATTACAGGAACA 1440
          I  K  I  D  E  P  L  E  G  S  E  D  R  I  I  T  I  T  G  T
clone48  ...........................................................* 720

L29769   CAGGACCAGATACAGAACGCACAGTATTTGCTGCAGAACAGTGTGAAGCAGTATGCAGATGTTGAAGGATTCtaatgcaagatatttcctt 1530
          Q  Q  Q  I  Q  N  A  Q  Y  L  L  Q  N  S  V  K  Q  Y  A  D  V  E  G  F  Stop
clone48  ..................................................................................------------ 759

L29769   ctttttatAGTGTGAAGCAGTATTCTGGAAAGTTTTTCtaagactagtgaagaactgaaggagtcctgcatcttttttttttttttatctg 1620
clone48  ---------S  V  K  Q  Y  S  G  K  F  F  Stop***************************************_***************** 839
```

# Figure 5

40 de la page

40

**Figure 5.** Sequence comparison between clone 48 and two isoforms of mouse hnRNP K cDNAs previously reported. **A:** Alignment of nucleotide sequences. Designations of reported isoforms used are Genbank accession numbers. Underlined sequences of clone 48 are adapters and identical nucleotides among the three sequences are asterisked. Termination codons (TAA) are shaded. Note that clone 48 has a gapped region (dashed line) to the other two sequences, which results in an alternative amino acid sequence in the carboxyl terminus of the clone as shown in **B.**

# Figure 6



forebrain  cerebellum  muscle  heart  thymus  spleen  liver  kidney   Size in Kb

A    ←6.2
     ←3.9
     ←2.8

B    ←6.6
     ←5.0
     ←3.6

C    ←1.9
     ←0.9

D    ←3.6
     ←2.6
     ←1.9

E    ←6.2
     ←3.9
     ←2.8

F    ←5.0
     ←3.6
     ←2.6
     ←1.9

42
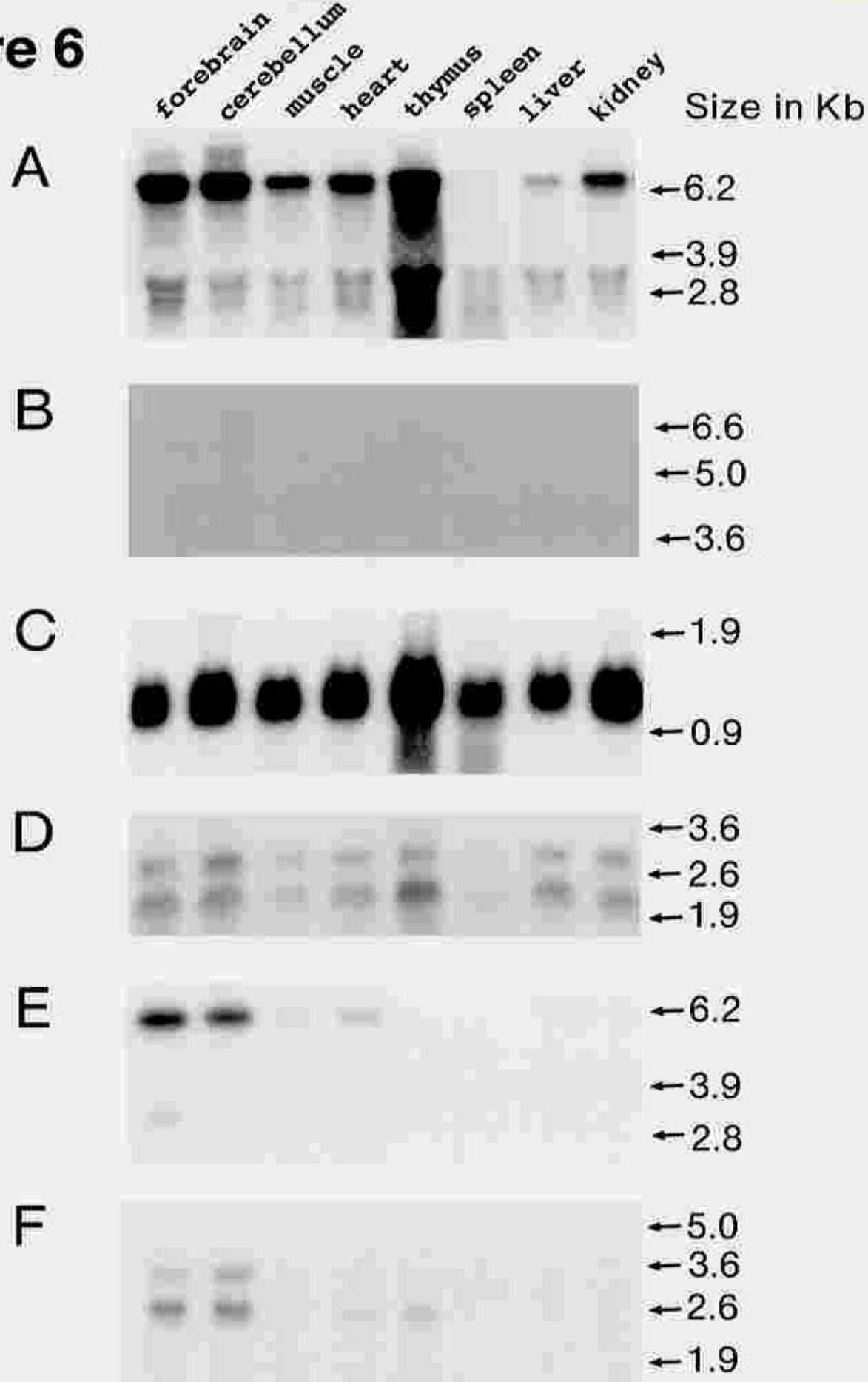
**Figure 6.** Northern blot analysis of the five cDNA clones and Fyn in adult tissues. Total RNAs from the indicated mouse tissues were hybridized to the radiolabeled clone 27 in (**A**), clone 34 (**B**), clone 40 (**C**), clone 48 (**D**), clone 82 (**E**), and Fyn cDNA (**F**), respectively. Exposure times are as follows: 24 hr (**D**), 40 hr (**A, C, E,** and **F**) and 72 hr (**B**). Each membrane was reprobed with GAPDH cDNA fragment and uniformity of RNA quantities in each lane was verified (data not shown). The RNA makers are indicated at left of each membrane.

## Figure 7

| | E15 | P0 | P10 | P30 | P60 | Size in Kb |

**A**
← 6.6
← 5.0
← 3.6
← 2.6

**B**
← 9.5
← 6.2
← 3.9

**C**
← 1.4
← 1.0
← 0.6

**D**
← 2.6
← 1.9
← 1.4

**E**
← 6.6
← 5.0
← 3.6

**F**
← 3.6
← 2.6

44
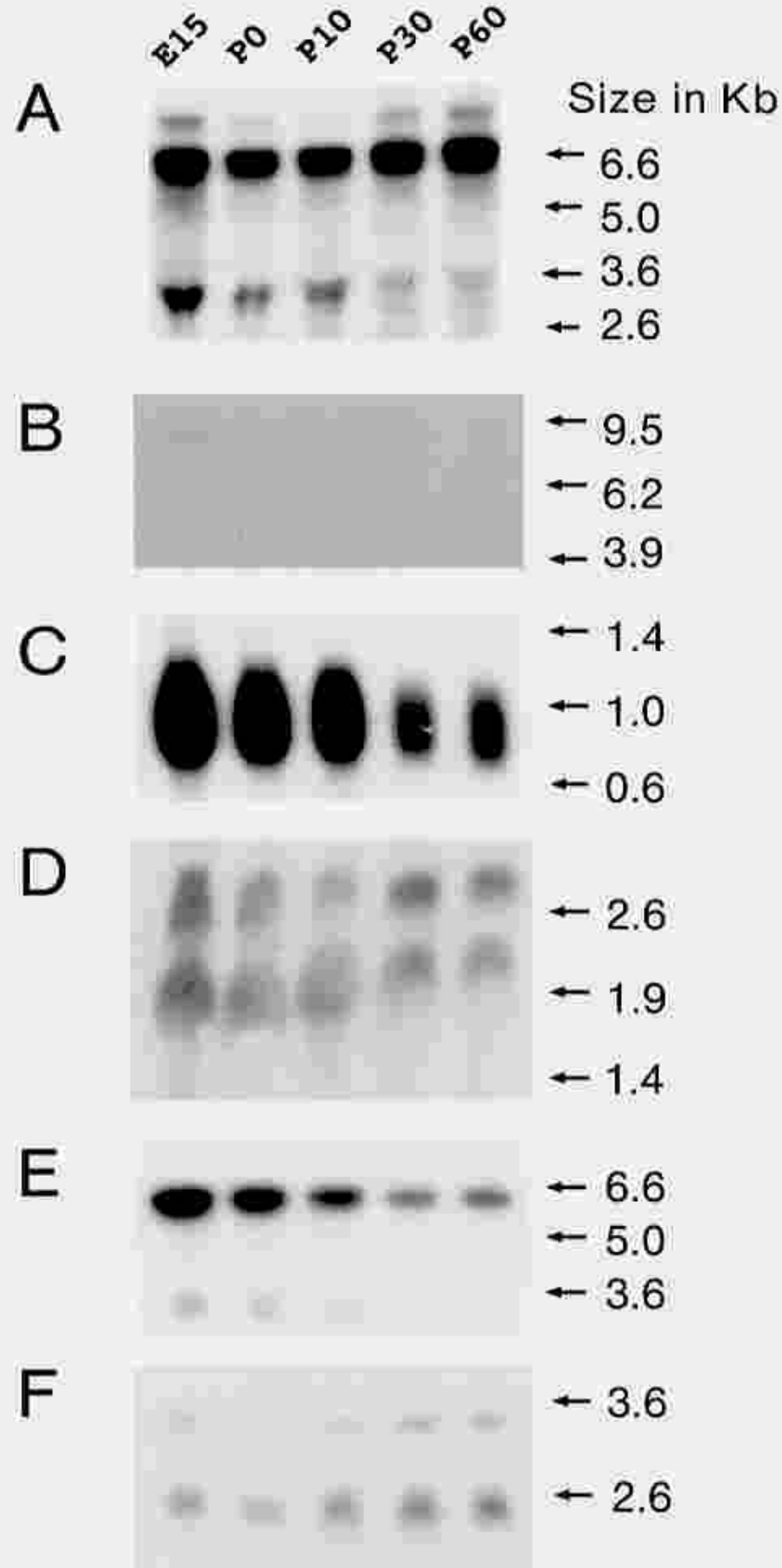
**Figure 7.** Developmental expression profile of the five cDNA clones and Fyn in brain. Total RNAs from the indicated ages (E15 to 8 weeks postnatal) were hybridized to the radiolabeled clone 27 in (**A**), clone 34 (**B**), clone 40 (**C**), clone 48 (**D**), clone 82 (**E**), and Fyn (**F**) cDNA, respectively. Exposure times are: 24 hr (**D**), 40 hr (**A, C, E,** and **F**) and 72 hr in (**B**). Each membrane was reprobed with GAPDH cDNA fragment and the uniformity of RNA quantities in each lane was verified (data not shown). The RNA makers are indicated at left of each membrane.

# Figure 8

A

liver pA+  brain pA+

9.5

6.2

3.9

Kb

B

φX HaeIII  muscle  testis  ovary  kidney  intestine  spleen  liver  lung  thymus  heart  brain

C

φX HaeIII  E13.5 whole  E17.5 head  P0  P4  P7  P14  adult

**Figure 8.** The expression profiles of Fap-1 mRNA. **A:** Comparison of the relative levels of expression between adult brain and liver by Northern analysis. The poly(A) RNA on nylon membrane was hybridized with clone 34 cDNA probe as those in Fig. 6 and 7. Exposure time was 40 hr. Molecular weight markers are indicated at right. RT-PCR analyses were performed on various adult tissues (**B**) and developmental stages of brain (**C**). Amplified products were separated on agarose gels with HaeIII digested φX174 DNA as a molecular weight marker (left). The predicted size of the amplified products was 795 base pairs long. No amplification was observed with 1 μg of genomic DNA as a template (data not shown).

```
GAATTCGGCACGAGGCGGCAAAATGGCGGACCGCTTCTCCCGCTTCAACGAGGACCGAGACTTTCAGGGTAATCACTTTGATCAGTATGA      90
AGAAGGACACTTGGAGATTGAACAAGCATCCCTTGACAAGCCCATAGAATCGCAGTTGGCCCATATTAGCTAAGCAGCCATGATCTGCCC     180
AGTTTTTAAAGCCAGGATAATATTGGACACCGCTTACTCCAGAAACATGGGTGGAAGCTGGGCCAGGGATTGGGAAAATCTCTTCAGGGG     270
AGAACAGACCCTATTCCAATCGTTGTCAAGTATGATGTCATGGGCATGGGTCGAATGGAAATGGAGCTGGATTATGCCGAGGATGCTACC     360
                                      M G M G R M E M E L D Y A E D A T      17

GAACGGCGTCGTGTCCTAGAAGTAGAGAAAGAAGACACAGAAGAGCTGAGGCAGAAGTACAAGGATTATGTTGACAAAGAAAAGGCAATT     450
E R R R V L E V E K E D T E E L R Q K Y K D Y V D K E K A I      47

GCAAAAGCCTTGGAAGATCTCAGAGCCAACTTTTACTGTGAACTTTGTGATAAACAGTATCAGAAACATCAGGAATTTGACAACCACATC     540
A K A L E D L R A N F Y C E L C D K Q Y Q K H Q E F D N H I      77

AATTCTTATGATCATGCACACAAGCAGAGATTAAAAGATCTCAAGCAGAGAGAGTTTGCCCGAAATGTCTCTTCAAGATCCCGCAAAGAC     630
N S Y D H A H K Q R L K D L K Q R E F A R N V S S R S R K D      107

GAGAAGAAACAAGAAAAAGCCCTTCGGCGGCTCCATGAGTTGGCAGAGCAAAGAAAACAAGCTGAATGTGCACCGGGAAGCGGGCCCATG     720
E K K Q E K A L R R L H E L A E Q R K Q A E C A P G S G P M      137

TTCCGACCAACCACGGTGGCTGTAGACGAGGATGGTGGAGAAGAGGATAAAGATGAGTCATCAACCAACAGTGGTGCAAGTGCTGTTTCT     810
F R P T T V A V D E D G G E E D K D E S S T N S G A S A V S      167

TCTTGTGGCTTTGGAGCCGACTTCTCCACAGATAAAGGGGGCTCCTTCACGTCAGTACAGATCACTAATACCACTGGACTGTCACAGGCT     900
S C G F G A D F S T D K G G S F T S V Q I T N T T G L S Q A      197

CCTGGCTTAGCCTCCCAAGGTATTAGCTTTGGCATTAAGAATAATCTGGGACCCCCACTGCAGAAATTGGGAGTATCATTTTCCTTTGCC     990
P G L A S Q G I S F G I K N N L G P P L Q K L G V S F S F A      227

AAGAAAGCTCCGGTCAAACTTGAATCAATAGCATCCGTTTTCAAGGACCATGCAGAGGAAGGGAGCTCAGAAGATGGAACGAAGGCTGAT     1080
K K A P V K L E S I A S V F K D H A E E G S S E D G T K A D      257

GAGAAGAGCTCTGACCAAGGGGTGCAGAAGGTGGGAGATACTGATGGCACTGGTAATCTTGATGGAAAGAAAGAAGATGAAGACCCTCAG     1170
E K S S D Q G V Q K V G D T D G T G N L D G K K E D E D P Q      287

GATGGAGGGTCCCTTGCCTCAACACTGTCCAAGTTGAAAAGGATGAAACGGGAAGAAGGAACAGGGGCTACAGAGCCAGAATATTACCAC     1260
D G G S L A S T L S K L K R M K R E E G I G A T E P E Y Y H      317

TACATCCCCCCAGCACACTGCAAGGTCAAACCTAATTTCCCCTTCTTACTCTTTATGAGAGCCAGTGAACAGATGGAAGGGGATCATAGT     1350
Y I P P A H C K V K P N F P F L L F M R A S E Q M E G D H S      347

GCACACTCAAAGAGTGCCCCCGAGAACAGAAAAAGCAGCTCTCCCAAGCCGCAAGGCTGTAGTAAGACAGCAGCAAGCCCAGGGGCAGAA     1440
A H S K S A P E N R K S S S P K P Q G C S K T A R S P G A E      377

AGAACAGTGAGTGAAGCTTCTGAGCTGCAAAAACCAAGCCGCTGTGCCTGGGCCTTCAGAGCCTGGAGGCAAAACTGAAACAAAGAAAGGC     1530
R T V S E A S E L Q K E A A V A G P S E P G G K T E T K K G      407

TCCGGAGGAGGGAAAGATGAGCAGAGTGTAGAGAGTAGGGAGACGTCAGAGAGCCCGATGTGTGAGTCCAATCCTAAAGACATTTCTCAG     1620
S G G G K D E Q S V E S R E T G E S P M C E L N P K D I S Q      437

GCCACCCCAGCAACAAAAGCAGGCCAGGGACCCAAGCATCCTACTGGTCCATTCTTTCCAGTTTTAAGCAAGGATGAAAGCACTGCCCTC     1710
A T P A T K A G Q G P K H P T G P F F P V L S K D E S T A L      467

CAGTGGCCATCAGAACTACTCATTTTCACCAAAGCAGAGCCTTCCATCTCCTACAGTTGTAATCCTTTATACTTTGACTTTAAACTTTCA     1800
Q W P S E L L I F T K A E P S I S Y S C N P L Y F D F K L S      497

AGAAACAAAGATGCTAAAGCTAAAGGGACAGAAAAGCCAAAAGATGTCGCAGGCTCCTCAAAGGATCATCTCCAGAGCCTTGATCCTAGA     1890
R N K D A K A K G T E K P K D V A G S S K D H L Q S L D P R      527

GAACCGAATAAAAGCCAGGAAGAGGAGCAGGATGTAGTGCTCTCTTCAGAAGGCAGAGTGGATGAACCTGCATCAGGGGCTGCCTGTAGC     1980
E P N K S Q E E E Q D V V L S S E G R V D E P A S G A A C S      557

AGCCTGAACAAGCAGGAGCCTGGGGGTAGCCATATGTCAGAAACTGAAGACACTGGGAGAAGCCATCCTAGCAAGAAAGAACCATCAGGC     2070
S L N K Q E P G G S H M S E T E D T G R S H P S K K E P S G      587

AAGTCTGACAGACACAAGAAGAAAAAGAAAACACAAAAAATCCAGCAAGCACAAACGTAAACACAAGGCTGACACGGAAGAGAAGAGTTCT     2160
K S D R H K K K K K H K K S S K H K R K H K A D T E E K S S      617

AAGGCAGAGTCTGGGGAGAAATCTAAGAAGCGCAAGAAACGAAAACGGAAGAAGAACAAATCATCAGCCGCAGCTGATTCTGAACGCGGA     2250
K A E S G E K S K K R K K R K R K K K N K S S A A A D S E R G      647

CCCAAATCAGAACCTCCTGGAAGCGGCAGCCCCGGCACCACCGAGAAGGCGGCGCCGAGCTCAAGATGATTCCCAGCGGAGATCCCTTCCT     2340
P K S E P P G S G S P A P P R R R R R A Q D D D S Q R R S L P      677
                                48
```

```
GCTGAAGAAGGAAACAGTGGCAAGAAGGATGATGGTGGGGGTGGTAGCAGTTGCCAAGATCACAGTGGGAGGAAACACAAAGGTGAACCA  2430
A  E  E  G  N  S  G  K  K  D  D  G  G  G  G  S  S  C  Q  D  H  S  G  R  K  H  K  G  E  P      70

CCAACTTCCTCCTGCCAGCGGAGAGCTAACACCAAACATAGCAGCCGGTCCAGCCATCGGAGCCAACCCAGTAGTGGTGATGAGGATAGT  2520
P  T  S  S  C  Q  R  R  A  N  T  K  H  S  S  R  S  S  H  R  S  Q  P  S  S  G  D  E  D  S      73

GATGATGCTTCCTCACACCGACTGCACCAGAAGTCTCCATCCCAGTACAGTGAGGAGGAGGAAGAGGAGGAGGAGGAAGAAGAGGAGGAG  2610
D  D  A  S  S  H  R  L  H  Q  K  S  P  S  Q  Y  S  E  E  E  E  E  E  E  E  E  E  E  E  E      76
                                                 -  -  -  -  -  -  -  -  -  -  -  -  -  -

GAAGAAGAGGAGGAAGATGAAGACTCCGGTAGTGAGCATTCTCGTAGCCGCTCTCGGTCTGGCCATCGCCATTCCTCACATCGTTCCTCC  2700
E  E  E  E  E  D  E  D  S  G  S  E  H  S  R  S  R  S  R  S  G  H  R  H  S  S  H  R  S  S      79
-  -  -  -  -  -

CGGCGCTCTTATTCTAGCAGCTCTGATGCCTCTTCAGACCAGAGCTGCTATAGTAGACAGCACAGTTACTCTGATGATAGCTATAGTGAC  2790
R  R  S  Y  S  S  S  S  D  A  S  S  D  Q  S  C  Y  S  R  Q  H  S  Y  S  D  D  S  Y  S  D      82

TATAGCGACCGATCACGAAGGCACTCTAAGCGCTCTCACGATTCAGATGATTCAGACTATACCAGCTCCAAACACAGGTCTAAACGACAC  2880
Y  S  D  R  S  R  R  H  S  K  R  S  H  D  S  D  D  S  D  Y  T  S  S  K  H  R  S  K  R  H      85

AAATACTCATCATCTGATGATGACTATAGCCTCAGTTGCAGCCAGTCCCGAAGCCGATCTCGGAGTCATACAAGGGAGCGATCAAGATCC  2970
K  Y  S  S  S  D  D  D  Y  S  L  S  C  S  Q  S  R  S  R  S  R  S  H  T  R  E  R  S  R  S      88

CGGGGTCGAAGCCGCAGTAGCAGCTGTAGTCGCAGTCGAAGCAAGAGGAGAAGTCGCAGCACCACAGCCCACAGCTGGCAGCGAAGCCGA  3060
R  G  R  S  R  S  S  S  C  S  R  S  R  R  S  K  R  R  S  R  S  T  T  A  H  S  W  Q  R  S  R   91

AGCTATAGCCGGGACCGGAGCCGCAGCACCAGGAGCCCTTCTCAGAGATCAGGCTCCAGAAAGGGCTCATGGGGTCATGAGAGCCCAGAG  3150
S  Y  S  R  D  R  S  R  S  T  R  S  P  S  Q  R  S  G  S  R  K  G  S  W  G  H  E  S  P  E      94

GAAAGGCGCTCTGGCCGCCGGGATTTCATTCGTTCAAAGATCTACCGCTCTCAATCCCCCCACTATTTCCAATCAGGTCGGGGGAGAAGGT  3240
E  R  R  S  G  R  R  D  F  I  R  S  K  I  Y  R  S  Q  S  P  H  Y  F  Q  S  G  R  G  E  G      97

CCTGGAAAGAAAGAAGATGGCAGAGGAGATGACAGTAAAGGAGCAGGCCTGCCCTCCCAGAATAGCAATACTGGCACAGGAAGGGGGTCA  3330
P  G  K  K  E  D  G  R  G  D  D  S  K  G  A  G  L  P  S  Q  N  S  N  T  G  T  G  R  G  S      100

GAAAGTGACTGCAGTCCTGAAGATAAGAATTCTGTTACTGCCAGACTGCTGCTACAGAAGATCCAGTCCAGGAAAGTGGAGAGGAAACCC  3420
E  S  D  C  S  P  E  D  K  N  S  V  T  A  R  L  L  L  E  K  I  Q  S  R  K  V  E  R  K  P      103

AATGTGTGCGAGGAGGTGCTGGCCACCCCTAATAAGGCTGGGCTCAAGTACAAGAACCCCCCACAAGGTTACTTTGGGCCTAAGCTCCCC  3510
N  V  C  E  E  V  L  A  T  P  N  K  A  G  L  K  Y  K  N  P  P  Q  G  Y  F  G  P  K  L  P      106

CCGTCTCTTGGTAATAAGCCTGTTCTTCCAATGATAGGGAAGCTTCCAGCTACCCGGAAGTCCAATAAGAAATGTGAAGAGTCTGGCTTA  3600
P  S  L  G  N  K  P  V  L  P  M  I  G  K  L  P  A  T  R  K  S  N  K  K  C  E  E  S  G  L      109

GAAAGGGGAGAAGAGCAGGAACATTCAGAGCCAGAAGAAGGGTCCCCAAGGAGTAGTGATGCTCCATTTGGGCATCAGTTCTCAGAGGAA  3690
E  R  G  E  E  Q  E  H  S  E  P  E  E  G  S  P  R  S  S  D  A  P  F  G  H  Q  F  S  E  E      112

GCAGCTGGTCCCTTATCAGACCCTCCCCCAGAACAGCCAAAGTCTGAAGAAGCTACTGCTGATCACTCTGTGGCTCCGGCTAGGCACCCA  3780
A  A  G  P  L  S  D  P  P  P  E  P  K  S  E  E  A  T  A  D  H  S  V  A  P  L  G  T      1155

GCCCACACTGACTGCTACCCTGGGGATCCAGCCATCTCCCATAACTACCTCCCGGACCCCAGTGATGGGGATACTCTGGAGTCTCTGGAT  3870
A  H  T  D  C  Y  P  G  D  P  A  I  S  H  N  Y  L  P  D  P  S  D  G  D  T  L  E  S  L  D     118

AGTGGCAGTCAACCAGGCCCTGTGGAATCCAGCTTGCTGCCTATAGCCCCAGACCTTGAGCACTTCCCCAATTATGCACCTCCCAGTGGG  3960
S  G  S  Q  P  G  P  V  E  S  S  L  L  P  I  A  P  D  L  E  H  F  P  N  Y  A  P  P  S  G    121

GAACCTAGTATTGAATCAACAGATGGGACTGAGGATGCCTCCTTGGCTCCTCTCGAGAGCCAGCCCATCACCTTCACCCCTGAGGAGATG  4050
E  P  S  I  E  S  T  D  G  T  E  D  A  S  L  A  P  L  E  S  Q  P  I  T  F  T  P  E  E  M    124

GAGAAGTACAGCAAGCTCCAGCAGGCTGCACAGCAGCACATCCAGCAGCAGCTTCTGGCCAAACAAGTGAAGGCCTTTCCAGCCTCCACC  4140
E  K  Y  S  K  L  Q  Q  A  A  Q  Q  H  I  Q  Q  Q  Q  L  L  A  K  Q  V  K  A  F  P  A  S  T  127

GCCCTAGCTCCAGCCACACCAGCCCTGCAGCCCATCCACATTCAGCAGCCAGCCACAGCCTCTGCCACCTCCATCACCACTGTTCAGCAT  4230
A  L  A  P  A  T  P  A  L  Q  P  I  H  I  Q  Q  P  A  T  A  S  A  T  S  I  T  T  V  Q  H    130

GCCATCCTACAGCATCATGCTGCAGCCGCTGCTGCCGCCATTGGCATTCACCCTCACCCCCACCCTCAGCCGCTTGCTCAAGTACATCAT  4320
A  I  L  Q  H  H  A  A  A  A  A  A  A  I  G  I  H  P  H  P  H  P  Q  P  L  A  Q  V  H  H    133

ATTCCCCAGCCCCCATCGTAACCCCTATTTCTTTGTCCCATCTCACTCACTCAATTATCCCTGGCCCACCCTGCCCACCTTTCTTGCTAGCCAC  4410
I  P  Q  P  P  H  L  T  P  I  S  L  S  H  L  T  H  S  I  I  P  G  H  P  A  T  F  L  A  S  H  136
```

49

```
CCTATCCATATAATTCCTGCCTCAGCCATCCATCCTGGGCCCTTCACCTTTCATCCCGTCCCACACGCTGCCCTCTACCCCACCCTGCTT  4500
 P  L  H  L  P  A  S  A  I  H  P  G  P  F  T  F  H  P  V  P  H  A  A  L  Y  P  T  L  L  1397

GCCCCACGGCCTGCTGCAGCAGCTGCCACAGCCCTCCATCTTCACCCACTGCTTCACCCCATCTTCTCAGGTCAGGACCTGCAGCACCCT  4590
 A  P  R  P  A  A  A  A  A  T  A  L  H  L  H  P  L  L  H  P  I  F  S  G  Q  D  L  Q  H  P  1427

CCCAGCCATGGGACTTGAGTTGGGAGTCGGAAGCCTGTGTTGAGCCAGGGAAGGTGGCCTGTTCTTCCCTTGGGGGTGTTGAGCCATTTA  4680
 P  S  H  G  T  *                                                                          1432

TACCAGCAGGTAAAAGCTGGAACGGGCAGTGTCTTGACAGCCAAAGAAATGAGTTTTGGCTTGCAGGGATGGATTTAGATTTGAGGTTTG  4770
CTTTGGGTTTACTATCAGGGATATTTACCCCCTTTCTATTAGTCCCTCCCTCCTGTGTTTCTAAGCTAGGGAACACCAGTAAGTGCTACC  4860
CACACTCCTCTGCAGCAACATCTCCAGACTGTCAAGTAAGCCTCCTTCCTCCCTCTACTTTCTGATCCTCTGCACTTCCCCTCTGTAAGT  4950
TCAAAACATTTTCCTCCTCTAATGGGTTGTCACCAAAGCACTTGCCATGGGCCCATCTGGTCCTGCCTTCTCAGTGCCACGAGAAAGGCC  5040
AGGCTGTCTCTTCAGTGGTGAAAACTCCACTTTTCGGCAGTGTCTTTGACTATCCGTTTGTGACACGTTATCCCAGGTGAGGTGGAAGAG  5130
GGAGTCTAACCTAGAACCAAAAGCAACTCAGGCTGCAATTTGATCCACTTTCCTTACAAACAGAGCCATGCATTTGGCACTACTGTCCTC  5220
ATGCCCCTCATCCCAGAAGAGGCATCTACATATCCCAGAAAAGAAGATCCTATCGTGAGTAGAAGCTACACTGAACATAGGGTCACCCAT  5310
GTGCTTCCCAGAAAGGAGCTCGTGTCTGAAGGGTATATTTCCTTCTGTTATGTTGATGTTTTCTTCTTAATTTATAGGATTTTTTAATGT  5400
AAAATTGCACTGAACTCTATAAACATAAATGCTGCTTTTTGTAGCTCATATTAAAAAGGTTCCGTATTTGTAGTATACTGCAATAATATT  5490
TTTTACATCCAGCTCTTTAAGTGATCCTTGTTCTGGTGGCTTTCTGTAAATATGTTTGCCCTAGTTGAATTAAGAAACTTTAAAGGTTAT  5580
AAAAGAAAAAAAATAAAATATTTGTTTTCTGCTAGATGCAAAATGACTAAGCATGTATATTTTATCACCCTCCTGTATTTTTTGAAGTT  5670
ATGAACTATATAAAAATGTTGACACTTTTTTTTGTTTTGTTTTGTTTGCTGGGACAGTGTTAAATAATTTGCAGTAGTTGGTTCATCCTG  5760
CACTGGGACTCTCGGGTTGATGAGCAGCAGGCGGCTGCTGCTGTCAGGAGGATGCACAGTTGCAGTGGTCTCCCTGACGTGCTTTGTTAT  5850
ACAGAGAAACCGTGGACTCCATGGAGGAGGGGAGGATTGGTTCAACAGCTTTATTCTCTGGAAGCAGTGACACTCGGTGATTGTCTGGGA  5940
AAAGGGGTCCTTTGGTTTCCCTGGGTACCTTGTTCCAGATAGTCTTATTTGCCTTCCTAGTATTTAGCCCCCTCTGACAATTTTTTAAAT  6030
GTGCCTTGTGGGTTGGCTAGACACTAAAATAGAAGTAGACCAGAGGATACTAGAGTTGGGACAGTTCTCTCTTTGTTGGGAGGTGAGCAC  6120
GTTTATAGGTAGAGAGATTTTTGGCTTCACAGTTCTGTGCCGTCTGCCCAGCTCGATTGCTATTTCAGCAGTAGAGGTTCCTTTTTTTTT  6210
TTTTCTCCAGTGACACTGACCAGAAGTCTGTCCATCCCTCTCCTCTTTGCCCTTCCCCCAGGCTCTCAGCCCTCCTTAACACTGACCCAA  6300
AGGTGCTTGTAATGCAGGTCTCATCTCCTTTGTTAGTGCTCTTCCTATGACTAGAACTGGTTAAGCAGTCAACTGACTGTTAACCCTGCT  6390
TGCCTAGGCCAAGGTTCTCTGTTCCTGCCTTCTCAAATTCCCAGCAGAGTGGGCTGTGTTAGGGAAGTGAGTGCTGGGCACTTGAGGCCC  6480
GTCCCTGCCTGAGCTTCTCCTACCTTCTGCTCATTGTCCCCTTACAAGTATCACTTTTAGGATAACTAGCCCTGGAAAAAACGGTATCTGT  6570
TTGAAGTGAAGGGACTGGTTCTTCTTGCTTTCAGCCCGTGTAAATATTTAAATAAAACAGTATTAACATTTTTGTGCTGCTTCTCGTTAG  6560
GTTGTAGATCTAGCCATGCTCTCGGCCGCCTCTGCCATCCTGGAGGTAGTTTTCCTTAACTTCCAGAATAGTGATTTTAAAAACTTTAAAAA  6750
AAAAAGGGGGGGGGATAACCTATCCTTACATACAAGCATAACAGATTGTATCTAACTTTATCACATATGATAGAGATGTATATGCTGTAA  6340
AGTGAGGGGAAGGAGCTTTCTAATAAACCAGTGCTTTGTGCAAACTTAAAAAAAAAAAAAAAAAAAAA  6306
```

## A



## B

Figure 9

50
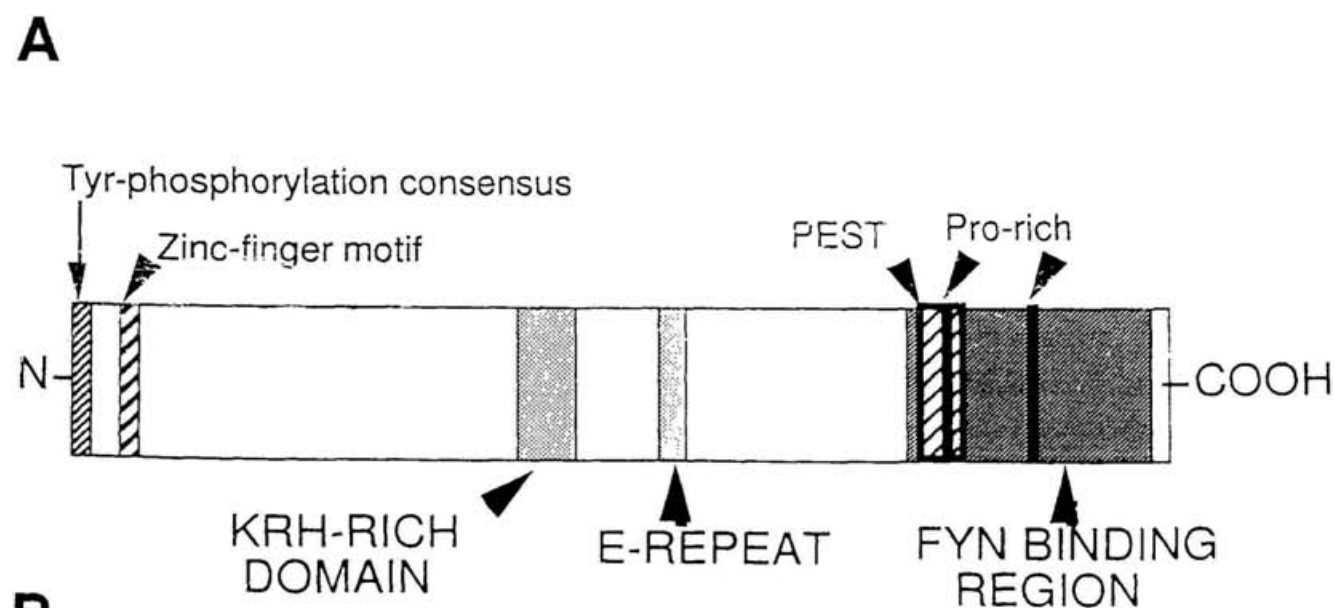
**Figure 9.** cDNA and deduced amino acid sequence of Fap-1 **A**: Sequence of the 6.9kb Fap-1 cDNA and the deduced amino acid sequence of the 1432 amino acid protein. The nucleotide and amino acid numbers are shown on the right. Two polyadenylation signals are underlined. The sequences which was obtained initially as two-hybrid clone 34 (predictable Fyn binding region) are shaded. Two potential SH3 binding sites (amino acid 1074-1077, -ProValLeuPro- and amino acid 1164-1167, -ProGlyAspPro-) are double-underlined. The two PEST sequences-like structures ( a.a 1099-1114 and a.a 1122-1141) are shown in boxes. The two charged amino acids clusters, the Glu-repeat and the Lys/Arg/His rich domain are each underlined with dashed lines. A zinc figer motif (a.a 60-82) and potential tyrosine phosphorylation site (a.a 5- 12) in the N-terminus of the protein are double underlined. **B**: Diagram of structual features of Fap-1.

```
AATTCCAGGAAGGAAGAGAGAGTGATGTCATCAGATTCCATGGACCAGCCCAGGAACCCCTGCGAGAACGAGCCACTGACCCCGGGTTAT   90
                      M  D  Q  P  R  N  P  C  E  N  E  P  L  T  P  G  Y    17

CATGGATTCCCAGCAAGGGATGGCCAGGGTAACCAGGAGCCAACAACAACTCCTGACGCAATGGTTCAGCCTTTTACTACCATCCCGTTC   180
 H  G  F  P  A  R  D  G  Q  G  N  Q  E  P  T  T  T  P  D  A  M  V  Q  P  F  T  T  I  P  F     47

CCACCACCTCCACAAAATGGAATTCCCACAGAATATGGAGTGCCACACACTCAGGACTATGCCGGCCAGACCAGTGAACATAACCTGACC   270
 P  P  P  P  Q  N  G  I  P  T  E  Y  G  V  P  H  T  Q  D  Y  A  G  Q  T  S  E  H  N  L  T     77

CTCTACGGGAGTACGCAGCCTCATGGAGAACAGAGTAGCAATTCACCCAGCAACCAGAATGGATCTCTCACGCAGACAGAAGGTGGAGCA   360
 L  Y  G  S  T  Q  P  H  G  E  Q  S  S  N  S  P  S  N  Q  N  G  S  L  T  Q  T  E  G  G  A    107

CAGACAGACGGACAACAGTCACAGACACAAAGTAGTGAAAAATTCAGAGAGTAAATCTACGCCCAAGCGACTACATGTCTCTAATATTCCC   450
 Q  T  D  G  Q  Q  S  Q  T  Q  S  S  E  N  S  E  S  K  S  T  P  K  R  L  H  V  S  N  I  P    137

TTCCGCTTCCGAGACCCTGACCTCCGGCAGATGTTTGGGCAGTTTGGCAAAATCCTAGATGTGGAAATAATCTTTAATGAGCGCGGCTCC   540
 F  R  F  R  D  P  D  L  R  Q  M  F  G  Q  F  G  K  I  L  D  V  E  I  I  F  N  E  R  G  S    167

AAGGGATTCGGGTTCGTAACTTTCGAGAATAGTGCTGATGCAGACAGGGCCAGGGAGAAATTGCACGGCACCGTGGTAGAGGGCCGTAAA   630
 K  G  F  G  F  V  T  F  E  N  S  A  D  A  D  R  A  R  E  K  L  H  G  T  V  V  E  G  R  K    197

ATCGAGGTGAATAATGCAACAGCACGGGTCATGACCAACAAGAAGATGGTCACGCCATATGCAAATGGCTGGAAGTTAAGCCCAGTAGTT   720
 I  E  V  N  N  A  T  A  R  V  M  T  N  K  K  M  V  T  P  Y  A  N  G  W  K  L  S  P  V  V    227

GGAGCTGTGTATGGTCCTGAGTTATATGCAGCATCCAGCTTTCAAGCTGATGTGTCCCTAGGCAATGAGGCGGCTGTGCCCTTGTCAGGA   810
 G  A  V  Y  G  P  E  L  Y  A  A  S  S  F  Q  A  D  V  S  L  G  N  E  A  A  V  P  L  S  G    257

AGAGGAGGCATCAACACTTACATCCCTCTAATCATTCCTGGCTTCCCTTACCCAACTGCAGCCACCACGGCAGCTGCTTTCCGAGGAGCC   900
 R  G  G  I  N  T  Y  I  P  L  I  I  P  G  F  P  Y  P  T  A  A  T  T  A  A  A  F  R  G  A    287

CATCTGAGGGGCAGAGGGCGGACAGTGTATGGTGCAGTCCGAGCGGTACCTCCAACAGCCATCCCCGCCTATCCAGGTGTGGTTTACCAG   990
 H  L  R  G  R  G  R  T  V  Y  G  A  V  R  A  V  P  P  T  A  I  P  A  Y  P  G  V  V  Y  Q    317

GACGGATTTTACGGTGCTGACCTCTATGGTGGATATGCAGCCTACAGATATGCACAGCCTGCTACTGCAACCGCAGCCACAGCTGCTGCA   1080
 D  G  F  Y  G  A  D  L  Y  G  G  Y  A  A  Y  R  Y  A  Q  P  A  T  A  T  A  A  T  A  A  A    347

GCCGCTGCAGCCGCTTACAGCGACGGTTACGGCAGGGTGTACACAGCTGACCCCTACCATGCCCTCGCCCCTGCCGCCAGCTATGGAGTT   1170
 A  A  A  A  A  Y  S  D  G  Y  G  R  V  Y  T  A  D  P  Y  H  A  L  A  P  A  A  S  Y  G  V    377

GGCGCTGTCGCGAGTTTGTACCGAGGTGGCTACAGCCGATTTGCCCCCTACTGAAGTGACGTGAGATCCCTGCAAATGGGACAGCCCCCA   1260
 G  A  V  A  S  L  Y  R  G  G  Y  S  R  F  A  P  Y                                          394

GTTCATGAGGCCTGGCTATTGCAATATTTACTAGTAGAGGACTCTATAGCAAGATGAAGAGGAAAACAAACAAAAAAGAGAGAATACTTT   1350
TTTATACCTCACTATGTTCTTTGAATATGTATTTTTTCCTTTAAATTTCTGCCTTTAATTCTTTTGTTCCAAAGATTGTGCTTTTTTTGT   1440
TTTGTTTTGTTTTTTTTGAATTTTTGTTTTTGTTTGGAGGGGGGGTTGGTTTTTTTGGGGGTTTTTTGTTTTGTTTTTTGTTTGGGAATTTT   1530
TTTGTTGTTTTTATTTTTGCTTTTTTTAAAACTGTGGTTAAAAAAAAGAAAAGATAGTGCATTCCACGTCTGTATGTCCACAGTTAGCT   1620
TATTGTCTTGGTCACAGAGAGACAGTCGAGGAGGAAGGAGCTGATACATGTATTTGATCAGAAGCCAGCAGACAGAATTACCAAAGTGG   1710
ATCTCGTGCTGATGGCTGGGGGGATAAGCAAAGTGGAGGCCAGCTCTGCAGCACTCTGCCCTCCTGGCCTGGCTGTGGCTGGAGGCTCTG   1800
ACTGGCTGTGGCTGGAATCCACATGCCTTTAGGTAGGAGTTTGTGCTTGCAAAACAGGAGAATTTAARRGATTCTTTTCTCTCCTCTTCC   1890
CTCCTGCCTTCATATTCTTTTTACAATCATTTACAGGCACAGCCTGAGAGTGTTTGGCATGGTTTTTGGAGTTACAGTGCTGGTATTTCC   1980
AAGCACGCTCTCAGACTGTGGATGAGAAACACCTCAGTACAACAGTGATCTTGGAGTGAGCTCCAGAATAGGAAAACACTTGTTTCTTTT   2070
CTTTTCAGAAACCTCAGCATTCCTGCTGGGCTGCTGCCCCTTTAATTGAGCCACTGTGCAACTCCTCTGTTCTTTTATTAACATGCTGCT   2160
CCCATGTCCTCAAGATTCAGGGATTTGGGATCCAGGAACCGAAGAATGAATAGTCAGTGGCCAGATTTCCCCCCAGGTATCTGGGCACAA   2250
TATGAGGATTAAAAGTACAAATTGTTCAGAGCCATTAAAGAGTGTGTAGAAGAGAAAAGGGGGGTGTTCAATCGATATCCCTTCGTTTTC   2340
TTGCAGCACGAAGTCAGTACAGAGTAAGAVAGGCATCAGAMAGCCTCTGCTGGGGCAGGGTAAGGMACCAAGCCCCTAGGCTGGCTGTCC   2430
TACCAGTGTACTGATGAGGACCAGTAAGGGCAGGCAGAGGCAGTTTGTCCTCACTCCAAAGCCATAGGAATTTCCATCTTCCAGAGCCTC   2530
TGTGAGTCTCCAGCCCTTGTGGTCACTGCATCCAGCTTCACAGTGTCCTCTAGAGTCACAGCCAAGCACCTGGACCCCTCGGGGAACGTT   2610
AAGAAGACAGGCCGTGGACTTTAAGCCTTTTAACCTTGGAAGGAGGATGTAAAATGGCAAGGGTCCTTCTGCAGGGGCCATGGCAGTGAA   2700
ACCTGGCTTCAGATTTGTTTCCAAAGGACTTCTTTCTCCCCCAAATCTGAGATGTATCCGAAAGGCAAGGGTCCCACAGAGGCTCCAAAA   2790
TGAAGGCATTAGGAAATGGGAATGTCAGCTGGCTCCCTGGTCAGTTCACCTGAGGGATTGCTGCTAAAGGGAGTAGTTCAGAGACATGAG   2880
CTTTGGTTTTGTCAGGGTCTGTTTTAGGGTTATTTTATTGGGGTTTGGTTCTTCTGTTGTTGTTGTCGTTGTTGTTGTTGTTGTTGTTGT   2970
TGTTGCCTCTGCCTCTACACATTTTTGACTGATAAATGATTCATGTCCCTGAATTAAAGTGGGCTGACCTTTGACTGCATCAAGCATNCT   3060
TTTGTAAGCAGGGAGGGCCTTCACCTGCAATTGACACATACCATCTCGCCCTCAGATCCTATCTCCCATGATACCTGAGGAAACAACATG   3150
GNAGCCAGAGACCTCAGTGTAGTCACTAGTTCTGGGGCAACATGTGGAATAGGAGACAAGACTGCCTGCCTGCTGGCCACTTATGGATTC   3240
CCTTCCCCCAAGCCAAAGTTTGGTTTGTTGTTGCTATTAGGACAGTTTTCATTGTATGTACATAAATATTTBAGTAAGGGGAAGGGATGG   3330
GCTGGGGCCTTTTGGAGTTCTAGAGAATGGGGGCAGAGGGGTTTTTCTTTCTCTTGTTTCTTTTTGAGAAAATTTACTGGCTAAAGGAAC   3420
AGACATTTTGGGTCTTTGTTTAATAAGGGCCATTCTCACAGAAAGGCAAGCCCTTCACCTAGTGGAGACGAGGCCAGTCTCTCAACAGTT   3510
GTATCCACAGATGGTAGGAAGGGAGAGAGAGTTGAAGAAGGCTTCTTCTTCCTAGACAGGAAAGGGTGGCCACGCACGGGTGAAGAGCAC   3600
```

52

```
AGTCCCATGGCCCATGCTGCAGGGCCTGAGGGGCCTTCATGCACTTCAGTTCAGCTGGTTCTGGCTGAGGCCTCCAAGGCCAGTCTGTTT   3690
GCCAGTGCATCTCTGCTGCGTGGACAAGGACACTCTCCGATGCCACTTCCTCAGGCCCTTTTCCAAGGAAGACTAGATGGAGTTGGCTTT   3780
TTATGCAAAGGTCGTTGGATTTCAGGGTGAGCACTGAGAGCTGGCTCCCCCAAGTGTCTGGATTTCACTATGTTGGCCTCATGGTTGATC   3870
CTCAAGGCCACTGTGCTGGCTTGTTGGACTGATTTAGGAAGCAGGCTAGAGGAGACGGGCATTTTATGACTCTGTTAGAGGATTCTGGGT   3960
GATGATTTCAAAAAGATGGCTGACGCTGAATCGCAGTGTGAGAAGGGTGTTGCAGTTTTGCAGGCCTCTGTGGCCTGGGCTTTCTATGC    4050
TGCGAATATTACTAGAAGCTGGATTCCTTTAAGGAGACTTGTAGTGAAAATGTTTGTCTGGATGTTCCTCTTTCAGCCCGGAGAGCTGCC   4140
TTAGGTCTTTTCTAGAATTTCTGTGTTACTTAGTTCAACCTCAGCCCAGAAGCACCATGCTCCTTGTCATTTTAAGGCTCTCAGGTTTCC   4230
AGGAACCTCCACCACCGCCACCCTGCAGCAGCCGCTGTCTGAACTGAGCCGTGTTTCCTTCCCTTGGAAACTGCCTGGCGCTTGGCTCTT   4320
AGTTTAAGGAGAGCTTGTGGCTTTTCCTCAAGCATGGACTCCTCGGGGCTGAACTGTTTGGTTTAGGCTGTGCTTTCTTTGAGGCACAGC   4410
GGGCAGGTCCGCGACCACCTCCGTCCTGTACCTAGTCATGTAGCTTCCAGACACTGAGATGAGAAGGATGTGAGCCTATGCAAGAAGAGC   4500
ATAGTGCGCCCCTCCCTGCGCCAGGCGAACAAGGCCAATCAAGGGCAGTCTGGAACTCGGGTGTGCCAGGAAGGGGGGGAGGGGGCGGGA   4590
AGCTGCCTGGCCAACCTTCCCCTCTGGCCCAGCCAGTTCTCCACTACTCAAACCTCTAGCATGGCAATGTTTGCAAAATTGCAGACTTTT   4680
CTCCCAGATATTAGGAGGAAAGGGACTTTGGGGGGTGGGGAAGGGGTCTTGGTGTTTAAAGCATAAGTTACCTGTTTGCACTGTTTTCAA   4770
ATAGGAAACAATAGTGGGCACGGTGAACATCTAATGTAAACGTGTGTGTTTTCTTTTGTCATGCTCTTGAGAATGTCTGATCATTTGTAGTA   4860
TACACCAGTGACACTTGATTCTCTGTTTCATAAAACACTATATTTTTGGAAATGTTGCTGTCCGGAAGCCTCTTCCCACCCATCCCCTCT   4950
CCTGTGTACCTCCTTTCTCCTTGCTTCCTCATCAGCCAGGCAGTTGCCATCCCTGGGCTAGAGCGGGAAACAGGGCCCCATCCCAAGCAG   5040
GCTCTGGGTGTCCTAAGCCAGCATGTGCCCTCTAGATTACTGAGTGCAGGACAGTCCCTTGGCAACTTTCCTTTCGAAAGGCTCACCAAC   5130
TACCACAGATTGCCAAACTCTTCAGCGAGGCAATAGAAGTGAGTCAGGGGGCGCGGGTCCCTGAAGTATTGTGGTTGGCTGTGGAAGCCA   5220
TGTTCTTATGCAAGTBACCGAAACTGTAGCCAGTTACAGTGTACTCAGGAAAAAACGGTAGATGAGTTCAGCCATGGCGGTGCTGGCTGG   5310
CAGGGATTGGTAACTGAGAACTGCCCATCAGCCAAAACTCACCHTTGCCCTTTGCAGAGACCTCCCACCTCGGGGGCAGCCCTCTGTACTT   5400
GCCAGCGCTGGGTGGGGCTGCTTTCACCCTGGCAGGCAAAAAGGGAACAGCCGATGCTCTTTCCTTCCCTCCCTCTTCCGGTTACCAATGC   5490
CAGATTCCCAGGGGCACGCGAGTTTTTGAGTTTTTCAGAAAAATGCTTTTGGTTTGTCTTTTCCGGGAACTGACGCGCTTTAGTCGCTGT   5580
CTGTCCCCATCAGAACTCCTAGCTTAGGTGGTTTCTTGTATTTTCTGTTAACAGTATTTCTTGTGTGCGTGCTTGTTGGGCAGCCCGTTT   5670
GGCTGCGTTATATAGCGAGTGATGAGCTGATCCTTTCCCCTAAGAGACATGACTTCCCTCCGTCAGCATTCAGCCTGTGAATAGCTAGAG   5760
CCACCCTGGGGCGAACACGTGAAAGTTAGAGCTGCAGAGTGAACTGAGGCTCTCAGGTTTTGAGTCTCTCTCTTTTCAAGGCTCAAAAGT   5850
ACCACAGAACCTCGCTCTTCTCCACACCCCTGTGGCAGGGAGTGGCCTGCCTGGCACATGTGGCTTCCCCATGCAGGACAGGCACCACTG   5940
GAACATGGGGGGCACATGCGGATGCTTGTCCAGCTGACCACCCCGCTGTGTCTCCTTCCCTCCCACCATCCTCTCCCAGAATCACCTCCT   6030
AGGGCACCCGAGCTGCTTGCCCAGGGTCCTGTTTCCCTGCTAACTCCAGAGAAGCACCCAGGGCGTTGTGACAGTCTCTAACTCCCTCCC   6120
TCCCTCCCTCCCTCTTCCTCCTTAAGAATCATATTGTATAGTAGCTTTCAGACCATACAGTATTCATTGGGTTACYCCTATTATCAAGTA   6210
GCTGGAATTGTGAAGGTTGAGTAGTTAGATCTTTAGCTKTTATTCATTATTTTTTTTTATTACTCTCCATGTGTATAAATTATTGATCATG   6300
TTGCTGGCTVTTATAAACTCTTAGCAAAGGAGWCGCCCTGCCTCAGCCTTTGCAAATGGTAATGAAGCACTGTTTTTAAATAAAAGAGAG   6390
AAACACCAAAAAAAAAA                                                                         6407
```

## A



## B

Figure 10

53

**Figure 10.** cDNA and deduced amino acid sequence of Fap-2 **A**: Nucleotide and deduced amino acid sequence of Fap-2. A polyadenylation signal and an internal poly-A stretch are underlined. Sequences in the carboxyl terminus corresponding to the Fyn binding region which was isolated initiallty as two hybrid clone 82 are shaded. The potential SH3 binding site (-ProAlaTyrPro-) and SH2 binding site (-TyrAlaGlnPro-) are double-underlined. The RBD sequence is boxed. **B**: An alignment between two RBDs of Fap-2 with C.elegans fox-1 and the comparison of Fap-2 corresponding region with the RNP consensus, taken from Burd and Drefuss (1994). The conserved amino acids between the two proteins are shaded. The identical residues in Fap-2 with the RNP consensus are indicated by asterisks. Spacing between RNP2 and RNP1 is variable at several points among menbers of the RNP family.

| Protein | Amino Acids | Sequence |
|---|---|---|
| clone27 | 326-334 | YMVPPLPEEPP |
| clone34 | 62- 73 | PLSDPPPEEPKS |
|  | 89-100 | AHTDCYPGDPAI |
| clone48 | 81- 92 | SPRRGPPPPPPG |
|  | 101-112 | ARNLPLPPPPPP |
| clone82 | 28- 39 | YIPLIIPGFPYP |
|  | 67- 78 | VPPTAIPAYPGV |
| 3BP1 | 266-277 | APTMPPPLPPVP |
| PI3Kp85 | 91-102 | RPPRPIPVAPGS |
| dynamin | 784-795 | RAPAVPPARPGS |
| mSos1 | 1148-1159 | DEVPVPPPVPPR |
| Sam68 | 292-301 | GAAPPPPPVPRG |
|  | 328-339 | TVTRGVPPPPTV |
|  | 352-361 | IQRIPLPPTPAP |
| phage library |  | XXXRPLPPXPXX |
| bead-bound peptide library |  | RX-PX-P |
|  |  | -PX-PXR |

# Figure 11

**Figure 11**. Alignment of SH3 binding motifs. Proline core sequences of the isolated cDNA clones and the reported binding sites for Fyn SH3 are presented in single-letter code with the amino acid position numbers. The bacteriophage library displayed consensus binding sequence from the Src, Fyn, Lyn, and phosphatidilinositol 3-kinase (PI3K) SH3 domains. Screening of bead-bound peptide library for Src and PI3K SH3 ligands revealed two classes of consensus sequences differing from the position of an arginine residue. Positions of the dashed line indicate a preference for valine, leucine, and isoleucine residues. Completely conserved Pro residues are shaded. In the case of 3BP1, the amino acid position numbers count from the beginning of the partial cDNA sequence given by Cicchetti et al. (1992). PI3Kp85; human phosphatidylinositol-3' kinase p85 subunit. dynamin; rat GTPase dynamin. mSos1; mouse homologue of *drosophila* Son of sevenless, the guanine nucleotide exchange factor. Sam68; mouse tyrosine-phosphorylated protein, a mitotic substrate of Src.