BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **235,** 794–798 (1997)
ARTICLE NO. RC976875

Cloning of a SH3 Domain-Containing Proline-Rich Protein, p85SPR, and Its Localization in Focal Adhesion

Woo Kyun Oh, Jae Cheal Yoo, Donggyu Jo, Young Hwa Song, Moon Gyo Kim,* and Dongeun Park[1]

Department of Life Science, Kwang-Ju Institute of Science and Technology, Kwang-Ju 506-303, Korea; and
**Laboratory of Cellular and Molecular Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892*

Received May 20, 1997

A mouse thymus cDNA expression library was screened with monoclonal antibody (mAb), B16-5 which binds to common epitope in SH3 domains of phospholipase C-γ 1 (PLC-γ 1) and Nck. We have determined the complete nucleotide sequence of one of several positive clones. The 4,172 bp cDNA clone (GenBank Accession No. U96634) encodes a SH3 domain-containing protein of 646 amino acids. Besides the SH3 domain, the predicted protein has a proline-rich region, nuclear localization signals, and leucine zipper motifs. The expressed protein in Sf9 insect cell exhibits a polypeptide of 85 kDa on SDS-PAGE. The protein is widely distributed in rat tissue with an especially high level of expression in brain and testis. Interestingly, the specific antibodies detected four related proteins of different size (75, 85, 90 and 105 kDa) in brain. In A431 cell, p85SPR is enriched at focal adhesion points indicating that the protein may interact with protein(s) in focal complexes.

Many biological processes require specific protein-protein interactions, and these interactions are often achieved through modular domains which serve to mediate protein-protein associations. Src-homology 2 and 3 (SH2 and SH3) are two such domains that appear to play a critical role in the formation of signaling complexes (1,2). SH2 domains bind with high affinity to tyrosine phosphorylated proteins, whereas SH3 domains bind to proline-rich sequences of proteins. SH3 is a small domain containing about 50 amino acid residues. It was first identified as a conserved sequence in the N-terminal non-catalytic part of Src protein tyrosine kinases (3). A number of proteins which are involved in tyrosine kinase signaling, have been shown to contain SH3 domains in addition to SH2 domain. The current view is that SH3 domains function in these proteins, in part, as protein-binding modules, and that they are involved in linking signals transmitted by protein-tyrosine kinases from the cell surface to downstream effector proteins. However, SH3 domains are not exclusively characteristic of signaling proteins. Both in single-cell and higher eukaryotes, SH3 is present in a very large group of proteins. In yeast, SH3 domains are found in proteins involved in morphogenesis, such as Bem-1 (4), the actin-binding protein Abp1 (5), and a protein involved in mating, Fus1 (6). In higher eukaryotes, SH3 domains are found in diverse proteins such as the activators of the human neutrophil respiratory burst oxidase, p47- and p67-phox (7), putative transcriptional regulator Vav (8), a cytoskeletal protein in red blood cells, spectrin (9), and a recently identified proline-rich serine/threonine kinase, SPRK (10). The role of SH3 domains in these proteins are not clear yet. A recent study shows that the SH3 domains of PLC-γ 1 and Grb2 are responsible for the intracellular targeting of these proteins to the microfilament network and to membrane ruffles, respectively (11), and the SH3 domain of Nck binds to p21-activated kinase (Pak1) and relocates it to the membrane (12). It is therefore that SH3 domains may be involved in targeting signaling proteins to their site of action in the plasma membrane or other subcellular compartments.

Previously, we and others reported that the mAb B16-5, which binds to the SH3 domains of PLC-γ 1, also recognizes another SH3 domain-containing protein, Nck (13,14). From these results, it was suggested that the mAb B16-5 binds to the conserved motif in SH3 domains. To search for novel SH3 domain-containing proteins, we have screened a mouse thymus cDNA expression library using mAb B16-5. We report here the isolation of a cDNA encoding a 85kDa SH3 domain-containing proline-rich protein. This protein contains a large proline-rich region, two potential leucine zipper motifs, and two putative nuclear localization signals. We called this protein p85SPR for "85 kDa SH3 do-

[1] To whom correspondence should be addressed. Fax: 82-62-970-2484. E-mail: depark@eunhasu.kjist.ac.kr.

0006-291X/97 $25.00

main-containing proline-rich protein." Immunofluorescence studies revealed that the protein colocalizes with structures resembling focal complexes.

MATERIALS AND METHODS

cDNA cloning and sequencing. A λZAP (Stratagene) cDNA library was constructed using mRNA isolated from thymi of 4-6 weeks old 129/SVJ mice (Jackson laboratory). To clone SH3 domain-containing proteins, the library was screened with the mAb B16-5, which was reported to bind to a conserved sequence in SH3 domains of PLC-γ 1 and Nck (13, 14). From positive clones in primary screen, clones encoding PLC-γ 1 were eliminated by second-round screening with mAb F7-2 which binds specifically to PLC-γ 1 (13). The mAbs B16-5 and F7-2 were kindly provided by Dr. Sue Goo Rhee, Laboratory of Cell Signaling, National Institutes of Health, USA. Nucleotide sequencing of the isolated clones was performed on automatic DNA sequencer (377A, Perkin Elmer). GenBank and EMBL Databases were screened with the BLAST program to search for the homologous proteins.

Expression of p85SPR in insect cells. The baculovirus transfer vector pVL1393 (PharMingen) was used in the construction of a recombinant baculovirus. The coding region of a mouse p85SPR cDNA was obtained as a product of a polymerase chain reaction using two primers (Forward: 5′-CTCCCGGGCTTCAGAGCCAGTACCGAAGCT-3′, Reverse: 5′-ATGAATTCGCTCTCGCTGTGCTGCTCCACT-3′) containing Sma I and EcoR I restriction site respectively. The resulting 2.84 kb fragment was digested with Sma I and EcoR I restriction enzymes and then subcloned into pVL1393 that had been digested with the same enzymes. Sf9 cells were grown in Grace's insect medium (GibcoBRL) supplemented with 10% fetal bovine serum. Recombinant baculovirus was generated by co-transfecting Sf9 cells with the recombinant transfer vector described above and BaculoGold DNA(PharMingen). The ability of recombinant viral clones to direct the expression of the p85SPR was examined by immunoblotting with mAb B16-5.

Generation of GST-p85SPR fusion proteins and antibodies. To subclone the SH3 domain of p85SPR into the pGEX 4T-1 vector, the gene was amplified with synthetic primers that contained EcoR I or Xho I site (Forward: 5′-GGGGGAATTCCGAAGCTTGGACATGAC-TGA-3′, Reverse: 5′-GGGGCTCGAGTGGGAGGGCTCTTCAAG-GGT-3′). The recombinant pGEX plasmid was introduced into *E. coli* DH5 α and the fusion protein was induced with IPTG. The expressed protein was purified by Glutathione affinity column. The GST-fusion form of full-length p85SPR protein was generated similarly. Rabbits were immunized with the purified fusion proteins, GST-p85SPR SH3 or GST-full-length p85SPR, four times in three-week intervals.

Immunoblot analysis of p85SPR in rat tissues. Various rat tissues were homogenized in 150 mM NaCl, 10 mM Tris-HCl (pH 7.4), 1 mM EDTA, 1 mM PMSF. The homogenates were centrifuged at 1,000 × g for 10 min to remove nuclei and tissue debris. The supernatants were diluted in an appropriate amount of RIPA buffer [150 mM NaCl, 10 mM Tris-HCl (pH 7.4), 1% Triton X-100, 1% deoxycholate, 0.1% SDS, and 1 mM EDTA]. 80 μg proteins from each tissue

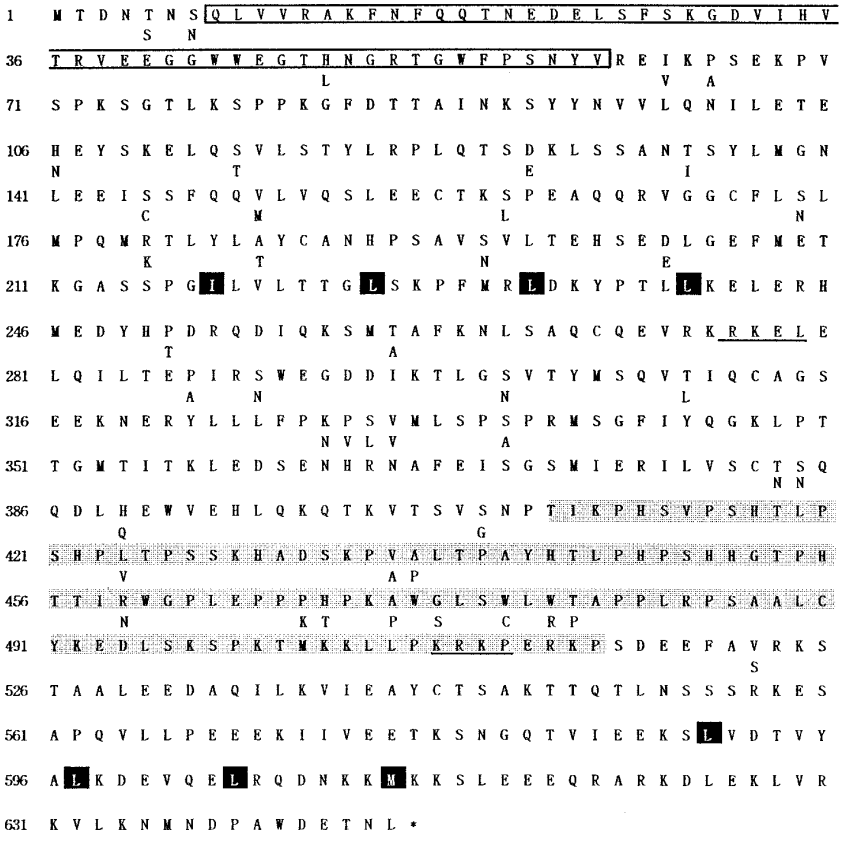

FIG. 1. Deduced amino acid sequence of p85SPR cDNA. The SH3 domain in amino-terminal region is boxed. The proline-rich region is shaded. The underlined residues are putative nuclear localization signals. The heptad repeat of hydrophobic residues in leucine zipper motifs are denoted on white-on-black. Amino acid residues that are different in human homolog, KIAA0142, are shown under the p85SPR sequence.

were resolved by SDS-PAGE, transferred to nitrocellulose membranes, and probed with anti-p85SPR antibodies.

Cell culture and immunofluorescence staining. A431 cell were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Cells were grown for 48 h on coverslips coated with fibronectin. Cells on coverslips were rinsed in Dulbecco's phosphate-buffered saline, pH7.5 (DPBS), fixed in 3.7% paraformaldehyde for 10 min at room temperature. Cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min at room temperature and stained by double labeling with affinity-purified anti-p85SPR antibodies followed by anti-paxillin mAb (Transduction Laboratories). Tetramethyl Rhodamine Isothiocyanate(TRITC)- or Fluorescein Isothiocyanate(FITC)-conjugated second antibodies (Jackson Immunology) were incubated for 40 min at room temperature. The cells were mounted in mounting solution(Sigma), and observed with a fluorescence microscope(Leica).

RESULTS AND DISCUSSION

Isolation and Sequence of cDNA Clone for p85SPR

Fourteen positive clones that reacted with mAb B16-5 but not with mAb F7-2, were isolated from 10^6 screened plaques. Partial DNA sequencing of these clones and GenBank Database search revealed that nine of them encode tyrosine kinase Fyn, one encodes SH2/SH3 adaptor protein, Nck, and the others encode two unknown proteins. One of the unknown clones contained a SH3 domain close to it's 5′-end, and selected for further characterization. This cDNA clone is 4,172 nucleotides long and contains a single open reading frame of 646 amino acids with a calculated molecular mass of 73,145 daltons (Fig.1). The putative translational initiation codon(ATG), the termination codon(TAG), and the polyadenylation signal(AATAAA) are present at the nucleotide position of 98, 2036, and 2854, respectively.

Amino Acid Sequence Analysis of p85SPR

Searching of the GeneBank Database with the predicted amino acid sequence of p85SPR identified a human cDNA clone KIAA0142 (Accession No D63476) as the closest homolog of p85SPR. Both p85SPR and KIAA0142 are composed of 646 amino acids, and the

A

```
KIAA0006 EEQIVTWLISLGVLESPKKTICDPEEFLKSSLKNGVVLCKLINRLMPGSVEKFCLDPQTEADCINNINDFLKGCATLQVEIFDPDDLYSGVNFSKVLSTL  100

p85SPR                                                                MTDNTNSQLVVRAKFNFQQTNEDELSFSKGDVIHVTRVEEGGWWEGTHN   49
KIAA0006 LAVNKATEDQLSERPCGRSSSLSAANTSQTNPQGAVSSTVSGLQRQSKTVE**E*GSH**I*D*R***D********VC***E*Y*************L*  200

p85SPR   GRTGWFPSNYVREIKPSEKPVSPKSGTLKSPPKGFDTTAINKSYYNVVLQNILETEHEYSKELQSVLSTYLRPLQTSDKLSSANTSYLMGNLEEISSFQQ  149
KIAA0006 ***************S**R*L***AV-------***E*APLT*N**T*******D**K**A*****L*V*******SNNN**TVEVTS*I**F**VCT***  294

p85SPR   VLVQSLEECTKSPEAQQRVGGCFLSLMPQMRTLYLAYCANHPSAVSVLTEHSEDLGEFMETKGASSPGILVLTTGLSKPFMRLDKYPTLLKELERHMEDY  249
KIAA0006 T*C*A****S*F**N*HK****L*****HFKSM***********N***Q**DE*EQ***NQ********I***N********E**V***Q********T  394

p85SPR   HPDRQDIQKSMTAFKNLSAQCQEVRKRKELELQILTEPIRSWEGDDIKTLGSVTYMSQVTIQCAGSEEKNERYLLLFPKPSVMLSPSPRMSGFIYQGKLP  349
KIAA0006 ***H***L*AIV***T*MG***DL****Q******S***QA***E***N**N*IF****MV*YGAC***E****M**SNVLI***A***********I*  494

p85SPR   TTGMTITKLEDSENHRNAFEISGSMIERILVSCTSQQDLHEWVEHLQK----------------------QTKVTSVSNPTIKPHSVPSHTLPSHPLTPSS  428
KIAA0006 IA*TVV*R*DEI*GNDCT***T*NTV***V*H*NNN**FQ**L*Q*NRLIRGPASCSSLSKTSSSSCSAHSSFS*TGQ*R---------------------  574

p85SPR   KHADSKPVALTPAYHTLPHPSHHGTPHTTIRWGPLEPPPHPKAWGLSWLWTAPPLRPSAALCYKE-------DLSKSPKTMKKLLPKRKPERKPSDEEFA  521
KIAA0006 --------------------------------******QII*P*S**C*RP**********G***RMSYILKES********F*H***T*****E**YV  642

p85SPR   VRKSTAALEEDAQILKVIEAYCTSAKTRQTLNSSSRKESAPQVLLPEEEKIIVEETKSNGQTVIEEKSLVDTVYALKDEVQELRQDNKKMKKSLEEEQRA  621
KIAA0006 I**********************NFQ*GHG**T**D*I**********L*I***R*****IM****************R**K*E**R**QC****LKS  742

p85SPR   RKDLEKLVRKVLKNMNDPAWDETNL                                                                              646
KIAA0006 *R*******RL**QTDECIRG*SSSKTSILP                                                                        773
```

B

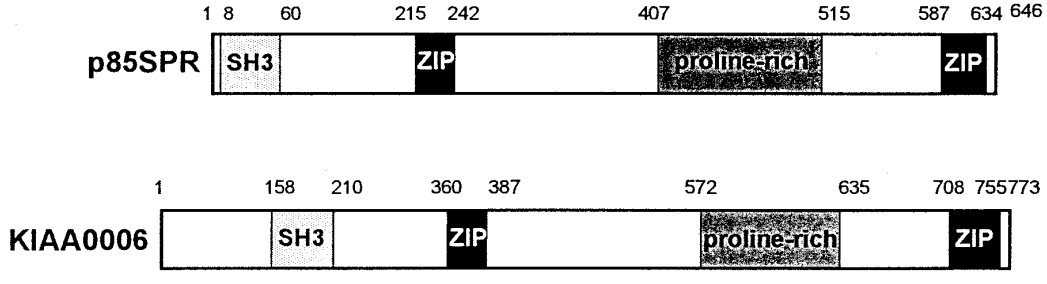

FIG. 2. Amino acid alignment of p85SPR and KIAA0006. A, Asterisks in KIAA0006 sequence represent identical residues with p85SPR and dashes indicate gaps that were introduced into the sequence to optimize alignment. The overall identity is about 60% (388/646). B, Block diagram comparison of p85SPR and KIAA0006.

```
p85SPR    QLVVRAKFNFQQTNEDELSFSKGDVIHVTRVEEGGWWEGTH--NGR-TGWFPSNYV
PLC-γ 1   KCA*K*L*DYKAQR****T*T*SAI*QNVEKQD****R*DY--G*KKQLWF*****
Nck       NMPAFV***YMAER*****LI**TKVI*MEKCSD***R*SY--**Q-I**L*C***
Fyn       VTLFV*LYDYEART**D***H**EKFQILNSS**D***ARSLTT*E-**YI*****
```

FIG. 3. Sequence alignment of SH3 domains. SH3 domains of p85SPR, PLC-γ 1, Nck (the second SH3 domain), and Fyn kinase were aligned. Residues showing identity with p85SPR are marked with "*" and dashes indicate the gaps.

overall identity is 93%. This cDNA was found in random sequencing of cDNA clones from human immature myeloid cell line KG-1. Another cDNA clone, KIAA0006 (Accession No. D25304) encodes a protein of 773 amino acids and overall amino acid sequence identity to p85SPR is 60% (Fig.2A). Amino acid 8-60 of p85SPR contains the SH3 concensus sequence (15). Because cDNA clones for p85SPR, Fyn, and Nck were obtained through screening with mAb B16-5, which binds to the SH3 domain of PLC-γ 1, the amino acid sequences of the SH3 domains in those proteins were compared. As expected, the SH3 domains in these proteins show a high degree of homology. The amino acid identities in the SH3 domain of PLC-γ 1, Nck (the second SH3 domains), and Fyn with p85SPR are 43, 43, and 40% respectively (Fig.3). In addition to the SH3 domain, p85SPR has a proline-rich region at amino acid 407-

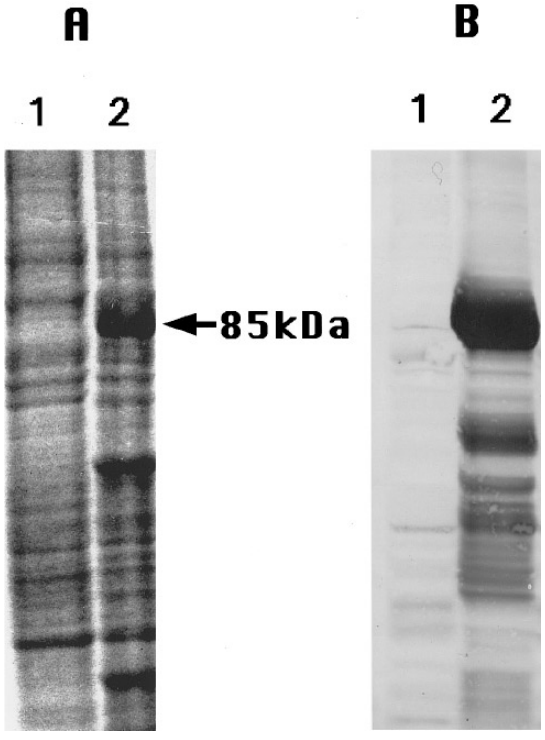

FIG. 4. Overexpression of p85SPR in Sf9 cells. Control (lanes 1) and p85SPR expressing (lanes 2) Sf9 cell extracts were separated on an 8% SDS-gel, and the gel was stained with Coomassie Brilliant Blue (panel A) or analyzed by immunoblotting with mAb B16-5 (panel B). The 85 kDa protein is indicated by an arrow.

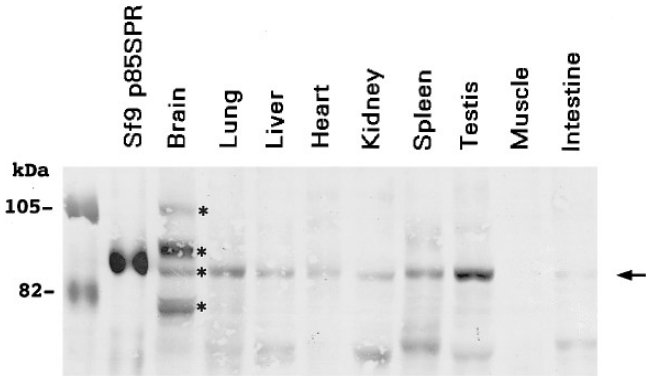

FIG. 5. Distribution of p85SPR protein in various rat tissues. Tissues were prepared as described under "Materials and Methods." Sf9 p85SPR is the partially purified p85SPR from Sf9 cells. An arrow indicates p85SPR and asterisks indicate the 75, 85, 90, and 105 kDa immuno-reactive protein in brain.

515. In this region, there are 23 proline residues among 109 amino acids (21.1%). The ligands for SH3 domains that have been identified thus far are short continuous proline-rich sequences. The minimum concensus se-

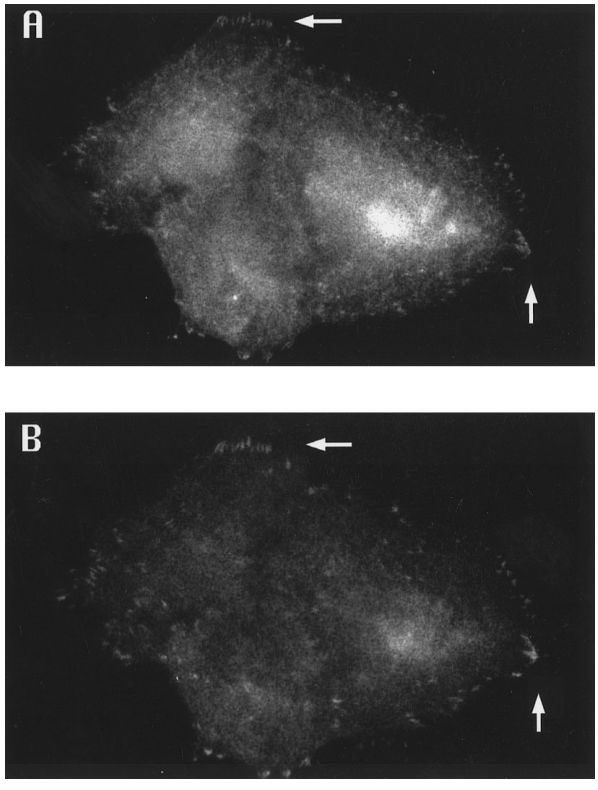

FIG. 6. Localization of p85SPR in focal complexes. Cells cultured for 48 h were stained by double-labeling with anti-p85SPR antibodies (A) followed by anti-paxillin mAb (B). Second antibodies were TRITC-conjugated donkey anti-rabbit and FITC-conjugated donkey anti-mouse immunoglobulins in A and B respectively. p85SPR clearly colocalized (arrows) to characteristic peripheral focal complex structures with paxillin.

quence for the defined SH3 domain-binding site was found to be PXXP. Two such motifs were found in p85SPR. These motifs include the sequence PSHPHTP which encompasses amino acids 420-426 and PLEPPPHP at 462-469 (Fig.1). Therefore, it is possible that the SH3 domain of p85SPR may bind to its proline-rich region in either an intramolecular or intermolecular fashion. It is also possible that the proline-rich sequence of p85SPR bind to the SH3 domains of other proteins. Besides SH3 domain and proline-rich region, p85SPR also has two putative nuclear localization signals, and leucine zipper motifs (Fig.1). The polybasic putative nuclear localization signals are located at amino acid 274-277 and 507-511. A distinctive feature of leucine zipper motif is the existence of nonaromatic hydrophobic amino acids such as leucine, isoleucine, valine or methionine in every 7 residues. p85SPR possesses two heptad repeats of nonaromatic hydrophobic amino acids characteristic of leucine zipper motifs. The first zipper motif is located at amino acid 215-242, and the second at amino acid 587-634. The second leucine zipper motif is near the carboxyl terminal region of the protein, and followed by many positively-charged basic amino acids such as lysine and arginine. The block diagram of p85SPR and comparison with KIAA0006 are shown in Fig.2B. SH3 domain, proline-rich region, leucine zipper motifs in KIAA0006 are co-linear with p85SPR.

Overexpression and Tissue Distribution of p85SPR

p85SPR was expressed in Sf9 cells using a recombinant baculovirus. In SDS-PAGE of infected Sf9 cell extract, a 85kDa recombinant protein appeared as a major protein band that also reacted with mAb B16-5 in immunoblots (Fig.4). The discrepancy between predicted (73kDa) and apparent molecular weight of this protein could be due to post-translational modifications or to an unusual conformation of the protein. We have used antibodies raised against GST-fusion forms of p85SPR to investigate the distribution of the protein in various rat tissues by immunoblotting. The protein is widely distributed in rat tissues (Fig.5). Interestingly, four different proteins (75, 85, 90 and 105 kDa) were detected in brain tissue. These four proteins were recognized by antisera raised against the GST-p85SPR SH3 domain or against the GST-full-length p85SPR. They were also immunoprecipitated by both antisera (data not shown). These observations suggest that these four proteins are closely related to each other. The 85kDa form is expressed ubiquitousely, and most abundant in testis. The 75, 90, and 105kDa forms were found primarily in brain, and the 90kDa form was the major form in brain. It remains unclear that these different proteins are isoforms of p85SPR or alternatively spliced variants.

Cellular Localization of p85SPR

Subcellular localization of p85SPR was examined in A431 cell by immunofluorescence staining. We observed dispersed cytoplasmic distribution of the protein. However, the protein is also enriched in punctate structures at the edge of cells which resembled focal complexes. No such staining was seen after treatment with preimmune sera (data not shown). The presence of p85SPR in focal complexes was confirmed by double labeling with anti-p85SPR antibodies and anti-paxillin mAb (Fig. 6). Treatments with growth factors such as epidermal growth factor and platelet-derived growth factor did not change the localization pattern of the protein (data not shown). Localization of p85SPR in focal adhesion points suggests that the protein may interact with protein(s) in focal complexes possibly through its SH3 domain. The functional significance of the presence of this protein in focal complexes is now under study.

ACKNOWLEDGMENT

This work was supported by grants from the Korea Science and Engineering Foundation through the Hormone Research Center (HRC-97-0103) and from the Genetic Engineer Research Fund of the Ministry of Education, Republic of Korea.

REFERENCES

1. Koch, C. A., Anderson, D., Noran, M. F., Ellis, C., and Pawson, T. (1991) *Science* **252,** 668–674.
2. Mayer, B. J., and Baltimore, D. (1993) *TICB* **3,** 8–13.
3. Mayer, B. J., Jamaguchi, M., and Hanafusa, H. (1988) *Nature* **332,** 272–275.
4. Chenevert, J., Corrado, K., Bender, A., Pringle, J., and Herskowitz, I. (1992) *Nature* **356,** 77–79.
5. Drubin, D. G., Mulholiand, Z. Z., and Botstein, D. (1990) *Nature* **343,** 288–290.
6. Trueheart, J., Boeke, J. D., and Fink, G. R. (1987) *Mol. Cell. Biol.* **7,** 2316–2328.
7. Leto, T. L., Lomax, K. J., Volpp, B. D., Nunoi, H., Sechler, J. M. G., Nauseef, W. M., Clark, R. A., Gallin, J. I., and Malech, H. L. (1990) *Science* **248,** 727–730.
8. Bustelo, X. R., Ledbetter, J. A., and Barbacid, M. (1992) *Nature* **356,** 68–71.
9. Sahr, K. E., Laurila, P., Kotula, L., Scarpa, A. L., Coupal, E., Leto, T. L., Linnenbacj, A. J., Winkelmann, J. C., Speicher, D. W., Marchesi, V. T., Curtis, P. J., and Forget, B. G. (1990) *J. Biol. Chem.* **265,** 4434–4443.
10. Gallo, K. A., Mark, M. R., Scadden, D. T., Wang, Z., Gu, Q., and Godowski, P. J. (1994) *J. Biol. Chem.* **269,** 15092–15100.
11. Bar-Sagi, D., Rotin, D., Batzer, A., Mandiyan, V., and Schlessinger, J. (1993) *Cell* **74,** 83–91.
12. Lu, W., Katz, S., Gupta, R., and Mayer, B. J. (1997) *Curr. Biol.* **7,** 85–94.
13. Park, D., and Rhee, S. G. (1992) *Mol. Cell. Biol.* **12,** 5816–5823.
14. Meisenhelder, J., and Hunter, T. (1992) *Mol. Cell. Biol.* **12,** 5843–5856.
15. Musacchio, A., Gibson, T., Lehto, V.-P., and Saraste, M. (1992) *FEBS* **307,** 55–61.