Volume 327, number 2, 165–171 FEBS 12745 July 1993

Identification of an isoform with an extremely large Cys-rich region of PC6, a Kex2-like processing endoprotease

Tsutomu Nakagawa[a,*], Kazuo Murakami[a], and Kazuhisa Nakayama[b]

[a]*Institute of Applied Biochemistry,* and [b]*Institute of Biological Sciences and Gene Experiment Center, University of Tsukuba, Tsukuba, Ibaraki 305, Japan*

Received 31 May 1993

In the previous study [1993, J. Biochem. (Tokyo) 113, 132–135] we identified PC6, a member of the Kex2 family of processing endoproteases. In this study, we identified another cDNA encoding an isoform of PC6, and designated it as PC6B and redesignated the originally identified PC6 as PC6A. PC6B had a very large Cys-rich region consisting of 22-times repeats of a Cys-rich motif, and a putative transmembrane domain which is not present in PC6A. A PC6B transcript was found mainly in the intestine, while PC6A transcripts were in various tissues. These results suggest distinct roles of PC6A and PC6B in endoproteolytic processing of precursor proteins.

cDNA sequence; Intestine; Northern blot analysis; Pro-protein cleavage

1. INTRODUCTION

Production of a variety of eukaryotic regulatory peptides and proteins often involves endoproteolytic processing of larger, biologically inactive precursor proteins [1]. Research on processing endoproteases have advanced with investigation of the Kex2 protease of the yeast *Saccharomyces cerevisiae*. It is a Ca^{2+}-dependent serine protease with a bacterial subtilisin-like catalytic domain, and is responsible for the processing of pro-α-factor and pro-killer toxin at pairs of basic amino acids [2]. Recently, a number of Kex2 homologues in higher eukaryotes have been identified by cDNA cloning (for review, see [3,4]): in mammals, furin, PC2, PC1/3, PC4, PACE4, and PC6 [5–15]; and in *Drosophila*, Dfurin1 and dKLIP-1, which appear to be generated via alternative splicing of the same primary transcript, and Dfurin2 [16–18]. Northern blot and in situ hybridization analyses have revealed that the expression of PC2 and PC1/3 is restricted to neuroendocrine tissues and cell lines [8–12], and that of PC4 is restricted to spermatogenic cells [13,19]. By contrast, the transcripts of furin, PACE4, and PC6 have been detected in a variety of tissues and cell lines [7,14,15,20].

PC6 is the mammalian Kex2 homologue which we have most recently identified by cDNA cloning [15]. Although it is expressed ubiquitously, the level of its expression is highest in gastrointestinal tissues. The PC6 protein deduced from the cDNA sequence shows a high similarity to other Kex2-like members in its NH_2-terminal half, which is thought to be essential for its endoproteolytic activity. On the other hand, the COOH-terminal half contains a relatively long Cys-rich region which is highly homologous only to that of PACE4 and Dfurin2.

In the previous study [15], we have isolated 7 (named B-1 to B-7) and 6 (named I-1 to I-6) cDNA clones of PC6 from mouse brain and intestine libraries, respectively, and determined the sequence of the B-1 cDNA. However, during the cloning process of the cDNAs from the mouse intestine cDNA library, we noticed that the inserts of two (I-5 and I-6) of the six PC6 cDNA clones showed restriction patterns which are similar but not identical to those of the rest of the intestine clones and of the 7 brain clones. In this study, we characterized the newly identified PC6 isoform. We tentatively named the new PC6 isoform as PC6B, and renamed the previously identified one as PC6A.

2. MATERIALS AND METHODS

2.1. *cDNA cloning and sequencing of PC6B*

Procedures for cloning of PC6 cDNAs from the mouse intestine cDNA library were described previously [15]. Both strands of the insert of one (I-6) of the 6 clones, which showed restriction patterns somewhat different from those of the rest (Fig. 1), were sequenced using a Sequenase kit (US Biochemical Corp.).

2.2. *Northern blot analysis*

Five μg of poly(A)$^+$ RNAs from mouse intestine and brain were electrophoresed on an agarose gel and blotted onto a GeneScreen*Plus* membrane (Du Pont-New England Nuclear) as described previously [15]. The blot was hybridized with a ^{32}P-labeled cDNA probe specific for PC6A (the *Fok*I fragment covering nucleotide residues 2711–2820 of the B-1 cDNA) or PC6B (the *Bgl*I fragment covering residues 4,289–4,554 of the I-6 cDNA), or a 5′ probe (the *Not*I–*Bam*HI fragment covering residues 1–380 of the B-1 cDNA) (see Fig. 1), and washed as described previously [15].

Correspondence address: K. Nakayama, Institute of Biological Sciences, University of Tsukuba, Tsukuba, Ibaraki 305, Japan.

*A fellow of the Japanese Society for the Promotion of Science for Japanese Junior Scientists.

3. RESULTS AND DISCUSSION

In the previous study [15], we have isolated 7 (B-1 to B-7) and 6 (I-1 to I-6) cDNA clones of PC6 from mouse brain and intestine libraries, respectively. Since the inserts of all the cDNA clones showed essentially the same restriction patterns, although their lengths were different from each other, we have determined the sequence of the B-1 cDNA. However, further careful analysis revealed that the restriction patterns of the I-5 and I-6 cDNAs, which we have neglected in the previous study since they do not contain the translation initiation codon, were somewhat different from those of the rest. Therefore, we determined the entire sequence of the I-6 cDNA, which was longer than the I-5 cDNA.

Fig. 2 shows the 5,208-nucleotide sequence of the I-6 cDNA and the deduced amino acid sequence. The 5′-terminus of the I-6 cDNA corresponded to the nucleotide residue 1,057 of the B-1 cDNA. The sequence from the 5′-terminus to the residue 1,645 of the I-6 cDNA was the same as the corresponding sequence of the B-1 cDNA. We believe that the nucleotide sequence of the lacking 5′-terminal part of the I-6 cDNA is identical with that of the B-1 cDNA based on the Northern blot data (see below). However, the I-6 cDNA showed no significant homology with the B-1 cDNA from that point to the 3′-terminus at the nucleotide level. We tentatively designate the protein encoded by the I-6 cDNA as PC6B and, in consequence, redesignated that encoded by the B-1 cDNA as PC6A. These data suggest that PC6A and PC6B mRNAs are generated via alternative splicing of the same primary transcript.

If the lacking 5′-terminal sequence of the I-6 cDNA was identical with that of the B-1 cDNA, the PC6B protein was assumed to consist of 1,877 amino acids (Figs. 2 and 3). The unique sequence of the I-6 cDNA encoded a region of the PC6B protein very rich with Cys residues, which extended the Cys-rich region about four times as large as that of PC6A (schematically shown in Fig. 3). The Cys-rich region consisted of a 22 times-repeated stretch of ~50 amino acid residues with a particular Cys motif. The consensus sequence of the Cys motif was $CX_2CX_3CX_2CX_{5-7}CX_2CX_{10-15}CX_{3-5}C$, and the motifs were separated from each other by a stretch of 10–16 amino acid residues (Fig. 4: see page 171). In PC6A and PACE4, five repeats were present, and in Dfurin2, ten repeats are present (Fig. 4). Furin and dKLIP-1 have two shortened repeats. Thus, the Cys-rich region of PC6B is much larger than those of other Kex2 family members. It is likely that the multiple repeats has been generated by duplication of the Cys motif unit(s) in gene evolution. In conclusion, the conservation of the particular Cys motif in various Kex2-like endoproteases suggest a functional role of the Cys-rich region. The region could be implicated in stabilization or intracellular localization of these proteases, since our previous deletion analysis has shown that the Cys-rich region of furin is not essential for its endoproteolytic activity [21].

Another structural feature is that PC6B contains a stretch of hydrophobic amino acids as a putative transmembrane domain near the COOH-terminus, like furin, dKLIP-1, and Dfurin2 (Figs. 2 and 3). No COOH-terminal hydrophobic stretch is present in PC6A. Recently, it has been shown that furin is localized in the Golgi compartments as a membrane associate form [22,23], while PC2 and PC3, neither of which has a hydrophobic transmembrane anchor, are present in the post-Golgi compartments [24–26]. Therefore, PC6A and PC6B might show different intracellular localizations. Immunocytochemical and/or cell fractionation analyses will be required to address this issue.

In the previous study [15], we have examined the distribution of PC6 transcripts by Northern blot analysis of total cellular RNAs from various mouse tissues, and detected a ~3.5-kb transcript in most of the examined tissues. By contrast, in the intestine, the other band of ~5.5 kb has been found in addition to the ~3.5-kb band. To examine whether these bands were derived from

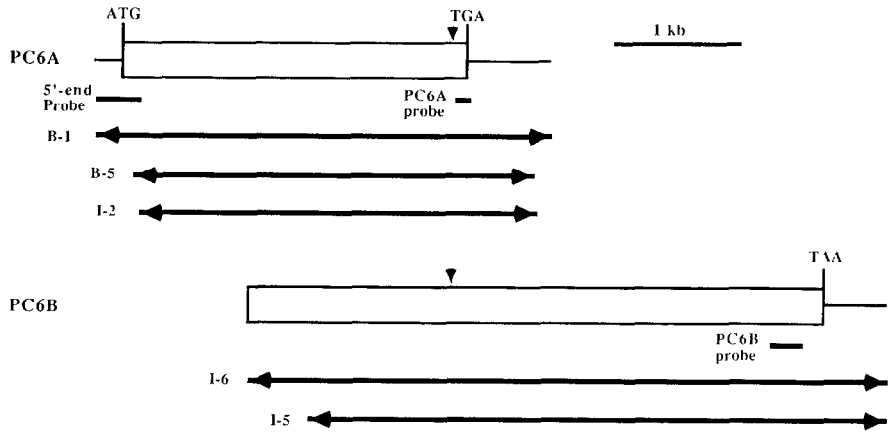

Fig. 1. Schematic representation of PC6A and PC6B cDNA clones The open reading frames in the cDNAs are represented by open boxes The relative positions of the different cDNA inserts and of the cDNA probes used for Northern blot analysis are indicated. The arrowhead represents the diverging point of the PC6A and PC6B cDNAs.

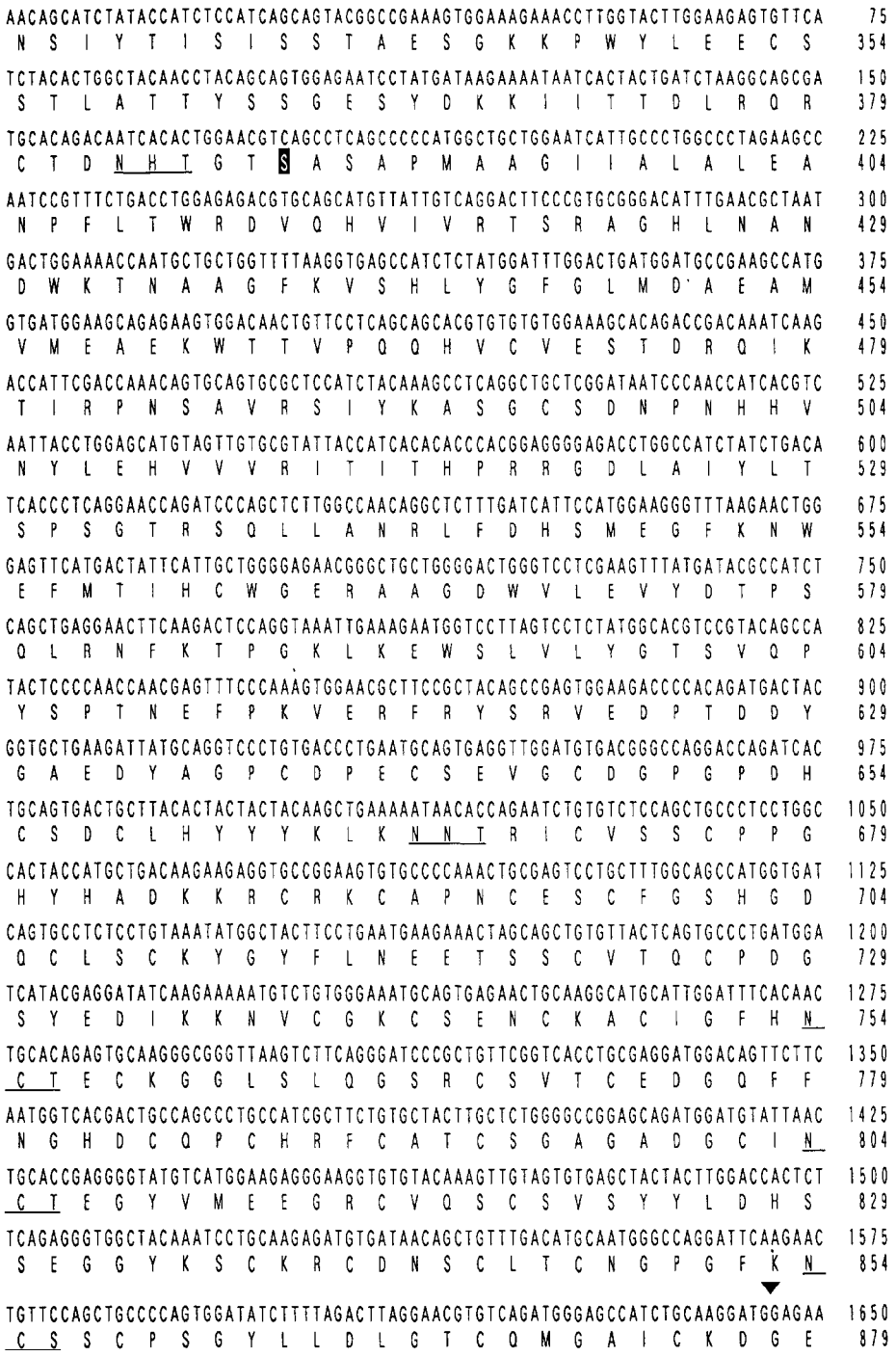

Fig. 2 Nucleotide and predicted amino acid sequences of PC6B The arrowhead represents the diverging point of the PC6A and PC6B cDNAs. The active site Ser residue is shown in a dark box The potential transmembrane domain and the potential N-glycosylation sites are shadowed and underlined, respectively. *(Continued on pages 168, 169.)*

PC6A or PC6B transcripts and to obtain clearer data, we here performed analysis of poly(A)$^+$ RNAs from mouse intestine and brain using a PC6A- or PC6B-specific cDNA fragment, or a 5′-terminal fragment as a probe. As shown in Fig. 5, the PC6A-specific probe detected ~5.5- and ~3.5-kb bands in the brain. In the intestine, the other band of ~6.5 kb was also detected. When hybridized with the PC6B-specific probe, only a ~6.5-kb band was found in the intestine, and no band was detectable in the brain. In contrast, to our surprise, very confusing results were obtained using the 5′-end probe which must be able to detect both the PC6A and PC6B transcripts: while only 2 bands of ~5.5 and ~3.5 kb were detected in the brain, in the intestine 5 bands of

```
TACATTGATGACCAAGGCCACTGCCAAACCTGTGAAGCCTCATGTGCCAAGTGCTGGGGACCAACTCAGGGAGAC 1725
 Y  I  D  D  Q  G  H  C  Q  T  C  E  A  S  C  A  K  C  W  G  P  T  Q  E  D   904

TGTATCAGCTGCCCCGTAACAAGGGTCTTGGATGATGGTCGCTGTGTTATGAACTGTCCTTCCTGGAAGTTCGAA 1800
 C  I  S  C  P  V  T  R  V  L  D  D  G  R  C  V  M  N  C  P  S  W  K  F  E   929

TTTAAGAAGCAATGCCATCCCTGCCACTACACTTGTCAAGGATGCCAAGGTAGTGGCCCTTCCAACTGCACCTCC 1875
 F  K  K  Q  C  H  .  C  H  Y  T  C  Q  G  C  Q  G  S  G  P  S  N  C  T  S   954

TGTAGAGCAGACAAGCATGGTCAGGAGCGCTTCCTGTACCACGGGGAATGTCTGGAGAACTGCCCTGTGGGCCAT 1950
 C  R  A  D  K  H  G  Q  E  R  F  L  Y  H  G  E  C  L  E  N  C  P  V  G  H   979

TATCCTGCCAAGGGACACACCTGCCTACCCTGCCCAGACAACTGTGAGCTTTGCTACAACCCACACATCTGCAGT 2025
 Y  P  A  K  G  H  T  C  L  P  C  P  D  N  C  E  L  C  Y  N  P  H  I  C  S  1004

CGATGCATGAGTGGCTATGTCATCATTCCCCCCAACCACACCTGCCAGAAGCTGGAGTGCAGACAAGGTGAATTC 2100
 R  C  M  S  G  Y  V  I  I  P  P  N  H  T  C  Q  K  L  E  C  R  Q  G  E  F  1029

CAGGATTCTGAGTATGAAGAATGCATGCCCTGTGAGGAAGGATGTCTGGGATGCACTGAGGATGATCCAGGAGCC 2175
 Q  D  S  E  Y  E  E  C  M  P  C  E  E  G  C  L  G  C  T  E  D  D  P  G  A  1054

TGTACCTCGTGTGCTACAGGATATTACATGTTTGAGCGGCATTGCTATAAAGCCTGTCCTGAGAAGACCTTCGGT 2250
 C  T  S  C  A  T  G  Y  Y  M  F  E  R  H  C  Y  K  A  C  P  E  K  T  F  G  1079

GTGAAATGGGAATGCAGGGCCTGTGGTACTAACTGTGGCAGCTGCGACCAACATGAGTGCTACTGGTGTGAGGAG 2325
 V  K  W  E  C  R  A  C  G  T  N  C  G  S  C  D  Q  H  E  C  Y  W  C  E  E  1104

GGCTTCTTTCTCTCCGGTGGCAGCTGTGTGCAGGATTGTGGCCCTGGCTTCCATGGAGACCAAGAGTTGGGAGAA 2400
 G  F  F  L  S  G  G  S  C  V  Q  D  C  G  P  G  F  H  G  D  Q  E  L  G  E  1129

TGCAAGCCCTGCCACCGAGCCTGTGAGACTTGCACTGGCTCGGGCTACAACCAATGCAGCAGCTGTCAAGAAGGG 2475
 C  K  P  C  H  R  A  C  E  T  C  T  G  S  G  Y  N  Q  C  S  S  C  Q  E  G  1154

TTGCAGCTATGGCATGGGACGTGTCTCTGGTCCACCTGGCCTCAGGTGGAAGGCAAGGACTGGAATGAAGCCGTG 2550
 L  Q  L  W  H  G  T  C  L  W  S  T  W  P  Q  V  E  G  K  D  W  N  E  A  V  1179

CCCACTGAAAAGCCATCTTTGGTGAGGAGTCTGCTGCAGGATCGACGAAAGTGGAAAGTTCAAATCAAAAGAGAT 2625
 P  T  E  K  P  S  L  V  R  S  L  L  Q  D  R  R  K  W  K  V  Q  I  K  R  D  1204

GCAACGAGCCAGAATCAACCTTGTCACTCTTCTTGTAAAACCTGCAATGGATCTCTCTGCGCTTCATGTCCCACA 2700
 A  T  S  Q  N  Q  P  C  H  S  S  C  K  T  C  N  G  S  L  C  A  S  C  P  T  1229

GGTATGTACCTGTGGCTGCAGGCCTGTGTTCCTTCCTGTCCCCAAGGCACTTGGCCATCAGTCACCAGTGGCAGC 2775
 G  M  Y  L  W  L  Q  A  C  V  P  S  C  P  Q  G  T  W  P  S  V  T  S  G  S  1254

TGTGAAAAGTGTTCCGAGGACTGTGTCTCCTGCTCCGGTGCCGACCTTTGCCAACAGTGCCTGAGCCAGCCGGAC 2850
 C  E  K  C  S  E  D  C  V  S  C  S  G  A  D  L  C  Q  Q  C  L  S  Q  P  D  1279

AACACTCTGCTTCTCCATGAGGGCAGGTGCTACCACAGTTGCCCAGAGGGCTTTTATGCAAAAGATGGTGTTTGT 2925
 N  T  L  L  L  H  E  G  R  C  Y  H  S  C  P  E  G  F  Y  A  K  D  G  V  C  1304

GAACACTGTAGTTCCCCTTGCAAAACATGCGAAGGAAATGCCACCAGCTGTAACTCTTGTGAAGGAGACTTCGTC 3000
 E  H  C  S  S  P  C  K  T  C  E  G  N  A  T  S  C  N  S  C  E  G  D  F  V  1329

CTAGACCATGGGGTGTGCTGGAAAACTTGCCCTGAAAAGCACGTGGCCGTGGAAGGAGTCTGCAAGCACTGTCCA 3075
 L  D  H  G  V  C  W  K  T  C  P  E  K  H  V  A  V  E  G  V  C  K  H  C  P  1354

GAGAGGTGCCAGGACTGCATCCACGAGAAAACTTGCAAAGAGTGCATGCCTGACTTCTTTCTATACAATGACATG 3150
 E  R  C  Q  D  C  I  H  E  K  T  C  K  E  C  M  P  D  F  F  L  Y  N  D  M  1379

TGCCATCGTTCCTGTCCCAAGAGCTTCTACCCTGACATGCGCCAGTGTGTCCCCTGCCACAAAAACTGTCTGGAG 3225
 C  H  R  S  C  P  K  S  F  Y  P  D  M  R  Q  C  V  P  C  H  K  N  C  L  E  1404

TGCAATGGCCCCAAGGAAGATGACTGTAAGGTCTGTGCTGATACTTCTAAGGCTCTCCACAATGGATTGTGCTTG 3300
 C  N  G  P  K  E  D  D  C  K  V  C  A  D  T  S  K  A  L  H  N  G  L  C  L  1429
```

Fig 2 continued *(Continued on page 169)*

~6.5, ~5.5, ~3.5, ~3.0, and ~2.2 kb were detected. Based on these data, we speculate that: (i) PC6B is encoded by the ~6.5-kb transcripts detected in the intestine with the PC6B-specific probe; (ii) PC6A is encoded not only by ~5.5-, and ~3.5-kb transcripts in the brain but also by the ~6.5-kb transcript in the intestine with the PC6A-specific probe (the ~6.5-kb transcript detected using the PC6A probe may be different from that using the PC6B probe, since the relative intensity of the ~6.5-kb bands to others observed using the 5′ probe is much higher than that using the PC6A probe and since no cDNA clones encoding PC6B was obtained from the brain library); and (iii) the ~3.0- and ~2.2-kb transcripts in the intestine probably encode PC6 proteins other than PC6A and PC6B. Thus, alternative splicing of the primary transcript and the differences in length of the 3′ and/or 5′ untranslated regions appear to give rise to these multiple transcripts of PC6. Since our previous [15] and present studies

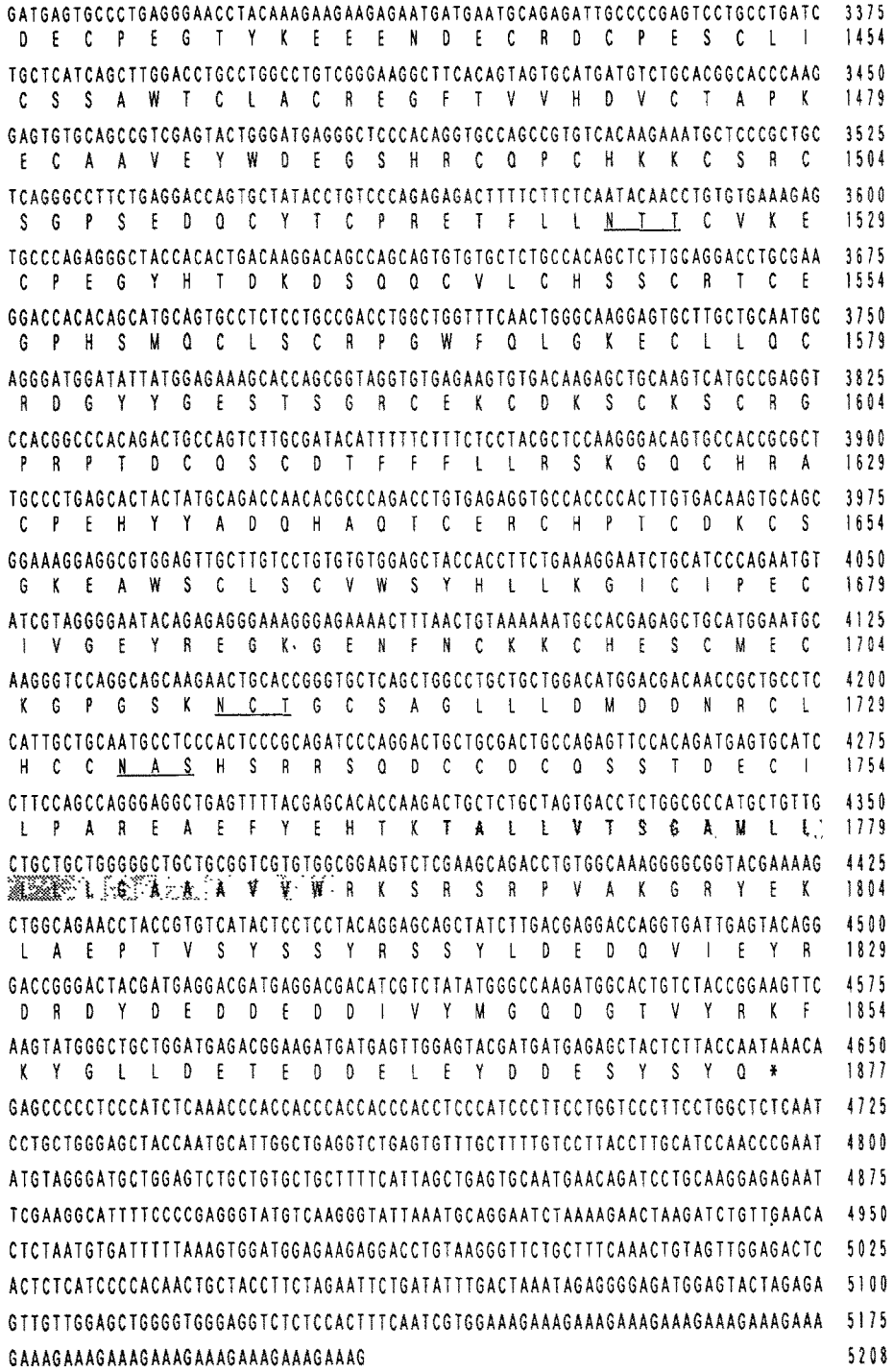

Fig. 2 continued

identified the PC6 cDNAs corresponding only to the ~3.5- and ~6.5-kb transcripts, further cloning and sequence analyses will be required to define the molecular basis for the observed differences among these multiple transcripts. Our preliminary data suggest that the ~2.2-kb transcript encodes a shorter isoform of PC6 which may be catalytically inactive as is the case with PACE4.1 [14].

In summary, we conclude from this study that more than two isoforms of PC6 are present with different tissue distributions, and are generated via alternative splicing of the same primary transcript. In view of the differences between PC6A and PC6B in the structure and in the tissue distribution, their roles in endoproteolytic processing of precursor proteins may be different from each other. They could show different intracellular

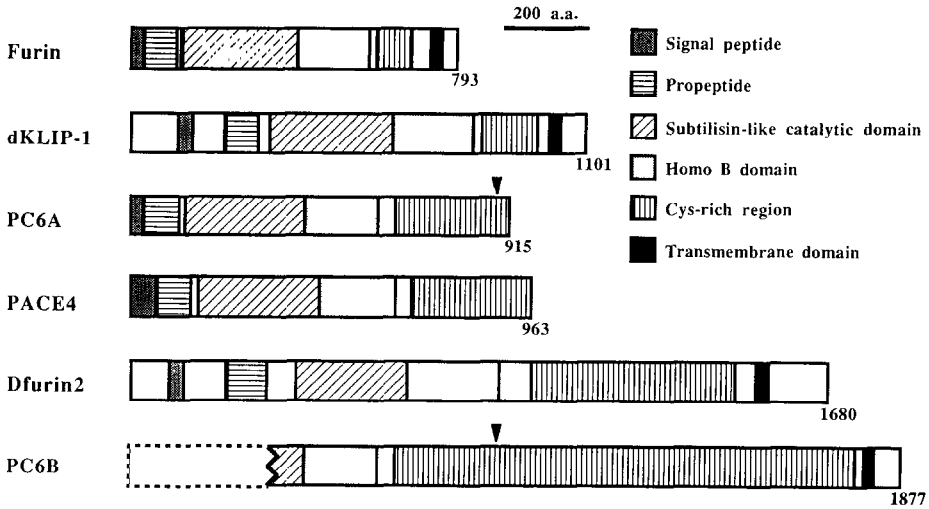

Fig. 3. Schematic representation of protein domains of mouse furin, PC6A and PC6B, human PACE4, and *Drosophila* dKLIP-1 and Dfurin2. The arrowhead represents the diverging point of the PC6A and PC6B.

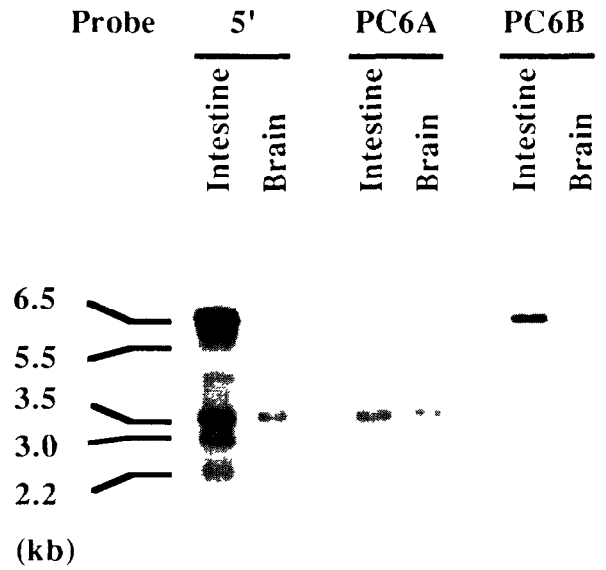

Fig. 5. Northern blot analysis of mouse intestine and brain poly(A)$^+$ RNAs using a cDNA probe specific for PC6A or PC6B, or a 5′-end probe. Experimental details are described in Section 2.

localizations and could cleave different substrate precursors. To address these issues, experiments are underway in our laboratory.

Acknowledgements: We would like to thank Dr. K. Yanagisawa for his encouragement. This work was supported in part by grants from the Ministry of Education, Science and Culture of Japan, the Special Research Project on Circulation Biosystem in University of Tsukuba, the Uehara Memorial Foundation, the Katoh Bioscience Foundation, the Mochida Memorial Foundation for Medical and Pharmaceutical Research, Sankyo Co., Ltd., and Chichibu Cement Co., Ltd.

REFERENCES

[1] Docherty, K. and Steiner, D.F. (1982) Annu. Rev. Physiol. 44, 625–638
[2] Fuller, R.S., Sterne, R.E. and Thorner, J. (1988) Annu. Rev. Physiol. 50, 345–362.
[3] Seidah, N.G. and Chrétien, M. (1992) Trends Endocrinol. Metab. 3, 133–140.
[4] Steiner, D.F., Smeekens, S.P., Ohagi, S. and Chan, S.J. (1992) J. Biol. Chem. 267, 23435–23438.
[5] Van den Ouweland, A.M.W., Van Duijnhoven, H L P., Keizer, G.D., Dorssers, L.C.J. and Van de Ven, W.J.M. (1990) Nucleic Acids Res. 18, 664.
[6] Misumi, Y., Sohda, M. and Ikehara, Y. (1990) Nucleic Acids Res. 18, 6719.
[7] Hatsuzawa, K., Hosaka, M., Nakagawa, T., Nagase, M , Shoda, A., Murakami, K. and Nakayama, K. (1990) J. Biol. Chem 265, 22075–22078
[8] Smeekens, S.P. and Steiner, D F. (1990) *J Biol. Chem. 265*, 2997–3000.
[9] Seidah, N.G., Gaspar, L., Marcinkiewicz, M., Mbikay, M. and Chrétien, M. (1990) DNA Cell Biol. 9, 415–424.
[10] Seidah, N.G., Marcinkiewicz, M., Benjannet, S., Gaspar, L., Beaubien, G., Mattei, M.G., Lazure, C., Mbikay, M. and Chrétien, M. (1991) Mol. Endocrinol. 5, 111–122.
[11] Smeekens, S.P., Avruch, A S., LaMendola, J., Chan, S.J. and Steiner, D.F. (1991) Proc. Natl. Acad. Sci. USA 88, 340–344.
[12] Nakayama, K., Hosaka, M., Hatsuzawa, K. and Nakayama, K. (1991) J. Biochem. (Tokyo) 109, 803–806.
[13] Nakayama, K., Kim, W.-S., Torii, S., Hosaka, M., Nakagawa, T., Ikemizu, J , Baba, T. and Murakami, K. (1992) J. Biol Chem. 267, 5897–5900.
[14] Kiefer, M.C., Tucker, J.E., Joh, R., Landsberg, K.E., Saltman, D. and Barr, P.J (1991) DNA Cell Biol. 10, 757–769.
[15] Nakagawa, T., Hosaka, M., Torii, S., Watanabe, T., Murakami, K. and Nakayama, K. (1993) J. Biochem. (Tokyo) 113, 132–135.
[16] Roebroek, A.J.M., Pauli, I.G.L., Zhang, Y. and Van de Ven, W J.M. (1991) FEBS Lett. 289, 133–137.
[17] Hayflick, J.S., Wolfgang, W.J., Forte, M.A. and Thomas, G. (1992) J. Neurosci 12, 705–717.
[18] Roebroek, A.J.M., Creemers, J.W.M., Pauli, I.G.L., Kurzik-Dumke, U., Rentrop, M., Gateff, E.A.F., Leunissen, J.A.M. and Van de Ven, W.J.M. (1992) J. Biol. Chem. 267, 17208–17215.
[19] Torii, S., Yamagishi, T., Murakami, K. and Nakayama, K. (1993) FEBS Lett 316, 12–16
[20] Schalken, J.A., Roebroek, A.J.M , Oomen, P.P C.A., Wagenaar, S.S., Debruyne, F.M.J., Bloemers, H P.J. and Van de Ven, W.J.M. (1987) J. Clin. Invest. 80, 1545–1549.

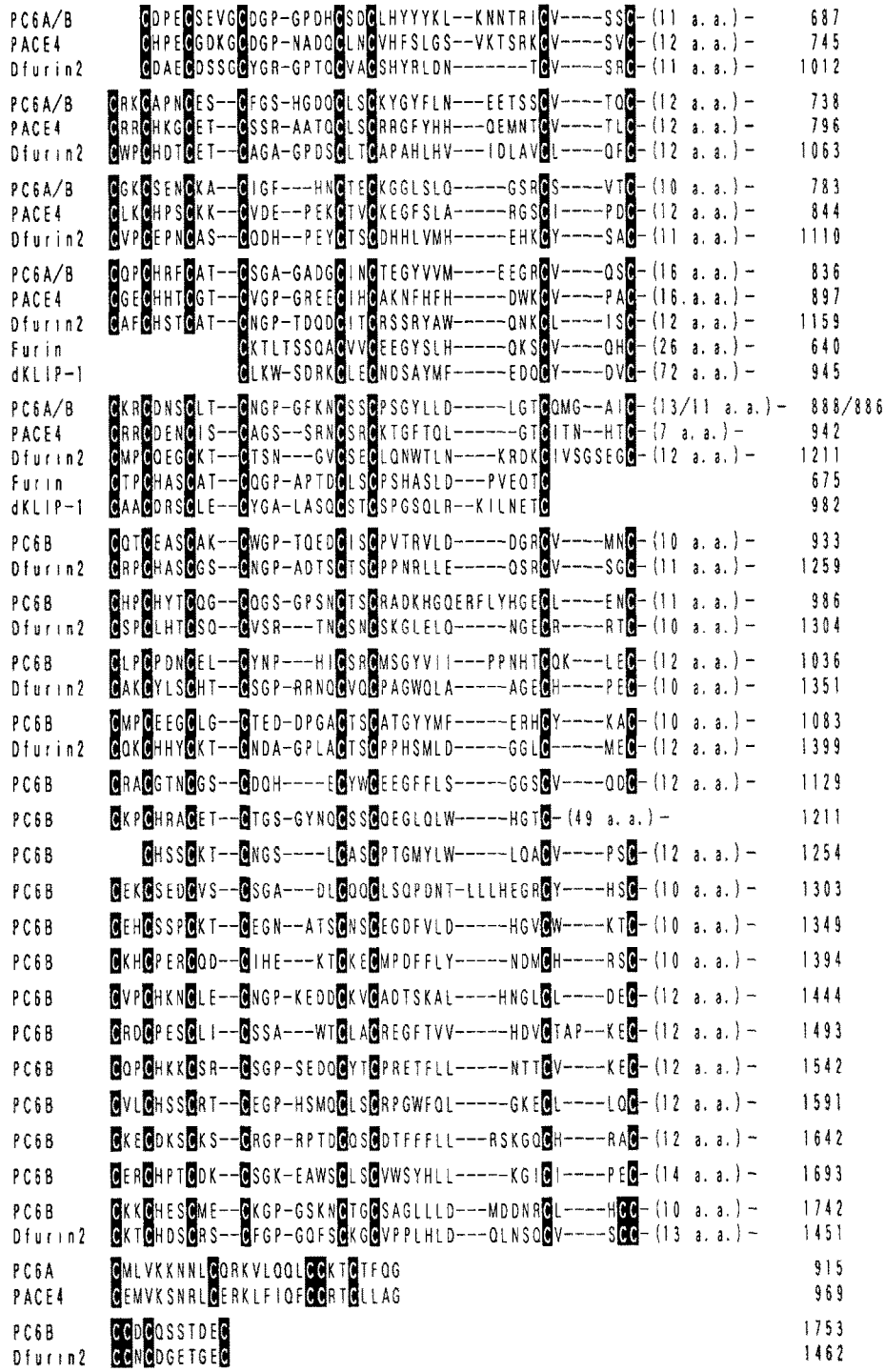

Fig 4. Amino acid alignment of the Cys-rich repeats of mouse furin, PC6A and PC6B, human PACE4, and *Drosophila* dKLIP-1 and Dfurin2. Cys residues are shown in dark boxes. Gaps introduced into the alignment are indicated by hyphens.

[21] Hatsuzawa, K., Murakami, K. and Nakayama, K. (1992) J. Biochem. (Tokyo) 111, 296–301.
[22] Bresnahan, P.A., Leduc, R., Thomas, L., Thorner, J., Gibson, H.L., Brake, A.J., Barr, P.J. and Thomas, G. (1990) J. Cell Biol. 111, 2851–2859.
[23] Misumi, Y., Oda, K., Fujiwara, T., Takami, N., Tashiro, K. and Ikehara, Y. (1991) J. Biol. Chem. 266, 16954–16959.
[24] Davidson, H.W., Rhodes, C.J. and Hutton, J.C. (1988) Nature 333, 93–96.
[25] Christie, D.L., Batchelor, D.C. and Palmer, D.J. (1991) J. Biol. Chem. 266, 15679–15683.
[26] Kirchmair, R., Egger, C., Gee, P., Hogue-Angeletti, R., Fischer-Colbrie, R., Laslop, A. and Winkler, H. (1992) Neurosci. Lett. 143, 143–145.