

# Molecular cloning of a novel mouse aspartic protease-like protein that is expressed abundantly in the kidney

Kiyoshi Mori<sup>a</sup>, Yoshihiro Ogawa<sup>a,\*</sup>, Naohisa Tamura<sup>a</sup>, Ken Ebihara<sup>a</sup>, Tomohiro Aoki<sup>a</sup>,  
Seiji Muro<sup>a</sup>, Shoichi Ozaki<sup>a</sup>, Issei Tanaka<sup>a</sup>, Kei Tashiro<sup>b</sup>, Kazuwa Nakao<sup>a</sup>

<sup>a</sup>Department of Medicine and Clinical Science, Kyoto University Graduate School of Medicine, Kyoto 606, Japan

<sup>b</sup>Department of Medical Chemistry, Kyoto University Graduate School of Medicine, Kyoto 606, Japan

Received 26 November 1996; revised version received 9 December 1996

**Abstract** By use of the signal sequence trap method, we isolated a cDNA encoding a novel aspartic protease-like protein from the mouse kidney, and termed it 'kidney-derived aspartic protease-like protein (KAP)'. The protein, a 419-amino-acid polypeptide with a 16-amino-acid signal sequence, had 47% identity with mouse cathepsin D, and its overall structure was closely related to known aspartic proteases. Northern blot analysis revealed that KAP mRNA is expressed at the highest level in the kidney, at a moderate level in the lung, and at low levels in the spleen and adipose tissue. In situ hybridization analysis demonstrated that the mRNA is expressed abundantly in the proximal straight tubule and slightly, but significantly, in the proximal convoluted tubule in the kidney. This intra-renal distribution differs distinctly from those of previously reported proteases such as cathepsins B, D, and H.

**Key words:** Aspartic protease; Cathepsin; Signal sequence trap; (Mouse kidney)

## 1. Introduction

Aspartic proteases (EC 3.4.23) are a group of endopeptidase of the pepsin family that share the same catalytic apparatus and usually function in acidic pH [1]. The best known sources of mammalian aspartic proteases are stomach (for gastric aspartic proteases; pepsin A, gastricsin, and chymosin), kidney (for renin), lysosome (for cathepsin D), endoplasmic reticulum (for cathepsin E). The functions of these enzymes include the digestion of proteins such as casein and clot milk (for gastric aspartic proteases), the biosynthesis of angiotensin I from angiotensinogen (for renin), and the degradation of intracellular and endocytosed proteins (for cathepsin D) [2]. The physiologic function of cathepsin E is currently unknown. All these aspartic proteases are synthesized as proenzymes (or zymogens) and are proteolytically activated to mature proteases [1,2].

To identify novel proteins that are involved in the physiologic functions of the kidney, we have tried to isolate cDNAs for novel proteins from the mouse kidney by use of the signal sequence trap method [3]. This method has enabled us to molecularly clone novel secretory proteins and type-I transmembrane proteins [4]. We here report molecular cloning of a

mouse cDNA encoding 'kidney-derived aspartic protease-like protein (KAP)'.

## 2. Materials and methods

### 2.1. Tissue preparation and RNA extraction

The whole kidney and other tissues were obtained from 8-week-old male BALB/c mice. Tissues were frozen in liquid nitrogen and stored at  $-84^{\circ}\text{C}$  until use. Total RNA extraction was carried out as described [5]. Poly(A)<sup>+</sup> RNA was purified using PolyATtract (Promega, Madison, WI).

### 2.2. Signal sequence trap

Signal sequence trap was performed as described [3,4] with slight modifications. Two micrograms of poly(A)<sup>+</sup> RNA from the mouse kidney was reverse transcribed by random hexamer priming using SuperScript II reverse transcriptase (Gibco BRL, Grand Island, NY), and deoxyadenosine (dA) tails were added at the 3'-end of the first strand cDNA. The second strand was synthesized with a specific primer containing polydeoxythymidine (dT) and *EcoRI* restriction site (5'-CGTGAATTCTGACTAACTGATTTTTTTTTTTTTTTT-3') and ligated with a *SacI* adaptor. The cDNA fragments of 300–700 bp in size were isolated by a 1.6% agarose gel electrophoresis and subjected to polymerase chain reaction (PCR). The reaction profile was as follows: denaturation at  $94^{\circ}\text{C}$  for 30 s, annealing at  $49^{\circ}\text{C}$  for 2 min, and extension at  $72^{\circ}\text{C}$  for 3 min for 25 cycles. The amplified fragments were digested with *EcoRI* and *SacI*, size-fractionated by agarose gel electrophoresis, and subcloned into the *EcoRI* and *SacI* sites of the pcDL-SR $\alpha$ -Tac(3') vector [6,7]. The expression plasmid library thus obtained was transfected to COS-7 cells by the lipofection method (Transfectam, Sepracor, Marlborough, MA), and the fusion proteins expressed on the cell surface were microscopically detected by immunostaining with anti-Tac antibody [8].

### 2.3. Rapid amplification of 3'-cDNA ends (3'-RACE)

The 3'-RACE experiment was performed using Marathon cDNA amplification kit (Clontech Inc., Mountain View, CA). One microgram of poly(A)<sup>+</sup> RNA from the mouse kidney was reverse transcribed by oligo (dT)<sub>15</sub> priming, using SuperScript II reverse transcriptase. After synthesis of the second-strand cDNA by the method of Gubler and Hoffman [9] and ligation with the Marathon cDNA adaptor, PCR was carried out using a 5'-gene-specific primer (5'-CATCTTGGCCCCAGCTCTAGTCC-3') and an adaptor primer (5'-CCATCCTAATACGACTCACTATAGGGC-3'). Aliquot of the reaction was further subjected to PCR using a downstream 5'-gene-specific primer (5'-ATCCGTGTCCCTCTTCAACG-3') and a nested adaptor primer (5'-ACTCACTATAGGGCTCGAGCGGC-3'). The 3'-RACE product was subcloned into the pCR II vector (Invitrogen Corp., San Diego, CA) for sequencing.

### 2.4. DNA sequencing

Nucleotide sequences were determined by the dideoxy chain-termination method using Dye Terminator Cycle Sequencing Kit, FS (Applied Biosystems Inc., Foster City, CA). All DNA sequences were confirmed by reading both DNA strands.

### 2.5. Northern blot analysis

Northern blot analysis was performed as described [10] with the <sup>32</sup>P-labeled *EcoRI*–*SacI* restriction fragment (nucleotide –171 to

\*Corresponding author. Fax: (81) (75) 771-9452.  
E-mail: ogawa@kuhp.kyoto-u.ac.jp

**Abbreviations:** cDNA, complementary DNA; cRNA, complementary RNA; PCR, polymerase chain reaction; 3'-RACE, rapid amplification of 3'-cDNA ends

+201) of mouse KAP cDNA as a probe. The blot was used to expose BAS-III imaging plate (Fuji, Kanagawa, Japan) for 20 h.

### 2.6. *In situ* hybridization analysis

**Sense and antisense  $^{35}\text{S}$ -CTP-labeled complementary RNA (cRNA)** were synthesized from the cloned mouse KAP cDNA fragment (nucleotide +70 to +1123), alkaline hydrolyzed to approximately 150 bp products, and used as probes. In situ hybridization analysis was performed as previously described [11]. In brief, 10  $\mu\text{m}$  cryosections from 8-week-old BALB/c mouse kidneys were mounted on poly-L-lysine-coated slides and stored at  $-84^{\circ}\text{C}$ . After fixation in phosphate-buffered saline with 4% paraformaldehyde and acetylation in 0.1 M triethanolamine with 0.25% acetic anhydride, hybridization was performed at  $57^{\circ}\text{C}$  for 8 h by adding  $^{35}\text{S}$ -labeled cRNA probe at a final concentration of  $1 \times 10^5$  cpm/ml in 50% formamide containing  $2 \times \text{SSC}$  ( $1 \times \text{SSC}$  is 150 mM NaCl, 15 mM sodium citrate), 10 mM Tris-HCl (pH 7.4), 100 mM dithiothreitol,  $1 \times$  Denhardt's solution, 0.2% sodium dodecyl sulfate, 10% dextran sulfate, 250  $\mu\text{g}/\text{ml}$  yeast tRNA, and 500  $\mu\text{g}/\text{ml}$  salmon sperm DNA. Slides were washed and treated with 20  $\mu\text{g}/\text{ml}$  RNase A, washed in  $0.1 \times \text{SSC}$  at  $60^{\circ}\text{C}$ , and dehydrated in graded ethanol containing 0.3 M ammonium acetate, and dried. Slides were apposed to Hyperfilm  $\beta$ -max film (Amersham International, Buckinghamshire, UK) for 13 h, or dipped into autoradiographic emulsion (NTB-2, Kodak, Rochester, NY), exposed at  $4^{\circ}\text{C}$  for 4 weeks, and counterstained with hematoxylin and eosin.

### 3. Results and discussion

### 3.1. Screening of mouse kidney cDNA library by signal sequence trap

Four thousand plasmid clones were screened, and 22 positive clones were obtained, all of which were subjected to sequencing. A single clone of 372 bp in size (E3A5, Fig. 1 thick box) contained a putative translation initiation codon preceded by an in-frame termination codon 12 bp upstream to it. The cDNA fragment encoded a 67-amino-acid peptide with

a putative 16-amino-acid signal sequence, as predicted by von Heijne's algorithm [12].

### 3.2. Isolation and sequence analysis of a full-length mouse KAP cDNA

To obtain the full-length cDNA clone, the 3'-RACE experiment was performed. A 1.2 kb fragment was obtained and sequenced. Nucleotide and deduced amino acid sequences of the full-length cDNA revealed that the 1479 bp nucleotide sequence includes the longest open reading frame, predicting a protein of 419-amino-acid residues with a relative molecular mass of 45.5 kDa (Fig. 1). The 5'-untranslated region contained a short tandem dinucleotide repeat ((CA)<sub>24</sub>). In the 3'-untranslated region, a potential polyadenylation signal [13] occurred 42 bp downstream of the translation termination codon.

The predicted protein possessed two conserved active sites of aspartic proteases [1,2]. Five potential N-linked glycosylation motifs were found, some of which may be glycosylated, mannose-6-phosphorylated, and may serve as lysosomal targeting signals [14,15].

### 3.3. Sequence comparison among mammalian aspartic proteases

Fig. 2 shows amino acid sequence comparison of KAP with six known mammalian aspartic proteases. Mouse KAP was 47%, 46%, 42%, 41%, 39%, 39% identical to mouse cathepsin D, human cathepsin E, mouse renal renin, human pepsin A, human gastricsin, and bovine chymosin, respectively (Fig. 2). The homology was observed throughout the length of the molecules, especially around two active site aspartic acid residues. All 7 cysteine residues of KAP, except for the one in the

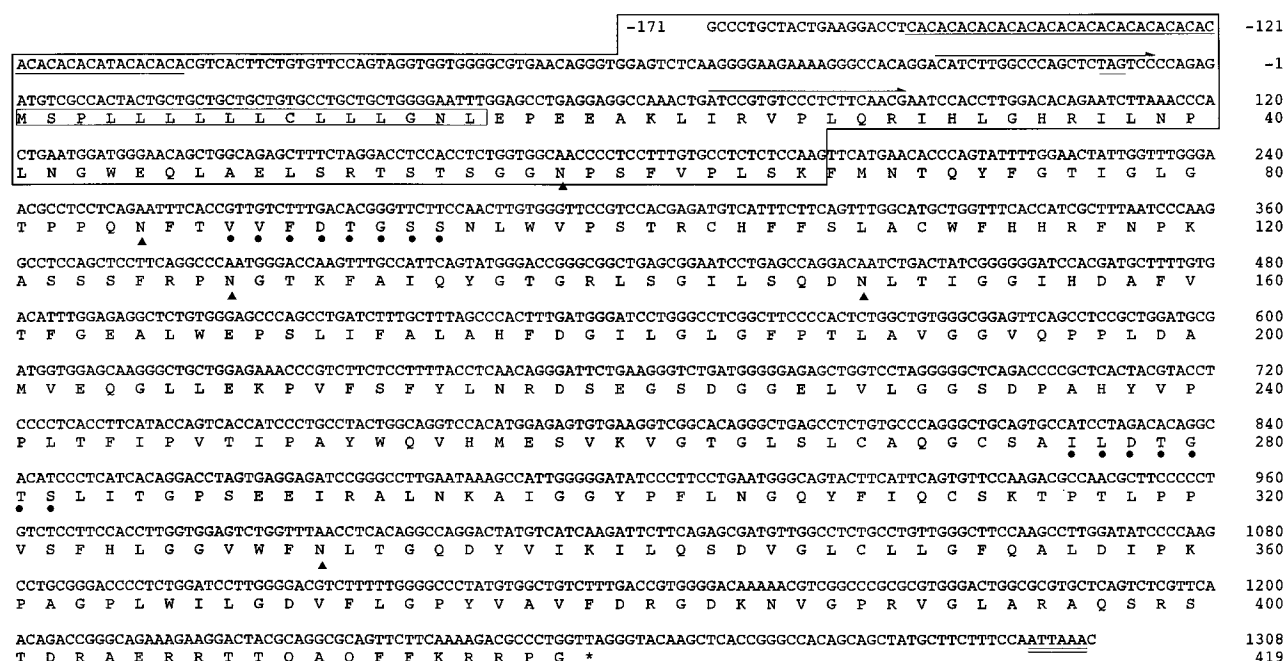


Fig. 1. Nucleotide and deduced amino acid sequences of mouse kidney-derived aspartic protease-like protein (KAP) cDNA. The 372 bp cDNA fragment (clone E3A5) obtained by the signal sequence trap method is indicated by *thick box*. Putative signal sequence is indicated by *thin box*. Two primer sequences used for 3'-RACE are indicated by arrows. Nucleotides and amino acids are numbered sequentially from the translation initiation site. The translation initiation codon is preceded by an in-frame termination codon (TAG, thin underline), and the termination codon is marked by an asterisk. A short tandem dinucleotide repeat ((CA)<sub>24</sub>) is marked by a thick underline. Potential polyadenylation site (AT-TAAA) is marked by double underline. Conserved amino acid residues around two active site aspartic acid residues are indicated by dots. Potential N-glycosylation sites are marked by triangles.

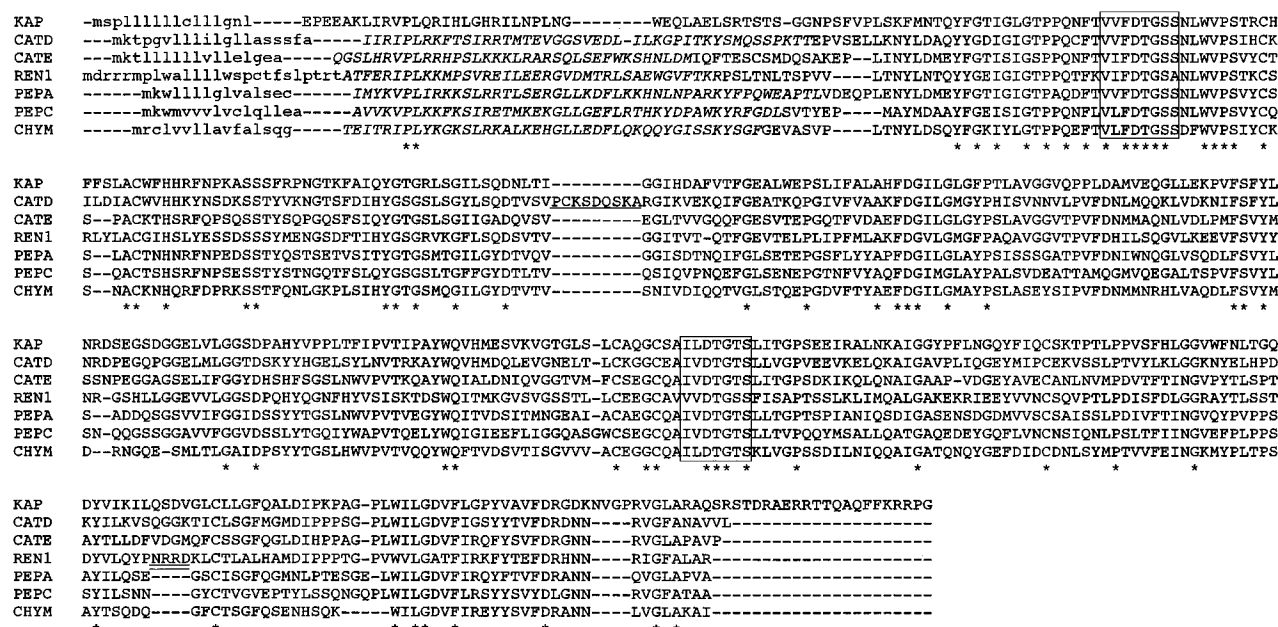


Fig. 2. Amino acid sequence comparison among mammalian aspartic proteases. Alignment of amino acid sequences of mouse kidney-derived aspartic protease-like protein (KAP) with those of mouse cathepsin D (CATD), human cathepsin E (CATE), mouse renal renin (REN1), human pepsin A (PEPA), human gastricsin (PEPC), and bovine chymosin (CHYM). Signal sequence and pro-peptide are shown in lower case and italic letters, respectively. Gaps are introduced for maximal alignment. Identical amino acids are indicated by asterisks. Conserved amino acid residues around active sites are boxed. An insertional sequence specific for cathepsin D is underlined. A 4-residue addition in renin containing the proteolytic processing site is double underlined. Amino acid sequences of CATD, CATE, REN1, PEPA, PEPC, and CHYM are from SWISS-PROT database accession number P18242, P14091, P06281, P00790, P20142, and P00794, respectively.

signal sequence, were in the same positions as those of other six aspartic proteases. These findings suggest that their three-dimensional structures and functions are closely related [14]. However, KAP is unique among aspartic proteases to have a longer C-terminus.

Aspartic proteases are known to be produced as proenzymes and the N-terminal pro-peptides are released upon activation [1,2]. KAP may also be produced as a proenzyme, but its cleavage site has yet to be determined. Cathepsin D exists as single-chain or two-chain enzyme in vivo [16], and the proteolytic processing of the single-chain to two-chain form takes place among an insertional sequence specific for cathepsin D (Fig. 2, underlined). However, KAP lacked this insertion, suggesting that it exists as single-chain enzyme like other aspartic proteases. A 4-residue addition in KAP at position 343–346 (Fig. 2, Gln–Ser–Asp–Val) was shared by cathepsin D, cathepsin E, and renin, which is the site for proteolytic processing for single-chain renin [16]. The known processing recognition Arg–Arg sequence in renin was changed to Ser–Asp in KAP, suggesting that KAP is not cleaved at this region.

### 3.4. Northern blot analysis

Northern blot analysis with the mouse KAP cDNA probe identified a 1.6 kb single hybridizing band found most abundantly in the kidney (Fig. 3). The mRNA was also expressed at a moderate level in the lung and at low levels in the spleen and adipose tissue. No significant amounts of mRNA were detected in other tissues. The restricted distribution of the KAP mRNA expression is in good contrast with the ubiquitous distribution of cathepsin D mRNA expression [1,2,17]. The KAP mRNA level in the kidney is as high as the cathepsin D mRNA level, and much higher than the cathepsin E

mRNA level [17,18], suggesting the importance of KAP in the kidney.

### 3.5. In situ hybridization analysis of KAP mRNA in the kidney

Fig. 4 depicts in situ hybridization analysis of mouse KAP mRNA expression in the kidney using <sup>35</sup>S-CTP-labeled cRNA

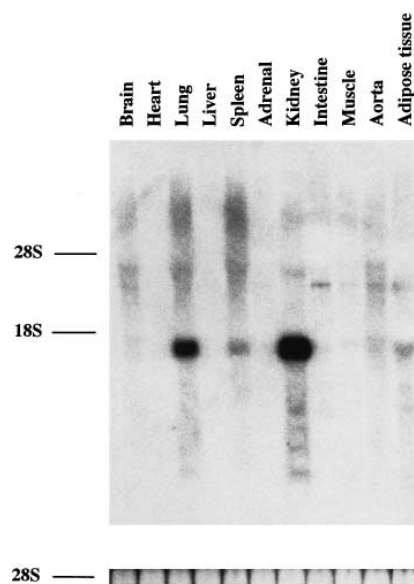


Fig. 3. Northern blot analysis of mouse KAP mRNA expression in tissues. Fifty micrograms of total RNA were loaded in each lane. The 28S ribosomal RNA bands visualized with ethidium bromide are shown in the bottom panel to verify the equivalent loading of RNA.

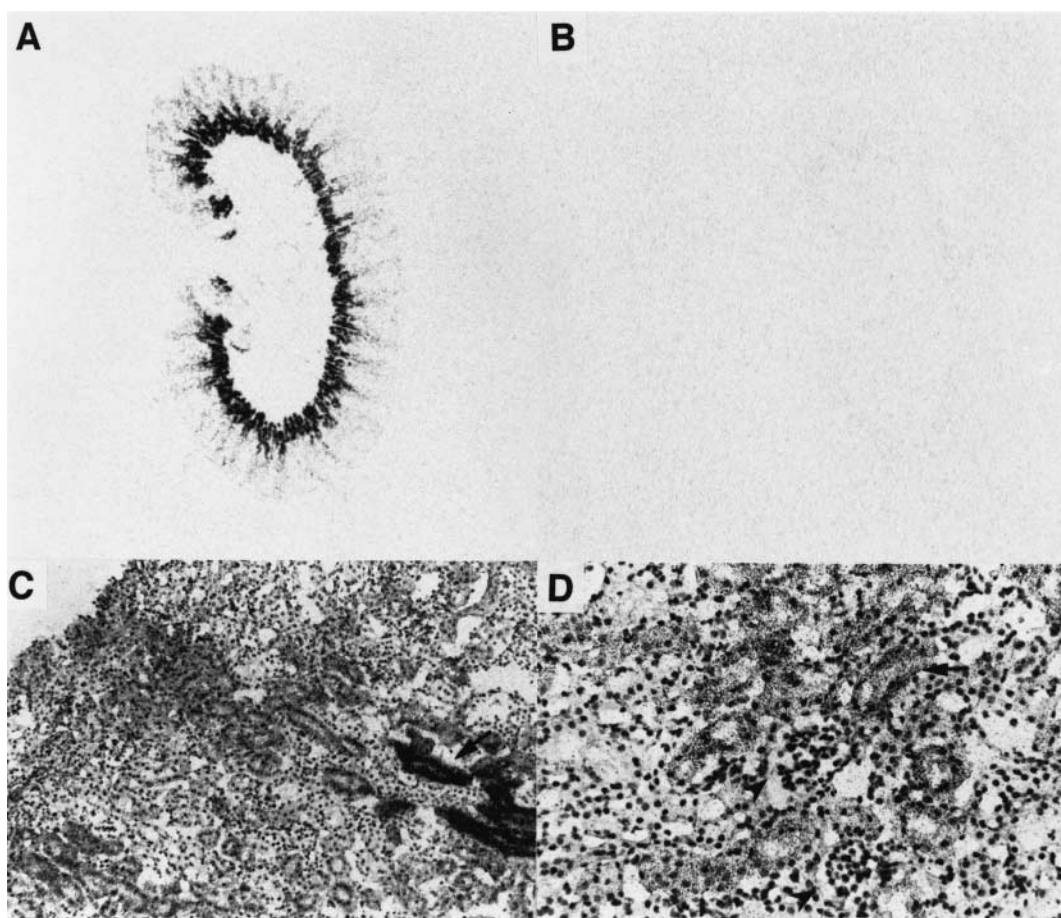


Fig. 4. In situ hybridization analysis of mouse KAP mRNA expression in the kidney. Ten micrometer sections were probed with  $^{35}\text{S}$ -CTP-labeled mouse KAP cRNA. A: Autoradiograph by the antisense probe, showing the overall distribution of the mRNA expression. B: Control experiment by the sense probe. C,D: Bright-field photomicrographs (Magnifications: C,  $\times 100$ ; D,  $\times 200$ ). Intense autoradiographic grains for KAP mRNA were restricted to the proximal straight tubule (arrow in C), and faint grains were to the proximal convoluted tubule (arrow in D), whereas no significant signals were observed in the glomerulus (arrowhead in D).

probe. At autoradiograph, strong hybridization signals were observed in the outer stripe of the outer medulla, and weak signals radiated toward the cortex (Fig. 4A). Little background signals were seen in sections hybridized with the sense cRNA probe (Fig. 4B). At photomicrograph, the intense labeling in the outer stripe was confined to the proximal straight tubule (Fig. 4C), which was confirmed by periodic acid Schiff reaction for staining of the brush border membrane (data not shown). The weak labeling in the cortex was detected in the proximal convoluted tubule, whereas no significant signals were found in the glomerulus (Fig. 4D). Previous immunohistochemical studies have revealed the intra-renal localizations of various proteases in rats [19–21]; cathepsin D is present most abundantly in the cortical collecting tubule, and cathepsin B and H (cysteine proteases) are mainly in the proximal convoluted tubule (S1 segment). The characteristic distribution of KAP suggests that KAP plays different roles from those of known proteases produced in the kidney.

### 3.6. Conclusion

In conclusion, we isolated a cDNA encoding a novel aspartic protease-like protein from the mouse kidney. The present study will lead to the better understanding of the mechanisms underlying the physiologic function of the kidney.

**Acknowledgements:** We thank Prof. T. Honjo, Department of Medical Chemistry, Kyoto University Graduate School of Medicine for encouragement, Prof. T. Uchiyama, Virus Research Institute, Kyoto University for anti-Tac antibody, Prof. T. Nakano, Research Institute for Microbial Diseases, Osaka University, Dr. T. Kuwahara, Department of Nephrology, Saiseikai Nakatsu Hospital, and Dr. T. Nakamura, Department of Medical Chemistry, Kyoto University Graduate School of Medicine for discussions. The authors also acknowledge Mr. T. Okazaki and Mr. N. Isse for their excellent technical assistance. This work was supported in part by research grants from the Japanese Ministry of Education, Science and Culture, the Japanese Ministry of Health and Welfare.

The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL/ DDBJ databases with accession No. D88899.

### References

- [1] Tang, J. and Wong, R.N.S. (1987) *J. Cell. Biochem.* 33, 53–63.
- [2] Szecsi, P.B. (1992) *Scand. J. Clin. Lab. Invest.* 52 (Suppl. 210), 5–22.
- [3] Tashiro, K., Tada, H., Heilker, R., Shirozu, M., Nakano, T. and Honjo, T. (1993) *Science* 261, 600–603.
- [4] Tashiro, K., Nakano, T. and Honjo, T. (1996) *Methods Mol. Biol.* 69, 203–219.
- [5] Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.
- [6] Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M. and Arai, N. (1988) *Mol. Cell. Biol.* 8, 466–472.

- [7] Nikaido, T., Shimizu, A., Ishida, N., Sabe, H., Teshigawara, K., Maeda, M., Uchiyama, T., Yodoi, J. and Honjo, T. (1984) *Nature* 311, 631–635.
- [8] Uchiyama, T., Broder, S. and Waldmann, T.A. (1981) *J. Immunol.* 126, 1393–1397.
- [9] Gubler, U. and Hoffman, B.J. (1983) *Gene* 25, 263–269.
- [10] Mori, K., Tanaka, I., Kotani, M., Miyaoka, F., Sando, T., Muro, S., Sasaki, Y., Nakagawa, O., Ogawa, Y., Usui, T., Ozaki, S., Ichikawa, A., Narumiya, S. and Nakao, K. (1996) *J. Mol. Med.* 74, 333–336.
- [11] Sugimoto, Y., Namba, T., Shigemoto, R., Negishi, M., Atsushi, I. and Narumiya, S. (1994) *Am. J. Physiol.* 266, F823–F828.
- [12] Von Heijne, G. (1986) *Nucl. Acids Res.* 14, 4683–4690.
- [13] Wickens, M. (1990) *Trends Biochem. Sci.* 15, 277–281.
- [14] Metcalf, P. and Fusek, M. (1993) *EMBO J.* 12, 1293–1302.
- [15] Nakayama, K., Hatsuzawa, K., Kim, W.-S., Hashiba, K., Yoshino, T., Hori, H. and Murakami, K. (1990) *Eur. J. Biochem.* 191, 281–285.
- [16] Shewale, J.G. and Tang, J. (1984) *Proc. Natl. Acad. Sci. USA* 81, 3703–3707.
- [17] Faust, P.L., Kornfeld, S. and Chirgwin, J.M. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4910–4914.
- [18] Holycross, B.J., Saye, J., Harrison, J.K. and Peach, M.J. (1992) *Hypertension* 19, 697–701.
- [19] Yokota, S., Tsuji, H. and Kato, K. (1985) *J. Histochem. Cytochem.* 33, 191–200.
- [20] Yokota, S., Tsuji, H. and Kato, K. (1986) *J. Histochem. Cytochem.* 34, 891–897.
- [21] Yokota, S., Tsuji, H. and Kato, K. (1986) *Histochemistry* 85, 223–230.