J. Biochem. 125, 566-573 (1999)

Two Subunits of the Insect 26/29-kDa Proteinase Are Probably Derived from a Common Precursor Protein[1]

Yasuyuki Fujimoto, Ayako Kobayashi, Shoichiro Kurata, and Shunji Natori[2]

Faculty of Pharmaceutical Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033

Received October 19, 1998; accepted December 9, 1998

We previously identified the 26/29-kDa proteinase in the hemocytes of *Sarcophaga peregrina* (flesh fly) that appears to participate in elimination of foreign proteins in this insect [*Eur. J. Biochem.* 209, 939-944 (1992)]. Here, we report the cDNA cloning of this proteinase. The cDNA encodes a protein which includes both the 26- and 29-kDa subunit, strongly suggesting that the both subunits are derived from a single precursor protein. The 26- and 29-kDa subunit located at the amino-terminal and carboxyl-terminal of the precursor protein. The 29-kDa subunit itself appeared to be a proteinase, for this subunit had 52% sequence identity with *Sarcophaga* cathepsin L, while 26-kDa subunit had no significant similarity. We also showed that 26/29-kDa proteinase was insensitive to specific inhibitors of cathepsin L. These results indicate that this proteinase is a novel member of the papain family. We isolated similar cDNAs from *Drosophila melanogaster* and *Periplaneta americana* (cockroach), suggesting that this proteinase is conserved in a wide variety of insects and participates in their defense mechanisms.

Key words: defense mechanism, heterodimer, insect, papain family proteinase, processing.

Injection of a foreign protein into the abdominal cavity of *Sarcophaga peregrina* larvae has been demonstrated to result in preferential degradation of the foreign protein in the hemolymph (*1*). This suggests that insects are able to discriminate between foreign and endogenous proteins. Subsequently, we purified and characterized three cysteine proteinases from *Sarcophaga* (*2-4*). Of these, two were identified as *Sarcophaga* cathepsin B and cathepsin L by isolating their respective cDNAs (*3, 5*). These two proteinases were shown to play crucial roles in the development of this insect (*3, 6-9*). The other proteinase, purified from larval hemocytes, appeared to participate in the defense mechanism of this insect (*10*). This proteinase was shown to be secreted from the hemocytes into the hemolymph when the larvae were injected with foreign cells such as sheep red cells, suggesting that it is released by larval hemocytes when they interact with foreign cells (*4*). This proteinase could therefore possibly be responsible for the preferential degradation of foreign proteins injected into the larval body cavity.

This proteinase was a heterodimer of two subunits with molecular masses of 26 and 29 kDa, respectively (*4*). This 26/29-kDa proteinase hydrolyzed Z-Arg-Arg-MCA (a substrate for cathepsin B) and Z-Phe-Arg-MCA (a substrate for cathepsin B and cathepsin L), but in contrast to *Sarcophaga* cathepsin B, the 26/29-kDa proteinase did not hydrolyze Suc-Leu-Leu-Val-Tyr-MCA (a substrate for chymotrypsin) (*4*). Moreover, *Sarcophaga* cathepsin B and cathepsin L are single proteins and have no subunit structure.

To gain more insight into the structure of the 26/29-kDa proteinase, we isolated its cDNA. We found that both the 26- and 29-kDa subunits are derived from a common precursor protein encoded in this cDNA and that the 29-kDa subunit is a proteinase with significant sequence similarity to *Sarcophaga* cathepsin L. We identified homologues of this cDNA in *Drosophila melanogaster* and *Periplaneta americana*, suggesting that this proteinase is conserved in a wide variety of insects.

MATERIALS AND METHODS

Proteinases and Proteinase Assay—*Sarcophaga* 26/29-kDa proteinase, cathepsin B and cathepsin L were purified as described previously (*2-4*). The proteinase assay was performed in 0.2 ml 50 mM sodium phosphate, pH 6.0, containing 0.05 mM substrate (Z-Arg-Arg-MCA for the 26/29-kDa proteinase, Suc-Leu-Leu-Val-Tyr-MCA for cathepsin B, and Z-Phe-Arg-MCA for cathepsin L), 100 mM NaCl, 1 mM EDTA, 5 mM 2-mercaptoethanol, and the test sample. After incubation for 10 min at 27°C, the reaction was terminated by adding 0.2 ml 20% aqueous acetic acid. Fluorescence was measured using fluorescence spectrophotometer F-3000 (Hitachi, Japan) with excitation

[1] This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan and by CREST (Core Research for Evolutional Science and Technology) of the Japan Science and Technology Corporation (JST). The novel nucleotide sequence data reported in this paper have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases with the accession numbers AB011375, AB011376, and AB011377.

[2] To whom correspondence should be addressed. Tel: +81-3-5684-2973, Fax: +81-3-3813-5099

Abbreviations: Z, carbobenzoxy; MCA, 4-methyl-coumaryl-7-amide; Boc, *t*-butyloxycarbonyl; Suc(OMe), *N*-methoxysuccinyl; Suc, succinyl; Bz, benzoyl; Pyr, L-pyroglutamyl; Ac, acetyl.

and emission wavelengths of 380 and 460 nm, respectively.

Determination of Partial Amino Acid Sequences of the 26/29-kDa Proteinase—Purified proteinase (20 μg) was subjected to SDS-PAGE and the two subunits were separated. After staining the gel with Coomassie Blue, the band corresponding to each subunit was removed and the protein present in the gel slip was electrophoretically eluted. The protein was precipitated by adding ethanol to the eluate and the resulting SDS-free protein was digested with lysyl-endopeptidase. The resulting peptides were subjected to HPLC (Gilson, France) using a Synchropak RP-P (C18) column (4.6×250 mm, SynChrom, USA). Elution was performed using a linear gradient of 0 to 60% acetonitrile in 0.05% trifluoroacetic acid. The purified peptides were sequenced on a protein sequencer (PPSQ-10, Shimadzu, Japan).

Determination of Amino- and Carboxyl-Terminal Amino Acid Sequences of 26-kDa and 29-kDa Subunit—To determine the amino-terminal sequences, the two subunits were electrophoretically transferred onto a polyvinylidene difluoride membrane (Immobilon P, Millipore, USA) after SDS-PAGE, and the filter containing each subunit was applied directly to the protein sequencer. To determine the carboxyl-terminal sequence of 26-kDa subunit, the two subunits were electrophoretically transferred onto a Teflon membrane after SDS-PAGE. The membrane was stained with sulforhodamine B and band of 26 kDa was applied directly to a carboxyl-terminal protein sequencer (G1009A, Hewlett Packard, USA).

Purification of 26- and 29-kDa Subunit—For mass spectrometry, arginylendopeptidase-digestion and deglycosylation experiment, we isolated 26- and 29-kDa subunit. For this, purified 26/29-kDa proteinase was denatured by 6 M guanidine hydrochloride, then the 26- and 29-kDa subunits were separated by HPLC under the same conditions described above except that the concentration of trifluoroacetic acid was 0.1%. Arginylendopeptidase-digestion was performed using 50 μg of purified 26-kDa subunit and 1.3 μg of arginylendopeptidase at 34°C for 15 h. The resulting peptides were purified by HPLC in the same way as above mentioned. Molecular mass of peptides were determined by MALDI TOF mass spectrometry using Voyager Elite (PerSeptive Biosystems, USA).

Cloning and Sequencing of the cDNA for the 26/29-kDa Proteinase—Sarcophaga, Drosophila, and *Periplaneta* cDNA for the 26/29-kDa proteinase were cloned from relevant pBluescript plasmid library [libraries from *Sarcophaga* larvae (*11*) and *Periplaneta* fat body (*12*)], by screening colonies of *Escherichia coli* carrying the library using PCR (nested PCR), with a portion of each colony as a template. For the *Sarcophaga* clone, primers were synthesized according to the sequence information obtained from the peptides derived from the 26-kDa subunit. The primers for the first amplification were 5′-AATCCACCC(CA)G(ATGC)TGGGA(TC)C-3′ and 5′-GTGCCGAA(ATGC)GG(AG)TG(TC)TC(AG)TT-3′, and those for the second amplification were 5′-TGGGACCCCAA(TC)TA(TC)AT(ACT)GT-3′ and 5′-TTCACCATGCC(ATGC)CC(AG)TA(AG)TA-3′. For isolation of the *Drosophila* clone, the primers for the first amplification were the T3 promoter on the vector and 5′-GCATCAATAGCAAC(AGC)GA(AGT)A(CT)(AG)GG-3′, and those for the second amplification were the SK promoter on the vector and 5′-TTCACCATGCC(ATGC)-CC(AG)TA(AG)TA-3′. The primers for the *Periplaneta* clone were the same as those for the *Drosophila* clone. DNA sequencing was performed using a DNA sequencer (373A, Applied Biosystems, USA) with a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems, USA). The nucleotide sequences of both strands were determined.

RESULTS

Structural Analysis of the 26/29-kDa Proteinase—As the 26/29-kDa proteinase is a heterodimer of two subunits, we tried to isolate the cDNAs for each subunit. For this, we separated the two subunits and determined the amino acid sequences of several peptides derived from each subunit. The sequences of nine peptides derived from the 26-kDa subunit and seven derived from the 29-kDa subunit were determined. The peptides from the 26-kDa subunit were VTXLQLXGTSEDPVQVQAILPNAK, NVYTLWVRYK, XPHYPASRMPIPVRYEMR, GTLYIPYAEIAEPFYAWYXK, TNPPRWDPNYIVK, XYQLANEHPFGTSLK, NTRRSRIDYYGGMVK, FRLEETIGDK, and LIGTETFLGFNXDK, while those from the 29-kDa subunit were XXXX TYWGNDQYILMSAXK, NVXLVAPITGFFXVTPNDPMALK, XGLXELDHAVLAVGYGTINGEDYWLVK, DTVPDQYDWRLYGAVTPVK, XQSXXGSXWXFGTIXHLXXAFFLK, NNXGVMTMPTYVEM, and XMMEMGGVPTEEEYGPYLGQDGY. We also determined the amino-terminal amino acid sequences of both subunits. These were TNPPRWD (26-kDa subunit) and DTVPDQYDXRLYGAV (29-kDa subunit). In these sequences, X indicates an unidentified residue.

To isolate the cDNA for the 26-kDa subunit, we then designed 4 degenerated oligodeoxyribonucleotide primers corresponding to the peptides TNPPRWDPNYIVK, XYQLANEHPFGTSLK, and FRLEETIGDK. We screened a *Sarcophaga* cDNA library by PCR using these primers, and finally isolated one cDNA clone. The nucleotide sequence and deduced amino acid sequence of the protein encoded by this cDNA are shown in Fig. 1. Surprisingly, all sixteen peptides derived from both the 26- and 29-kDa subunits were included in this protein, which consists of 550 amino acid residues. This strongly suggests that both the 26- and 29-kDa subunits are derived from a single precursor protein.

The 26-kDa subunit was located at the amino-terminal end and the 29-kDa subunit at the carboxyl-terminal end of this precursor protein. Based on amino acid sequencing, the amino-terminal residues of the 26- and 29-kDa subunits were assigned as the Thr at position 20 and the Asp at position 329, respectively. Thus, the 19-amino acid sequence starting from the first Met is likely to be a signal peptide. Moreover, there must be an intervening sequence between the carboxyl-terminal residue of the 26-kDa subunit and the amino-terminal residue of the 29-kDa subunit, because the molecular mass of the sequence located between the Thr at position 20 and the Arg at position 328 was calculated to be 36.5 kDa, which is significantly more than 26 kDa. Indeed, molecular mass of 26-kDa subunit was determined to be 24.3 kDa by mass spectrometry.

To deduce the carboxyl-terminal residue of the 26-kDa subunit, we determined the amino acid composition of the purified 26-kDa subunit. We then determined a putative

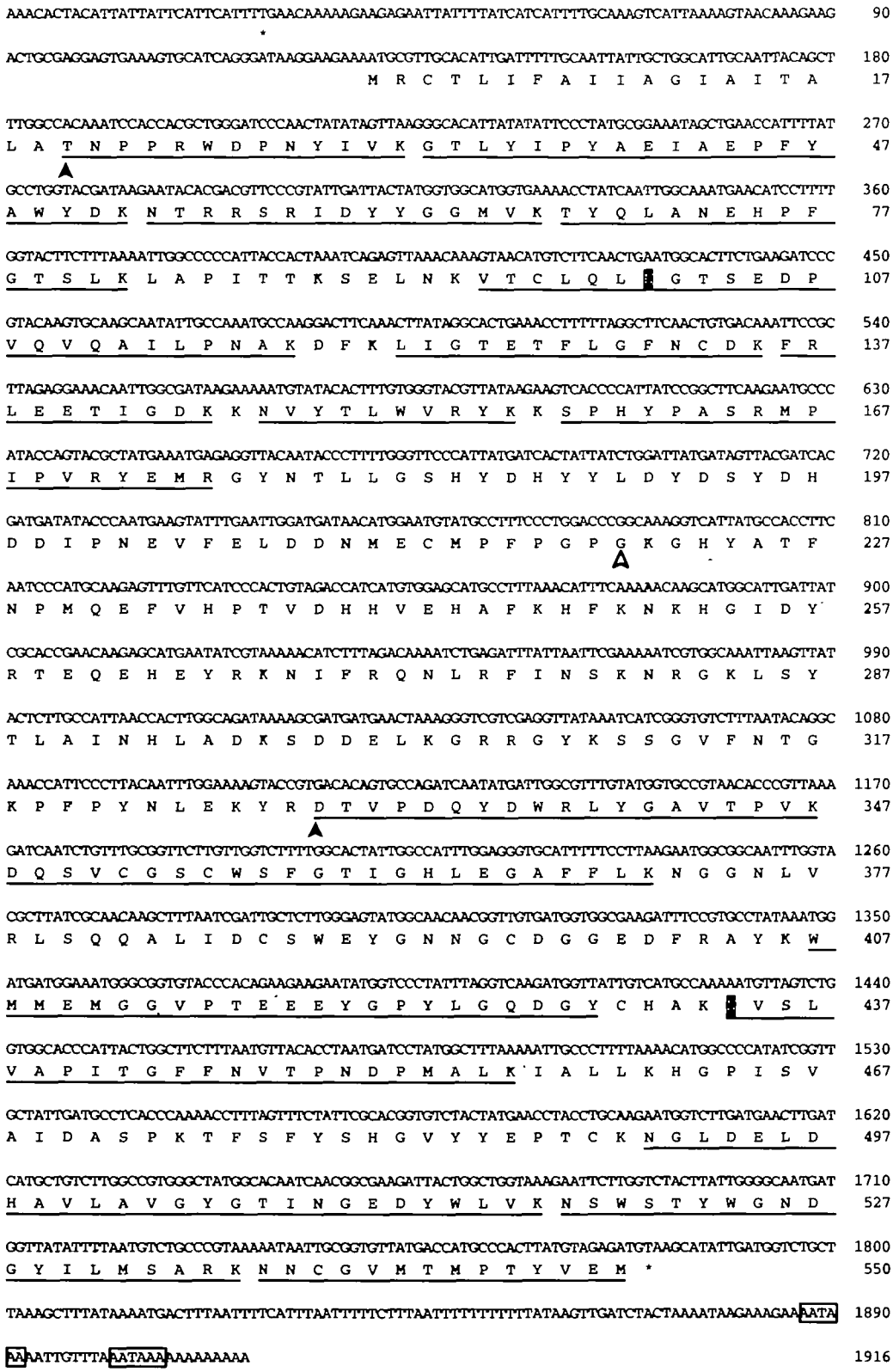

Fig. 1. **The nucleotide sequence of the cDNA encoding the *Sarcophaga* 26/29-kDa proteinase.** The deduced amino acid sequence is shown below the nucleotide sequence. The solid arrowheads show the amino-terminal residues of the 26- and 29-kDa subunits. The open arrowhead shows the deduced carboxyl-terminal residue of the 26-kDa subunit. Possible *N*-glycosylation sites are shown in closed boxes. The underlined sequences were determined separately by isolating each peptide. The asterisks indicates the termination codons and the polyadenylation signals are boxed. The numbers on the right correspond to the nucleotide and amino acid numbers.

carboxyl-terminal residue for the 26-kDa subunit in the precursor protein that gave maximal matching between the observed and calculated amino acid compositions of the 26-kDa subunit. This residue appeared to be the Gly at position 220. Assuming that this Gly is the carboxyl-terminal residue of the 26-kDa subunit, a comparison of the calculated and observed amino acid compositions are shown in Fig. 2. They are coincided well.

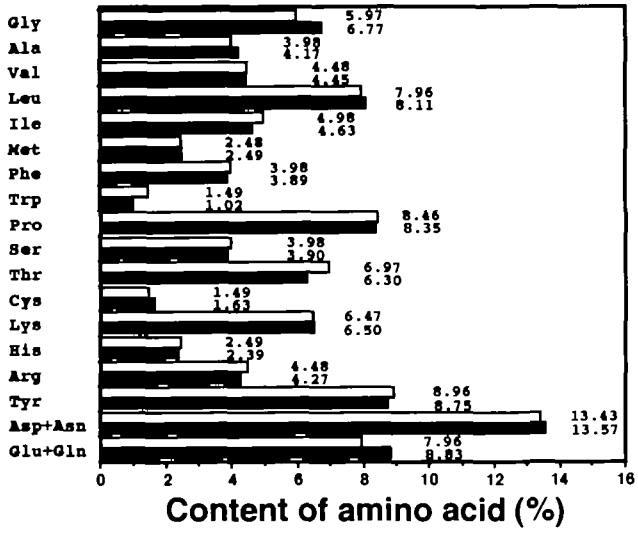

Fig. 2. **Comparison of the observed and calculated amino acid compositions of the 26-kDa subunit.** The amino acid composition of the purified 26-kDa subunit was determined. Then, assuming that the Gly at position 220 is the carboxyl-terminal residue of the 26-kDa subunit, the amino acid composition of the putative 26-kDa subunit (the region from the Thr at position 20 to the Gly at position 220) was calculated. The open and solid bars represent the calculated and determined amino acid compositions, respectively.

In order to confirm further that this Gly is the carboxyl-terminal residue, we digested the purified 26-kDa subunit with arginylendopeptidase, separated the resulting peptides by HPLC and obtained a peptide with the molecular mass of 5,272 Da. This value was consistent with the molecular mass of the peptide spanning the region from the Gly at position 176 to the Gly at position 220, which was calculated to be 5,276 Da. The 15 amino-terminal residues of the peptide was determined as GYNTLLGSHYDHYYL, which coincided with the sequence from the Gly at position 176 (position 175 is Arg). Considering the specificity of arginylendopeptidase, we concluded that this peptide was derived from the carboxyl-terminal of the 26-kDa subunit, and that its carboxyl-terminal residue is the Gly at position 220. Furthermore, we determined three residues of carboxyl-terminal end of the 26-kDa subunit as -GXG-COOH by using a carboxyl-terminal protein sequencer (X is unidentified residue), supporting the above conclusion.

The 29-kDa subunit was found to have significant sequence similarity with the mature cathepsin L enzymes of various species. It showed greatest similarity with the mature cathepsin L of *Sarcophaga*, 52% of the amino acids being identical. Moreover, six Cys residues and the three catalytic residues of the papain family of proteinases (Cys, His, and Asn) were also conserved in the 29-kDa subunit, indicating that it is a proteinase. However, no significant sequence similarity was found between the 26-kDa subunit and other proteins. As shown in Fig. 3, the precursor protein of the 26/29-kDa proteinase had an overall structural similarity to *Sarcophaga* procathepsin L except that the former contained an insert of the 26-kDa protein between the signal peptide and the prosegment.

There is one possible *N*-glycosylation site in each of the 26- and 29-kDa subunits (Fig. 1). To determine whether these subunits are glycosylated, we isolated them separate-

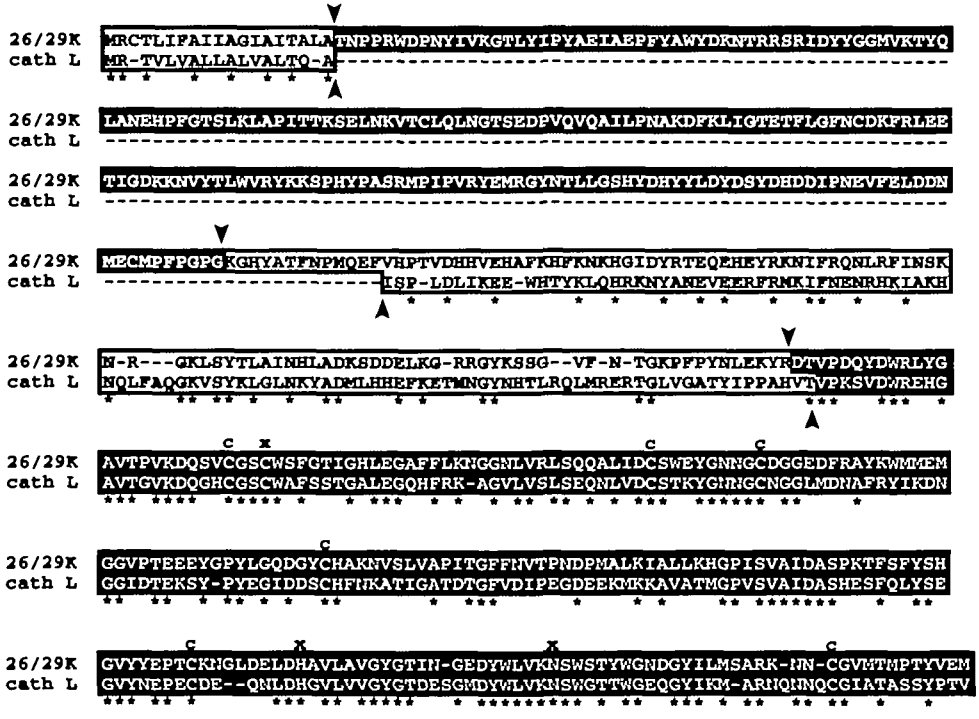

Fig. 3. **Comparison of the amino acid sequences of *Sarcophaga* 26/29-kDa proteinase and cathepsin L.** Signal sequences and prosegments are boxed. The 26-kDa subunit and proteinase regions are shown in gray and black, respectively. The asterisks indicate identical amino acids. The c and x markers indicate cysteine residues and three active site residues conserved in the papain family of proteinases. The arrowheads indicate cleavage sites. Gaps have been introduced to obtain maximal matching.

ly, treated them with *N*-glycanase, then subjected them to SDS-PAGE. As is evident from Fig. 4, both the 26- and 29-kDa subunits migrated faster after treatment with *N*-glycanase, indicating that they are glycosylated in their natural form. The polysaccharide chain in the 29-kDa subunit may be larger than that in the 26-kDa subunit, because the difference in mobility of the 29-kDa subunit before and after *N*-glycosylase treatment was greater than that of the 26-kDa subunit.

Although we revealed that the structure of 26/29-kDa proteinase was closely related to that of cathepsin L, it preferentially hydrolyzed Z-Arg-Arg-MCA, the specific substrates for cathepsin B (*4*). Therefore, we examined substrate specificity of this proteinase precisely. As shown in Fig. 5, the 26/29-kDa proteinase hydrolyzed Z-Arg-Arg-MCA and Pro-Phe-Arg-MCA with almost the same efficiency. It also hydrolyzed Z-Phe-Arg-MCA, Z-Val-Lys-Met-MCA, and Boc-Leu-Thr-Arg-MCA significantly, but efficiency of hydrolysis was much less. No papain family proteinase that shows such a wide range substrate specificity has been reported. Thus, the 26/29-kDa proteinase was clearly different from cathepsin L and cathepsin B.

Previously, we demonstrated that this proteinase was inhibited by various cysteine protease inhibitors such as leupeptin, antipain, chymostatin, and E-64 like cathepsin L and cathepsin B (*4*). Therefore, we examined the effect of specific inhibitors of cathepsin L and cathepsin B on the activity of the 26/29-kDa proteinase. As summarized in Table I, the 26/29-kDa proteinase was totally insensitive to all of these inhibitors, suggesting that it is a novel type of the papain family.

The 26/29-kDa Proteinases of Other Insects—To deter-

Fig. 4. **The change in the mobility of the two subunits on electrophoresis after deglycosylation.** The 26- and 29-kDa subunits were each treated with endoglycosidase F at 37°C for 24 h, and then subjected to SDS-PAGE. Lane 1, purified 26/29-kDa proteinase; lanes 2 and 3, purified 29- and 26-kDa subunits, respectively; lanes 4 and 5, *N*-glycanase-treated 26- and 29-kDa subunits, respectively. Each lane contained 0.5 μg protein. The gel was calibrated with four molecular mass markers. Their positions on the gel are shown on the left in kDa.

TABLE I. **Effects of inhibitors on *Sarcophaga* 26/29-kDa proteinase.** Percentage residual activity of *Sarcophaga* cathepsin B, cathepsin L, and 26/29-kDa proteinase was determined in the presence of various inhibitors specific for cathepsin B (CA-074) and cathepsin L (CLIK-40 and CLIK-46) (*29, 30*). N.T., not tested.

Inhibitor (μM)		Substrate and proteinase			
		Z-Phe-Arg-MCA			Z-Arg-Arg-MCA
		Cathepsin B	Cathepsin L	26/29-kDa proteinase	26/29-kDa proteinase
CA-074	0	100	100	100	100
	0.01	18	91	96	101
	0.1	2.8	91	102	94
	1.0	0.19	54	89	89
CLIK-40	0.01	109	7.1	93	103
	0.1	85	1.6	92	95
	1.0	80	0.20	63	67
CLIK-46	0.01	106	26	N.T.	N.T.
	0.1	99	1.1	88	93
	1.0	85	0.20	65	78

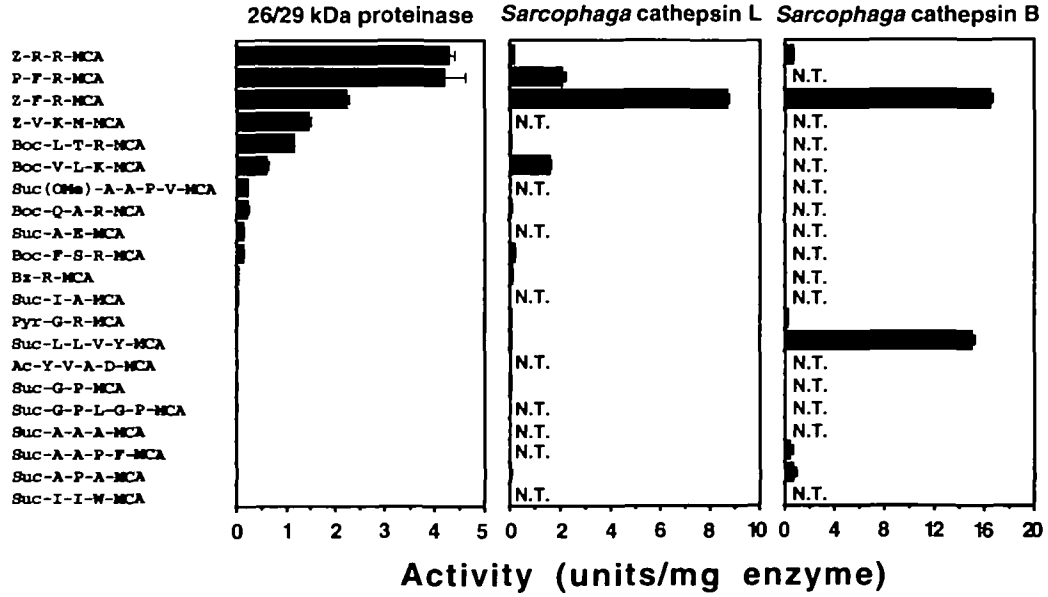

Fig. 5. **Substrate specificity of the 26/29-kDa proteinase.** Specific activity of purified *Sarcophaga* 26/29-kDa proteinase was determined for various peptidyl-MCA substrates and compared with that of purified *Sarcophaga* cathepsin L and cathepsin B. One unit of proteinase activity was defined as the amount that catalyzed the hydrolysis of 1 μmol substrate in 10 min at 27°C. N.T., not tested.

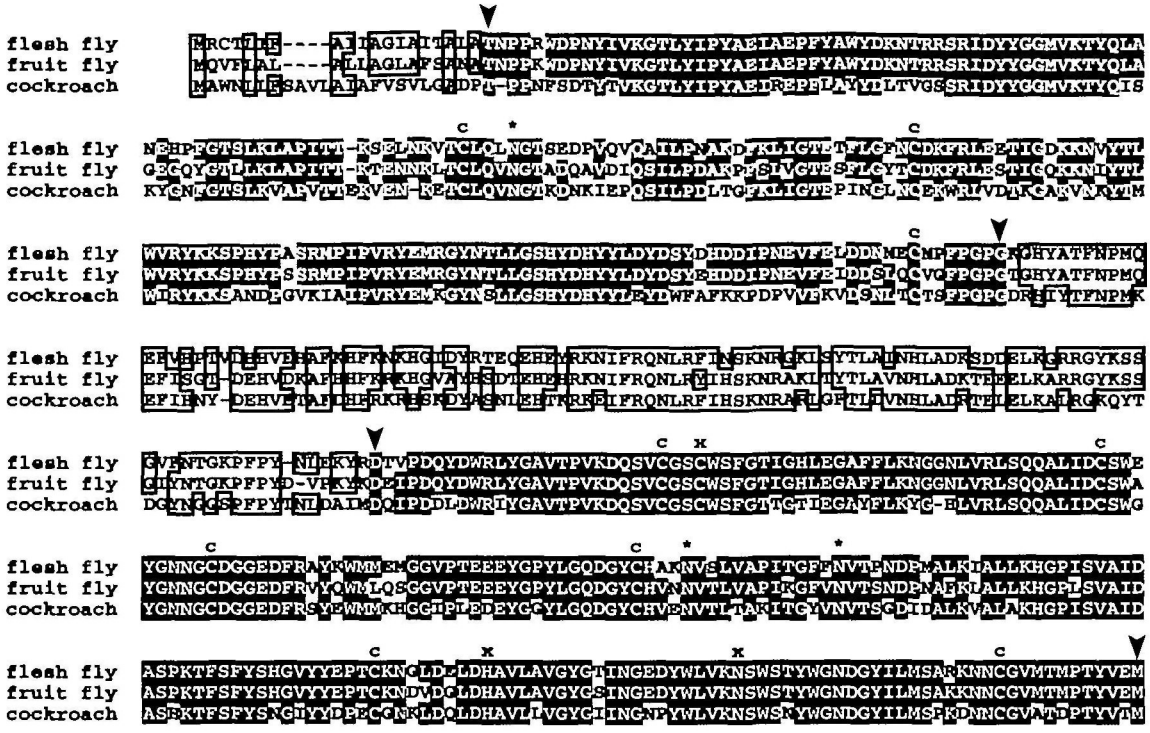

Fig. 6. **Comparison of amino acid sequences of 26/29-kDa proteinases from three insect species.** The 26/29-kDa proteinase cDNAs were isolated from *Drosophila* and *Periplaneta*, and the amino acid sequences of the precursor proteins encoded by these cDNAs were compared with that of *Sarcophaga*. Amino acid residues conserved in two or three insect species are boxed. The arrowheads indicate the amino-terminal and carboxyl-terminal residues of the 26- and 29-kDa subunits. The c markers indicate conserved cysteine residues. The x markers indicate three residues forming active sites. The asterisks indicate possible *N*-glycosylation sites. Gaps have been introduced to obtain maximal matching.

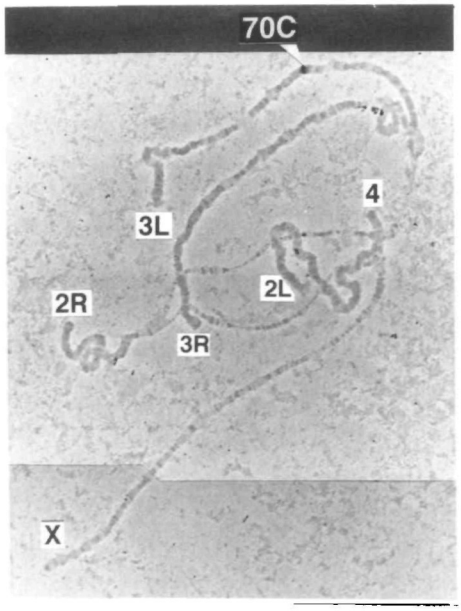

Fig. 7. **Chromosome mapping of the *Drosophila* 26/29-kDa proteinase gene.** Polytene chromosomes from the salivary glands of *D. melanogaster* were hybridized with digoxigenin-labeled *Drosophila* 26/29-kDa proteinase cDNA. The hybridized regions were detected by an alkaline phosphatase-conjugated anti-digoxigenin antibody. The arrowhead indicate a signal detected in 70C locus of the left arm of the third chromosome.

mine whether the 26/29-kDa proteinase is present in other insects, we tried to isolate its cDNA from cDNA libraries for the fruit fly (*D. melanogaster*) and the cockroach (*P. americana*) using PCR. We isolated cDNAs very similar to that of *Sarcophaga* 26/29-kDa proteinase, and these contained full length open reading frames. As shown in Fig. 6, these clones encoded the 26/29-kDa proteinase homologues of *Drosophila* and *Periplaneta*, respectively. The overall amino acid identity between the *Sarcophaga* and *Drosophila* proteins was 80%, and that between the *Sarcophaga* and *Periplaneta* proteins was 63%. This proteinase is thus present in a wide variety of insect hemocytes and may participate in their defense mechanisms like the 26/29-kDa proteinase of *Sarcophaga*.

In situ hybridization with the polytene chromosome of the *Drosophila* salivary gland revealed that the *Drosophila* 26/29-kDa proteinase gene is located at position 70C on the left arm of the third chromosome (Fig. 7). This locus is clearly different from that of *Drosophila* cathepsin L, which is located at position 50C on the right arm of the second chromosome (*13*).

DISCUSSION

The 26/29-kDa proteinase is a hemocyte proteinase of *Sarcophaga* that participates in the defense mechanism of this insect. In this study, we suggested that both its 26- and 29-kDa subunits are derived from a single precursor protein. So far, in the papain family, cathepsin C is the only enzyme known to be produced *via* this type of processing.

The precursor of cathepsin C contains a novel sequence (unique region) in its amino-terminal region, and this precursor is processed into the proteinase subunit and a unique subunit (*14, 15*). However, as judged by amino acid sequence identity, the 26/29-kDa proteinase differs from cathepsin C, but it is similar to cathepsin L (Fig. 8).

Although the 26/29-kDa proteinase shows a close resemblance to cathepsin L, it preferentially hydrolyzes Z-Arg-Arg-MCA and Pro-Phe-Arg-MCA. This is a novel feature in an enzyme belonging to the papain family. Interestingly, *Sarcophaga* cathepsin B hydrolyzes Z-Arg-Arg-MCA poorly (*4*). In *Sarcophaga*, Z-Arg-Arg-MCA hydrolyzing activity is mainly due to 26/29-kDa proteinase.

As mRNA for 26/29-kDa proteinase has been detected in both holometabolous (*endopeterygota*) and hemimetabolous (*exopeterygota*) insects, this proteinase seems to be conserved in various insect species. In a phylogenetic analysis based on the amino acid similarity of the mature proteinases in the papain family, 26/29-kDa proteinases were found to form an independent branch in the neighborhood of cathepsin L family (Fig. 8). Thus, 26/29-kDa proteinases seem to constitute a novel subfamily in papain family proteinases.

As reported previously, 26/29-kDa proteinase behaves uniquely when *Sarcophaga* larvae are injected with sheep red cells. More than 80% of the enzyme is released into hemolymph, and this is the only hemocyte proteinase so far known to be released into the hemolymph under these conditions (*4*). In the hemolymph, the 26/29-kDa proteinase appears to discriminate foreign and endogenous proteins in some way, and preferentially digests foreign protein. As the 29-kDa subunit can be regarded as a proteinase, we assume that the 26-kDa subunit participates in the recognition of the foreign substance to be degraded by the 29-kDa subunit. There seems to be no covalent linkage between the 26- and 29-kDa subunits, because they are easily separated on SDS-PAGE even under nonreducing conditions.

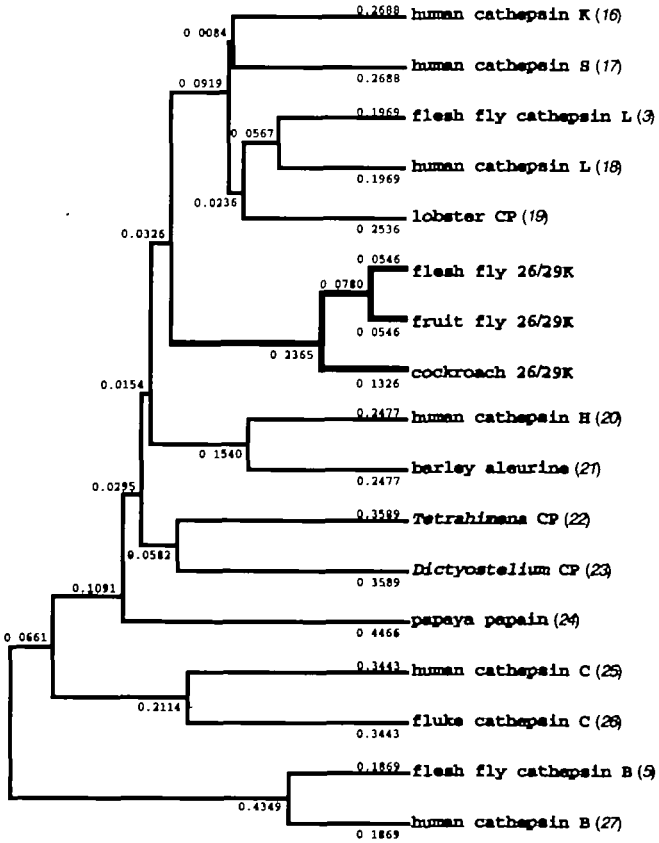

Fig. 8. **The position of insect 26/29-kDa proteinases in the phylogenetic tree of papain family proteinases.** The amino acid sequences of mature proteinase regions were compared using UPGMA method (*28*). The numbers in the tree denote genetic distances.

We would like to thank Dr Nobuhiko Katunuma (Tokushima Bunri University) for cathepsin L inhibitors, CLIK40 and CLIK46.

REFERENCES

1. Suzuki, T. and Natori, S. (1985) Preferential degradation and elimination of foreign protein introduced into the abdominal cavity of *Sarcophaga peregrina* larvae. *Comp. Biochem. Physiol.* **81**, 191-193
2. Kurata, S., Saito, H., and Natori, S. (1992) Purification of a 29-kDa hemocyte proteinase of *Sarcophaga peregrina*. *Eur. J. Biochem.* **204**, 911-914
3. Homma, K., Kurata, S., and Natori, S. (1994) Purification, characterization, and cDNA cloning of procathepsin L from the culture medium of NIH-Sape-4, an embryonic cell line of *Sarcophaga peregrina* (flesh fly), and its involvement in the differentiation of imaginal discs. *J. Biol. Chem.* **269**, 15258-15264
4. Saito, H., Kurata, S., and Natori, S. (1992) Purification and characterization of a hemocyte proteinase of *Sarcophaga*, possibly participating in elimination of foreign substances. *Eur. J. Biochem.* **209**, 939-944
5. Takahashi, N., Kurata, S., and Natori, S. (1993) Molecular cloning of cDNA for the 29 kDa proteinase participating in decomposition of the larval fat body during metamorphosis of *Sarcophaga peregrina* (flesh fly). *FEBS Lett.* **334**, 153-157
6. Kurata, S., Komano, H., and Natori, S. (1989) Dissociation of *Sarcophaga peregrina* (flesh fly) fat body by pupal hemocytes *in vitro*. *J. Insect. Physiol.* **35**, 559-565
7. Kurata, S., Saito, H., and Natori, S. (1990) Participation of hemocyte proteinase in dissociation of the fat body on pupation of *Sarcophaga peregrina* (flesh fly). *Insect Biochem.* **20**, 461-465
8. Kurata, S., Saito, H., and Natori, S. (1992) The 29-kDa hemocyte proteinase dissociates fat body at metamorphosis of *Sarcophaga*. *Dev. Biol.* **153**, 115-121
9. Homma, K. and Natori, S. (1994) Identification of substrate proteins for cathepsin L that are selectively hydrolyzed during the differentiation of imaginal discs of *Sarcophaga peregrina*. *Eur. J. Biochem.* **240**, 443-447
10. Natori, S. (1998) Relation between insect defense proteins and development of the flesh fly, *Sarcophaga peregrina* in *Molecular Mechanisms of Immune Responses in Insects* (Brey, P.T. and Hultmark, D., eds.) pp. 245-260, Chapman & Hall, London
11. Kunieda, T., Kurata, S., and Natori, S. (1997) Regeneration of *Sarcophaga* imaginal discs in vitro: implication of 20-hydroxyecdysone. *Dev. Biol.* **183**, 86-94
12. Kawasaki, K., Kubo, T., and Natori, S. (1996) Presence of the *Periplaneta* lectin-related protein family in the American cockroach *Periplaneta americana*. *Insect Biochem. Mol. Biol.* **26**, 355-364
13. Matsumoto, I., Watanabe, H., Abe, K., Arai, S., and Emori, Y. (1995) A putative digestive cysteine proteinase from *Drosophila melanogaster* is predominantly expressed in the embryonic and larval midgut. *Eur. J. Biochem.* **227**, 582-587
14. Nikawa, T., Towatari, T., and Katunuma, N. (1992) Purification and characterization of cathepsin J from rat liver. *Eur. J. Biochem.* **204**, 381-393
15. Dolenc, I., Turk, B., Pungercic, G., Ritonja, A., and Turk, V.

(1995) Oligomeric structure and substrate induced inhibition of human cathepsin C. *J. Biol. Chem.* **270**, 21626-21631
16. Inaoka, T., Bilbe, G., Ishibashi, O., Tezuka, K., Kumegawa, M., and Kokubo, T. (1995) Molecular cloning of human cDNA for cathepsin K: novel cysteine proteinase predominantly expressed in bone. *Biochem. Biophys. Res. Commun.* **206**, 89-96
17. Shi, G.P., Munger, J.S., Meara, J.P., Rich, D.H., and Chapman, H.A. (1992) Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. *J. Biol. Chem.* **267**, 7258-7262
18. Gal, S. and Gottesman, M.M. (1988) Isolation and sequence of a cDNA for human pro-(cathepsin L). *Biochem. J.* **253**, 303-306
19. Laycock, M.V., MacKay, R.M., Di Fruscio, M., and Gallant, J.W. (1991) Molecular cloning of three cDNAs that encode cysteine proteinases in the digestive gland of the American lobster (*Homarus americanus*). *FEBS Lett.* **292**, 115-120
20. Fuchs, R. and Gassen, H.G. (1989) Nucleotide sequence of human preprocathepsin H, a lysosomal cysteine proteinase. *Nucleic Acids Res.* **17**, 9471
21. Rogers, J.C., Dean, D., and Heck, G.R. (1985) Aleurain: a barley thiol protease closely related to mammalian cathepsin H. *Proc. Natl. Acad. Sci. USA* **82**, 6512-6516
22. Karrer, K.M., Peiffer, S.L., and DiTomas, M. (1993) Two distinct gene subfamilies within the family of cysteine protease genes. *Proc. Natl. Acad. Sci. USA* **90**, 3063-3067
23. Williams, J.G., North, M.J., and Mahbubani, H. (1985) A developmentally regulated cysteine proteinase in *Dictyostelium discoideum*. *EMBO J.* **4**, 999-1006
24. Cohen, L.W., Coghlan, V.M., and Dihel, L.C. (1986) Cloning and sequencing of papain-encoding cDNA. *Gene* **48**, 219-227
25. Paris, A., Strukelj, B., Pungercar, J., Renko, M., Dolenc, I., and Turk, V. (1995) Molecular cloning and sequence analysis of human preprocathepsin C. *FEBS Lett.* **369**, 326-330
26. Butler, R., Michel, A., Kunz, W., and Klinkert, M.Q. (1995) Sequence of *Schistosoma mansoni* cathepsin C and its structural comparison with papain and cathepsins B and L of the parasite. *Protein Pept. Lett.* **2**, 313-320
27. Ritonja, A., Popovic, T., Turk, V., Wiedımann, K., and Matchleidt, W. (1985) Amino acid sequence of human liver cathepsin B. *FEBS Lett.* **181**, 169-172
28. Sokal, R.R. and Michener, C.D. (1958) A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* **28**, 1409-1438
29. Murata, M., Miyashita, S., Yokoo, C., Tamai, M., Hanada, K., Hatayama, K., Towatari, T., Nikawa, T., and Katunuma, N. (1991) Novel epoxysuccinyl peptides. Selective inhibitors of cathepsin B, in vitro. *FEBS Lett.* **280**, 307-310
30. Katunuma, N. (1997) Molecular mechanisms of bone collagen degradation in bone resorption. *J. Bone Miner. Met.* **5**, 1-8