Purification, sequence and antibacterial activity of two novel sapecin homologues from *Sarcophaga* embryonic cells: similarity of sapecin B to charybdotoxin

Katsushige YAMADA and Shunji NATORI*
Faculty of Pharmaceutical Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Two sapecin homologues were purified from the culture medium of NIH-Sape-4, an embryonic cell line of *Sarcophaga peregrina* (flesh fly). These homologues contained six cysteine residues with exactly the same disulphide pairings as those in sapecin. The amino acid sequence of one of them, sapecin C, was also very similar to that of sapecin. The other homologue, sapecin B, was less similar to sapecin but showed significant similarity to charybdotoxin, an inhibitor of K^+ channels isolated from a scorpion venom. Like sapecin, these homologues repressed the growth of various Gram-positive bacteria.

INTRODUCTION

Dipteran insects have the ability to induce antibacterial proteins in response to injection of bacteria or body injury [1–5]. Sapecin, consisting of 40 amino acid residues, is one of these antibacterial proteins of *Sarcophaga peregrina* (flesh fly), and has been shown to have potent bactericidal activity against Gram-positive bacteria and also against Gram-negative bacteria at higher concentrations [6]. It was first purified from the culture medium of NIH-Sape-4, an established embryonic cell line of *Sarcophaga* [6]. Subsequently, the same protein was found to be synthesized by haemocytes and secreted into the haemolymph when the body wall of *Sarcophaga* larvae was injured [7]. Similar proteins have been isolated from two other dipteran insects. *Phormia terranova* (phormicin) [5] and *Apis mellifera* (royalicin) [8].

Sapecin has the novel character of interacting with cardiolipin in the bacterial plasma membrane, and the susceptibility of bacteria to sapecin depends on the content of cardiolipin in their plasma membrane [9]. Northern-blot analysis with sapecin cDNA revealed that the sapecin gene in the haemocytes was activated when the larval body wall was pricked with a hypodermic needle, which would facilitate bacterial infection [7]. The sapecin gene is also activated transiently in the embryonic and pupal stages of *Sarcophaga* without any outside stimulus, suggesting that sapecin participates in development of this insect [7]. In fact, we recently found that sapecin has growth-factor-like activity, stimulating proliferation of NIH-Sape-4 cells *in vitro* [10]. Therefore it is likely that sapecin is a double-functional molecule, active in both protection from bacterial infection and stimulation of embryonic cell proliferation.

In the present study, we purified two novel sapecin homologues from the culture medium of NIH-Sape-4, and determined their amino acid sequences and intermolecular disulphide bridges. The amino acid sequence of one of them was found to have significant similarity to that of charybdotoxin, a toxin acting on K^+ channels isolated from the venom of a scorpion [11,12].

MATERIALS AND METHODS

Cells and culture medium

The *Sarcophaga peregrina* embryonic cell line NIH-Sape-4 was cultured in M-M medium at 25 °C as described previously [13]. Briefly, cells were inoculated at a density of 5×10^5 cells/ml into 2 litre of M-M medium and cultured in a spinner flask with constant stirring at 45 rev./min. for about 5 days. When the cell density reached 4×10^6 cells/ml, the culture medium was collected by centrifugation (1500 rev./min, 15 min) and stored at −20 °C. This medium was used as the starting material for the purification of sapecin homologues.

Assay of antibacterial activity

During purification of sapecin homologues, we monitored their antibacterial activity with *Staphylococcus aureus* IFO 12732 as an indicator bacterium. Bacterial cells grown in antibiotic medium (Difco) were collected in the exponential phase of growth and suspended in 10 mM sodium phosphate buffer (pH 6.0) containing 130 mM NaCl (buffer A). The A_{650} of the cell suspension was adjusted to 0.3. Test samples were diluted serially with buffer A containing 0.2% (w/v) BSA, and portions (200 μl) of diluted samples were incubated with antibiotic medium (190 μl) and *S. aureus* suspension (10 μl) at 37 °C for 180 min. The mixture was then rapidly chilled, and its A_{650} was measured. One unit of antibacterial activity is defined as the amount causing 50% inhibition of bacterial growth relative to the control.

H.p.l.c.

The sample was applied to a column (250 mm × 4.1 mm) of Synchropak RP-P (C_{18}) connected to an LC-6AD system (Shimadzu Co., Tokyo, Japan), and material was eluted with a linear gradient of 15–35% solution B [0.05% (v/v) trifluoroacetic acid in acetonitrile] in solution A (0.05% trifluoroacetic acid in water) at a flow rate of 2 ml/min. The absorbance of the eluate at 280 nm was monitored.

SDS/PAGE

Electrophoresis on SDS/polyacrylamide slab gels was carried out by the method of Laemmli [14]. The stacking gel (3% acrylamide) was about 2 cm long and the separating gel (15% acrylamide) was about 7 cm long. Proteins (about 2 μg/lane)

* To whom correspondence should be addressed.
The amino acid sequence data presented have been deposited with the PIR International Sequence Data Bank and are available under the accession numbers JU 0224 (sapecin B) and JU 0225 (sapecin C).

were denatured by heating them in 1 % SDS solution containing 2 % (v/v) 2-mercaptoethanol for 20 min at 75 °C. After electrophoresis, the gels were stained by the method of Fairbanks et al. [15].

Determination of amino acid sequences and amino acid compositions of sapecin homologues

Samples of each protein (about 10 μg) separated by h.p.l.c. were lyophilized and subjected to automated protein sequence analysis in a gas-phase protein sequencer model 477A (Applied Biosystems Inc., Foster City, CA, U.S.A.). Pyridylethylation of cysteine residues in purified proteins was carried out essentially as described by Fullmer [16]. For analysis of the amino acid composition, purified protein was hydrolysed in 6 M HCl containing 4 % thioglycollic acid and the resulting amino acids were determined in an amino acid analyser model 835 (Hitachi, Tokyo, Japan).

Determination of disulphide pairs in the sapecin homologues

Disulphide pairs in one of the sapecin homologues, sapecin C, were determined by exactly the same method as used for their determination in sapecin [17]. Briefly, 100 μg of sapecin C was dissolved in 200 μl of 0.1 M Mes buffer (pH 6.5). Then 20 μl of 60 mM $CaCl_2$ and 1 μl of 1 mg/ml thermolysin were added, and the mixture was incubated for 19 h at 37 °C. The resulting peptides were analysed by reversed-phase h.p.l.c., and the amino acid sequences of peptides containing cysteine residues were determined.

Determination of disulphide pairs in the other sapecin homologue, sapecin B, was performed according to the flow chart shown in Figure 5. For this, we first digested a sample of 100 μg with 5 μg of endopeptidase Lys-C in 260 μl of 25 mM Tris/HCl buffer (pH 6.8) containing 1 mM EDTA for 4 h at 35 °C. As sapecin B contained six cysteine residues and two lysine residues, the digestion product gave a single peak on h.p.l.c. with a slightly different retention time from that of native sapecin B. This product was digested further with 15 μg of endopeptidase Glu-C in 300 μl of 50 mM phosphate buffer (pH 7.0) for 18 h at 35 °C. As sapecin B contained one glutamate residue, two peptides (fragment 1 and fragment 2) were obtained from the digest by h.p.l.c. Each peptide was digested with pepsin in 0.05 M HCl for 3 h at 25 °C, and the resulting digestion products were subjected to h.p.l.c. Two short peptides (fragment 3 and fragment 4) containing cysteine residues were identified by this procedure. One of these peptides (fragment 4) was further cleaved into two peptides (fragment 5 and fragment 6) containing cysteine residues by treatment with Tos-Phe-$CHCl_2$-treated trypsin. The amino acid sequences of the three peptides containing cysteine residues obtained in this way were determined, and the positions of disulphide pairs in sapecin B were assigned from the results.

Peptides containing cysteine residues were identified by measuring their absorptions at 412 nm after their reactions with 5,5′-dithiobis-(2-nitrobenzoate), by the method of Odani and Ikenaka [18]. Enzymes were purchased from Boehringer-Mannheim Gmbh.

RESULTS

Purification of two sapecin homologues from the culture medium of NIH-Sape-4

Previously, we purified sapecin from the culture medium of NIH-Sape-4 [6]. For this, we used *Escherichia coli* as an indicator bacterium. After purification of sapecin, we tested its antibacterial spectrum, and found that in general it had more effect on Gram-positive bacteria than on Gram-negative bacteria. As other antibacterial proteins of *Sarcophaga* such as sarcotoxin I and sarcotoxin II have several isoforms with very similar amino acid sequences [19,20], we thought that sapecin might also have various isoforms and that these might be released into the culture medium of NIH-Sape-4, but not be detected with *E. coli* which is not a sensitive indicator. In fact, two sapecin homologues have been purified from the haemolymph of immunized larvae of *Phormia terranova*, another dipteran insect [5].

To detect sapecin homologues, we reinvestigated the antibacterial proteins in the culture medium of NIH-Sape-4 using *S. aureus* IFO 12732 as an indicator bacterium. The procedures used were almost the same as those we used previously [6]. Briefly, about 900 ml of culture medium was loaded on to a column of CM-cellulose, and proteins with antibacterial activity were eluted with 10 mM phosphate buffer containing 520 mM NaCl. This fraction was heated for 10 min at 100 °C, and denatured protein was removed by centrifugation. The resulting clear supernatant was concentrated to about 5 ml by ultrafiltration (Diaflo, UF disc type YM10, Amicon), and further fractionated on a column of Sephadex G-50. As shown in Figure 1, the antibacterial activity was eluted as a single peak. Fractions containing antibacterial activity were pooled and subjected to reversed-phase h.p.l.c. Six clear peaks, numbered I to VI, were eluted as shown in Figure 2. Each peak was collected separately and lyophilized. Peaks IV, V and VI contained significant antibacterial activity, whereas no activity was detected in peaks I, II and III.

The retention time of peak IV protein coincided well with that of sapecin, so we concluded that the material in this peak was sapecin. This conclusion was confirmed by determination of the amino acid sequence of this material; its amino acid sequence coincided completely with that of sapecin. The proteins in peaks V and VI were thought to be novel sapecin homologues as they inhibited the growth of *S. aureus*. At this stage, each protein was almost homogeneous and gave a single band on SDS/PAGE, as shown in Figure 3. The results of a typical purification are shown in Table 1. The amounts of peak-V and peak-VI proteins were always about one-seventh of that of sapecin. The specific activities

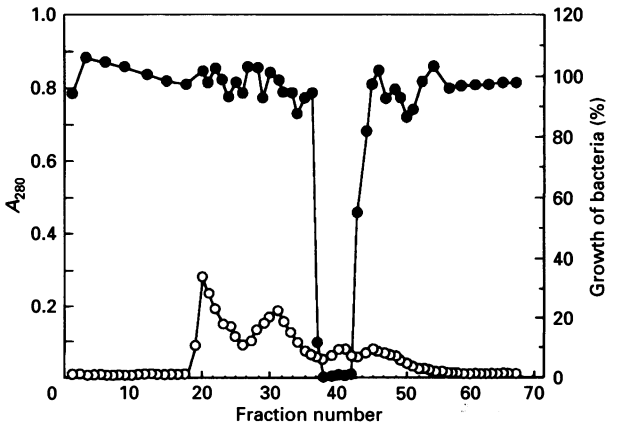

Figure 1 Chromatography of antibacterial activity on a column of Sephadex G-50

Concentrated solution (about 5 ml) was fractionated on a column (1.5 cm × 120 cm) of Sephadex G-50. Fractions were collected, and the antibacterial activity of 20 μl of each fraction was tested using *S. aureus*. Antibacterial activity was measured as inhibition of bacterial growth. ○, A_{280}; ●, bacterial growth.

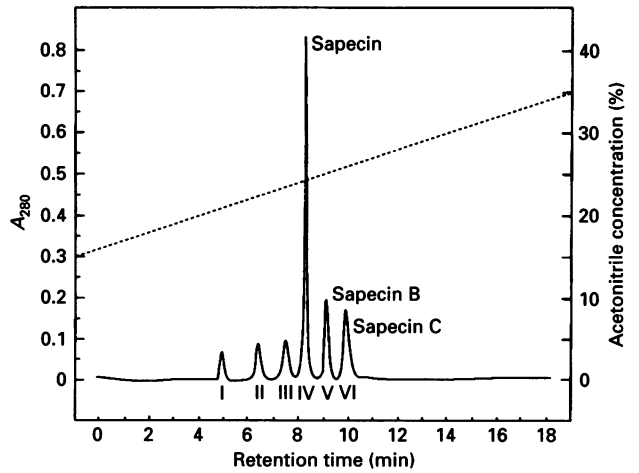

Figure 2 Reversed-phase h.p.l.c. of the Sephadex G-50 fraction

The active fraction from Sephadex G-50 was fractionated in a h.p.l.c. system with a Synchropak RP-R (C_{18}) column. Material was eluted with a linear gradient of 15–35% acetonitrile. ——, A_{220}; – – – –, acetonitrile concentration (%).

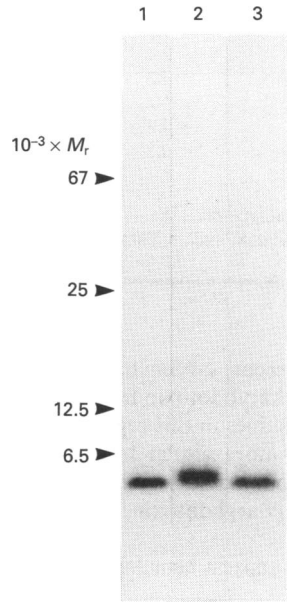

Figure 3 SDS/PAGE of purified proteins

Electrophoresis was carried out with 2 μg of purified protein per lane under denaturing conditions. Lane 1, peak IV protein; lane 2, peak V protein; lane 3, peak VI protein. Arrows indicate the positions of the following marker proteins: BSA (67000), α-chymotrypsinogen (25000), cytochrome c (12500) and aprotinin (6500).

Table 1 Summary of purification of sapecin-like proteins

About 900 ml of culture medium was used as starting material. Protein was determined by the method of Lowry et al. [21].

Purification step	Total activity (units)	Total protein (mg)	Specific activity (units/mg)	Yield (%)
Culture medium	6675	3258	2	100
CM-cellulose	4247	26.8	159	63.6
Heat-treatment	3406	18.3	186	51.0
Concentration	2912	15.8	184	43.6
Sephadex G-50	1986	5.3	375	29.8
H.p.l.c.				
Sapecin	1540	0.53	2917	23.1
Sapecin B	58.3	0.07	833	0.9
Sapecin C	206.6	0.08	2583	3.1

of sapecin and the peak-VI protein were almost the same, but that of peak-V protein was clearly less.

Structural analyses of peak-V and peak-VI proteins

We determined the amino acid sequences of the peak-V and peak-VI proteins. The peak-VI protein consisted of 40 amino acid residues, like sapecin, its amino acid sequence differing from that of sapecin in nine amino acid residues (78% identity), as shown in Figure 4. Therefore, the peak-VI protein is clearly a sapecin homologue. The peak-V protein consisted of 34 amino acid residues, and the similarity of its amino acid sequence to that of sapecin was much less, being only 41% identity. However, like sapecin, it contained six cysteine residues, so we concluded that it was also a sapecin homologue. Therefore we named peak-V protein and peak-VI protein sapecin B and sapecin C respectively.

We determined the disulphide pairs in sapecin B and C. As the primary structure of sapecin C was very similar to that of sapecin, we determined its disulphide pairs by exactly the same method as we used to determine those in sapecin [17]. However, this method was not appropriate for determination of the disulphide pairs in sapecin B. Therefore, to determine them, we successively treated sapecin B with endopeptidase Lys-C, endopeptidase Glu-C, pepsin and trypsin, according to the flow chart shown in Figure 5, and finally obtained three peptides each containing a cysteine pair. The amino acid sequences of these peptides indicated the positions of the disulphide pairs in sapecin B. The pairings of cysteine residues in both sapecin B and sapecin C were the same as those in sapecin. Namely, the pairings of

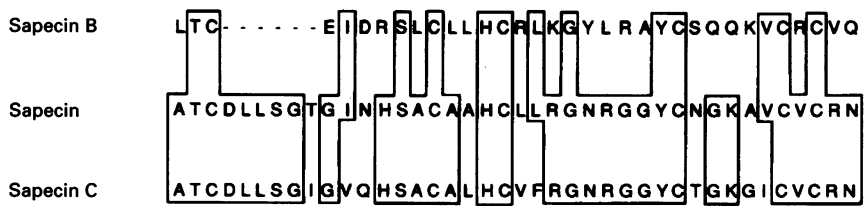

Figure 4 Comparison of amino acid sequences of sapecin and its homologues

Identical amino acids in sapecin and its homologues are boxed. A gap was introduced to obtain maximal sequence similarity.

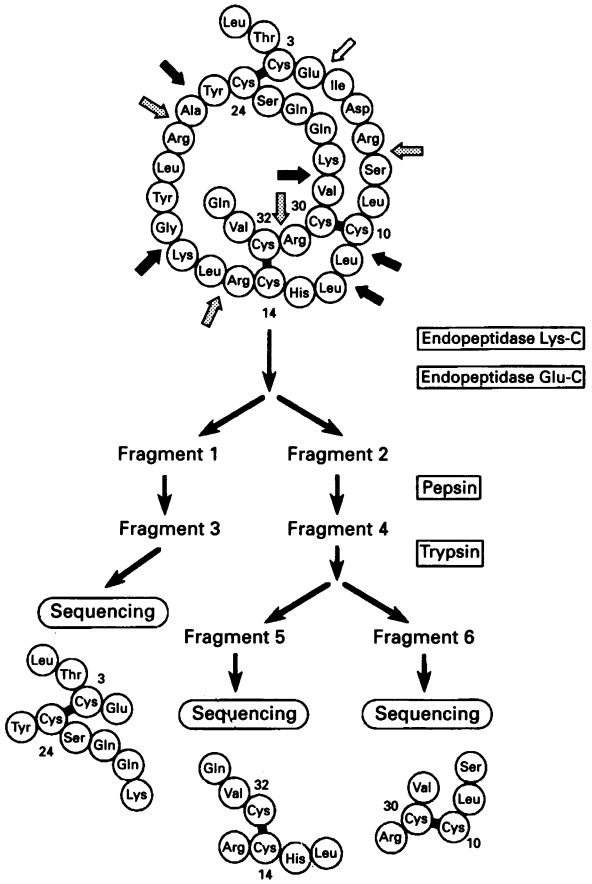

Figure 5 Flow chart of the determination of the disulphide pairs of sapecin B

Points of cleavage by the various proteinases are shown by arrows. Numbers indicate the positions of cysteine residues from the N-terminal amino acid residue. Cleavage sites are: ➡, endopeptidase Lys-C; ⇨, endopeptidase Glu-C; ➡, pepsin; ⇨, trypsin. Fragments 3, 5 and 6 were sequenced.

the six cysteine residues numbered from the N-terminus to the C-terminus were 1–4, 2–5 and 3–6 (see Figure 5).

Recently, we presented a model of sapecin based on its ^{1}H n.m.r. [22]. Bontems et al. [23] pointed out that sapecin has the same structural motif as that in charybdotoxin, a toxin isolated from the venom of a scorpion. Although the pairings of the three disulphide bonds in sapecin and charybdotoxin are identical, the numbers of amino acid residues between the first and second

Table 2 Antibacterial activities of sapecin, sapecin B and sapecin C

Antibacterial activities were measured as described in the text unless otherwise mentioned and antibacterial activities are expressed as concentrations causing 50% inhibition of bacterial growth relative to the control. N.T., not tested.

Bacterial strain	Antibacterial activity (μg/ml)		
	Sapecin	Sapecin B	Sapecin C
Escherichia coli K-12	> 20.0	> 20.0	> 20.0
Proteus mirabilis IFO 3849	> 20.0	> 20.0	> 20.0
Klebsiella pneumoniae ATCC21316	> 20.0	> 20.0	> 20.0
Staphylococcus aureus ATCC6538B	0.5	1.6	0.6
Staphylococcus epidermidis IFO 3762	1.0	7.0	1.2
Staphylococcus aureus MR3626*	N.T.	N.T.	4.6
Staphylococcus aureus MR1550*	N.T.	N.T.	6.6
Staphylococcus aureus MR1587*	N.T.	N.T.	0.9
Staphylococcus aureus MR3636*	N.T.	N.T.	5.4
Streptococcus mutans 6515†	4.5	6.6	2.6
Streptococcus sanguis ST3†	1.5	3.8	0.9
Streptococcus salivarius HHT†	2.6	3.5	0.7
Streptococcus bovis ATCC27960†	5.2	0.9	2.4
Bacillus megaterium IAM 1166	< 0.2	< 0.2	< 0.2
Bacillus circulans IFO 3967	< 0.2	< 0.2	< 0.2
Corynebacterium glutamicum ATCC13059	< 0.2	< 0.2	< 0.2

*Methicillin- and cephem-resistant strain [24].
†Cultured with brain heart infusion medium (Difco).

cysteine residues are different: sapecin has 12 amino acid residues in this region, whereas charybdotoxin has only 5 [12]. As sapecin B has six amino acid residues in this region, its overall structure is expected to be much more similar to that of charybdotoxin. Moreover, significant amino acid sequence similarity was found between sapecin B and charybdotoxin, as shown in Figure 6.

Antibacterial activities of sapecin homologues

We examined the antibacterial activities of the sapecin homologues against various bacteria. As shown in Table 2, the antibacterial spectra of the two sapecin homologues were similar to that of sapecin, all inhibiting growth of various Gram-positive

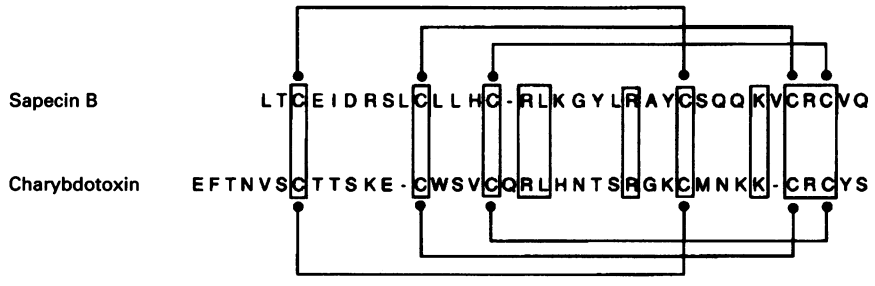

Figure 6 Comparison of amino acid sequences of sapecin B and charybdotoxin

Identical amino acids are boxed. Gaps were introduced to obtain maximal sequence similarity. Disulphide bridges are indicated.

bacteria at concentrations of less than 10 μg/ml. However, the homologues at 20 μg/ml did not affect the growth of the Gram-negative bacteria tested. It is particularly noteworthy that sapecin C inhibited the growth of methicillin-resistant *S. aureus*, because only a few antibiotics with this effect are known. Bacteria that cause tooth decay, *Streptococcus mutans*, *Streptococcus sanguis* and *Streptococcus salivarius*, were also sensitive to sapecin and its homologues. These results suggest that sapecin and its homologues are potentially interesting proteins for the development of useful antibacterial drugs.

DISCUSSION

This paper describes the purification of two sapecin homologues, sapecin B and sapecin C, from the culture medium of NIH-Sape-4. Previously, we purified sapecin from the same culture medium, but we did not detect these sapecin homologues [6]. This may have been because the contents of these homologues in the medium were only one-seventh of that of sapecin and because we used *E. coli* as an indicator bacterium instead of *S. aureus*.

The primary structure of sapecin C is very similar to that of sapecin, 31 of 40 amino acid residues being identical. Thus it is readily conceivable that the two are derived from a common ancestral gene. This situation was the same as that of the homologues of other antibacterial proteins of this insect: we previously reported very high similarities in the amino acid sequences of three members of the sarcotoxin I family [25] and three members of the sarcotoxin II family [20]. However, the similarity between sapecin and sapecin B was clearly less than that between sapecin and sapecin C: only 14 of 34 amino acid residues were found to be identical when the sequences were aligned to give maximum matching.

On the other hand, significant similarity was found between sapecin B and charybdotoxin: on alignment to give maximum matching, 11 of 34 amino acid residues were identical. Charybdotoxin is a small protein which was purified from the venom of a scorpion and shown to inhibit K^+ channels [11]. It contains a circular motif consisting of Cys-Trp-Ser-Val-Cys and Cys-Arg-Cys in which the cysteine residues of both sides are linked by disulphide bonds [23]. A similar motif consisting of Cys-Leu-Leu-His-Cys and Cys-Arg-Cys is present in sapecin B. It is not yet known whether sapecin B affects K^+ channels, but possibly sapecin B is not a member of the sapecin family but a member of the charybdotoxin-like protein family regulating cationic channels. To test this possibility, it may be useful to examine expression of the sapecin B gene in a haemocytes when the body wall of *Sarcophaga* larvae is injured by pricking with a hypodermic needle because the sapecin gene is significantly activated under these conditions, and if sapecin B belongs to the sapecin family, its gene should be activated like the sapecin gene; if not, it may not be activated [7].

Sapecin and its homologues had effects on all Gram-positive bacteria tested, but their effective concentrations differed depending on the bacterial strain. For instance, *S. epidermidis* was more sensitive to sapecin and sapecin C than to sapecin B, whereas *Strep. bovis* was most sensitive to sapecin B. Probably, multiple sapecin homologues are needed for a potent defence system against Gram-positive bacteria. It is noteworthy that sapecin C was effective against four methicillin-resistant *S. aureus* strains tested. Sapecin and sapecin C should have a similar effect. We have previously shown that the primary target of sapecin is the bacterial membrane, and that the content of cardiolipin in the membrane is important for susceptibility to sapecin [9]. The mode of action of sapecin is totally different from that of β-lactam antibiotics, so sapecin and its homologues should provide a clue for designing novel antibacterial drugs that are effective against methicillin-resistant *S. aureus*.

This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan and by a Grant-in-Aid BMP 92-IV-2 (Bio Media Program) from the Ministry of Agriculture, Forestry and Fishery of Japan.

REFERENCES

1 Natori, S. (1977) J. Insect Physiol. **23**, 1169–1173
2 Okada, M. and Natori, S. (1983) Biochem. J. **211**, 727–734
3 Ando, K., Okada, M. and Natori, S. (1987) Biochemistry **260**, 7174–7177
4 Dimarcq, J.-L., Keppi, E., Dunbar, B., Lambert, J., Reichhart, J. M., Hoffmann, D., Rankine, S. M., Fothergill, J. E. and Hoffmann, J. A. (1988) Eur. J. Biochem. **171**, 17–22
5 Lambert, J., Keppi, E., Dimarcq, J.-L., Wicher, C., Reichhart, J. M., Dunbar, B., Lepage, P., van Dorsselaer, D. A., Hoffmann, J. A., Fothergill, J. and Hoffmann, D. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 262–266
6 Matsuyama, K. and Natori, S. (1988) J. Biol. Chem. **263**, 17112–17116
7 Matsuyama, K. and Natori, S. (1988) J. Biol. Chem. **263**, 17117–17121
8 Fujiwara, S., Imai, J., Fujiwara, M., Yaeshima, T., Kawashima, T. and Kobayashi, K. (1990) J. Biol. Chem. **265**, 11333–11337
9 Matsuyama, K. and Natori, S. (1990) J. Biochem. (Tokyo) **107**, 514–518
10 Komano, H., Homma, K. and Natori, S. (1991) FEBS Letts. **289**, 167–170
11 Miller, C., Moczydlowski, E., Latorre, R. and Philipps, M. (1985) Nature (London) **313**, 316–318
12 Gimenez-Gallego, G., Navia, M. A., Reuben, J. P., Katz, G. M., Kaczorowski, G. J. and Garcia, M. L. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 3329–3333
13 Komano, H., Kasama, E., Nagasawa, E., Nakanishi, Y., Matsuyama, K., Ando, K. and Natori, S. (1987) Biochem. J. **248**, 217–222
14 Laemmli, U. K. (1970) Nature (London) **227**, 680–685
15 Fairbanks, G., Steck, T. L. and Wallach, D. F. H. (1971) Biochemistry **10**, 2606–2617
16 Fullmer, C. S. (1984) Anal. Biochem. **124**, 336–339
17 Kuzuhara, T., Nakajima, Y., Matsuyama, K. and Natori, S. (1990) J. Biochem. (Tokyo) **107**, 514–518
18 Odani, S. and Ikenaka, T. (1973) J. Biochem. (Tokyo) **74**, 697–715
19 Kanai, A. and Natori, S. (1989) FEBS Letts. **258**, 199–202
20 Kanai, A. and Natori, S. (1990) Mol. Cell. Biol. **10**, 6114–6122
21 Lowry, O. H., Rosebrough, N. L., Farr, A. L. and Randall, R. J. (1951) J. Biol. Chem. **193**, 265–275
22 Hanzawa, H., Shimada, I., Kuzuhara, T., Komano, H., Kohda, D., Inagaki, F., Natori, S. and Arata, Y. (1990) FEBS Letts. **269**, 413–420
23 Bontems, F., Roumestand, C., Gilquin, B., Menez, A. and Toma, F. (1991) Science **254**, 1521–1523
24 Murakami, K., Nomura, K., Doi, M. and Yoshida, T. (1987) Antimicrob. Agents Chemother. **31**, 1307–1311
25 Okada, M. and Natori, S. (1985) J. Biol. Chem. **260**, 7174–7177

Received 10 August 1992/8 October 1992; accepted 13 October 1992