Amino acid sequence of PR-39

Isolation from pig intestine of a new member of the family of proline-arginine-rich antibacterial peptides

Birgitta AGERBERTH[1], Jong-Youn LEE[2], Tomas BERGMAN[3], Mats CARLQUIST[4], Hans G. BOMAN[2], Viktor MUTT[1] and Hans JÖRNVALL[3]

Departments of [1] Biochemistry II and [3] Chemistry I, Karolinska Institutet, Stockholm, Sweden,
[2] Department of Microbiology, Stockholm University, Sweden, [4] Karo Bio AB, Huddinge, Sweden

We recently isolated from pig intestine and characterized a 31-residue antibacterial peptide named cecropin P1 with activity against *Escherichia coli* and several other Gram-negative bacteria. The isolation involved a number of batch-wise steps followed by several chromatography steps. The continued investigation of these antibacterial peptides has now yielded another antibacterial peptide with high activity against both *E. coli* and *Bacillus megaterium*. Amino acid analysis showed a very high content of proline (49 mol%) and arginine (26 mol%), an intermediate level of phenylalanine and low levels of leucine, tyrosine, isoleucine, and glycine. The primary structure was determined by a combination of Edman degradation, plasma desorption mass spectrometry and C-terminal sequence analysis by carboxypeptidase Y degradation using capillary zone electrophoresis for detection of liberated residues. The calculated molecular mass was 4719.7 Da, which is in excellent agreement with 4719 Da obtained by plasma desorption mass spectrometry. The peptide was named PR-39 (proline-arginine-rich with a size of 39 residues). The lethal concentration of the peptide was determined against six Gram-negative and four Gram-positive strains of bacteria.

The small intestine has been a rich source for the isolation of physiologically active peptides (Mutt, 1986). Since the stomach and the upper part of the small intestine, during normal healthy conditions, contain few bacteria, we predicted that these organs should produce antibacterial peptides. A search in different pools, generated during the isolation of intestinal hormones, revealed several fractions with activity against both *Escherichia coli* and *Bacillus megaterium*, two suitable assay organisms. From one such porcine fraction we isolated a peptide, cecropin P1, with an unexpected similarity to the insect cecropins (Lee et al., 1989). We have continued our search for new antibacterial factors from pig intestine and have now characterized a 39-residue peptide exceptionally rich in proline and arginine, and consequently named PR-39. Cecropins (Boman and Hultmark, 1987) and defensins (Lehrer and Ganz, 1990; Lehrer et al., 1991) were the first antibacterial peptides discovered but, since the finding of the magainins in 1987 (Zasloff, 1987), the field is rapidly growing and a review of the subject appeared recently (Boman, 1991).

MATERIALS AND METHODS

Materials

CM-cellulose (carboxymethyl-cellulose CM 23) was from Whatman (Kent, UK), Sephadex G-25 fine from Pharmacia

---

Correspondence to B. Agerberth, Department of Biochemistry II, Karolinska Institutet, Box 60400, S-104 01 Stockholm, Sweden

Abbreviations. CM-cellulose, carboxymethyl-cellulose; PR-39, 39-residue proline-arginine-rich peptide.

Enzymes. Carboxypeptidase Y (EC 3.4.16.1); trypsin (EC 3.4.21.4).

(Uppsala, Sweden). Trypsin (treated with L-(tosylamido 2-phenyl)ethyl chloromethane was obtained from Worthington (Freehold, NJ, USA), and carboxypeptidase Y from Boehringer Mannheim (Mannheim, FRG).

Antibacterial assays

Activity was recorded by the inhibition zone assay in thin agarose plates seeded with *Escherichia coli* D21 (Hultmark et al., 1983). A standard curve was made from known amounts of synthetic cecropin A; 1000 units was defined as an activity equal to that of 1 μg of cecropin A. The chromatographic fractions were lyophilized and a sample of each fraction was redissolved in water, giving solutions with 10 mg/ml of which 3 μl were used for each assay. For comparison of activities against different bacteria, lethal concentrations (LC values) were calculated (Hultmark et al., 1983) and are given together with the diameter of the maximum inhibition zone.

Purification

The starting material was a concentrate of thermostable intestinal peptides prepared from the upper part of the porcine small intestine (Chen et al., 1988; Mutt, 1959). It was dissolved in water, treated with ethanol and the peptides which were in neutral aqueous ethanol were adsorbed onto alginic acid, eluted with 0.2 M HCl and precipitated by saturation with NaCl (Lee et al., 1989). The precipitate, dissolved in 0.2 M acetic acid, was chromatographed on Sephadex G-25 fine. The eluate corresponding to fraction II in Mutt (1978) was precipitated by saturation with NaCl. The precipitate was

dissolved in water, reprecipitated with NaCl at pH 4.0 and extracted with methanol (50 ml/g). The mixture was filtered, the solution brought to pH 7.5 and precipitated material discarded. After adjustment to pH 2.7, ether (3 vol.) was added. The precipitate that formed was recovered by suction filtration, dissolved in water, and reprecipitated with NaCl to saturation. After a desalting step, the material was chromatographed on CM-cellulose at pH 8.0 (Said and Mutt, 1972). The fraction indicated by a bar, denoted post-secretin, was rechromatographed on CM-cellulose in 22.5 mM sodium phosphate pH 6.4 with a linear gradient 0–0.3 M NaCl as described (McDonald et al., 1983). The final material, eluted by 0.2 M HCl, contained antibacterial activity, later found to be due to PR-39. We first tried to purify this material by reverse-phase HPLC but found that the method was unsuitable. It was, however, possible to purify it by conventional chromatography on CM-cellulose.

Structural analysis

The purity of the peptide was ascertained by thin-layer chromatography on silica gel plates (Riedel de Haen, Hannover, FRG) in 1-butanol/acetic acid/water/pyridine (15:3:12:10, by vol.) and by capillary electrophoresis (Bergman et al., 1991). Cleavage with trypsin was performed in 1% ammonium bicarbonate for 6 h at 37°C with an enzyme/substrate ratio of 2–3:5. Tryptic fragments were separated by HPLC on Vydac 218 TP (5 µm; 4.6 × 250 mm; Separations Group, Hesperia, CA, USA).

Sequencer degradations were carried out with an Applied Biosystems 470A instrument, utilizing a Hewlett Packard 1090 HPLC for identification of phenylthiohydantoins with a sodium acetate/acetonitrile gradient as described (Kaiser et al. 1988). For C-terminal sequence analysis, the peptide was cleaved with carboxypeptidase Y in pyridine acetate pH 6.0 at 37°C (1.6 µg protease/nmol peptide). At the indicated time the reaction was terminated by boiling. Aliquots were analyzed by capillary zone electrophoresis with a Beckman P/ACE system 2000 (Bergman et al., 1991). Capillaries were of fused silica (Beckman), inner diameter 75 µm, total length 57 cm and 50 mM phosphate pH 2.5 was employed as electrolyte. Samples were injected by pressure for 20 s.

Total compositions were determined with Beckman 121 M and LKB 4151 Alpha Plus amino acid analyzers after hydrolysis in evacuated tubes with 6 M HCl/0.5% phenol at 110°C for 20 h.

Molecular masses were determined by plasma desorption mass spectrometry using a BioIon 20 mass spectrometer (BioIon, Uppsala, Sweden, now Applied Biosystems). Samples were dissolved in 0.1% trifluoroacetic acid containing 50% ethanol, applied to aluminium foil covered by nitrocellulose and dried in a stream of nitrogen. Data were accumulated for 15 min at 15 kV acceleration voltage. Hydrogen and sodium or nitrous oxide were used as internal standards for calibration. About 10 nmol of the peptide was used for methyl esterification (Ørskov et al., 1989) and an aliquot was analyzed by plasma desorption mass spectrometry.

RESULTS

Isolation of the antibacterial peptide from pig intestine

The purification starts from a crude peptide fraction as described in Materials and Methods following the scheme outlined in Fig. 1. From the penultimate step (CM-cellulose
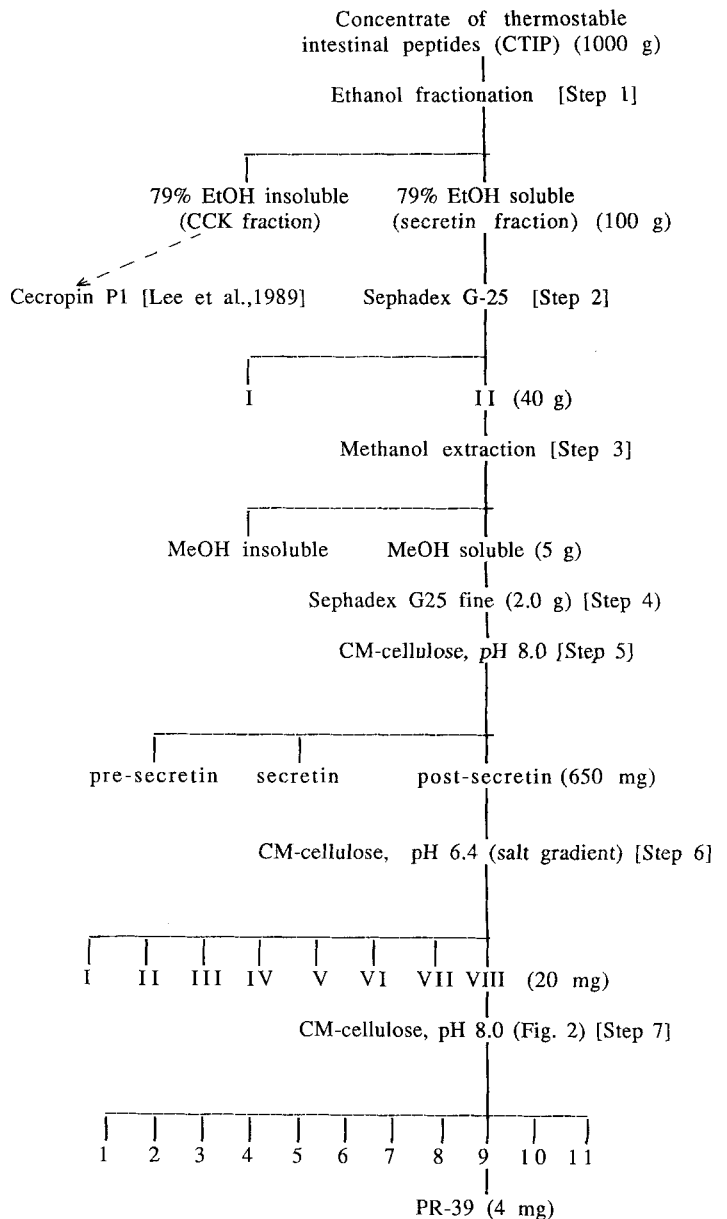

Fig. 1. Purification scheme for the porcine antibacterial peptide PR-39. Details of the different precipitations are given in Materials and Methods.

chromatography at pH 6.4), fraction VIII (30 mg) was dissolved in 1 ml water and adjusted to pH 8.0 with a few drops 0.2 M ammonia. The clear solution was applied to a column with CM-cellulose (0.6 × 40 cm) and the peptides were eluted step-wise with 0.02 M, 0.05 M, 0.1 M, 0.2 M, 0.3 M, 0.4 M, 0.5 M ammonium bicarbonate at pH 8.0, and finally with 0.2 M ammonia. The early steps gave only small peaks but the 0.2–0.4 M steps gave higher peaks as seen from Fig. 2. The total eluate, divided into 11 fractions, was lyophilized and assayed for antibacterial activity against *E. coli*. Fraction VI contained 40 units/µg, while fraction VII contained 340 units/µg, and fraction VIII slightly less. Maximum specific activity, 700 units/µg, was found in fraction IX (Fig. 2) and corresponded to the latter part of the peak eluted with 0.4 M ammonium bicarbonate. The material in this peak was found to be essentially homogeneous by the following criteria: only one spot was detected with ninhydrin on silica gel thin-layer
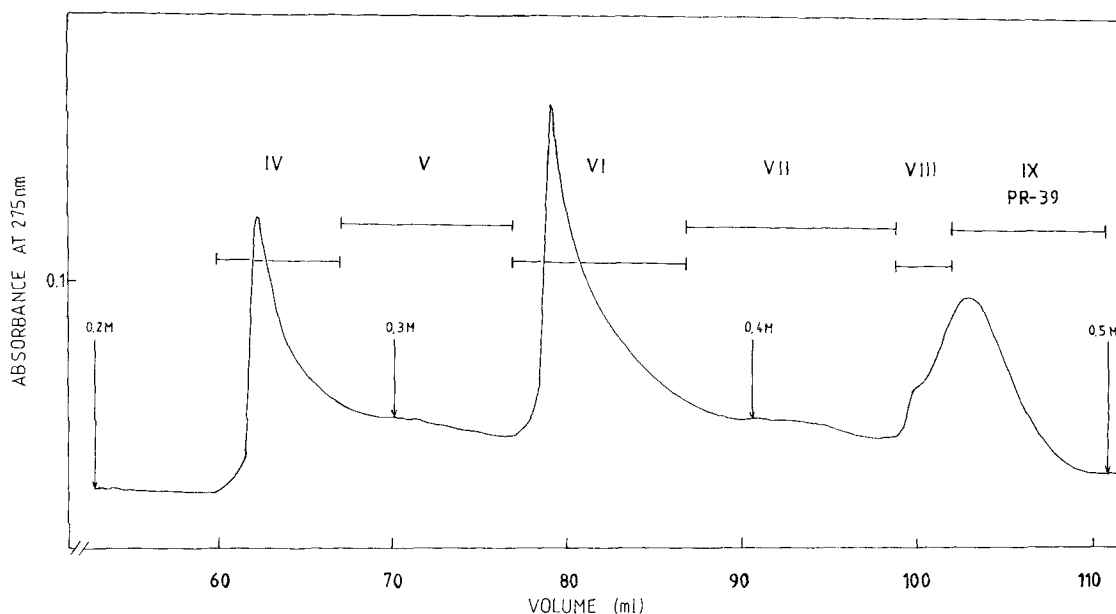
Fig. 2. The final step in the purification of PR-39, rechromatography on CM-cellulose. The starting material was fraction VIII from step 6 in Fig. 1. Elution was stepwise with different concentrations of ammonium bicarbonate as indicated by arrows. The eluate was pooled as shown by the bars. PR-39 was present in pure form in the latter part of the last peak eluted by 0.4 M ammonium bicarbonate.
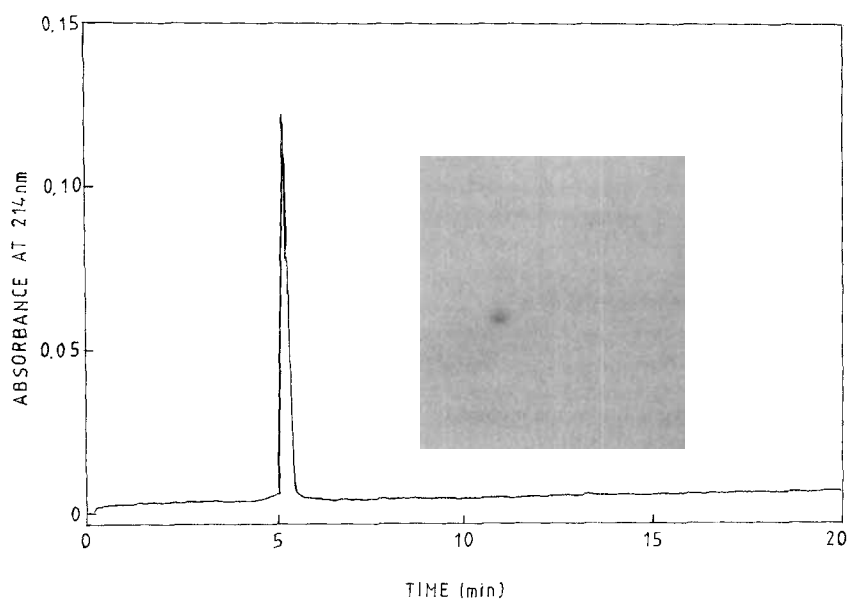
Fig. 3. Capillary zone electrophoresis of PR-39 with an insert showing thin-layer chromatography on a silica gel plate. Electrophoresis conditions were 25 kV, 25 °C in 50 mM phosphate pH 2.5; the amount of PR-39 applied was 10 pmol; 2 nmol PR-39 was applied to the silica gel plate.

chromatography, only one symmetric peak obtained with capillary zone electrophoresis (Fig. 3) and structural analysis gave one amino acid sequence.
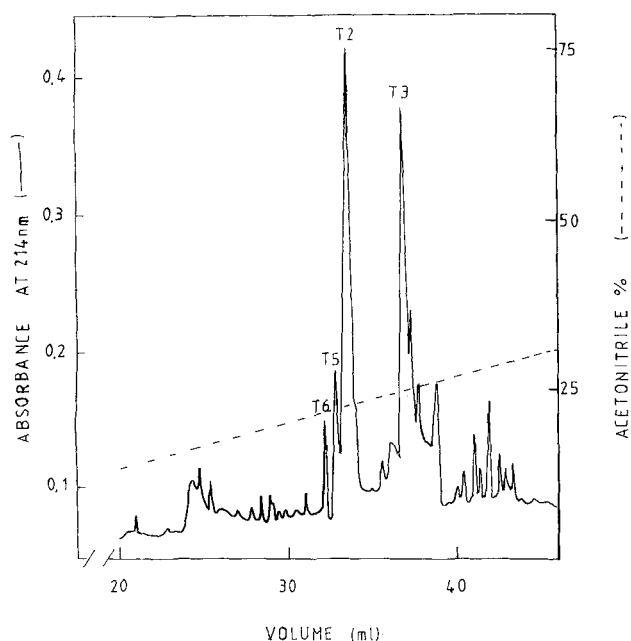
Determination of the primary structure

Table 1 shows the amino acid composition of PR-39 revealing it to contain only seven different amino acids and being very rich in proline (49 mol%) and arginine (26 mol%). The remaining five amino acids (tyrosine, leucine, phenylalanine, isoleucine and glycine) are predominantly hydrophobic. The frequent repeats of a few residues complicated degradations and several attempts at direct sequence analysis failed to

establish the entire primary structure. The only enzyme giving a fairly stoichiometric cleavage of the peptide was trypsin, but only at a high enzyme/substrate ratio $(2-3:5)$. The tryptic fragments were separated on reverse-phase HPLC and the material corresponding to four peaks (T2, T3, T5, and T6; Fig. 4) were analyzed for amino acid composition and sequence. Fig. 5 depicts the sequence deduced from several degradations of the intact peptide and tryptic fragments. Both the isolation and structure of T6 and the sequencer degradation of intact PR-39 show that residue 37 is arginine. Furthermore, mass spectroscopy indicated a molecular mass of 4719 Da compatible with a C-terminal at position 39. If so, tryptic digestion should liberate a dipeptide, which would be
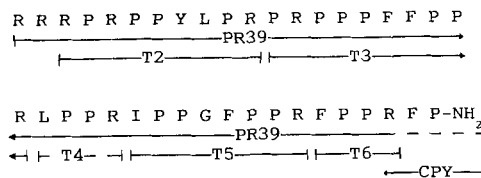
Table 1. Amino acid composition of PR-39. Values shown are molar ratios from acid hydrolysis and, within parentheses, those from sequence analysis.

Amino acid	Amount
	mol/mol
Pro	18.8 (19)
Arg	9.6 (10)
Phe	4.7 (5)
Leu	1.9 (2)
Gly	1.4 (1)
Ile	0.9 (1)
Tyr	1.0 (1)

Fig. 5. Amino acid sequence of PR-39. Residues 1−37 of intact PR-39 and its tryptic fragments T2, T3, T5 and T6 were obtained by direct sequence analysis (solid lines). The C-terminus was deduced from digestion with carboxypeptidase Y (solid line). The remaining regions (dashed lines) were proven by total composition and mass spectrometry. Composition of PR-39: Pro 48.7 mol%; Arg 25.6 mol%; Phe 12.8 mol%; Leu 5.1 mol%; Tyr, Ile, Gly 2.6 mol% each.

Antibacterial spectra for PR-39

Purified PR-39 was assayed for antibacterial activity against six Gram-negative and four Gram-positive strains of bacteria. Table 3 shows that lethal concentrations of 0.3 μM were obtained, both with *E. coli* K12 and a pig pathogenic strain of the same bacterium. This is also the activity previously found for porcine cecropin P1 (Lee et al., 1989) and for the insect cecropins A and B (Boman and Hultmark, 1987). Of the other Gram-negative organisms *Salmonella typhimurium* and *Acinetobacter calcoaceticus* were also quite sensitive to PR-39, while *Pseudomonas aeruginosa* and *Proteus vulgaris* were resistant. Two Gram-positive bacteria, *Bacillus megaterium* and *Streptococcus pyogenes*, were both sensitive, while *Bacillus subtilis* was somewhat sensitive and *Staphylococcus aureus* quite resistant. Our results are in good agreement with the data for two bovine proline-arginine-rich peptides from neutrophils, which acted on *E. coli* and *S. typhimurium*, while *P. vulgaris* and *Ps. aeruginosa* were found to be resistant (Gennaro et al., 1989).

Fig. 4. Reverse-phase HPLC separation of tryptic fragments obtained from PR-39. Elution was by a 90-min gradient of acetonitrile (0 − 60%; 1 ml/min). Both solvents contained 0.1% trifluoroacetic acid.

difficult to detect on HPLC. We established the C-terminal residues by amino acid analysis after digestion with carboxypeptidase Y. Fig. 6 shows the separation and Table 2 the amounts of four liberated components, prolinamide (Pro-NH₂), Pro, Phe and Arg, as detected by capillary zone electrophoresis. This method was found to be useful, rapid and provided the necessary high sensitivity (Bergman et al., 1991). The results with carboxypeptidase Y show that a non-stoichiometric deamidation of a C-terminal Pro-NH₂ can occur, explaining the release of both Pro and Pro-NH₂ in amounts equimolar to those of the penultimate Phe. Consequently, residues 38 and 39 are concluded to be Phe and Pro-NH₂, respectively, and Arg to occupy position 37. This structure corresponds to a molecular mass of 4719.7 Da, a value in good agreement with 4720 Da and 4718 Da obtained by two different plasma desorption mass spectrometer instruments. Methylation of PR-39 did not change in molecular mass (which is consistent with a blocked C-terminus) while a parallel methylation experiment with secretin produced the expected mass increase.

DISCUSSION

The determination of the primary structure of PR-39 was complicated, largely because of the monotonous sequence and the limited sensitivity of the peptide to enzymic degradation. The C-terminal portion and the C-terminal amide were especially difficult to establish and the final remedy was capillary zone electrophoresis for direct detection of liberated Pro-NH₂ (Bergman et al., 1991). Notably, difficulties have been reported also in the analysis of two proline-arginine-rich antibacterial peptides from bovine neutrophils (Frank et al., 1990). Other proline-rich proteins are often structural, as in the case of collagen or the secretory protein corresponding to the Balbiani rings of *Chironomus* (Wieslander et al., 1984). Proline-rich proteins have also been isolated from parotid saliva (Kauffman et al., 1991), but their function is not yet known. A common feature of most proline-rich proteins is the presence of repeated motifs.

Many of the antibacterial peptides so far isolated fall into one of four chemically distinct groups, the cecropins, the defensins, the magainins and the proline-rich peptides. The cecropins may be considered as two-helix molecules (Holak et al., 1988), the magainins as one-helix compounds (Marion et al., 1988), both being devoid of cysteine. The defensins on the other hand have three intramolecular disulfide bridges in 29 − 34 residues (Selsted and Harwig, 1989) and have a crystal structure with dimeric β-sheets (Hill et al., 1991).

Several proline-rich antibacterial peptides have recently been characterized. One group, isolated from honey bees,

Table 2. Digestion of PR-39 with carboxypeptidase Y. Amino acids liberated during the reaction were analyzed by capillary zone electrophoresis (Bergman et al., 1991). The amino acids were detected underivatized and differ greatly in absorbance at 200 nm; Pro-NH₂ and Pro are therefore more difficult to quantify at very low amounts, especially in relation to Phe, explaining the uncertain and low Pro and Pro-NH₂ estimates at early times (20 s and 5 min).

Amino acid	Amount liberated after incubation for					
	20 s	5 min	15 min	30 min	1 h	2 H
	pmol					
ProNH₂	<1	<1	2	3	6	8
Pro	<5	<5	21	28	42	42
Phe	7	14	20	27	49	48
Arg	5	9	14	28	33	48
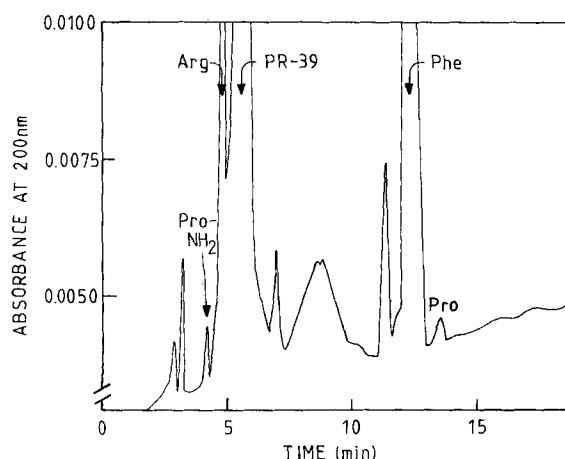
Fig. 6. Analysis of C-terminal residues in PR-39 by capillary zone electrophoresis after digestion with carboxypeptidase Y. The liberated amino acids were underivatized and differed greatly in absorbance at 200 nm. The amounts liberated at different times are given in Table 2. The two first peaks and those at 7—9 min are artifacts present in controls.

Table 3. Inhibition zone assay of PR-39 with nine different bacteria. The maximum concentration of PR-39 was 2.5 mM; 3 μl was applied to each well (3 mm in diameter) in thin agarose plates seeded with test bacterium. Lethal concentrations, calculated as described (Hultmark et al., 1983), were obtained from dilution series as the lowest concentration killing the respective bacteria. *E. coli* Bd2221/75 is a clinical isolate, pathogenic for piglets (obtained from Olof Söderlund, SVA, Uppsala).

Organism	Strain	Maximum	Lethal conc.
		mm	μM
E. coli K12	D21	18.3	0.3
E. coli	Bd2221/75	17.6	0.3
S. typhimurium	LT2	15.2	1
P. vulgaris	Pv11	3.5	300
Ps. aeruginosa	OT97	4.5	200
A. calcoaceticus	Ac11	12.5	3
B. megaterium	Bm11	17.4	0.3
B. subtilis	Bs11	10.3	15
Staph. aureus	Cowan I	3.7	200
Strep. pyogenes	—	14.9	2

contains about 29 mol% Pro and only about 11 mol% basic residues (Casteels et al., 1989, 1990). A second group, isolated from bovine neutrophils (Gennaro et al., 1989), contains as much as 45 and 47 mol% Pro and Arg (23 and 28 mol%) as the only basic residue. The structure of these two peptides (Frank et al., 1990) and our results in Fig. 5 show that the members of this group of peptides have very similar and qualitatively related amino acid compositions. We therefore suggest that they should be called PR peptides with a number indicating the number of residues (in case of an equal number of residues a letter will have to be added in the future). For comparison, Fig. 7 gives the amino acid sequences of our PR-39 and the two bovine peptides, Bac5 and Bac7, which may now be named as PR-42 and PR-59. There is no obvious overall homology among the three peptides, and they differ with respect to minor components. In the larger of the bovine peptides (PR-59) the motif Pro-Arg-Pro is repeated 12 times, while it is absent in the shorter peptide (PR-42). In PR-39 this motif is present twice. Furthermore, the smaller bovine PR-42 contains 18 of 19 Pro residues as Pro-Pro repeats, while in the porcine PR-39, 14 of 19 Pro residues are in such repeats. In the larger of the bovine peptides, there is only one pair of Pro-Pro repeats. Thus, from one aspect (the P-R-P motif), our peptide is related to the larger of the bovine peptides; in the

```
PR-39:        RRRPRPPYLP  RPRPPPFFPP  RLPPRIPPGF  PPRFPPRFP-NH2

Bac5 (PR-42): RFRPPIRRPP  IRPPFYPPFR  PPIRPPIFPP  IRPPFRPPLR  FP

Bac7 (PR-59): RRIRPRPPRL  PRPRPRPLPF  PRPGPRPIPR  PLPFPRPGPR

              PIPRPLPFPR  PGPRPIPRP
```

Fig. 7. Comparison of PR-39 (from Fig. 5) and two peptides, Bac5 and Bac7, from Frank et al. (1990). The sequence of Bac7 is continued in the last line.

other (the P-P motif) to the smaller of the bovine peptides. We believe that the repeats may be a chemical necessity for the functional conformation of these peptides.

Solution NMR has been used to determine the helical structures both of cecropins (Holak et al., 1988) and magainins (Marion et al., 1988). It would also be interesting to study the three-dimensional structure of proline-arginine-rich peptides since their extreme and repetitive compositions are likely to produce special conformational properties. Moreover, solid-state NMR spectroscopy was recently used to study the orientation of two amphipathic peptides in a lipid bilayer (Bechinger et al., 1991). Magainin, an antibacterial frog

peptide, was found to lie in the plane of the bilayer, while a helical segment of the nicotinic acetylcholine receptor was found to span the membrane, perpendicular to the bilayer. It would be interesting if this methodology can be applied also to the proline-arginine-rich peptides with antibacterial activity.

REFERENCES

Bechinger, B., Kim, Y., Chirlian, L. E., Gesell, J., Neuman, J.-M., Montal, M., Tomich, J., Zasloff, M. & Opella, S. J. (1991) *J. Biomol. NMR 1*, 167 – 174.

Bergman, T., Agerberth, B. & Jörnvall, H. (1991) *FEBS Lett. 283*, 100 – 103.

Boman, H. G. (1991) *Cell 65*, 205 – 207.

Boman, H. G. & Hultmark, D. (1987) *Annu. Rev. Microbiol. 41*, 103 – 126.

Casteels, P., Ampe, C., Jacobs, F., Vaeck, M. & Tempst, P. (1989) *EMBO J. 8*, 2387 – 2391.

Casteels, P., Ampe C., Riviere, L., Van Damme, J., Elicone, C., Fleming, M., Jacobs, F. & Tempst, P. (1990) *Eur. J. Biochem. 187*, 381 – 386.

Chen, Z.-w., Agerberth, B., Gell, K., Anderson, M., Mutt, V., Östenson, C.-G., Efendic, S., Barros-Söderling, J., Persson, B. & Jörnvall, H. (1988) *Eur. J. Biochem. 174*, 239 – 245.

Frank, R. W., Gennaro, R., Schneider, K., Przybylski, M. & Romeo, D. (1990) *J. Biol. Chem. 265*, 18871 – 18874.

Gennaro, R., Skerlavaj, B. & Romeo, D. (1989) *Infect. Immun. 57*, 3142 – 3146.

Hill, C. P., Yee, J., Selsted, M. E. & Eisenberg, D. (1991) *Science 251*, 1481 – 1485.

Holak, T. A., Engström, Å., Kraulis, P. J., Lindeberg, G., Bennich, H., Jones, T. A., Gronenborn, A. M. & Clore, G. M. (1988) *Biochemistry 27*, 7620 – 7629.

Hultmark, D., Engström, Å., Andersson, K., Steiner, H., Bennich, H. & Boman, H. G. (1983) *EMBO J. 2*, 571 – 576.

Kaiser, R., Holmquist, B., Hempel, J., Vallee, B. L. & Jörnvall, H. (1988) *Biochemistry 27*, 1132 – 1140.

Kauffman, D. L., Bennick, A., Blum, M. & Keller, P. J. (1991) *Biochemistry 30*, 3351 – 3356.

Lee, J.-Y., Boman, A., Sun, C., Andersson, M., Jörnvall, H., Mutt, V. & Boman, H. G. (1989) *Proc. Natl Acad. Sci. USA 86*, 9159 – 9162.

Lehrer, R. I. & Ganz, T. (1990) *Blood 76*, 2169 – 2181.

Lehrer, R. I., Ganz, T. & Selsted, M. E. (1991) *Cell 64*, 229 – 230.

Marion, D., Zasloff, M. & Bax, A. (1988) *FEBS Lett. 227*, 21 – 26.

McDonald, T. J., Jörnvall, H., Tatemoto, K. & Mutt, V. (1983) *FEBS Lett. 156*, 349 – 356.

Mutt, V. (1959) *Ark. Kemi 15*, 69 – 74.

Mutt, V. (1978) in *Gut hormones* (Bloom, S. R., ed.), pp. 21 – 27, Churchill Livingstone, Edinburgh.

Mutt, V. (1986) *Chem. Scr. 26 B*, 191 – 207.

Said, I. S. & Mutt, V. (1972) *Eur. J. Biochem. 28*, 199 – 204.

Selsted, M. E. & Harwig, S. (1989) *J. Biol. Chem. 264*, 4003 – 4007.

Wieslander, L., Höög, C., Höög, J.-O., Jörnvall, H., Lendahl, U. & Daneholt, B. (1984) *J. Mol. Evol. 20*, 304 – 312.

Zasloff, M. (1987) *Proc. Natl Acad. Sci. USA 84*, 5449 – 5453.

Ørskov, C., Bersani, M., Johnsen, A. H., Højrup, P. & Holst, J. J. (1989) *J. Biol. Chem. 264*, 12826 – 12829.