Prophenin-1, an exceptionally proline-rich antimicrobial peptide from porcine leukocytes

Sylvia S.L. Harwig[a], Vladimir N. Kokryakov[b], Kristine M. Swiderek[c], Galina M. Aleshina[b], Chengquan Zhao[a], Robert I. Lehrer[a,*]

[a]*Department of Medicine, UCLA School of Medicine, Los Angeles CA 90095, USA*
[b]*Department of General Pathology and Pathophysiology, Institute for Experimental Medicine, 19736 St. Petersburg, Russian Federation*
[c]*Division of Immunology, Beckman Research Institute, City of Hope, Duarte CA 91010, USA*

Received 10 February 1995

Abstract We purified and characterized an unusual antimicrobial peptide, prophenin-1 (PF-1), from porcine leukocytes. The peptide had a mass of 8,683 and contained 79 residues, including 42 (53.2%) prolines and 15 (19.0%) phenylalanines. Its N-terminal 60 residues consisted of three perfect and three nearly perfect repeats of a decamer, FPPPNFPGPR. Prophenin-1 was encoded on a cathelin-containing precursor and showed substantially more activity against *E. coli*, a Gram-negative bacterium, than against *Listeria monocytogenes*, a Gram-positive organism, in vitro.

Key words: Antimicrobial; Cathelin; Peptide; Porcine leukocyte; Proline-rich

1. Introduction

Endogenous antimicrobial peptides play important roles in host defense against microbial infection [1]. Although these peptides show marked diversity in structure and antimicrobial spectrum, the precursors of many antibiotic peptides of porcine [2–6], bovine [7–10] and rabbit [11,12] origin contain a highly conserved preproregion that is homologous to cathelin, a putative cysteine-proteinase inhibitor originally isolated from pig leukocytes [13,14]. Several of these cathelin-associated peptides have been isolated and characterized, including porcine PR-39 [15] and protegrins PG-1, 2 and 3 [16]; bovine Bac 5 [17,18], indolicidin 9 [19] and cyclic dodecapeptide [20]; and rabbit p15 [12] and CAP 18 [11]. Others – such as porcine PMAP-36 [5], PMAP-23 [6] and PG-4 – have been prepared for study by chemical synthesis based on sequence translations from cloned cDNA.

We recently purified a remarkable proline and phenylalanine-rich antimicrobial peptide, 'prophenin-1', from porcine leukocytes. Its characterization, primary sequence, antimicrobial activity and relationship to the constellation of cathelin-associated antibiotics are described below.

2. Materials and methods

2.1. Preparation of acid extracts of porcine leukocytes and size exclusion chromatography

An acid extract of porcine leukocytes was prepared and ultrafiltered through a YM-5 filter as previously described [16]. The filtrate was concentrated and subjected to size exclusion chromatography on a 2.5 × 117 cm BioGel P-10 column. Prophenin-1 (PF-1) and several closely related molecules ('Prophenin-2') emerged at an elution volumn of ~250–284 ml (43.6–49.4% of column bed volume) and were concentrated for subsequent purification.

2.2. Purification and antimicrobial testing of prophenins

Concentrated prophenin containing fractions were subjected to reversed phase (RP)-HPLC purification on a 4.6 × 250 mm Vydac C18 column using a linear gradient of acetonitrile that contained 0.1% trifluoroacetic acid and increased in acetonitrile by 1% min^{-1}. HPLC fractions were tested for antimicrobial activities against *E. coli* ML-35, *L. monocytogenes*, strain EGD and *C. albicans* strain 820 by previously described gel overlay and radial diffusion techniques [21]. Further purification involved preparative continuous acid urea gel electrophoresis elution [22] followed by RP-HPLC.

2.3. Biochemical analysis

Tricine-SDS-PAGE [23] and AU-PAGE [24] were performed in minigel formats. Amino acid composition was determined by the Picotag technique [25]. Approximately 1 nmol of each prophenin was sequenced (Porton 2090 sequencer). Peptide concentration was determined by amino acid analysis, bicinchoninic acid [26] 'BCA' protein assay reagent (Pierce Chemical Co., Rockford, Il), and by the spectrophotometric method (A_{215}/A_{225}) of Waddell [27].

Approximately 2 nmol of Prophenin-1 was digested with TPCK trypsin (enzyme/substrate mass ratio of 1/50) in 20 μl of 0.1 M NH_4HCO_3, pH 8.15 at 37°C for 4 hours. Digestion was quenched by addition of glacial HOAc (final 25% v/v). After removing an aliquot (50 pmol) for LC-Electrospray Mass Spectrometric (LC-ESI-MS) measurement, the digest was dried, and redissolved in 0.1% TFA for RP-HPLC purification. The amino acid composition of each peak was determined, and the peak that corresponded to C-terminal fragment was also sequenced. Similar digestion conditions were used later to generate large amount of tryptic fragments for biological functional studies.

2.4. LC-ESI-MS

LC-mass spectrometric measurements were performed at the mass spectrometry core facility at the Beckman Research Institute of the Coty of Hope (Duarte, CA) on a quadrupole TSQ-700 mass spectrometer (Finnigan-MAT, San Jose, CA) equipped with an electrospray ion source operating at atmospheric pressure, as reported previously [28]. Mass scans were acquired every three seconds in a mass range from 500–2,000 and the data collection was monitored using both base peak and the reconstructed ion current profile. Spectra were generated by averaging the scans containing the peak, and the mass assignments were made using the Finnigan MAT BIOMASS data reduction software.

3. Results

3.1. Purification of prophenins

Prophenins were eluted at 48% acetonitrile on RP-HPLC, and were shown to contain at least two major components on AU-PAGE. The more cationic band was named 'Prophenin-1', the other was named 'Prophenin-2'. They were subsequently purified by preparative continuous acid urea PAGE followed

*Corresponding author. Department of Medicine, UCLA Center for the Health Sciences, 10833 LeConte Avenue, Los Angeles, CA 90095-1678, USA. Fax: (1) (310) 206 8766.

SSDI 0014-5793(95)00210-3

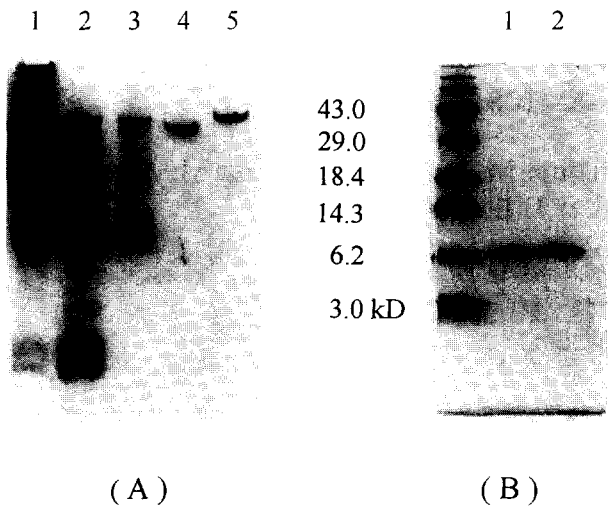

Fig. 1. Polyacrylamide gel electrophoresis of prophenins. (A) Coomassie blue stained AU-PAGE (12.5%): lane 1, ~30 μg of porcine leukocyte acid extract; lane 2, YM-5 filtrate, ~15 μg; lane 3, pooled P-10 column fractions that contained prophenins, ~5 μg; lane 4, prophenin-1, 1 μg; lane 5, prophenin-2, 1 μg. (B) Silver stained tricine SDS-PAGE (16.5%): lane 1, prophenin-1, 1 μg; lane 2, prophenin-2, 1 μg.

by desalting on C18 RP-HPLC, and appeared homogeneous by acid urea (AU)-PAGE on 12.5% acrylamide gels and by SDS-PAGE on 16.5% acrylamide gels (Fig. 1).

3.2. Biochemical characterization

The name *prophenin* was selected to reflect the abundance of *pro*line (57.1%) and *phen*ylalanine (19.0%) residues found by amino acid analysis of the purified peptide, as shown in Table 1. Primary sequence analysis of 1 nmol of Prophenin-1 was conducted with an initial yield of 57.6%, and a repetitive yield

Decamer	Primary Sequence and Residue Number									
I	A_1	F	P	P	P_5	N	V	P	G	P_{10}	R
II		F	P	P	P_{15}	N	F	P	G	P_{20}	R
III		F	P	P	P_{25}	N	F	P	G	P_{30}	R
IV		F	P	P	P_{35}	N	F	P	G	P_{40}	R
V		F	P	P	P_{45}	N	F	P	G	P_{50}	P
VI		F	P	P	P_{55}	I	F	P	G	P_{60}	W
VII		F	P	P	P_{65}	P	P	F	R	P_{70}	P
		P	F	G	P_{75}	P	R	F_{78}	P_{79}		

Fig. 2. Amino acid sequence of prophenin-1. The mature peptide sequence was determined by a combination of automated Edman degradation, tryptic mapping and direct LC-ESI-MS analysis of the tryptic digest. Res. 78 and 79 were assigned based on mass difference. The sequence was confirmed by cDNA cloning ([30], Zhao et al., unpublished). The conserved residues in the polydecameric portion of the molecule are shaded. The trypsin cleavage sites are indicated by double underlining.

of 95.98%. Remarkably, this provided unambiguous identification of all but 3 of the first 65 residues (residues 51, 58 and 61) shown in Fig. 2. A related peptide, 'Prophenin-2', showed slightly slower migration on AU-PAGE gels (Fig. 1) and had a similar composition on amino acid analysis (data not shown). However, since at least two residues were seen at each sequence cycle, we failed to obtain a definitive sequence for this peptide mixture.

The tryptic digest of Prophenin-1 eluted as three broad peaks (Fig. 3), which we designated as Fragments 1, 2 and 3 in the order of their elution from the column. Amino acid composition analysis (Table 1) indicated that fragment 1 represented residues 1–11, fragment 2 represented residues 12–21. The integrated area of fragment 2 was ~3 times greater that of fragment 1 (results not shown), consistent with the initial sequence data shown in Fig. 2 which demonstrates 3 exact repeats of this motif between residues 12 and 41. Fragment 3, which contained an Ile residue and two additional Arg residues, therefore repre-

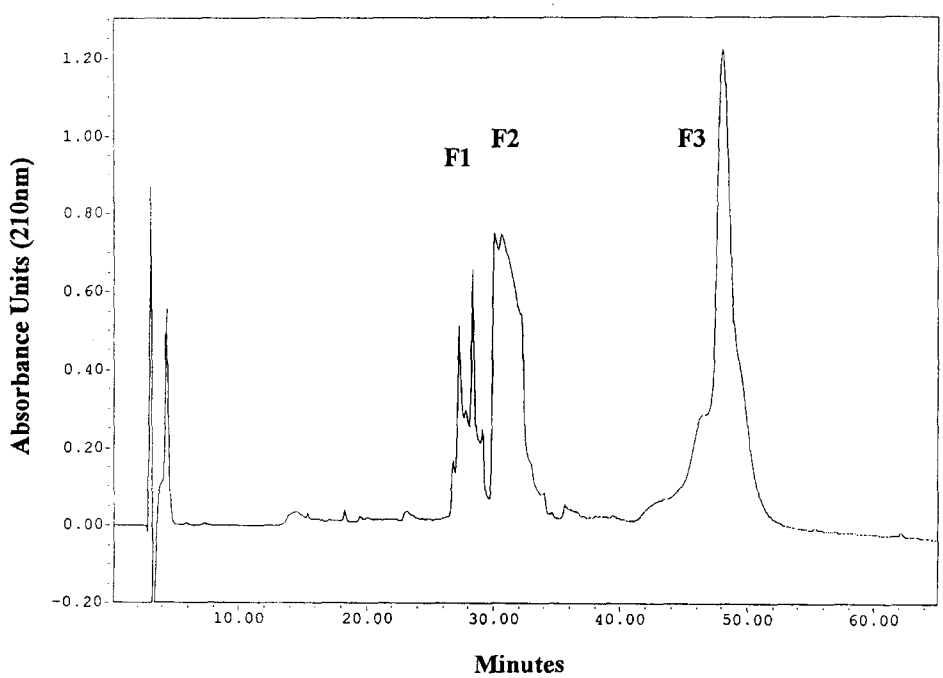

Fig. 3. RP-HPLC purification of prophenin-1 tryptic digest. Approximately 10 nmol of the digest was purified on a 4.6 × 250 mm Vydac C18 column using a linear gradient of acetonitrile that contained 0.1% TFA and increased in acetonitrile concentration by 1% per min.

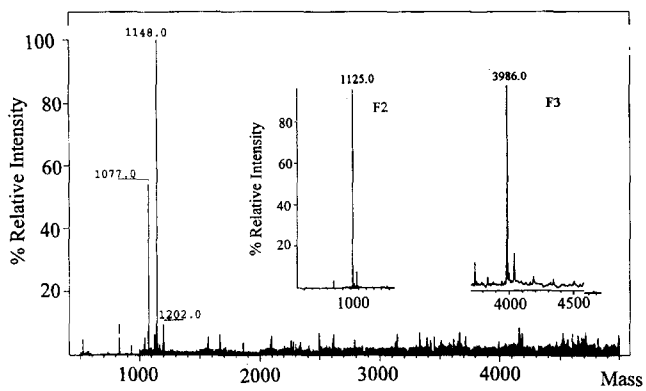

Fig. 4. Deconvoluted mass spectra of prophenin-1 tryptic fragments obtained during direct LC-ESI-MS analysis of ~30 pmol of tryptic digest. The main figure shows tryptic fragment 1 (F1). The insets show tryptic fragments 2 (F2) and 3 (F3).

sented residue 42 to nearby the C terminus. Microsequencing of Fragment 3 provided sequence data through residue 76, showing the presence of Arg at residue 69 followed by a proline. Another Arg was assigned at residue 77 based on the amino acid composition of this fragment, and was confirmed later by mass measurements of this fragment.

3.3. LC-ESI-MS analyses

LC-ESI-MS measurement yielded an average mass value of 8,683.0 for Prophenin-1. When 'Prophenin-2' was similarly examined (data not shown), we discerned the presence of at least two closely related peptides with average mass values of 8,807.0 and 8,794.0.

We subjected 30 pmol of Prophenin-1 tryptic digest to LC-ESI-MS analysis, and obtained three broad peaks in both the base peak and the reconstructed ion current profiles of the mass scan (data not shown). The reconstructed ion current profile appeared almost identical to RP-HPLC chromatogram (Fig. 3) during purification of the digestion mixture. The deconvoluted mass spectra for each tryptic fragment are shown in Fig. 4. Fragment 1, 2, and 3 had average mass values of 1,148.0, 1,125.0 and 3,986.0, respectively. These values agreed with the sequence assignments up to residue 77. In addition to 1,148.0, a mass value of 1,077.0 was also seen in the deconvoluted mass spectrum of fragment 1. The difference (71 mass units) was most likely due to the presence of *des*-ala PF-1, resulting from enzymatic (elastase-mediated?) cleavage after the peptide's solitary alanine residue at position 1. This was also reflected by the slightly lower value of alanine in the amino acid composition of this fragment (Table 1).

The calculated average mass for res. 1–77 is 8,437.01, 245.99 mass units smaller than measured mass (8,683.0) of Prophenin-1. This mass difference strongly suggests another Pro and Phe on the C-terminus of this peptide. This dipeptide, generated during trypsin digestion, would not have been easily recovered during RP-HPLC purification of the tryptic digest. The order Phe^{78}-Pro^{79} was assigned based on the successful cleavage after Arg^{77} by trypsin, since proteolytic enzymes are often cryptic if the specific cleavage sites are immediately followed by a proline residue.

3.4. Quantitation of prophenin-1

PF-1 showed anomalous behavior in two commonly used methods for measuring protein concentration. When tested in the BCA assay, it failed to generate any color development. The

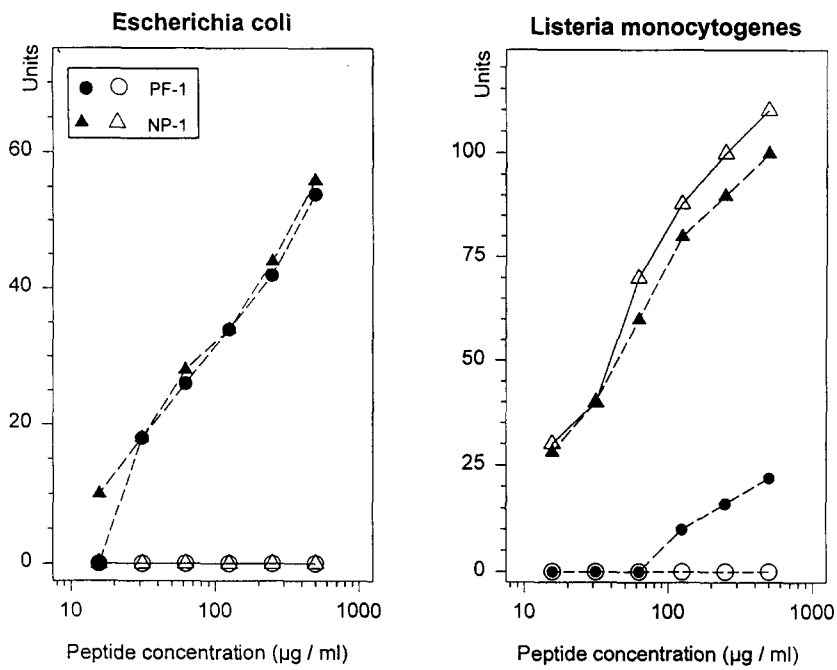

Fig. 5. Antimicrobial activity of purified prophenin-1. Purified prophenin-1 and rabbit defensin NP-1 were tested against *E. coli* ML-35p and *L. monocytogenes* EGD under low (solid symbols) and high (open symbols) ionic strength conditions by radial diffusion assay. Low ionic strength conditions were obtained by preparing underlay gels that contained 1% agarose+0.3 mg/ml trypticase soy broth powder in 10 mM sodium phosphate buffer (pH 7.4), and high ionic strength conditions were obtained by supplementing this underlay agar with 0.1 M NaCl. Concentrations of both peptides were determined by amino acid analysis.

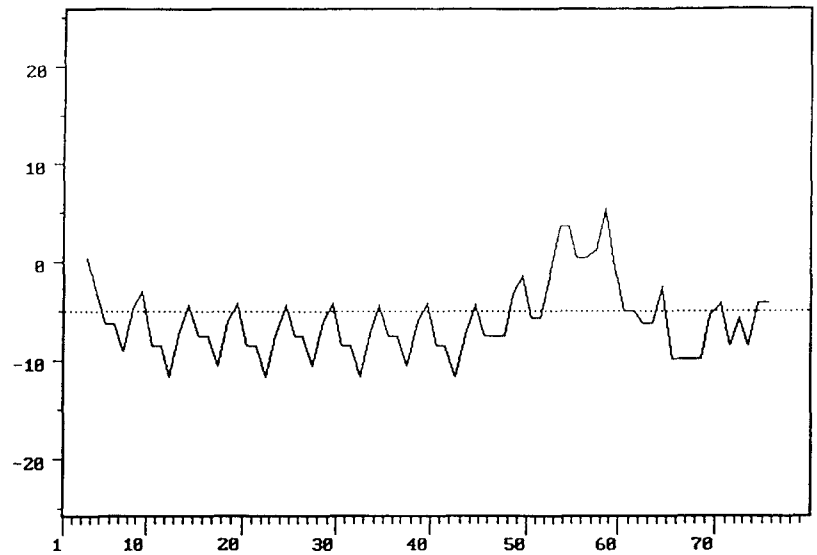

Fig. 6. Hydropathic index of prophenin-1. Computation was done using an interval of 7 amino acids. Note the unusual repetetive pattern of the molecule's N-terminal half.

spectrophotometric method of A_{215}/A_{225} consistently gave values that were ~4.5× higher than the values we obtained by amino acid analysis, using isoleucine as an internal standard.

3.5. Antimicrobial activities

PF-1 demonstrated potent bactericidal activity against *Escherichia coli* ML-35, but showed little activity against *L. monocytogenes* and was inactive against *C. albicans*. As shown in Fig. 5, its activity against *E. coli* resembled that of NP-1, a potent defensin [29] produced by rabbit leukocytes. Prophenin tryptic fragments F1 and F2 were inactive against *E. coli* at concentrations up to 500 μg/ml while F3 retained partial microbicidal activity at >125 μg/ml (data not shown).

4. Discussion

We were able to establish the first 77 residues sequence of the mature prophenin-1 by a combination of biochemical techniques: direct microsequencing, tryptic peptide mapping and direct LC-ESI-MS analysis of the tryptic digest. The last two C-terminal residues, Phe and Pro, were assigned based on the mass difference, and was confirmed by the deduced amino acid sequence of clone 12 which we have independently confirmed by PCR cloning of bone marrow and genomic DNA (data not shown).

Pungercar et al. [30] recently obtained from porcine bone marrow and sequenced two cDNA clones, C12 and C6, that encoded virtually identical proline and arginine rich antimicrobial peptides at the C-terminus of a cathelin-containing prepropeptide. The translated sequence of clone C12 included a 30 residue, N-terminal signal peptide followed by a 198 residue protein whose N-terminal region was 84% identical to porcine cathelin in primary sequence. The C-terminal half was noted to be rich in proline, arginine and phenylalanine residues and to show similarities to a previously isolated Pro-Arg rich group of antimicrobial peptides, including porcine PR-39 and bovine bactenecins, Bac 5 and Bac 7. The prophenin-1 peptide sequence reported in this paper is nearly identical to the C-terminal 79 residue stretch near the C-terminus of clone 12, except for the Phe for Val substitution at residue 17 of the mature peptide which corresponds to the residue present in their incompletely sequenced clone 6.

The sequence of mature prophenin-1 indicates that its precursor undergoes cleavage after the glutamine residue (res. # 144 as per Pungercar) rather than only at the # 126 valine residue postulated by Pungercar et al., which otherwise would

Table 1
Amino acid composition of Prophenin-1 (PF-1) and its tryptic fragments

Amino acid	PF-1 mol % obs. (seq.)	PF-1 Residues obs. (seq.)	Fragment 1 Residues obs. (seq.)	Fragment 2 Residues obs. (seq.)	Fragment 3 Residues obs. (seq.)
Asx	5.9 (6.3)	4.6 (5)	0.9 (1)	1.0 (1)	1.4 (1)
Glx	0.0 (0.0)	0.0 (0)	0.1 (0)	0.0 (0)	0.3 (0)
Ser	0.0 (0.0)	0.0 (0)	0.1 (0)	0.0 (0)	0.3 (0)
Gly	9.4 (8.9)	7.3 (7)	1.1 (1)	1.1 (1)	3.5 (3)
Arg	9.0 (7.6)	7.0 (6)	1.1 (1)	1.1 (1)	2.2 (2)
Ala	1.1 (1.3)	0.9 (1)	0.6 (1)	0.0 (0)	0.3 (0)
Pro	57.1 (53.2)	44.5 (42)	5.0 (5)	4.9 (5)	20.3 (21)
Val	1.4 (1.3)	1.1 (1)	0.9 (1)	0.0 (0)	0.0 (0)
Ile	0.9 (1.3)	0.7 (1)	0.0 (0)	0.0 (0)	0.8 (1)
Phe	15.2 (19.0)	11.9 (15)	1.2 (1)	1.9 (2)	5.9 (7)
Trp	N.D. (1.3)	N.D. (1)	N.D. (0)	N.D. (0)	N.D. (1)

result in a 100 amino-acid mature peptide, instead of the 79 residue peptide described herein. Pungercar's prediction was based on the observation that post-translational processing to the final mature form of other cathelin associated antimicrobial peptides, such as bovine indolicidin [7] and Bac 5 [9], often involved cleavage after the valine residue. It remains to be determined if post-translational processing of the PF-1 precursor results in the liberation of an intact arginine and tryptophan-rich 18-mer (RRFPWWWPFLRRPRLRRQ), or if these residues are released by piecemeal hydrolytic steps. The other difference between the mature PF-1 peptide and the structure predicted by Pungercar et al is found at the C-terminus, where the last three residues (GRR) are removed. The presence of Gly-X-X (X = basic amino acid) sequence at the carboxyl ends of many precursors to amidated peptides has been identified as the universal combined amidation/proteolytic cleavage recognition site [31]. Therefore, prophenin-1 most likely also exists as the amidated form, as do several other cathelin-encoded peptides of porcine (protegrin and PR-39) and bovine leukocytes (indolicidin).

The prophenin-1 peptide sequence is remarkable for the 5 nearly perfect tandem repeats of a proline-rich decamer, FPPPNFPGPR. The fidelity of this repetitive motif is broken only by the conservative substitution of Val for Phe in decamer 1 and by the substitution of Pro^{51} for Arg^{51} at the end of the fifth decamer. The latter substitution marks the beginning of the peptide's only relatively hydrophobic domain, evident in the hydropathy plot shown in Fig. 6. The functional significance of the proline-rich tandem repeats and the hydrophobic domain are currently under investigation, and will be reported in a later communication.

Acknowledgements: The authors are indebted to Dr. Audree Fowler of the UCLA protein sequencing core facility for expertly obtaining nearly all of the complete sequence information of such a challenging peptide. We also thank Leonie Tan and Olga V. Shamova for their assistance in the purification of these peptides, and Elena Panyutich and Yoon Cho for performing the antimicrobial assays. This work was supported, in part, by Research Grants, AI 22839 and CA 33572 from the National Instiitutes of Health.

References

[1] Boman, H.G. (1991) Cell 65, 205–207.
[2] Storici, P. and Zanetti, M. (1993) Biochem. Biophys. Res. Commun. 196, 1363–1368.
[3] Storici, P. and Zanetti, M. (1993) Biochem. Biophys. Res. Commun. 196, 1058–1065.
[4] Zhao, C., Liu, L. and Lehrer, R.I. (1994) FEBS Lett. 346, 285–288.
[5] Storici, P., Scocchi, M., Tossi, A., Gennaro, R. and Zanetti, M. (1994) FEBS Lett. 337, 303–307.
[6] Zanetti, M., Storici, P., Tossi, A., Scocchi, M and Gennaro, R. (1994) J. Biol. Chem. 269, 7855–7858
[7] Del Sal, G., Storici, P., Schneider, C., Romeo, D. and Zanetti, M. (1992) Biochem. Biophys. Res. Commun. 187, 467–472.
[8] Storici, P., Del Sal, G., Schneider, C. and Zanetti, M. (1992) FEBS Lett. 314, 187–190.
[9] Zanetti, M., Del Sal, G., Storici, P., Schneider, C. and Romeo, D. (1993) J. Biol. Chem. 268, 522–526.
[10] Scocchi, M., Romeo, D. and Zanetti, M. (1994) FEBS Lett. 352, 197–200.
[11] Larrick, J. W., Morgan, J. G., Palings, I., Hirata, M. and Yen, M. H. (1991) Biochem. Biophys. Res. Commun. 179, 170–175.
[12] Levy, O., Weiss, J., Zarember, K., Ooi, C.E. and Elsbach, P. (1993) J. Biol. Chem. 268, 6058–6063.
[13] Kopitar, M., Ritonja, A., Popovic, T., Gabrijelic, D., Krizaj, I. and Turk, V. (1989) Biol. Chem. Hoppe-Seyler 370, 1145–1151.
[14] Ritonja, A., Kopitar, M., Jerala, R., and Turk, V. (1989) FEBS Lett. 255, 211–214.
[15] Agerberth, B., Lee, J., Bergman, T., Carlquist, M., Boman, H., Mutt, V. and Jörnvall, H. (1991) Eur. J. Biochem. 202, 849–854.
[16] Kokryakov, V.N., Harwig, S.S.L., Panyutich, E.A., Shevchenko, A.A., Aleshina, G.M., Shamova, O.V., Korneva, H.A. and Lehrer, R.I. (1993) FEBS Lett. 327, 231–236.
[17] Gennaro, R., Skerlavaj, B. and Romeo, D. (1989) Infect. Immunol. 57, 3142–3146.
[18] Frank, R.W., Gennaro, R., Schneider, K., Przybylski, M., and Romeo, D. (1990) J. Biol. Chem. 265, 18871–18874.
[19] Selsted, M.E., Novotny, M.J., Morris, W.L., Tang, Y.Q., Smith, W. and Cullor, J.S. (1992) J. Biol. Chem. 267, 4292–4295.
[20] Romeo, D., Skerlavaj, B., Bolognesi, M. and Gennaro, R. (1988) J. Biol. Chem. 263, 9573–9575.
[21] Lehrer, R.I., Rosenman, M., Harwig, S.S.L., Jackson, R. and Eisenhauer, P.B. (1991) J. Immunol. Methods 137, 167–173.
[22] Harwig, S.S.L., Chen, N.P., Park, A.S.K. and Lehrer, R.I. (1993) Anal. Biochem. 208, 382–386.
[23] Schagger, H., and von Jagow, G. (1987) Anal. Biochem. 166, 368–379.
[24] Panyim, S. and Chalkley, R. (1969) Arch. Biochem. Biophys. 130, 337–346.
[25] Bidlingmeyer, R.A., Cohen, S.A. and Tarvin, T.L. (1984) J. Chromatogr. 336, 93–104.
[26] Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J. and Klenk, D.C. (1985) Anal. Biochem. 150, 76–85.
[27] Waddell, W.J. (1956) J. Lab. Clin. Med. 48, 311–314.
[28] Harwig, S.S.L., Swiderek, K.M., Kokryakov, V., Lee, T.D. and Lehrer, R.I. (1994) in: Techniques in Protein Chemistry V, (Crabb, J., Ed.) pp. 81–88.
[29] Lehrer, R.I., Lichtenstein, A.K. and Ganz, T. (1993) Annu. Rev. Immunol. 11, 105–128.
[30] Pungercar, J., Strukelj, B., Kopitar, G., Renko, M., Lenarcic, B., Gubensek, F. and Turk, V. (1993) FEBS Lett. 336, 284–288.
[31] Mains, R.E., Eipper, B.A., Glembotski, C.C. and Dores, R.M. (1983) Trends Neurosci. 6, 229–235.