cDNA sequences of three sheep myeloid cathelicidins

Luigi Bagella[a], Marco Scocchi[a], Margherita Zanetti[a,b,]*

[a]*Laboratorio Nazionale Consorzio Interuniversitario Biotecnologie, (CIB), AREA Science Park, Padriciano, I-34012 Trieste, Italy*
[b]*Dipartimento di Scienze e Tecnologie Biomediche, University of Udine, I-3310 Udine, Italy*

Received 18 October 1995

Abstract Several myeloid antimicrobial peptide precursors have been shown to consist of a N-terminal proregion similar to a protein named cathelin and a structurally varied C-terminal antimicrobial domain. Proteins with these features have been named cathelicidins. In this paper we report the cDNA sequences of three ovine cathelicidins of 155, 160 and 190 residues, respectively, with cationic C-terminal sequences corresponding to putative antimicrobial domains. These are structurally varied and include a Cys-rich sequence of 12 residues, which is similar to the bovine antimicrobial cyclic dodecapeptide, a novel 29 residue sequence named SMAP-29 with a possible α-helical conformation, and a 60 residue sequence named Bac7.5, which appears to be a new member of the Pro- and Arg-rich group of mammalian antimicrobial peptides.

Key words: Antimicrobial peptide; Cathelicidin; Myeloid cell; Innate immunity

1. Introduction

Antimicrobial peptides are a widespread means of defense used by animals and plants against pathogens [1]. In general, different animal species are equipped with different arrays of antimicrobial peptides, which cover a broad spectrum of antimicrobial activity [1,2]. In mammals, several such peptides have been isolated from mucosal epithelia and from professional phagocytes [1–5]. They are synthesized as precursors (prepropeptides) from which the mature peptides, which are in general cationic and amphipathic molecules, are released by proteolytic processing.

It has recently been shown that a variety of myeloid antimicrobial peptides are located C-terminal to precursors characterized by a highly conserved pre-proregion [6–19]. The corresponding proforms are stored in the neutrophil granules [20,21]. The C-terminal region of these proforms exhibits antimicrobial activity after proteolytic removal of a propiece [23–29] which is similar to a leukocyte protein named cathelin [22]. The antimicrobial domains show significant diversity in size (12–100 residues), structure and spectrum of antimicrobial activity. Transcripts encoding proteins with these features have been identified in human [18,19,21], bovine [7–9,13], porcine [10–12,14–17] and rabbit [6,30] myeloid cells. Due to the presence of a common cathelin-like domain, these molecules have been named cathelicidins [31]. The antimicrobial peptides derived from cathelicidins include the Cys-rich cyclic dodecapeptide [23] and protegrins [26], the Pro- and Arg-rich Bac5 and Bac7 [24], prophenin [27] and PR-39 [28], the Trp-rich indolicidin [25] and PMAP-23 [15], the α-helical PMAP-36 [14], PMAP-37 [17], CAP18(106–142) [29] and FALL-39/human CAP18 [18–19].

In order to characterize the cathelicidin family further, we investigated the diversity of cathelicidin mRNAs in sheep myeloid cells by selectively amplifying those transcripts containing cathelin-like sequences. We report here the cDNA sequences of three ovine cathelicidins showing structurally varied C-terminal putative antimicrobial domains, which are discussed in relation to congeners identified in other species.

2. Materials and methods

2.1. RNA extraction and Northern analysis

Total RNA was extracted from sheep bone marrow cells with guanidinium thiocyanate [32]. Northern analysis was performed as described [7]. Synthetic antisense oligonucleotides ^{32}P-labelled using standard protocols and derived from the 3′ cDNA sequences of the ovine dodecapeptide (5′-CAGATCCAGTAGCTTGAGGC-3′), SMAP-29 (5′-AGTTGGGCCATACTTCTTCA-3′) and Bac7.5 (5′-AGTGCTAACCTTGATGTT-3′), were used to hybridize Northern blots of sheep bone marrow total RNA.

2.2. cDNA cloning, sequencing and sequence analysis

The experimental conditions to obtain the 3′ and 5′ ends of the ovine dodecapeptide, SMAP-29 and Bac7.5 cDNAs were as described in [11]. For the 3′ end amplification, sheep bone marrow mRNA was first reverse transcribed using the antisense primer adaptor 5′-TCGGATCCCTCGAGAAGC(T)$_{18}$-3′. The 3′ ends of the three cDNAs were then amplified at once by PCR, using the sense oligonucleotide 5′-CGCGAATTCTGTGAGCTTCAGGGTG-3′, derived from the cathelin-like sequence of the cathelicidin cDNAs, and the antisense primer adaptor 5′-TCGGATCCCTCGAGAAGCTT-3′. For each sheep cathelicidin, the 5′ end cDNA amplification was performed by using a sense primer 5′-CAAGAATTCGGAGACTGGGGACCATG-3′ derived from the conserved 5′ untranslated sequence of previously described cathelicidins [7–9], and antisense primers derived from cDNA sequences unique to each sheep cathelicidin, namely: 5′-GACGAATTCGAAAACCCTTAGGACTC-3′ (sheep dodecapeptide), 5′-AATGAATTCCTACCCAGTCTTCGAA-3′ (SMAP-29), and 5′-CGCGGATCCCTTCCCAATGATTATCACA-3′ (Bac7.5). Amplified cDNAs were cloned in Bluescript KS$^+$ vector (Stratagene, San Diego, CA) and sequenced on both strands with deazaguanosine and automated fluorescent DNA sequencing (EMBL fluorescent DNA sequencer, Heidelberg, Germany). For each different cDNA, at least 6 clones generated from different preparations of RNA were completely sequenced in both directions. Sequence analysis and homology searches were carried out using programs from the IG suite version 5.4 (IntelliGenetics Inc.) and from the GCG suite. Secondary structure of SMAP-29 was predicted using the Peptidestructure (GCG) [33], the nnpredict [34] and the PHD network [35] prediction methods.

*Corresponding author. Laboratorio Nazionale CIB, AREA Science Park, Padriciano 99, I-34012 Trieste, Italy. Fax: (39) (40) 398990.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries with the following accession numbers: L46852 (Bac7.5), L46853 (s-dodecapeptide), L46854 (SMAP-29).

0014-5793/95/$9.50 © 1995 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(95)01285-0

3. Results and discussion

3.1. Cloning of sheep cathelicidin cDNAs

We have used a molecular biological approach based on the high conservation of the preproregion of cathelicidins to identify transcripts encoding molecules containing a cathelin-like proregion and a C-terminal putative antimicrobial domain in sheep myeloid cells.

To this aim, sheep bone marrow mRNAs were amplified by PCR, using a sense oligonucleotide primer derived from the conserved prosequence of bovine [7–9] and porcine [10,11,14,15,17] cathelicidins, and an oligo-dT primer, so that the amplified products would include the 3′ sequences corresponding to putative antimicrobial domains. The 3′ ends of three different cDNAs were amplified at once by using this method. Northern analysis of sheep bone marrow RNA using antisense oligonucleotide probes derived from the unique 3′ cDNA sequences revealed two transcipts of approximately 0.6 kb, and one of approximately 0.9 kb (not shown). The 5′ ends of these cDNAs were then amplified by using antisense primers derived from the unique 3′ regions, and a sense oligonucleotide primer based on the conserved 5′ non-coding region of previously identified cathelicidins [7–9].

3.2. Features of the predicted sequences

The three full length cDNA sequences and the deduced amino acid sequences, obtained from the overlapping 5′ and 3′ cDNA ends, are shown in Fig. 1. The predicted polypeptides show 75–81% identity in the pre-prosequences, and display the characteristic features of cathelicidins. These include a putative 29 residue signal peptide and a cathelin-like prosequence of 101–114 residues, including four invariant cysteine residues. A putative proteolytic cleavage site for elastase, commonly observed at the end of a cathelin-like domain, precedes a cationic C-terminal sequence of 12–60 residues.

The 155 residue cathelicidin (Fig. 1A) has a predicted mass of 17649 Da and pI of 7.6. Analysis of the sequence suggests that this is the sheep homologue of the precursor of the bovine cyclic dodecapeptide [8], an antimicrobial peptide which is active against gram-negative and gram-positive bacteria [23]. The cyclic dodecapeptide is the smallest known animal antimicrobial peptide and has so far been identified only in bovine neutrophils. The molecule is amphipathic, with a hydrophilic region formed by the N- and C-terminal regions brought together by a disulfide bond, and a hydrophobic region formed by residues involved in the g turn [23]. The ovine and bovine prepropeptides share 97% overall identity at the protein level, and alignment of the C-terminal antimicrobial domains (Fig. 2A) shows that the four conservative substitutions of hydrophobic residues in the sheep dodecapeptide do not alter the amphipathic character of the molecule.

The 160 residue cathelicidin (Fig. 1B) has a calculated mass of 17813 Da and pI of 10. The 29 residue signal peptide is followed by a 102 residue cathelin-like domain and a 29 residue cationic C-terminal region with a putative amidation signal at the C-terminus [36]. The C-terminal sequence is unique and has been termed SMAP-29 from 'sheep myeloid antimicrobial peptide of 29 residues'. Sequence comparison with the other mammalian antimicrobial peptides indicates that the highest level of identity (34%) is with the sequence of the porcine PMAP-36 [14]. Structure prediction analysis of SMAP-29 indicates that

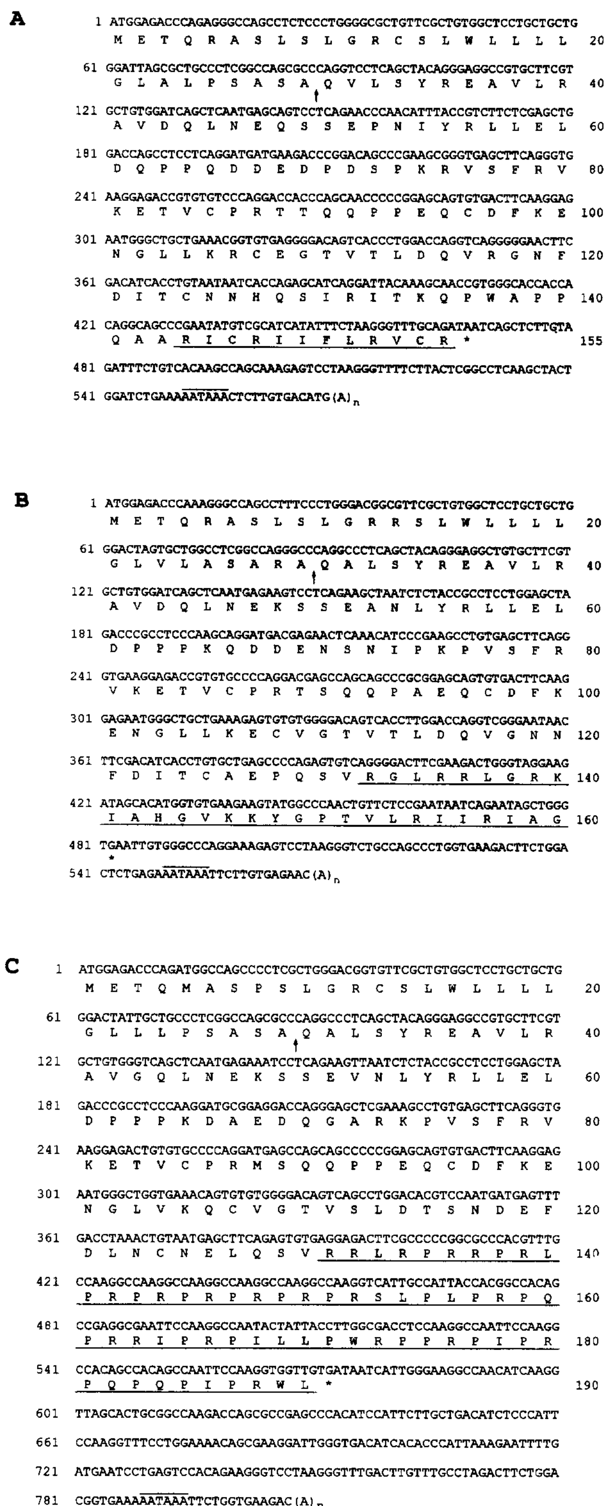

Fig. 1. Nucleotide and deduced amino acid sequences of the precursors of the sheep cyclic dodecapeptide (A), SMAP-29 (B) and Bac7.5 (C). The nucleotide sequences are numbered from the left, the amino acid sequences are numbered from the first methionine on the right. The arrows show the respective putative cleavage sites for signal peptidase. The putative antimicrobial sequences are underlined. Asterisks mark the stop codons. Polyadenylation signals are overlined.

the peptide may assume a N-terminal α-helical, and a C-terminal extended conformation. Although full elucidation of the

```
A   s-dodecapeptide   RICRIIFLRVCR
                      |:|||:::||||
    b-dodecapeptide   RLCRIVVIRVCR


B   Bac 7.5           RRLRPRRPRLPRPRPRPRPRPRSLPLPRPQPRRIPRPILLPWRPPRPI......PRPQPQPIPRWL
                      ||:||| |||||||||      ||:||| || ||||: :|   ||||      ||| | |||| |
    Bac 7             RRIRPRPPRLPRPRPRP......LPFPRPGPRPIPRPLPFPRPGPRPIPRPLPFPRPGPRPIPRPL
```

Fig. 2. Alignment of the sheep dodecapeptide (s-dodecapeptide) with the bovine dodecapeptide (b-dodecapeptide) (A), and of Bac7.5 with the bovine Bac7 (B). The amino acid sequences of s-dodecapeptide and of Bac7.5 are deduced from cDNA. The sequence of b-dodecapeptide is from [8,23], the sequence of Bac7 is from [13,37]. :, indicates identical residues; :, indicates similar residues. Gaps are indicated by dots.

structure of SMAP-29 requires structural analysis of the peptide, it is nonetheless worth noting that a helical wheel projection of the N-terminal region encompassing residues 1–18 (Fig. 3) would result in an α-helix with a clearcut amphipathic arrangement typical of membrane-perturbing agents. The helix would be characterized by the symmetric placing of Gly residues at the interface between polar and hydrophobic regions. The presence of Gly and Pro at positions 18 and 19 suggests a loop/turn at this point.

The 190 residue cathelicidin (Fig. 1C) has a calculated mass of 21831 Da and pI of 11.7, and shows a Pro- and Arg-rich C-terminal domain of 60 residues following a 29 residue signal peptide and a 101 residue cathelin-like prosequence. This protein is 85% identical to the precursor of the bovine Bac7 [13], a 60 residue antimicrobial peptide characterized by a highly repetitive Pro- and Arg-rich sequence [37]. Comparison of the nucleotide sequences (not shown) shows 90% identity in the 5′ pre-proregions (nt 1–390), 83% with a two gap insertion in the sequences corresponding to the antimicrobial domain (nt 391–570) and 93% in the 3′ non-coding regions (nt 571–807). The high level of identity suggests that the ovine and bovine precursors are derived from a common ancestor gene, from which the precursor of porcine PR-39 [10,38] may also have originated. Also, a decreased nucleotide identity in the region encompassing nt 391–570 indicates that the rate of mutation in the sequence corresponding to the antimicrobial domain is higher than in the remaining parts of the molecule. In fact, a fast evolving antimicrobial domain is a feature common to other genes encoding defense peptides [1] and appears to be an important requisite for molecules that have to keep up with highly variable and rapidly changing pathogens. Compared to bovine Bac7, the putative ovine antimicrobial peptide, which has been termed Bac7.5 based on the mol wt, shows a higher content of Arg (33% versus 28%) and a lower content of Pro (38% versus 45%). The two sequences (Fig. 2B) show a short stretch of similar residues at the N-terminus, most of them in common to the sequence of PR-39 [10,28]. This region is followed in the sequence of Bac7 by three tandemly repeated tetradecamers which are not present in the sequence of Bac7.5, although short regions of similarity, including several PRP repeats, can be noted. PRP repeats, and other proline-rich repeats are often observed in the Pro-and Arg-rich antimicrobial peptides [27,28,37] and are likely to influence their structure. Recent studies indicate that this type of peptide may adopt conformations similar to the polyproline type II helix [39–40]. Considering the sequences of the ovine Bac7.5 and the bovine Bac7 from an evolutionary point of view, it may be speculated that duplication of the tetradecamer in Bac7 occurred after the two species diverged, in view of the nearly complete conservation of the three repeats at the nucleotide level [13]. Further evolutionary changes likely contributed to the structural diversity of the two sequences, in addition to the duplication event, and analysis at the gene level will help understand the genetic mechanisms which promoted these changes. Furthermore, it is becoming apparent that a study of the homology of the cathelicidin pre-proregions within a species and among species may provide valuable information on evolutionary relationships. Thus, it is also important that a large number of sequences be collected from several species.

Fig. 3. Helical wheel projection of SMAP-29(1–18). Charged residues are shown in bold and hydrophobic residues are in italics.

Acknowledgements: We are grateful to Prof. M. Galeotti, Dipartimento di Scienze della Produzione Animale, University of Udine, for providing sheep ribs as a staring material for RNA preparation. We also thank Prof. D. Romeo and Dr. A. Tossi from the Dipartimento di Biochimica of the University of Trieste, and Prof. R. Gennaro from the Dipartimento di Scienze e Tecnologie Biomediche of the University of Udine, for useful discussions. This work was supported by grants from the National Research Council and from the Italian Ministry for University and Research.

References

[1] Boman, H.G., Marsh, J. and Goode, J.A. (1994) Antimicrobial Peptides, Ciba Found. Symp. No. 186, John Wiley and Sons, Chichester.

[2] Boman, H.G. (1995) Annu. Rev. Immunol. 13, 62–92.
[3] Lehrer, R.I, Lichtenstein, A.K. and Ganz, T. (1993) Annu. Rev. Immunol. 11, 105–128.
[4] Weiss, J. (1994) Current Opin. Hematology 1, 78–84.
[5] Selsted, M.E. and Ouellette, A.J. (1995) Trends Cell Biol. 5, 114–119.
[6] Larrick, J.W., Morgan, J.G., Palings, I., Hirata, M. and Yen, M.H. (1991) Biochem. Biophys. Res. Commun. 179, 170–175.
[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] Storici, P. and Zanetti, M. (1993) Biochem. Biophys. Res. Commun. 196, 1058–1065.
[11] Storici, P. and Zanetti, M. (1993) Biochem. Biophys. Res. Commun. 196, 1363–1368.
[12] Pungercar, J., Strukelj, B., Kopitar, G., Renko, M., Lenarcic, B., Gubensek, F. and Turk, V. (1993) FEBS Lett. 336, 284–288.
[13] Scocchi, M., Romeo, D. and Zanetti, M. (1994) FEBS Lett. 352, 197–200.
[14] Storici, P. Scocchi, M., Tossi, A. Gennaro, R. and Zanetti, M. (1994) FEBS Lett. 337, 303–307.
[15] Zanetti, M., Storici, P., Tossi, A., Scocchi, M. and Gennaro, R. (1994) J. Biol. Chem. 269, 7855–7858.
[16] Zhao, C., Liu, L. and Lehrer, R.I. (1994) FEBS Lett. 346, 285–288.
[17] Tossi, A., Scocchi, M., Zanetti, M., Storici, P. and Gennaro, R. (1995) Eur. J. Biochem. 228, 941–946.
[18] Agerberth, B., Gunne, H., Odeberg, J., Kogner, P., Boman, H.G. and Gudmundsson, G.H. (1995) Proc. Natl. Acad. Sci. USA 92, 195–199.
[19] Larrick, J.W., Hirata, M., Balint, R.F., Lee, J., Zhong, J. and Wright, S.C. (1995) Infect. Immun. 63, 1291–1297.
[20] Zanetti, M., Litteri, L., Gennaro, R., Horstmann, H. and Romeo, D. (1990) J. Cell Biol. 111, 1363–1371.
[21] Cowland, J.B., Johnsen, A.H. and Borregaard, N. (1995) FEBS Lett. 368, 173–176.
[22] Ritonja, A., Kopitar, M., Jerala, R. and Turk, V. (1989) FEBS Lett. 255, 211–214.
[23] Romeo, D., Skerlavaj, B., Bolognesi, M. and Gennaro, R. (1988) J. Biol. Chem. 263, 9573–9575.
[24] Gennaro, R., Skerlavaj, B. and Romeo, D. (1989) Infect. Immun. 57, 3142–3146.
[25] 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.
[26] 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.
[27] Harwig, S.S.L., Kokryakov, V.N., Swiderek, K.M., Aleshina, G.M., Zhao, C. and Lehrer, R.I. (1995) FEBS Lett. 362, 65–69.
[28] Agerberth, B., Lee, J.-Y., Bergman, T., Carlquist, M., Boman, H.G., Mutt, V. and Jornvall, H. (1991) Eur. J. Biochem. 202, 849–854.
[29] Larrick, J.W., Hirata, M., Shimomura, Y., Yoshida, M., Zheng, H., Zong, J. and Wright, S.C. (1993) Antimicrob. Agents Chemother. 37, 2534–2539.
[30] Levy, O., Weiss, J., Zarember, K., Ooi, C.E. and Elsbach, P. (1993) J. Biol. Chem. 268, 6058–6063.
[31] Zanetti, M., Gennaro, R. and Romeo, D. (1995) FEBS Lett., in press.
[32] Chomczinski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[33] Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymol. 47, 45–148.
[34] Kneller, D.G., Cohen, F.E. and Langridge, R. (1990) J. Mol. Biol. 214, 171–182.
[35] Rost, B. and Sander, C. (1993) J. Mol. Biol. 232, 584–599.
[36] Mains, R.E., Eipper, B.A., Glembotski, C.C. and Dores, R.M. (1983) Trends Neurosci. 6, 229–235.
[37] Frank, R., Gennaro, R., Schneider, K., Przybylski, M. and Romeo, D. (1990) J. Biol. Chem. 265, 18871–18874.
[38] Gudmundsson, G.H., Magnusson, K.P., Chowdhary, B.P., Johansson, M., Andersson, L. and Boman, H.G. (1995) Proc. Natl. Acad. Sci. USA 92, 7085–7089.
[39] Cabiaux, B., Agerberth, B., Johansson, J., Homble, F., Goormaghtigh, E. and Ruysschaert, J.-M. (1994) Eur. J. Biochem. 224, 1019–1027.
[40] Raj, P.A. and Edgerton, M. (1995) FEBS Lett. 368, 526–530.