Volume 314, number 2, 187–190 FEBS 11857 December 1992

cDNA sequence analysis of an antibiotic dodecapeptide from neutrophils

Paola Storici[a], Giannino Del Sal[a,b], Claudio Schneider[b] and Margherita Zanetti[a]

[a]*Department of Biochemistry, Biophysics and Macromolecular Chemistry, University of Trieste, 34127 Trieste, Italy* and [b]*National Laboratory of the Interuniversity Consortium for Biotechnology (CIB), AREA Science Park, Padriciano, 34012 Trieste, Italy*

Received 20 October 1992

The full-length cDNA of a neutrophil antibiotic dodecapeptide has been cloned by reverse transcription/PCR from bovine bone marrow RNA. This peptide was originally isolated from bovine neutrophils, and shown to exert a potent antimicrobial activity in vitro on both *Escherichia coli* and *Staphylococcus aureus*. The cDNA codes for a polypeptide of 155 amino acid residues with a predicted mass of 17,629 Da and a pI of 8.03. The deduced sequence comprises a putative signal peptide of 29 amino acids, a 114 residue pro-region, and a carboxy-terminal dodecapeptide corresponding to the mature antibiotic. The pro-sequence displays extensive identity to corresponding regions of other structurally unrelated antibiotic peptides of bovine neutrophils recently cloned.

Neutrophil; Antibiotic; Cyclic dodecapeptide; cDNA

1. INTRODUCTION

A central role of neutrophils in the initial host response to microbial challenge is testified by a massive recruitment of these cells at the infection site, followed by rapid and effective pathogen inactivation. Killing is ensured by the production of highly reactive oxygen derivatives [1] combined with the release of granule-associated cytotoxic proteins into phagocytic vacuoles [2,3]. Oligo- and polypeptides with diverse antimicrobial spectra have been isolated from neutrophils of various animal species [4–7]. In particular, from granule extracts of bovine neutrophils, four cationic peptides have been purified and characterized. Two of them, associated as inactive proforms to the large granules [8,9], are known as Bac5 and Bac7, and exhibit bactericidal activity mainly, but not exclusively, on Gram-negative organisms [10,11]. In addition to these, a cyclic dodecapeptide, previously named bactenecin [12], and a Trp-rich tridecapeptide amide named indolicidin [13], exert a potent antibacterial action on both *Escherichia coli* and *Staphylococcus aureus*. We report here the nucleotide sequence of a full-length cDNA encoding the precursor form of the cyclic dodecapeptide. The cDNA cloning of Bac5 [14] and indolicidin [15] has recently shown that these antibiotics share highly similar pro-sequences. A homologous pro-sequence is also predicted to precede the sequence of the dodecapeptide. Although the mature forms of all these peptides appear structurally unrelated, they are likely to derive from a common gene family also including members isolated from different animal species [16,17].

Correspondence address: M. Zanetti, Dipartimento di Biochimica, Biofisica and Chimica delle Macromolecole, via A. Valerio 38, Università di Trieste, 34127 Trieste, Italy. Fax. (39) (40) 676 3691.

2. EXPERIMENTAL

2.1. *cDNA cloning*

Total RNA was extracted from bovine bone marrow cells with guanidinium thiocyanate [18]. The rapid amplification of cDNA ends, or RACE, [19] was used to obtain both the 3′ and 5′ ends of the dodecapeptide cDNA. To generate the 3′ end, the first strand cDNA was synthesized using 1 μg of total RNA and 1.5 ng of primer adaptor 5′-TCGGATCCCTCGAGAAGC(dT18)-3′ and Moloney murine leukemia virus (MMLV) reverse transcriptase (Bethesda Research Laboratories, Gaithersburg, MD) (1 h at 40°C). The reaction mixture (20 μl) was then heated to 95°C and 80 μl of the PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$) were added with 50 pmol of both the upstream oligonucleotide 5′-CGCGAATTCAAAGCCTGTGAGCTTC-3′ and the downstream primer adaptor 5′-CGAGCTCGGATCCCTCGAGAAGCTT-3′. PCR was carried out with 2.5 U of *Taq*I polymerase (Perkin-Elmer-Cetus, Norwalk, CT) for 30 cycles of denaturation at 94°C for 2 min, annealing at 55°C for 1 min and polymerization at 72°C for 2 min.

The RACE to the 5′ end was carried out using the oligonucleotide 5′-CAGATCCAGTAGCTTGAGGC-3′ derived from the 3′ end sequence of the dodecapeptide. Single-stranded cDNA was tailed using terminal deoxynucleotidyl transferase (Bethesda Research Laboratories) and dGTP. Amplification conditions were as above, using an upstream primer complementary to the dGTP tail (dC15), and a downstream primer 5′-GACGAATTCGAGTAAGAAAACCCTTA-3′.

2.2. *cDNA sequencing and sequence analysis*

The amplified cDNA was cloned in Bluescript KS⁺ vector (Stratagene, San Diego, CA). Sequencing was performed on both strands with the dideoxy chain-termination method [20]. Regions with high G+C content were also sequenced in parallel with deazaguanosine and automated fluorescent DNA sequencing (EMBL fluorescent DNA sequencer, Heidelberg, Germany). Analysis of the DNA sequence was conducted with the aid of the IntelliGenetics Suite version 5.4 (IntelliGenetics Inc., Mountain View, CA), and the homology search was

carried out on the Swiss-Prot database The hydropathy plot was produced by the method of Kyte and Doolittle [21] with a width of 9 residues.

2.3. *Northern analysis*

Total RNA (10 μg) was separated by electrophoresis on 0.41 M formaldehyde/1% agarose gels, transferred to a nylon membrane (Zeta-Probe GT, Bio-Rad, Richmond, CA) by the downward alkaline capillary method [22], crosslinked with UV-Stratalinker (Stratagene), and hybridized (60°C, 0.5 M NaPi, 7% SDS, 1 mM EDTA, 16 h) with a cDNA restriction fragment, ^{32}P labelled by random primer synthesis (Pharmacia-LKB, Uppsala, Sweden).

2.4. *In vitro transcription and translation*

1 μg of the Bluescript KS$^+$ plasmid containing the full-length cDNA insert was linearized and transcribed with T7 RNA polymerase in the presence of the cap analogue (Pharmacia) In vitro translation was conducted using a rabbit reticulocyte lysate (Novagene, Madison, WI). Translation products were analyzed by 15% SDS-PAGE

3. RESULTS AND DISCUSSION

3.1. *Cloning of the cyclic dodecapeptide*

Cloning of the cyclic dodecapeptide from bovine bone marrow cells was approached by reverse transcription followed by PCR (RACE protocol) [19]. Using an upstream primer identical to a sequence of a Bac5 cDNA clone [14], and a downstream oligo-dT primer, three products of different size were amplified (not shown). By sequence analysis these were found to encode the precursor forms of Bac5 [14], indolicidin [15], and the cyclic dodecapeptide, respectively. The three fragments revealed homologous 5′ regions, incomplete at their 5′ ends. Northern analysis with a ^{32}P-labelled restriction fragment corresponding to a sequence unique to the dodecapeptide, revealed the presence of a corresponding mRNA of about 0.6 kb in bovine bone marrow RNA (Fig. 1), but not in bovine heart, kidney, liver, lung, small intestine, spleen, and stomach (not shown).

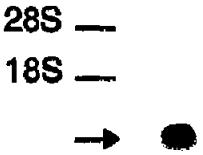

Fig. 1. Northern analysis. A Northern blot of total RNA from bovine bone marrow cells was probed with a ^{32}P-labelled restriction fragment corresponding to nucleotides 419–536 of the dodecapeptide cDNA sequence. This recognizes a mRNA of about 0.6 kb

The full-length cDNA of the cyclic dodecapeptide was obtained by the 5′ RACE [19]. Amplified products were cloned in Bluescript vector. Since the sequences analyzed were found to differ at two positions, 24 clones generated from different preparations of RNA were completely sequenced in both directions. A guanine was found to replace adenine-395 in the translated region of 18 of the clones, and two additional adenine nucleotides were found inserted at position 513 of the 3′-untranslated region of 9 random clones (Fig. 2).

3.2. *Features of the predicted sequence*

The full-length cDNA of the cyclic dodecapeptide (Fig. 2) shows a putative translation start codon at position 13 and a stop codon at position 478. The poly(A) tail is preceded by a consensus signal. The sequence predicts a protein (pre-prododecapeptide) of 155 amino acid residues, with a calculated mass of 17,629 Da, confirmed by in vitro translation of the transcript (not shown), and an overall pI of 8.03. No *N*-linked glycosylation sites are present in the predicted sequence. A hydrophobic stretch (amino acid residues 1–29) typical of a signal sequence, including a consensus signal peptid-

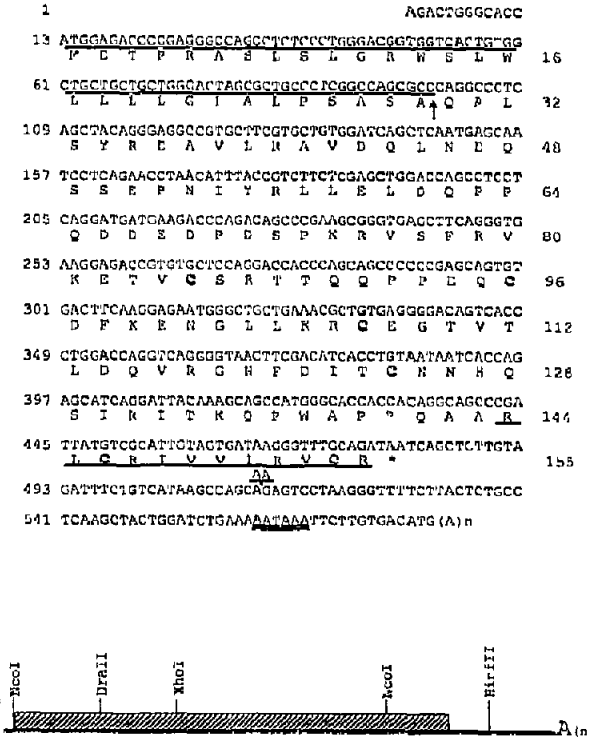

Fig. 2. Nucleotide and deduced amino acid sequence of the cyclic dodecapeptide precursor, and partial restriction endonuclease map of the cDNA. The nucleotide sequence is numbered on the left. The amino acid sequence is numbered from the first methionine on the right. The signal sequence is overlined, the arrow indicates the putative cleavage site for signal peptidase. The sequence of the mature antibiotic is underlined, and the stop codon is marked with an asterisk. The polyadenylation signal is double underlined. Cysteines are in bold. A schematic representation of the full-length cDNA of the dodecapeptide is shown at the bottom.

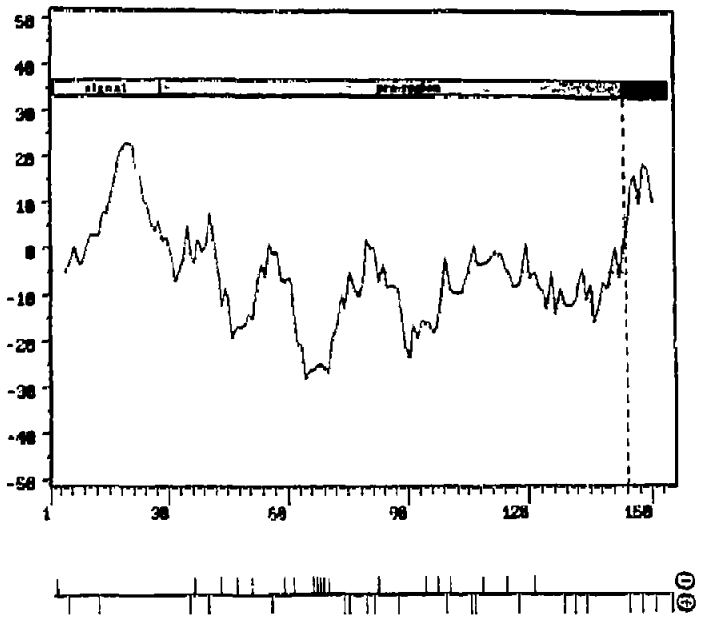

Fig. 3. Hydropathy plot [21] of the cyclic dodecapeptide precursor. Linear diagram of the deduced amino acid sequence, with the amino-terminal on the left and the carboxyl-terminal on the right. A dashed vertical line shows the cleavage site for the maturation of the cyclic dodecapeptide. The lower panel indicates the distribution of basic (+) and acidic (−) residues.

ase cleavage site (Ala-X-Ala) [23] is present at the amino-terminus of the protein (Fig. 3). The putative signal peptide is followed by a 114 amino acid long hydrophilic pro-sequence, corresponding to residues 30–143, which is characterized by a cluster of negatively charged amino acids in the region between residues 59 and 71. The carboxy-terminal dodecapeptide (residues 144–155) corresponds to the mature antibiotic isolated from granule extracts of neutrophils, and exactly matches the sequence determined by protein sequencing [12]. The two cysteine residues at positions 146 and 154 are known to maintain the mature antibiotic in a cyclic structure by a disulfide bond [12]. In the conformation proposed [12], the peptide chain forms a bend at residue 7, with the hydrophobic residues clustered around the bend. This compact structure is reminiscent of defensins, a family of numerous small antimicrobial peptides with a characteristic cysteine motif forming three intramolecular disulfide bonds [24].

Unlike defensins, which are expressed by various cell types of mammals and insects [24], the bovine neutrophil has so far been the only source of the dodecapeptide [12]. Recently, another basic, cysteine-rich peptide, with functional similarities to both the dodecapeptide and the defensins, has been isolated from bovine tracheal mucosa [25]. The sequences of all these cysteine-rich peptides, as deduced from their mRNAs, lack the carboxy-terminal glycine required for amidation, a characteristic feature of several other antibacterial peptides [14,15,26,27], that may be required to increase their biostability. The amino-terminus of the dodecapeptide is preceded by an Ala residue at position 143. This is a likely site of proteolytic cleavage for elastase, which may be responsible for the removal of the pro-sequence from the precursor [9]. Due to base substitution, the clones analyzed predicted the alternative presence of Gln and Arg at position 128 of the pro-sequence, pointing to the existence of either a polymorphic, or an additional gene. Genomic cloning will resolve this issue, since neither possibility can be excluded at present.

3.3. *Observed similarity of the dodecapeptide precursor with other proteins*

The sequence of the deduced precursor is compared in Fig. 4 with that of Bac5 [14] and indolicidin [15], both isolated from bovine neutrophils. As expected from the cloning strategy, the pro-sequence of the dodecapeptide displays extensive identity to corresponding regions of

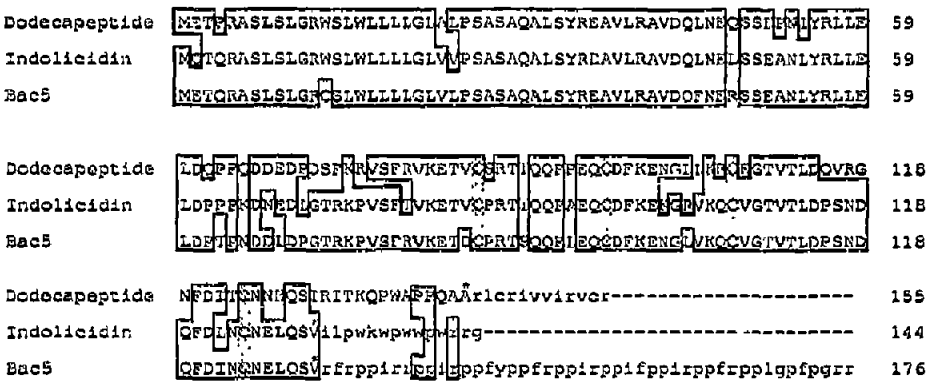

Fig. 4. Alignment of the predicted sequence of the cyclic dodecapeptide with other antibiotic peptides of bovine neutrophil. The amino acid sequence deduced from the cDNA encoding the dodecapeptide is aligned with the deduced sequences of indolicidin [15], and Bac5 [14]. Boxed residues reflect those common to the dodecapeptide and at least one other peptide. Aligned cysteines are shaded. Small letters indicate the sequences corresponding to the mature antibiotics, and asterisks show the carboxy-terminal residues of the pro-sequences. Sequences were aligned using the GENALIGN program (IntelliGenetics).

both Bac5 and indolicidin, including 4 invariant cysteine residues, that may offer possible structural constraints. Both structural and functional implications have been previously suggested [14] by the presence of homologous pro-regions in the sequences of proBac5 and proindolicidin. The evidence that the cyclic dodecapeptide also carries a homologous region further strengthens the hypothesis of a common pro-sequence acting as a carrier for defense peptides. In fact, in addition to the bovine antibiotics mentioned above, 2 homologous proteins from different animal species have been identified, providing evidence for a more widespread presence of this sequence: CAP18, an antibacterial protein from rabbit [17], and cathelin, an inhibitor of cathepsin L isolated from pig neutrophils [16]. These exhibit about 57% and 66% identity, respectively, to the pro-region of the dodecapeptide (not shown).

Both proBac 5 and proindolicidin carry a valyl residue (marked with an asterisk in Fig. 4) at the cleavage site for the proteolytic maturation of the respective antibiotics. The pro-sequence of the dodecapeptide assignes Ile at the same position, followed by a stretch of 13 additional residues (131–143 of the sequence) extending its carboxy-terminus. It is not known whether these residues are cleaved from the precursor during the intracellular maturation of the antibiotic. Immunochemical investigations with antibodies to the dodecapeptide will clarify this point. In particular, since the cyclic dodecapeptide has been the only proteolytic product so far isolated from neutrophils [12], it will be interesting to know whether the same precursor may produce more biologically relevant fragments.

Acknowledgements. We thank Dr. D. Romeo for helpful discussion. This work was supported by grants from the National Research Council (Progetto Finalizzato 'Biotecnologie e Biostrumentazione') and from the Italian Ministry for University and Research.

REFERENCES

[1] Klebanoff, S.J. (1988) in: Inflammation Basic Principles and Clinical Correlates (J.I. Gallin, I.M. Goldstein and R. Snydermann eds.) pp. 391–444, Raven Press, New York.
[2] Joiner, K.A., Ganz, T., Albert, J. and Rotrosen, D. (1989) J. Cell Biol. 109, 2771–2782.
[3] Zanetti, M., Litteri, L., Griffiths, G., Gennaro, R. and Romeo, D. (1991) J. Immunol. 146, 4295–4300.
[4] Elsbach, P. and Weiss, J. (1988) in: Inflammation: Basic Principles and Clinical Correlates (J.I. Gallin, I.M. Goldstein and R. Snydermann eds.) pp. 391–444 Raven Press, New York.
[5] Lehrer, R.I and Ganz, T. (1990) Blood 76, 2169–2181.
[6] Gennaro, R., Romeo, D., Skerlavaj, B. and Zanetti, M. (1991) in: Blood Cell Biochemistry (J.R. Harris ed.) pp. 335–368, Plenum, New York.
[7] Zasloff, M. (1992) Curr. Opin. Immunol. 4, 3–7.
[8] Zanetti, M., Litteri, L., Gennaro, R., Horstmann, H. and Romeo, D. (1990) J. Cell Biol. 111, 1363–1371.
[9] Scocchi, M., Skerlavaj, B., Romeo, D. and Gennaro, R., Eur. J. Biochem. (in press).
[10] Gennaro, R., Skerlavaj, B. and Romeo, D. (1989) Infect. Immun. 57, 3142–3146.
[11] Frank, R.W., Gennaro, R., Schneider, K., Przybylski, M. and Romeo, D. (1990) J. Biol. Chem. 265, 18871–18874.
[12] Romeo, D., Skerlavaj, B., Bolognesi, M. and Gennaro, R. (1988) J. Biol. Chem. 263, 9573–9575.
[13] Selsted, M.E., Novotnj, M.J., Morris, W.L., Tang, Y.-Q., Smith, W. and Cullor, J.S. (1992) J. Biol. Chem. 267, 4292–4295.
[14] Zanetti, M., Del Sal, G., Storici, P., Schneider, C. and Romeo, D. (1993) J. Biol. Chem. (in press).
[15] Del Sal, G., Storici, P., Schneider, C., Romeo, D. and Zanetti, M. (1992) Biochem. Biophys. Res. Commun. 187, 467–472.
[16] Ritonja, A., Kopitar, M., Jerala, R. and Turk, V. (1989) FEBS Lett. 255, 211–214.
[17] Larrick, J.W., Morgan, J.G., Palings, I., Hirata, M. and Yen, M. H. (1991) Biochem. Biophys. Res. Commun. 179, 170–175.
[18] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[19] Frohman, A.M., Dush, K.M. and Martin, G.R. (1988) Proc Natl. Acad Sci. USA 85, 8998–9002.
[20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[21] Kyte, J. and Doolittle, R.I. (1982) J. Mol. Biol. 157, 105–132.
[22] Chomczynski, P. (1992) Anal. Biochem. 201, 134–139.
[23] von Heijne, G. (1986) Nucleic Acids Res. 14, 4683.
[24] Lehrer, R.I., Ganz, T. and Selsted, M.E. (1991) Cell 64, 229–230.
[25] Diamond, G., Zasloff, M., Eck, H., Brasseur, M., Maloy, W.L. and Bevins, C.L. (1991) Proc. Natl. Acad. Sci. USA 88, 3952–3956.
[26] Agerberth, B., Lee, J., Bergman, T., Carlquist, M., Boman, H.G., Mutt, V. and Jornvall, H. (1991) Eur. J. Biochem. 202, 849–854.
[27] Boman, H.G., Boman, I.A., Andreu, D., Zong-qu Li, Merrifield, R.B., Schlenstedt, G. and Zimmermann, R. (1989) J. Biol. Chem. 264, 5852–5860.