Vol. 187, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 31, 1992 Pages 467-472

cDNA CLONING OF THE NEUTROPHIL BACTERICIDAL PEPTIDE INDOLICIDIN

Giannino Del Sal[1,2], Paola Storici[1], Claudio Schneider,[2] Domenico Romeo[1], and Margherita Zanetti[1,*]

[1]Department of Biochemistry, Biophysics & Macromolecular Chemistry, University of Trieste, 34127 Trieste, Italy

[2]National Laboratory of the Interuniversity Consortium for Biotechnology (CIB), AREA Science Park, Padriciano, 34012 Trieste, Italy

Received July 13, 1992

SUMMARY: A structurally novel, tryptophan-rich antimicrobial tridecapeptide amide, named indolicidin, has recently been purified from bovine neutrophils (Selsted *et al.*(1992) J. Biol. Chem. 267, 4292-4295). Here we describe the molecular cloning of this endoantibiotic, which is synthesised in bone marrow cells as a 144 amino acid residue precursor. The encoded protein has a predicted mass of 16479 Da and a pI of 6.51. A putative signal peptide of 29 amino acids precedes a 101 residue pro-region. The mature peptide is at the 3' end of the open reading frame. A glycine, not found in purified indolicidin, is present at the carboxyl terminus of the deduced sequence and is very likely involved in post-translational peptide amidation.

A major function of cytoplasmic granules of neutrophils is the delivery of several antimicrobial oligo- and poly-peptides to phagocytic vacuoles (1,2). The neutrophil "endoantibiotics" so far characterized in various animal species (3-6) do not exhibit common structural features, but are all very cationic. Furthermore, they show diverse antimicrobial spectra. The treatment of ingested microbes with a combination of granule endoantibiotics is thus likely to generate a situation full of potentialities for pathogen inactivation.

We have started the characterization of the repertoire of endoantibiotics of bovine neutrophils by resolving granule extracts into several active fractions (7). From these we have purified to homogeneity and sequenced three antimicrobial

* To whom correspondence should be addressed.

0006-291X/92 $4.00

peptides: an arginine-rich cyclic dodecapeptide, capable of killing both *E.coli* and *S.aureus* (8) and two proline- and arginine-rich polypeptides, Bac5 and Bac7, which appear to be active mainly, but not exclusively, on Gram-negative organisms (9,10).

More recently, Selsted and coworkers have purified an antibacterial tridecapeptide amide from bovine neutrophils, which is characterized by an unusual high proportion of tryptophan residues and therefore named indolicidin (11). In this report we describe the cloning and characterization of the full length cDNA that codes for indolicidin, which appears to be synthesised in bone marrow cells as pre-proform.

MATERIALS AND METHODS

cDNA generation.

Total RNA was extracted from bovine bone marrow cells with guanidinium thiocyanate (12). A primer derived from a previously characterized incomplete cDNA of Bac5†, was used to obtain the 3' end by the RACE protocol (13). Ten µg of total RNA was used to synthesise the first strand cDNA with 1.5 ng of primer adaptor 3'-(dT_{18})CGAAGAGCTCCCTAGGCT-5' and Moloney murine leukemia virus (MMLV) reverse transcriptase (BRL) (1h 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 pmoles of both the upstream oligonucleotide 5'-CGCGAATTCAAAGCCTGTGAGCTTC-3'(240-255 bp of Bac5 cDNA sequence) and the downstream primer adaptor 3'-TTCGAAGAGCTCCCTAGGCTCGAGC-5'. PCR was carried out with 2.5 units of *Taq* polymerase 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'-TCTGAACAAATCAGACACTTA-3' derived from the 3' end sequence of indolicidin. Single stranded cDNA was tailed using terminal deoxynucleotidyl transferase (BRL) and dGTP. Amplification was performed using the upstream primer complementary to the dGTP tail (dC_{15}), and a downstream primer 5'-ACGAATTCCGAGGCTCAGGATCTGACA-3'. Amplification conditions were as above.

cDNA sequencing.

The amplified cDNA was cloned in Bluescript SK^+ vector (Stratagene). Sequencing was performed on both strands with the dideoxy chain termination method (14). Regions with high G+C content were also sequenced in parallel with deazaguanosine and automated fluorescent DNA sequencing (EMBL fluorescent DNA sequencer). DNA sequence analysis was conducted with the aid of the IntelliGenetics Suite version 5.4, and homology search was carried out on the Swiss-Prot database.

Northern and primer extension analyses.

Total RNA (5 µg) was separated by electrophoresis on 6.7% formaldehyde/1% agarose gels, transferred to a nylon membrane

†Zanetti, M., Del Sal, G., Storici, P., Schneider, C., and Romeo, D., manuscript in preparation.

(GeneScreen-Plus, NEN), and crosslinked with UV-Stratalinker (Stratagene). The probe for hybridization (55°C, 1 M NaCl, 1% SDS, 16 h) was a 21 mer oligonucleotide derived from the 3' end sequence (nt 439-459 of the indolicidin cDNA sequence). For primer extension analysis 1 ng of the same oligonucleotide was annealed to 10 µg of total bovine bone marrow RNA for 30 min at 55°C. The extension reaction was performed by using MMLV reverse transcriptase (BRL) in 30 µl final volume for 1 h at 42°C. The extended product was analyzed with a 6% polyacrylamide denaturing gel.

***In vitro* transcription-translation.**

One µg of the Bluescript SK$^+$ plasmid containing the full length cDNA insert was linearized downstream with *Hind* III and transcribed with T3 RNA polymerase in the presence of the cap analogue (Pharmacia). *In vitro* translation was conducted with rabbit reticulocyte lysate (Novagene). Translation products were analysed by 15% SDS-PAGE.

RESULTS AND DISCUSSION

An oligonucleotide designed within the upstream region of a cDNA clone encoding a partial sequence of Bac5[†] was used as a primer to generate the missing 3' end. This caused amplification of two nucleotide sequences, that appeared to be the 3' ends of different cDNAs by sequence analysis. Bac5 was present in one of the two cDNAs, whereas the second one included the sequence correspondig to the recently discovered bactericidal tridecapeptide amide indolicidin (11). An oligonucleotide derived from this sequence was thus synthesised and used to probe a Northern blot of bovine bone marrow RNA. This probe recognized an mRNA of about 0.6 kb (Fig.1, A). The 5' end of this mRNA, was

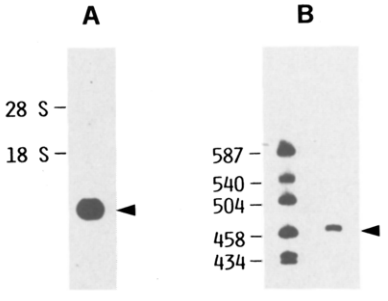

Figure 1. Northern and primer extension analyses of indolicidin mRNA. A: Total RNA from bovine bone marrow cells was probed with a [^{32}P]labelled oligonucleotide complementary to nt 439-459 of indolicidin cDNA. The messenger recognized is about 0.6 kb. B: The same oligonucleotide was used to prime RNA from bovine bone marrow cells. The size of the extended product, on the right, is about 460 nt. Size markers on the left.

```
  1 AGACTGGGGACCATGCAGACCCAGAGGGCCAGCCTCTCACTGGGG
                M  Q  T  Q  R  A  S  L  S  L  G   11

 46 CGGTGGTCGCTGTGGCTACTGCTGCTGGGACTAGTGGTGCCCTCG
    R  W  S  L  W  L  L  L  L  G  L  V  V  P  S   26

 91 GCCAGCGCCCAAGCCCTCAGCTACAGGGAGGCCGTGCTTCGTGCT
    A  S  A  Q  A  L  S  Y  R  E  A  V  L  R  A   41

136 GTGGATCAGCTCAATGAGCTGTCCTCAGAAGCTAATCTCTACCGC
    V  D  Q  L  N  E  L  S  S  E  A  N  L  Y  R   56

181 CTCCTGGAGCTAGACCCACCTCCCAAGGATAATGAAGATCTGGGC
    L  L  E  L  D  P  P  P  K  D  N  E  D  L  G   71

226 ACTCGAAAGCCTGTGAGCTTCACGGTGAAGGAGACTGTGTGCCCC
    T  R  K  P  V  S  F  T  V  K  E  T  V  C  P   86

271 AGGACGATTCAGCAGCCCGCGGAGCAGTGTGACTTCAAGGAGAAA
    R  T  I  Q  Q  P  A  E  Q  C  D  F  K  E  K  101

316 GGGCGGGTGAAACAGTGTGTGGGGACAGTCACCCTGGACCCATCC
    G  R  V  K  Q  C  V  G  T  V  T  L  D  P  S  116

361 AATGACCAGTTTGACCTAAACTGTAATGAGCTCCAGAGTGTCATC
    N  D  Q  F  D  L  N  C  N  E  L  Q  S  V  I  131

406 CTACCCTGGAAATGGCCATGGTGGCCTTGGCGCAGAGGTTGATGG
    L  P  W  K  W  P  W  W  P  W  R  R  G  *      144

451 AGAAGAGCTGTCAGATCCTGAGCCTCGGGAAGAGTCTTAAGTGTC

496 TGATTTGTTCAGATTCGGGCTTCTGGACAGTGAAAATAAATTCTT

541 GTGAAAACGGAAAAAAAAAAAAAAA
```

Figure 2. Nucleotide and deduced amino acid sequence of indolicidin. The nucleotide sequence is numbered on the left, from the 5' end of the mRNA. The amino acid sequence is numbered from the first methionine on the right. The signal peptide is overlined, and the arrow indicates the putative cleavage site for signal peptidase. The sequence of mature indolicidin is underlined, and the stop codon is marked with an asterisk. The polyadenylation signal is double underlined.

determined by primer extension using the same synthetic oligonucleotide, and shown to be about 440 nt upstream of the primer (Fig. 1, B). The 550 bp full length cDNA encoding indolicidin, shown in Fig.2, was then obtained by the RACE protocol to the 5' end.

Sequence analysis revealed that indolicidin is synthesised as 144 amino acid precursor that shares 51% and 74% identity, respectively, with CAP18 from rabbit (15) and cathelin from pig (16). The predicted mass of the precursor is 16479 Da, as confirmed by *in vitro* translation of the transcript (not shown), with a calculated pI of 6.51. The sequence predicts a hydrophobic 29 residue signal peptide, overlined in Fig.2, that is followed by a pro-region of 101 residues.

The deduced sequence of mature indolicidin, located at the carboxyl end of the precursor (underlined in Fig. 2), is comprised of 14 amino acid residues, including a carboxy-terminal glycine. Indolicidin as purified from neutrophils is a tridecapeptide ending with an amidated arginine (11). The C-terminal glycine thus appears to be lost by posttranslational processing that results in α amidation of the carboxyl end of indolicidin, with the nitrogen atom of the amide deriving from the missing residue (17).

Selsted *et al.* (11) did not provide evidence for the existence of a precursor of indolicidin, since they isolated only the mature tridecapeptide amide from granule extracts of bovine neutrophils. In contrast, Bac5 and Bac7, other bactericidal polypeptides of bovine neutrophils, have been found to be present as proforms in the neutrophil large granules (18), wherefrom they can be isolated as such. Their proteolytic maturation is carried out by elastase (19) that is contained in neutrophil primary granules, as an early event during phagocytosis (2). In this respect, it is interesting to note that valine, a specific cleavage site for elastase, is present at the C-terminus of the pro-sequence of indolicidin (residue 130 in Fig.2), thus suggesting that elastase might also be responsible for the maturation of proindolicidin. Whether this maturational step is carried out in the resting neutrophil is not clear. Since Selsted *at al.* (11) suggest that indolicidin is stored as mature peptide in a still unknown organellar location, it is possible that indolicidin and elastase co-localize to the primary granules, allowing for proteolytic maturation being carried out in this compartment. However, the possibility that the pro-region is retained and proindolicidin represents the neutrophil storage form, as in the case of probactenecins, cannot be ruled out. Antibodies to indolicidin will be useful in identification of its intracellular localization as well as investigations on its processing.

ACKNOWLEDGMENTS. We thank P. Edomi for performing the automated nucleotide sequence analysis. 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.Joiner, K.A., Ganz, T., Albert, J., and Rotrosen, D. (1989) J. Cell Biol. 109, 2771-2782.

2.Zanetti, M., Litteri, L., Griffiths, G., Gennaro, R., and Romeo, D. (1991) J. Immunol. 146, 4295-4300.
3.Lehrer, R.I., and Ganz, T. (1990) Blood 76, 2169-2181.
4.Spitznagel, J.K. (1990) J. Clin Invest. 86, 1381-1386.
5.Gennaro, R., Romeo, D., Skerlavaj, B., and Zanetti, M. (1991) In Blood Cell Biochemistry (J.R. Harris Ed.), pp.335-368, Plenum Publ.Co., New York.
6.Boman, H.G. (1991) Cell 65, 205-207.
7.Gennaro, R., Dolzani, L, and Romeo, D. (1983) Infect. Immun. 40, 684-690.
8.Romeo, D.,Skerlavaj, B., Bolognesi, M., and Gennaro, R. (1988) J. Biol. Chem. 263; 9573-9575.
9.Gennaro, R., Skerlavaj, B., and Romeo, D. (1989) Infect. Immun. 57, 3142-3146.
10.Frank, R.W., Gennaro, R., Schneider, K., Przybylski, M., and Romeo, D. (1990) J. Biol. Chem. 265, 18871-18874.
11.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.
12.Chomczinski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156-159.
13.Frohman, A. M., Dush, K. M., and Martin, G. R. (1988) Proc Natl. Acad. Sci. USA 85, 8998-9002.
14.Sanger, F., Nicklen, S., and Coulson, A., R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467.
15.Larrick, J. W., Morgan, J. G., Palings, I., Hirata, M., and Yen, M., H. (1991) Biochem. Biophys. Res. Commun. 179, 170-175.
16.Ritonja, A., Kopitar, M., Jerala, R., and Turk, V. (1989) FEBS Lettr. 255, 211-214.
17.Bradbury, A., F., and Smyth, D., G. (1991) Trends Biochem. Sci. 16, 112-115.
18.Zanetti, M., Litteri, L., Gennaro, R., Horstmann, H.,and Romeo, D. (1990) J. Cell Biol. 111, 1363-1371.
19.Scocchi, M., Skerlavaj, B., Romeo, D., and Gennaro, R., Eur. J. Biochem., accepted for publication.