FEBS Letters 352 (1994) 197–200

FEBS 14567

Molecular cloning of Bac7, a proline- and arginine-rich antimicrobial peptide from bovine neutrophils

Marco Scocchi[a], Domenico Romeo[b], Margherita Zanetti[a,c,*]

[a]*Laboratorio Nazionale Consorzio Interuniversitario Biotecnologie, AREA Science Park, Padriciano, I-34012 Trieste, Italy*
[b]*Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università di Trieste, I-34127 Trieste, Italy*
[c]*Dipartimento di Scienze e Tecnologie Biomediche, Università di Udine, Udine, Italy*

Received 4 August 1994

Abstract Bac7 is a 7 kDa proline- and arginine-rich antimicrobial peptide which was purified from bovine neutrophils. We have used PCR to clone the cDNA of Bac7 precursor, a polypeptide of 21,569 Da. This cDNA is highly conserved in the 5′ region, with respect to the corresponding region in the precursors of several other structurally unrelated myeloid antimicrobial peptides. Furthermore, a 148 nt non-coding region at the 3′ end is 75% homologous to a corresponding region of the cDNA of the precursor of PR-39, a porcine antibacterial peptide which is also proline- and arginine-rich.

Key words: cDNA; Antibacterial peptide; Bactenecin; Neutrophil

1. Introduction

Antimicrobial peptides and polypeptides are important mediators of innate immunity. They are widely distributed among living organisms and often show a marked inter-species diversity in structure and spectrum of activity [1–8]. In mammals, they appear to be located in mucosal epithelial cells [9–11] and in circulating phagocytes where they are thought to substantially contribute to the oxygen-independent defense system [1–7]. The antimicrobial peptides of phagocytes originate from precursors that are synthesized in bone marrow cells [5,6,12–22]. It has recently been shown that several such precursors share highly homologous N-terminal preprosequences of 128–130 residues [12–22]. In contrast to the conserved preprosequence, the antimicrobial peptides derived from these precursors, which include bovine Bac5 [23], indolicidin [24] and cyclic dodecapeptide [25], porcine PR-39 [26], protegrins [27], PMAP-23 [21], PMAP-36 [20] and C12 [19], and rabbit CAP18 [12], are structurally varied. Also belonging to this family is cathelin [28], a pig leukocyte protein [29] corresponding to only the prosequence of these precursors and possibly representing the N-terminal fragment of one of them.

In this paper we report the cDNA cloning of Bac7, a proline- and arginine-rich peptide originally known as 7 kDa bactenecin [23,30]. The mature peptide originates from the proteolytic processing of an inactive precursor [31] stored in the large granules of bovine neutrophils [32]. We show here that the precursor of Bac7 contains a cathelin-like domain and shares structural features with other precursors from this group.

2. Materials and methods

2.1. Purification and amino acid sequence analysis of proBac7

ProBac7 was purified from granule extracts of bovine blood neutrophils as described [31]. Purified proBac7 was incubated with endoproteinase Lys-C (Boehringer-Mannheim Biochemica), 20:1 ratio by mass, in 25 mM Tris-HCl, pH 8.5, 1 mM EDTA and 0.01% SDS for 15 h at 37°C. The incubation mixture was then subjected to 14% tricine SDS-PAGE [33] and electroblotted onto Biotrans polyvinylidene difluoride membranes (ICN Biomedicals, Costa Mesa, CA) for amino-terminal sequencing of proBac7 fragments. Sequence analysis was performed by automatic Edman degradation in a gas-phase sequencer 477A (Applied Biosystem, Foster City, CA).

2.2. RNA extraction and Northern analysis

Total RNA was extracted from bovine bone marrow cells with guanidinium thiocyanate [34]. For Northern analysis, the *Eco*RI restriction fragment of the 3′ end cDNA of Bac7 precursor (nt 395 to the polyadenylated tail) was ^{32}P-labelled by random primer synthesis (Pharmacia-LKB, Uppsala, Sweden) and used to hybridize a Northern blot of bovine bone marrow total RNA, as described [17].

2.3. cDNA cloning and sequencing, and sequence analysis

1 μg of total RNA was used to synthesize first strand cDNA using M-MLV reverse transcriptase and the oligonucleotide primer adaptor 5′-TCGGATCCCTCGAGAAGC(T18)-3′, as described [15]. For PCR amplification of Bac7 3′ end cDNA, 5 μl of the reaction mixture was added to 95 μl of PCR buffer (50 mM Tris-HCl, pH 8.3, 0.25 mg/ml BSA and 2 mM $MgCl_2$), containing 50 pmol of the antisense primer adaptor 5′-CGAGCTCGGATCCCTCGAGAAGCTT-3′ and 50 pmol of a set of degenerate (512-fold degeneracy) sense oligonucleotides 5′-CCGAATTCGCYCGYAAXCCYACYAG-3′ (Y = T,C,A,G; X = A,G,) deduced from the amino acid sequence of a proBac7 proteolytic fragment and corresponding to amino acid residues 72–77 in Fig. 1. Amplification was performed in a 1605 Air Thermo Cycler (Idaho Technology, Idaho Falls, ID) using 2.5 units of *Taq*I polymerase (Perkin Elmer-Roche Mol. Syst., Branchburg, NJ) for 30 cycles: 15 s at 94°C, 15 s at 50°C, and 30 s at 72°C. To obtain the 5′ end cDNA, the antisense oligonucleotide primer 5′-AGTGCTAACCTTGATGTT-3′, derived from the 3′ end sequence of Bac7 and complementary to nt 591–608 of the cDNA sequence, was used for reverse transcription. Amplification was performed using the antisense oligonucleotide 5′-CGCGGATCCGTTCCCAATGATTTTCACA-3′ (complementary to nt 569–588 of the cDNA sequence) and a sense oligonucleotide primer 5′-CAAGAATTCGGAGACTGGGGACCATG-3′ derived from a 5′ conserved sequence of preproBac5, including a short 5′ untranslated region [15]. Thirty amplification cycles were performed: 15 s at 94°C, 15 s at 55°C, and 30 s at 72°C. Amplified cDNA was cloned in Bluescript KS$^+$ vector (Stratagene, San Diego, CA), sequenced on both strands, and the sequence analized as described previously [15]. Homology searches were carried out on the GenBank database using programs from the GCG suite.

2.4. In vitro transcription–translation

The pBluescript KS$^+$ vector containing the full-length clone was linearized downstream with *Xho*I, and 1 μg of DNA was transcribed

*Corresponding author. Laboratorio Nazionale CIB, AREA Science Park, Padriciano 99, I-34012 Trieste, Italy. Fax: (39) (40) 398 979.

SSDI 0014-5793(94)00954-6

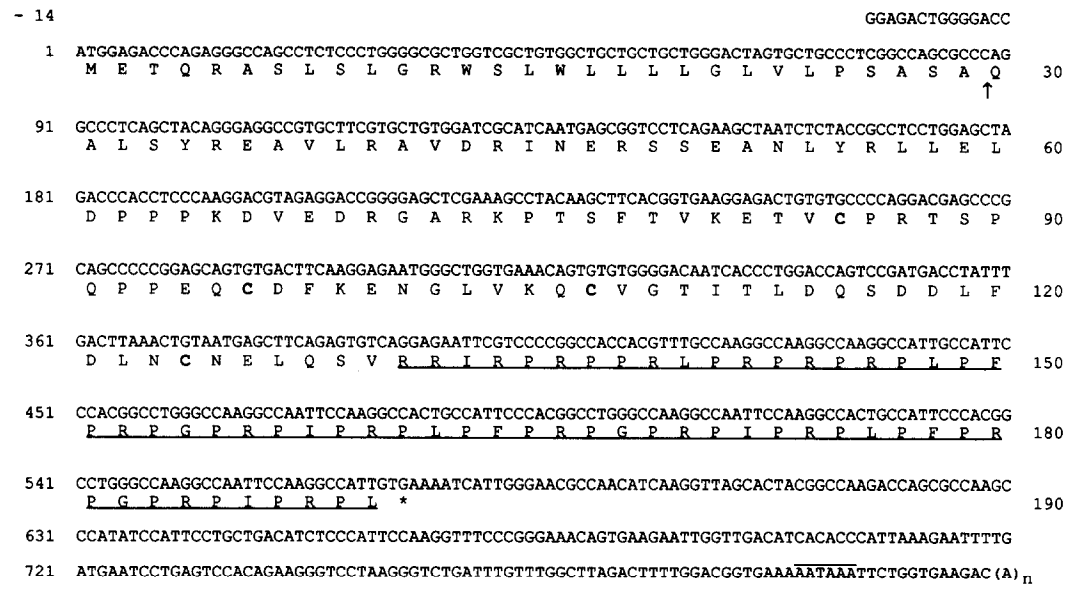

Fig. 1. Nucleotide and deduced amino acid sequence of Bac7. Numbering is on the left for nucleotides and on the right for amino acids. The end of the putative signal peptide is indicated by an arrow and the stop codon by an asterisk. The polyadenylation signal is overlined and the sequence of the mature Bac7 is underlined. Cysteines are in boldface type.

with T7 RNA polymerase (Stratagene) as described [15]. Translation reactions using rabbit reticulocyte lysate (Novagene, Madison, WI) in the presence of [^{35}S]methionine and [^{35}S]cysteine (Amersham International, Amersham, UK), membrane translocation analysis, and immunoprecipitation of the translated products, were performed as described [15]. The translation products were analyzed by 15% SDS-PAGE.

3. Results and discussion

A group of proline- and arginine-rich peptides with highly repetitive amino acid sequences have been identified among the myeloid-derived antimicrobial peptides, either through peptide purification or cDNA cloning. These include bovine Bac5 and Bac7 [30], and porcine PR-39 [18,26] and C12 [19]. cDNA cloning of the precursors of Bac5 [15], PR-39 [18] and C12 [19] revealed that they share highly homologous preprosequences and suggested that the 59 residue bovine Bac7 may also derive from a similar precursor. Rather unexpectedly, PCR amplification of reverse-transcribed myeloid mRNA, using an upstream oligonucleotide primer derived from a conserved region of the prosequence of the precursor of Bac5 and an oligo-dT downstream primer, while enabling us to identify several other bovine and porcine antimicrobial peptide precursors with similar preprosequences and structurally varied C-terminal antimicrobial domains [13,14,17,20,21], failed to amplify the cDNA sequence of Bac7.

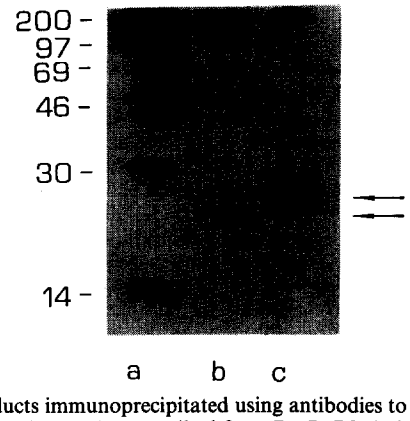

Fig. 2. Products immunoprecipitated using antibodies to Bac7 from in vitro-translated mRNA transcribed from Bac7 cDNA, in the presence (lane b) and in the absence (lane c) of a microsomal membrane fraction. (Lane a) Molecular weight standards.

We therefore decided to clone the cDNA of this precursor, using proBac7-specific primers deduced from its amino acid sequence. A partial amino acid sequence, obtained by microsequencing the amino-terminal regions (residues 66–82 and 106–128 in Fig. 1) of two major proteolytic fragments of purified proBac7, suggested the presence of a cathelin-like domain similar to those of the bovine congeners [13–15]. Based on a region of lower homology in this domain, a set of proBac7-specific, degenerate, sense oligonucleotides were synthesized and used for the 3′ end RACE of Bac7 precursor, which amplified a 3′ end cDNA (nt 214 to the polyadenylated tail in Fig. 1) including in its sequence an *Eco*RI restriction site (nt 395–400). The 3′ end restriction fragment was sequenced, and found to contain the almost complete sequence of the mature Bac7 and the 3′ non-coding region. When this fragment was used as a probe for Northern analysis of bovine bone marrow RNA, it recognized a transcript of approximatively 0.9 kb (not shown). To obtain the complete cDNA sequence of the precursor, the 5′ region was amplified with an antisense primer derived from the non-coding 3′ end of Bac7 and a sense primer from the non-coding 5′ end which is conserved within the bovine precursors of this family [13–15]. The amplified sequence, containing the complete open reading frame, was transcribed and translated in vitro, and the translation products, immunoprecipitated with antibodies to purified Bac7, were analyzed by SDS-PAGE and fluorography (Fig. 2), showing molecular masses similar to those previously reported for preproBac7 and pro-Bac7, respectively [32]. The nucleotide and deduced amino acid

```
BAC7           METQRASLSLGRWSLWLLLLGLVLPSASAQALSYREAVLRAVDRINERSSEANLYRLLELDPPPKDVEDRGARKPTSFTVKETVCPRTSPQPPEQ
INDOLICIDIN    -Q--------------------V------------------QL--L-----------------N--L-T---V------------IQ--A--
BAC5           ------------C-------------------------------QF-----------------T-N-DL-P-T---V--R----D-----Q--L--
DODECAPEPTIDE  ---P-----------------A-------------------QL--Q---P-I-------Q--Q-D--PDSP-RV--R------S--TQ-----

BAC7           CDFKENGLVKQCVGTITLDQSDDLFDLNCNELQSVRRIRPRPPRLPRPRPRPLPFPRPGPRPIPRPLPFPRPGPRPIPRPLPFPRPGPRPIPRPL
INDOLICIDIN    -----K-R-------V---P-N-Q----------ILPWKWPWWPWRRG
BAC5           ---------------V---P-N-Q--I--------RFRPPIRRPPIRPPFYPPFRPPIRPPIFPPIRPPFRPPLGPFPGRR
DODECAPEPTIDE  --------L-R-E--V----VRGN--IT--NH--IRITKQPWAPPQAARLCRIVVIRVCR
```

Fig. 3. Alignment of the precursor of Bac7 with other bovine myeloid precursors. Dashes denote identical residues. The C-terminal bioactive peptides are in italic.

sequence obtained from the two partially overlapping clones is shown in Fig. 1. A 570 nt open reading frame, with a stop codon at nt 571–573, predicts a 21,569 Da protein of 190 amino acid residues, with a pI of 11.7. The preproregion consists of a putative signal peptide (residues 1–29) and a cathelin-like domain (residues 30–130), similar to those reported for the other antimicrobial peptide precursors of this group [12–22]. The preproregion is 83–91% identical to the corresponding regions of the bovine (Fig. 3), and porcine [17–22] congeners. As for most of these precursors [13,15,17–22], the proregion (cathelin-like domain) ends with a C-terminal valine (residue 130), which is a cleavage site for elastase. Previous studies on the maturation of proBac7 showed that elastase liberates the mature peptide by cleavage of proBac7 at this site [31]. The deduced amino acid sequence of the mature peptide, corresponding to residues 131–190 (Fig. 1), matches the sequence obtained from Edman degradation, with an additional C-terminal leucine not identified by protein sequencing of the purified peptide [23], and lacks the C-terminal amidation signal which has often been observed in peptides derived from this group of precursors [13,15, 17–20,22].

```
BAC7  AGGAGAATTCGTCCCCGGCCACCACGTTTGCCAAGGCCAAGGCCAAGGCC 440
      ||||||  |||||||| || |||  ||||||||||||||||||    ||
PR39  AGGAGA...CGTCCCCGACCCCCATATTTGCCAAGGCCAAGGCCACCTCC 454

BAC7  ATTGCCATTCCCACGGCCTGGGCCAAGGCCAATTCCAAGGCCACTGCCAT 490
        ||      ||| |||     ||||||      ||| ||   |
PR39  GTTTTTCCCACCAAGGCTCCCACCAAGGATCCCACCAGGGTTCC..... 498

BAC7  TCCCACGGCCTGGGCCAAGGCCAATTCCAAGGCCACTGCCATTCCCACGG 540
PR39  ..................................................  499

BAC7  CCTGGGCCAAGGCCAATTCCAAGGCCATTGTGAAAATCATTGGGAACGCC 590
                                      *
PR39  ..................................................  499

BAC7  AACATCAAGGTTAGCACTACGGCCAAGACCAGCGCCAAGCCCATATCCAT 640
        || ||||||
PR39  ..CACCAAGGT.....................................  512

BAC7  TCCTGCTGACATCTCCCATTCCAAGGTTTCCCGGGAAACAGTGAAGAATT 690
                  |||||  ||| |||| ||||| |||| |||| | | |
PR39  ............TCCCA..CCACGGTTCCCCGGAAAACGGTGATGGAGT 542
                                          *

BAC7  GGTTGACATCACACCCATTAAAGAATTTTGATGAATCCTGAGTCCACAGA 740
      || ||| | || ||||||||||   ||||| ||||||||||| |||  ||
PR39  GGCTGATAACATACCCATTAAAAGCTTTTGGTGAATCCTGAGCCCAGGGA 592

BAC7  AGGGTCCTAAGGGTCTGATTTGTTTGGCTTAGACTTTTGGACGGTGAAAA 790
      |  |||||| || ||| |||  | ||||| |||||| || ||  ||||||
PR39  A AGTCCTA.GGATCTTATTGTTGTGGCTCAGACTTCTGAACCATGAAAA 640

BAC7  ATAAATTCTGGTGAAGAC(A)n 808
      ||||||||| |||||
PR39  ATAAATTCTTGTGAAACG(A)n 658
```

Fig. 4. Sequence comparison of the 3′ end cDNAs of Bac7 and PR-39, including the sequences encoding the mature peptides and the 3′ non-coding regions. The sequences corresponding to the mature peptides are underlined, and asterisks denote the respective stop codons.

The sequence between residues 131–145 (N-terminal region of the mature Bac7) shows identity over 13 residues to the corresponding N-terminal region of the mature porcine PR-39 [26], with one amino acid substitution at position 8 and a one residue gap between the second and third amino acid of mature PR-39. The remaining portion of the sequence of mature Bac7 is characterized by three tandemly repeated tetradecamers, PFPRPGPRPIPRPL, encompassing residues 149-190.

The similarity between Bac7 and PR-39 drops in the 146–190 region of Bac7, but a comparative analysis of their 3′ end cDNAs (Fig. 4) shows that the non-coding region of Bac7 from nt 661 to 808 is 75% identical to the PR-39 sequence encompassing nt 513–658, with the non-coding region of Bac7 retaining most of the sequence encoding the seven C-terminal amino acid residues, the stop codon and the non-coding region of PR-39. When compared to other precursors, the 3′ cDNA of Bac7 shows a lower, but significant similarity to the corresponding 3′ regions of Bac5 and C12, and a variable and in general more limited similarity to other members of this family (not shown). A certain degree of conservation (17–19 identical nt) was observed relative to the 24 nt sequence corresponding to the seven C-terminal amino acid residues and stop codon of PR-39 (nt 513–536 in Fig. 4), the C-terminal sequence of Bac5 and C12, and a corresponding sequence in the non-coding region of protegrins (not shown). These results follow the recently reported observation that the 3′ cDNAs of porcine protegrins share an almost complete identity with the 3′ cDNA of C12, the mature products of which have been named prophenins [22]. Based on this sequence analysis, which shows high 5′ and 3′ cDNA identity of protegrins and C12/prophenins, it has been suggested that the protegrin family arose by insertion of its exon into a gene originally specifying for C12/prophenins [22]. The observed similarity of Bac7 and PR-39 cDNAs may be explained by an analogous evolutionary event also including, in this case, duplication of the sequence encoding the repeated tetradecamer of Bac7. However, the mechanism by which these peptides are likely to have originated from a common ancestor gene will be better elucidated by the genomic

analysis of the precursors. Considering the extent of the homology in the conserved region of Bac7 and PR-39, one could speculate that the separation of these two precursors occurred early enough to allow for changes at several nucleotide positions. It is noteworthy that Northern and PCR analyses have failed to detect Bac7 in pig, or PR-39 in bovine bone marrow cells (not shown), suggesting that each peptide was selected for in each of these species.

Acknowledgements: We are indebted to Prof. S. Ronchi, from the Istituto di Fisiologia Veterinaria e Biochimica, Università di Milano, for performing the amino acid sequencing. 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 (MURST 60% and 40%).

References

[1] Boman, H.G. (1991) Cell 65, 205–207.
[2] Gennaro, R., Romeo, D., Skerlavaj, B. and Zanetti, M. (1991) in: Blood Cell Biochemistry, vol. 3 (Harris, J.R., ed.) pp. 335–368, Plenum, New York.
[3] Zasloff, M. (1992) Curr. Opin. Immunol. 4, 3–7.
[4] Elsbach, P. and Weiss, J. (1993) Curr. Opin. Immunol. 5, 103–107.
[5] Leher, R.I., Lichtenstein, A.K. and Ganz, T. (1993) Ann. Rev. Immunol. 11, 105–128.
[6] Gabay, J.E. and Almeida, R.P. (1993) Curr. Opin. Immunol. 5, 97–102.
[7] Weiss, J. (1994) Curr. Opin. Hematol. 1, 78–84.
[8] Cociancich, S., Bulet, P., Hetru, C. and Hoffmann, J.A. (1994) Parasitol. Today 10, 132–139.
[9] Ouellette, A.J., Greco, R.M., James, M., Frederick, D., Naftilan, J. and Fallon, J.T. (1989) J. Biol. Chem. 108, 1687–1695.
[10] Diamond, G., Zasloff, M., Eck, H., Brasseur, M., Maloy, W.L., and Bevins, C.L. (1991) Proc. Natl. Acad. Sci. USA 88, 3952–3956.
[11] Jones, D.E. and Bevins, C.L. (1992) J. Biol. Chem. 267, 23216–23225.
[12] Larrick, J.W., Morgan, J.G., Palings, I., Hirata, M. and Yen, M.H. (1991) Biochem. Biophys. Res. Commun. 179, 170–175.
[13] Del Sal, G., Storici, P., Schneider, C., Romeo, D. and Zanetti, M. (1992) Biochem. Biophys. Res. Commun. 187, 467–472.
[14] Storici, P., Del Sal, G., Schneider, C. and Zanetti, (1992) FEBS Lett. 314, 187–190.
[15] Zanetti, M., Del Sal, G., Storici, P., Schneider, C. and Romeo, D. (1993) J. Biol. Chem. 268, 522–526.
[16] Levy, O., Weiss, J., Zarember, K., Ooi, C.E. and Elsbach, P. (1993) J. Biol. Chem. 268, 6058–6063.
[17] Storici, P. and Zanetti, M. (1993) Biochem. Biophys. Res. Commun. 196, 1363–1368.
[18] Storici, P. and Zanetti, M. (1993) Biochem. Biophys. Res. Commun. 196, 1058–1065.
[19] Pungercar, J., Strukelj, B., Kopitar, G., Renko, M., Lenarcic, B., Gubensek, F. and Turk, V. (1993) FEBS Lett. 336, 284–288.
[20] Storici, P., Scocchi, M., Tossi, A., Gennaro, R. and Zanetti, M. (1994) FEBS Lett. 337, 303–307.
[21] Zanetti, M., Storici, P., Tossi, A., Scocchi, M. and Gennaro, R. (1994) J. Biol. Chem. 269, 7855–7858.
[22] Zhao, C., Liu, L.and Lehrer, R.I. (1994) FEBS Lett. 346, 285–288.
[23] Frank, R.W., Gennaro, R., Schneider, K., Przybylski, M. and Romeo, D. (1990) J. Biol. Chem. 265, 18871–18874.
[24] 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.
[25] Romeo, D., Skerlavaj, B., Bolognesi, M. and Gennaro, R. (1988) J. Biol. Chem. 263, 9573–9575.
[26] Agerberth, B., Lee, J.-Y., Bergman, T., Carlquist, M., Boman, H.G., Mutt, V. and Jornvall, H. (1991) Eur. J. Biochem. 202, 849–854.
[27] Kokryakov, V.N., Harwig, S.S.L., Panyutich, E.A., Shevchenko, A.A., Aleshina, G.M., Shamova, O.V., Korneva, H.A., and Leher, R.I. (1993) FEBS Lett. 327, 231–236.
[28] Ritonja, A., Kopitar, M., Jerala, R. and Turk, V. (1989) FEBS Lett. 255, 211–214.
[29] Kopitar, M., Ritonja, A., Popovic, T., Gabrijelcic, D., Krizaj, I. and Turk, V. (1989) Biol. Chem. Hoppe-Seyler 370, 1145–1151.
[30] Gennaro, R., Skerlavaj, B. and Romeo, D. (1989) Infect. Immun. 57, 3142–3146.
[31] Scocchi, M., Skerlavaj, B., Romeo, D. and Gennaro, R. (1992) Eur. J. Biochem. 209, 589–595.
[32] Zanetti, M., Litteri, L., Gennaro, R., Horstmann, H., and Romeo, D. (1990) J. Cell Biol. 111, 1363–1371.
[33] Schagger, H. and von Jagow, G (1987) Anal. Biochem. 166, 368–379.
[34] Chomczinski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.