Volume 303, number 1, 31–35 FEBS 11055 May 1992

Structure of the guinea pig neutrophil cationic peptide gene

Isao Nagaoka, Akimasa Someya, Kazuhisa Iwabuchi and Tatsuhisa Yamashita

Department of Biochemistry, Juntendo University, School of Medicine, Hongo, Bunkyo-ku, Tokyo 113, Japan

Received 19 March 1992; revised version received 6 April 1992

Guinea pig neutrophils contain the antimicrobial cationic peptides GNCP-1 and GNCP-2 in the granules. In this study, the GNCP gene was isolated, and the structure was characterized. Using cDNA probes, one phage clone was isolated from a guinea pig genomic library. The gene spanned >3 kb, and comprised three exons and two introns. Sequence analysis revealed that the gene encoded GNCP-2. Exon 1 mainly coded for the 5′ untranslated region, exon 2 coded for the prepro-peptide region of GNCP-2, and exon 3 coded for the mature peptide region of GNCP-2 and the 3′ untranslated region. Primer extension analysis indicated that the transcription initiation site was located to a thymidine residue, 93 bp upstream of the ATG initiation codon of GNCP-2 mRNA. A possible TATA box was located 24 bp upstream of the transcription start site. Interestingly, the pyrimidine-rich sequences identified in the promoter regions of the human neutrophil elastase and myeloperoxidase genes were also found in the 5′ flanking region of the GNCP-2 gene.

Neutrophil cationic peptide; Antimicrobial peptide; Gene structure; Primer extension; Guinea pig

1. INTRODUCTION

Neutrophils play a central role in protecting hosts against the microbial infection. Both oxidative and non-oxidative mechanisms are used by neutrophils to restrict infection [1–3]. Molecules such as H_2O_2, O_2^- and HOCl are examples of microbicidal agents that are derived from oxidative metabolism [1,3]. On the other hand, neutrophil granules contain antimicrobial proteins and peptides that contribute to the oxygen-independent host defense mechanism [2,3]. The most abundant of these antimicrobial components are low molecular-weight cationic peptides with potent microbicidal activities against bacteria, fungi and viruses [2–7]. In addition, the cationic peptides have been shown to display cytotoxic and histamine-releasing activities [5,7].

Recently we have purified the two cationic peptides GNCP-1 and GNCP-2 from guinea pig neutrophils, and found that GNCP-1 and GNCP-2 are single-chain peptides comprising 31 amino acid residues, which differ only by the substitution of an isoleucine (GNCP-1) for a leucine (GNCP-2) at position 21 [7]. Furthermore, we have isolated the cDNA clones for GNCPs, and shown that transcription of the GNCP gene and expression of GNCP mRNA are observed in bone marrow cells, but not in mature neutrophils [8]. Thus, the expression of the GNCP gene appears to be restricted to the maturing neutrophils in the bone marrow. However, the mechanism(s) controlling GNCP gene expression during neutrophil maturation is not clear. In this study, therefore, to begin to understand the regulation of the GNCP gene expression, we have isolated the genomic clone for GNCP and determined the gene structure.

The nucleotide sequence reported in this paper has been submitted to the DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number X63677.

Abbreviation: GNCP, guinea pig neutrophil cationic peptide

Correspondence address: Isao Nagaoka, Department of Biochemistry, Juntendo University, School of Medicine, Hongo, Bunkyo-ku, Tokyo 113, Japan. Fax: (81) (3) 3814-9300.

2. MATERIALS AND METHODS

2.1. *Screening of the guinea pig genomic DNA library*

A guinea pig genomic DNA library was constructed with partial *Sau*3AI digests of guinea pig liver DNA and the cloning vector EMBL3 (Promega), according to the manufacturer's instructions. Two million clones were plated on *Escherichia coli* strain LE392, and screened by plaque hybridization [9], using GNCP-1 and GNCP-2 cDNAs as probes [8]. The probes were labeled with [α-^{32}P]dCTP (3000 Ci/mmol, Amersham) using random hexanucleotide primers (Promega) [10]. Positive plaques were subjected to a second and third screening to isolate single plaques. Phage DNA was purified by the standard method of cesium chloride step gradient centrifugation [9].

2.2. *Restriction endonuclease mapping of the genomic DNA clone*

Phage DNA was digested with one or more restriction endonucleases and the resulting fragments were separated by agarose (1%) gel electrophoresis, denatured in alkali, and transferred to a nylon membrane (Hybond-N, Amersham) as described by Southern [11]. Filters were hybridized with ^{32}P-labeled GNCP-1 and GNCP-2 cDNA probes, and the sizes of the fragments containing exons were determined. The phage DNA was also mapped by the method of Rackwitz et al. [12], using partial restriction endonuclease digests of the phage DNA and ^{32}P-labeled synthetic 12-base oligonucleotides complementary to the cos sites of the phage DNA. The oligonucleotides were labeled with [γ-^{32}P]ATP (6000 Ci/mmol, Du Pont/NEN) at the 5′ end by T4 polynucleotide kinase (Promega).

2.3. *Subcloning and DNA sequencing*

Selected restriction fragments of the genomic DNA insert were subcloned into the plasmic vector Bluescript SK (−) (Stratagene). The

nucleotide sequences of the DNA fragments were determined from both directions with sequence specific oligonucleotide primers by the dideoxy chain-termination procedure [13], using the Sequenase DNA Sequencing Kit (United States Biochemical), or the Taq Dye Deoxy Terminator Cycle Sequencing Kit and the 373A DNA Sequencer (Applied Biosystems).

2.4. *Mapping the 5' end of the GNCP mRNA*

The location of the 5' end of the GNCP mRNA was identified by primer extension analysis [14]. Total cellular RNA was isolated from bone marrow cells by the guanydinium thiocyanate extraction method [15], and poly(A)$^+$ RNA was selected by oligo-dT cellulose column chromatography (Collaborative Research). The oligonucleotide primer (5'-GGTCTCAAGAGGCA-3') which was specific for GNCP-2 mRNA and complementary to nucleotides +11 to +24 of the GNCP-2 gene sequence (numbered according to Fig. 2), was labeled with [γ-^{32}P]ATP (6000 Ci/mmol, Du Pont/NEN) using T4 polynucleotide kinase. The labeled primer (2×10^5 cpm) was mixed with poly(A)$^+$ (10 μg) in 10 mM Tris-HCl, pH 8.3, 250 mM KCl and 1 mM EDTA (total volume 20 μl), heated to 60°C for 1 h, and then allowed to anneal at room temperature for 1.5 h. The reverse transcriptase reaction was performed in 20 mM Tris-HCl, pH 8.3, 75 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.25 mM EDTA, 0.25 mM dNTPs, 50 μg/ml actinomycin D and 10 U/μl M-MLV reverse transcriptase (Bethesda Research Laboratories) (total volume 80 μl) at 37°C for 1 h. After extraction with phenol/chloroform and ethanol precipitation, the products were separated on a 6% polyacrylamide, 8 M urea sequencing gel. The sequencing ladders were generated by the dideoxy chain-termination method utilizing the same oligonucleotide primer against the genomic DNA fragment, and used as size markers.

3. RESULTS AND DISCUSSION

3.1. *Cloning and sequencing of the GNCP gene*

By screening the guinea pig genomic library (2×10^6 independent clones) with the ^{32}P-labeled GNCP-1 and GNCP-2 cDNA probes, one hybridizing phage clone was identified, containing a 14 kb DNA insert. Restriction enzyme mapping by Southern analysis using the cDNA probe indicated that the exons of the GNCP genes dispersed over a 3 kb region (Fig. 1A). Then, the 3.2 kb *Hin*dIII–*Bam*HI fragment was completely sequenced (Fig. 1B and Fig. 2), and the sequence was compared to the previously published cDNA sequences for GNCP-1 and GNCP-2 (accession numbers X57705 and X63676, respectively, for DDBS/EMBL/GenBank) [8]. The gene proved to code for GNCP-2 cDNA: the gene encoded Gln^{-8} and Leu^{21} of GNCP-2, compared to Glu^{-8} and Ile^{21} of GNCP-1 [7,8]. The gene was composed of three exons interrupted by two introns. Each splice site agreed with the consensus splice site sequences in that the 5' end of the intron began with the dinucleotide GT, and the 3' end finished with the dinucleotide AG preceded by a pyrimidine rich sequence [16,17]. The length of each exon varied from 80 bp (exon 1) to 227 bp (exon 3). The length of the intron was 1217 bp between exon 1 and exon 2, and 561 bp between exon 2 and exon 3. The single, in-frame start codon for translation (Met^{-62}) was located in exon 2, and the single, in-frame stop codon (TAA) was located in exon 3. Exon 1 coded mainly the 5' untranslated region, exon 2 mainly coded for consensus sequence for translation initiation [18] and the prepro-peptide region, and exon 3 mainly encoded the mature peptide region and the 3' untranslated region. The gene ended with a polyadenylation signal and a site for poly (A) addition 15 bp 3' to the signal.

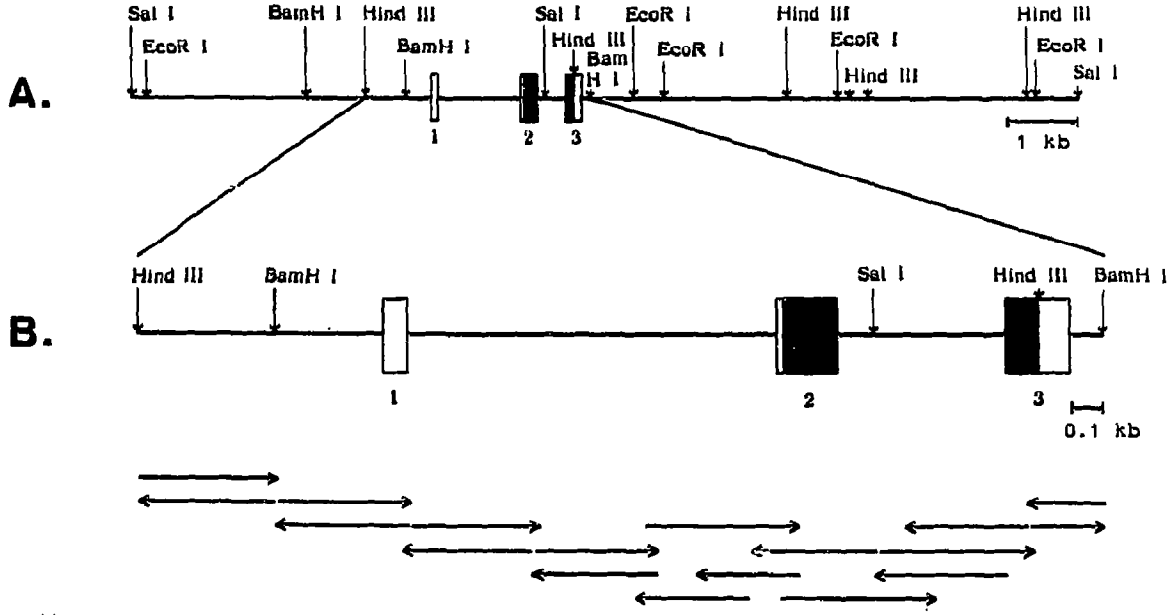

Fig. 1. Restriction map and schematic structure of the GNCP-2 gene. (A) Diagram of the GNCP-2 gene. Restriction sites and exons were determined with the 14 kb *Sal*I fragment insert of the phage clone obtained from the size-fractionated EMBL3 genomic library. Exons are shown by the numbered boxes. Filled areas indicate the coding regions and open areas depict the 5' and 3' untranslated regions of the GNCP-2 mRNA. (B) Diagram of the 3205 bp *Hin*dIII–*Bam*HI segment of the cloned genomic DNA that was completely sequenced. Horizontal arrows indicate the direction and length of sequencing reactions from the specific primers.

3.2. *Characterization of the transcription initiation site and the promotor region*

The transcription initiation site of the GNCP-2 gene was determined by primer extension analysis. Using poly(A)$^+$ RNA from bone marrow cells, one predominant primer extension product was observed (Fig. 3A). The extension product was aligned with sequencing ladders produced by dideoxy sequencing of the genomic clone using the same primer, and the major transcription start site was localized to a thymidine residue 10 bp upstream of the primer and 93 bp upstream of the initiation codon (ATG) of GNCP-2 mRNA.

Examination of the sequence of the 5′ flanking region revealed the presence of a TATA-like sequence (TAAATAT) at nucleotide −24 and two CAAT-like sequences at nucleotides −88 and −132 (Fig. 2). A glucocorticoid responsive element-like sequence (AAGACAGGAAGTCCT vs. AGAACANNNTGTTCT [19]) was also found upstream of the TATA box at −179 to −165. Interestingly, in a preliminary experiment the GNCP gene transcription rate analyzed by a transcription run-off assay [8] was observed to be up-regulated 2- to 3-fold by treatment of the bone marrow cells with 10 μg/ml dexamethasone (not shown).

The transcription of the genes for neutrophil granule components, such as elastase, myeloperoxidase, and cationic peptide, is suggested to be limited to a period of neutrophil maturation that takes place in the bone marrow [8,20–23]. In the 5′ flanking regions of the human neutrophil elastase and myeloperoxidase genes, pyrimidine-rich sequences (CCCCTTCTCCCCCTTTTCA for the neutrophil elastase gene and CCCCTTCCCCCCATTTCA for the myeloperoxidase gene) were identified, and these sequences are assumed to be a regulatory element for the two genes [24–26]. Interestingly, the pyrimidine-rich sequences are also found in the 5′ flanking region of the GNCP-2 gene at −67 to −52 and −473 to −433, although the pyrimidine-rich sequences of the GNCP-2 gene shared at most 60–70% homology with those of the human neutrophil elastase and myeloperoxidase genes. The role of the promotor region, including the glucocorticoid responsive element and the pyrimidine-rich sequences in the GNCP gene expression remains to be elucidated in the future.

Fig. 2. The nucleotide sequences of the GNCP-2 gene. The nucleotide sequences are numbered from the putative transcription initiation site marked with an asterisk. The sequence of the 5′ flanking region is numbered from −1 to −804 starting 5′ to the transcription initiation site. Nucleotides comprising exons are in upper-case letters, and those comprising introns are in lower-case letters. For the 5′ flanking region, the putative glucocorticoid responsive element is double-underlined; the pyrimidine-rich sequences found in the promotor regions of the human neutrophil elastase [24] and myeloperoxidase genes [25] are underlined; the TATA-like and CAAT-like sequences are boxed. The sequence of the complementary oligonucleotide used in the primer extension study is underlined in exon 1 (see Fig. 3). The polyadenylation signal (AATAAA) is underlined in the 3′ untranslated region. The 3′ end of exon 3 is deduced from the site for poly(A) addition in the GNCP-2 mRNA. For the coding exons, the deduced amino acid sequences are shown under the nucleotide sequences, and are numbered beginning at the N-terminal residue of the mature GNCP-2. Negative numbers (−62 to −1) refer to the putative signal peptide (−62 to −44) and pro-peptide (−43 to −1). The arrows indicate the predicted signal peptide and pro-peptide cleavage sites.

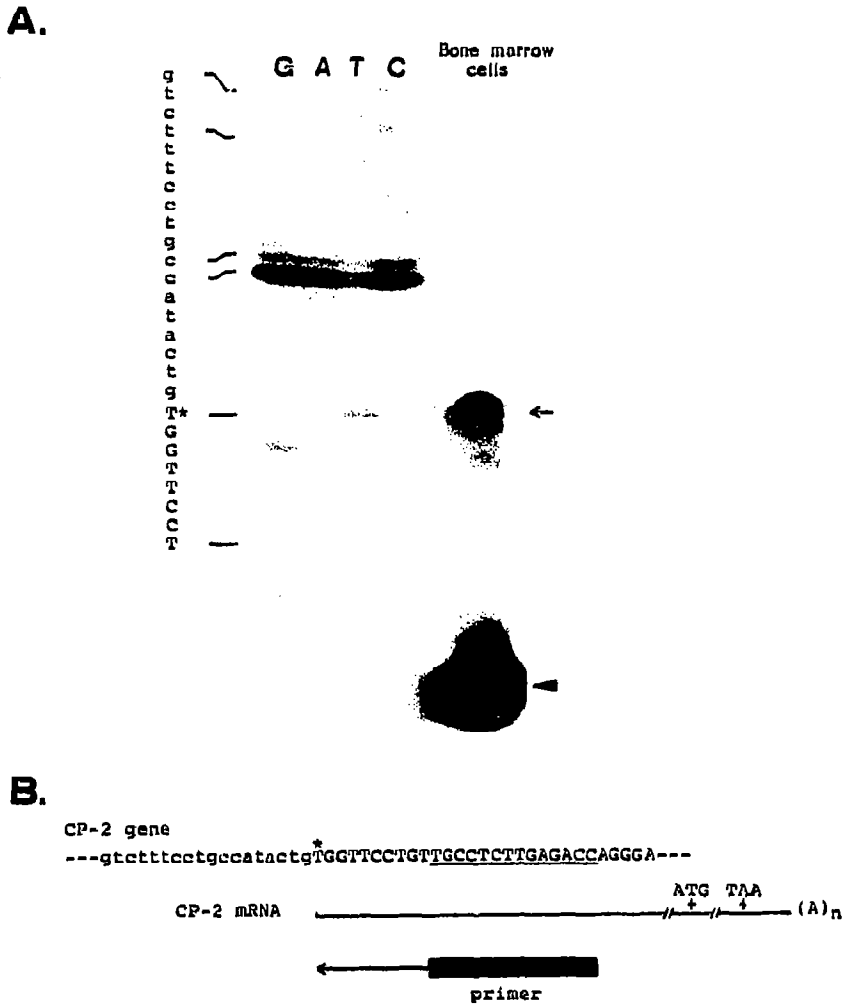

Fig. 3. Identification of the transcription initiation site for the GNCP-2 gene by primer extension analysis. (A) Primer extension evaluation of bone marrow cell mRNA. ^{32}P-labeled oligonucleotide primer (2×10^5 cpm) corresponding to the specific sequence of exon 1 of the GNCP-2 gene shown in Fig. 2 and Fig. 3B was annealed to poly(A)$^+$ RNA (10 μg), and extended by reverse transcriptase. The extended products were analyzed on a 6% polyacrylamide denaturing gel. The length of the products was determined by the comparison with sequencing ladders run alongside the extension reaction. The arrow indicates the major primer extension product, and the corresponding transcription initiation site is marked with an asterisk. The arrowhead indicates the labeled primer. (B) Structures of the GNCP-2 gene and GNCP-2 mRNA. The start codon (ATG), the stop codon (TAA) and the poly(A) tail are shown in GNCP-2 mRNA. The filled box indicates the primer, and the corresponding sequence is underlined in the GNCP-2 gene. The putative transcription initiation site is marked with an asterisk.

REFERENCES

[1] Klebanoff, S.J. (1988) in: Inflammation, Basic Principles and Clinical Correlates (J.I. Gallin, I.M. Goldstein and R. Snyderman, Eds.), Raven Press, New York, pp. 391–444.
[2] Elsbach, P. and Weiss, J. (1988) in: Inflammation, Basic Principles and Clinical Correlates (J.I. Gallin, I.M. Goldstein and R. Snyderman, Eds.) Raven Press, New York, pp. 445–470.
[3] Lehrer, R.I., Ganz, T., Selsted, M.E., Babior, B.M. and Curnutte, J.T. (1988) Ann. Intern. Med. 109, 127–142.
[4] Spitznagel, J.K. (1990) J. Clin. Invest. 86, 1381–1386.
[5] Lehrer, R.I. and Ganz, T. (1990) Blood 76, 2169–2181.
[6] Boman, H.G. (1991) Cell 65, 205–207.
[7] Yamashita, T. and Saito, K. (1989) Infect. Immun. 57, 2405–2409.
[8] Nagaoka, I., Someya, A., Iwabuchi, K. and Yamashita, T. (1991) FEBS Lett. 280, 287–291.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
[10] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[11] Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
[12] Rackwitz, H.-R., Zehetner, G., Frischauf, A.-M. and Lehrach, H. (1984) Gene 30, 195–200.
[13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[14] McKnight, S.L. and Kingsbury, R. (1982) Science 217, 316–324.
[15] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[16] Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383.
[17] Mount, S.M. (1982) Nucleic Acid Res. 10, 459–472.
[18] Kozak, M. (1984) Nucleic Acid Res. 12, 857–872.

[19] Faisst, S. and Meyer, S. (1992) Nucleic Acid Res. 20, 3–26.
[20] Kramps, J.A., van der Valk, P., van der Sandt, M.M., Lindeman, J. and Meijer, C.J.L.M. (1984) J. Histochem. Cytochem. 32, 389–394.
[21] Takahashi, H., Nukiwa, T., Basset, P. and Crystal, R.G. (1988) J. Biol. Chem. 263, 2543–2547.
[22] Bainton, D.F., Ullyot, J.L. and Farquahar, M.G. (1971) J. Exp. Med. 134, 907–934.
[23] Fouret, P., du Bois, R.M., Bernaudin, J.-F., Takahashi, H., Ferrans, V.J. and Crystal, R.G. (1989) J. Exp. Med. 169, 833–845.
[24] Takahashi, H., Nukiwa, T., Yoshimura, K., Quick, C.D., States, D.J., Holmes, M.D., Whang-Peng, J., Knutsen, T. and Crystal, R.G. (1988) J. Biol. Chem. 263, 14739–14747.
[25] Morishita, K., Tsuchiya, M., Asano, S., Kaziro, Y. and Nagata, S. (1987) J. Biol. Chem. 262, 15208–15213.
[26] Han, J., Unlap, T. and Rado, T.A. (1991) Biochem. Biophys. Res. Commun. 181, 1462–1468.