Promoter paper

Structural organization and expression of the gaegurin 4 gene of \textit{Rana rugosa}

So Yeon Kwon a, Bradley A. Carlson b, Jin Mo Park a, Byeong Jae Lee a,.*

\textit{a} Laboratory of Molecular Genetics, Institute of Molecular Biology and Genetics, Seoul National University, Seoul 151-742, South Korea

\textit{b} Section on the Molecular Biology of Selenium, Laboratory of Basic Research, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received 11 January 2000; received in revised form 17 March 2000; accepted 22 March 2000

Abstract

Gaegurin 4 (GGN4) is a member of the antimicrobial peptide subfamily isolated from the skin of \textit{Rana rugosa}. We cloned \textit{gDNA} encoding GGN4 to study its gene organization and regulation of expression. The GGN4 gene occurs in single copy in the \textit{R. rugosa} genome and contains a single intron of about 3.4 kb. The transcription start site is located 68 bases upstream of the translation initiation codon. The GGN4 gene was expressed both in \textit{Xenopus} kidney epithelial cells (A6) and in \textit{Xenopus} oocytes using the chloramphenicol acetyltransferase reporter gene system. The 5' flanking region of the GGN4 gene contains a dl binding site that is known to regulate acute phase immune response related gene expression in mammals and insects. The dl protein bound specifically to the GGN4 gene promoter region. Mutants that serially delete the 5' flanking region show that removal of the dl binding site inhibited GGN4 gene expression in both A6 cells and \textit{Xenopus} oocytes. From these results, we propose that expression of the GGN4 gene may be regulated by the region containing the dl element which plays a key role in the regulation of antimicrobial peptide genes in \textit{Drosophila} and mammals. ß 2000 Elsevier Science B.V.

Keywords: Gaegurin 4; Antimicrobial peptide; Promoter; Innate immunity; Nuclear factor xB/dl

The concept of innate immunity refers to the first line of host defense that serves to limit infection in the early hours after exposure to microorganisms [1]. Most vertebrates are armed with antimicrobial peptides to properly control invading microorganisms [2]. Endogenous antimicrobial peptides produced from granular glands within amphibian skin are fundamental contributors to innate immunity against microbial infections. A number of peptides from the skin of various amphibians have been found to have a broad spectrum of antimicrobial activity [3]. Six antimicrobial peptides, named gaegurins, were also isolated from the skin of a Korean frog \textit{Rana rugosa} [4]. Among them, gaegurin 4 (GGN4) is composed of 37 amino acids [3,4]. The GGN4 peptide manifests antimicrobial activity against Gram-positive and Gram-negative bacteria, fungi and protozoa, but has only a slight detrimental effect on human red blood cells [4]. The peptides are supposed to be stored in the secretory glands of epithelial tissues and released upon microbial infection. cDNA cloning studies revealed that the GGN4 precursor polypeptide has a unique tripartite structure: a putative signal sequence at the NH2-terminus followed by an acidic spacer region rich in glutamic and aspartic acids and a mature peptide at the COOH-terminus [5]. The GGN4 peptide is also known to be expressed highly in the skin like other amphibian peptide antibiotics [5,6].

Elucidation of the regulatory mechanisms controlling the rapid synthesis of these antimicrobial peptides is an important challenge in the field. Many immunity related genes are under the dual control of the rel containing transcription factor, nuclear factor xB (NF-xB), and an inhibitor, IxB [7]. Thus far, mammalian genes encoding antimicrobial peptides have been found to contain promoter motifs that potentially bind NF-xB-like factors [1,8], but no functional studies have yet been reported. In contrast, \textit{Drosophila} has become a model system for both in vitro and in vivo studies of molecular mechanisms elucidating NF-xB/IxB regulation [9-11]. In insects, this type of mechanism is used for control of both immunity and development [12]. Recent observations have high-
lighted similarities between pathogen recognition, signaling pathways, and effector mechanisms of innate immunity in *Drosophila* and mammals [1,8], suggesting that they have a common ancestry of these defenses.

Despite the fact that many antimicrobial peptides from various organisms have been isolated and their cDNA sequences identified [13], the genomic structure and the mechanism of expression are poorly understood, particularly in amphibians [14]. To further investigate gene organization of GGN4 and the regulation of its expression, we constructed a genomic library and isolated a clone encoding the GGN4 gene.

To isolate GGN4 genomic clones, a bacteriophage genomic DNA library (5 × 10⁵ independent clones) was constructed by cloning *R. rugosa* skeletal muscle DNA into the XhoI site of the λ FixII vector (Stratagene). The library was screened by hybridization with fragments of the GGN4 full-length cDNA [4] as a probe using the ECL direct nucleic acid labeling and detection system (Amersham).

Hybridization was carried out at 42°C for 16 h and membranes were washed four times with buffer containing 6 M urea at 42°C for 40 min. The clone containing the GGN4 gene was isolated after four rounds of plating and hybridization. The restriction map of the clone was obtained by digestion with several enzymes (Fig. 1A). The DNA from this clone was digested with *SalI/XhoI/EcoRI* and then hybridized with full-length cDNA or with 111 bp DNA encoding the entire mature peptide region that was produced by polymerase chain reaction (PCR) to identify exon and intron regions. Genomic DNA was digested with various restriction endonucleases, the DNA fragments subcloned separately into pBluescript II KS(+) and the sequences determined by automated fluorescent DNA sequencing (Perkin Elmer Applied Biosystem) as shown in Fig. 1A. The DNA region is 5910 bp which includes a 1938 bp 5' upstream region (Fig. 1B). By comparing the genomic sequence with the sequence of GGN4 cDNA, an intron was found to be located in the acidic prosequence. A canonical GT/AG exon–intron junction sequence is found at the extremities of the intron. The two exons consist of 74 bp and 163 bp, respectively, and are interrupted by a large intron of 3430 bp. The first exon encodes 25 amino acid residues, comprising the 22 residue signal peptide. Exon 2 contains the remaining coding region (55 amino acid residues), which includes part of the acidic peptides (residues 23–39) and one copy of the GGN4 peptide (residues 40–75). The 3' flanking region of the gene includes one polyadenylation site (AATAAA) located 66 bp downstream of the TAA stop codon.

The genomic organization of the GGN4 gene was analyzed by Southern blot analysis. Genomic DNA was extracted from *R. rugosa* skeletal muscle and subjected to digestion with various restriction endonucleases. Digested DNAs were electrophoresed on a 1% agarose gel, the fractionated DNA transferred to a nylon membrane and the membrane hybridized with the full-length cDNA probe. As shown in Fig. 2, only a single band hybridized with the probe with each restriction enzyme (a 5.0 kb, a 0.4 kb, a 1.6 kb and a 3.0 kb fragment with *EcoRI*, *SalI*, *MboI* and *TaqI*, respectively). Each restriction site exists in the sequenced genomic GGN4 clone (Fig. 1B). These results suggest that the GGN4 gene is encoded in single copy within the genome of *R. rugosa*.

To determine the GGN4 transcription start site of the gene, we performed primer extension analysis. Total RNA was extracted from *R. rugosa* skin. An 18-mer antisense oligonucleotide (5'-AAATAACAGGGATTTCTT-3') complementary to the first exon of the GGN4 gene was annealed to 1 or 5 µg of total RNA purified from *R. rugosa* skin and extended using Moloney murine leukemia virus (MMLV) reverse transcriptase (Promega) [15].
The results shown in Fig. 3 indicate that the transcription start point is located at 68 bp upstream of the initiation methionine codon (i.e., at position −68).

Sequence analysis of the 5′ flanking region and the immediate downstream position of the transcription start site shows that a typical TATA or initiator element was not present in the proximal promoter region where basal promoter elements are normally expected to occur. It is tempting to speculate that non-canonical basal promoter elements with no known sequence characteristics may direct assembly of the RNA polymerase II general transcription factors on the GGN4 promoter. The 5′ flanking region further upstream of the transcription start point contains several binding consensus sequences for transcription activators that are known to regulate inducible or tissue specific expression of immune response proteins in mammals and insects. For example, a dorsal binding site is located at position −320, a NF-IL6 binding site at −228, and a GATA-1 element at −399 and at −1298 (Fig. 1B). These elements have been implicated in induction of acute phase proteins in mammalian cells [1,8] and antifungal proteins in the fat body cells of fruit flies [16,17].

To localize transcription regulatory sites, a series of deletion mutants were prepared and their promoter activities were determined with the chloramphenicol acetyltransferase (CAT) reporter system (Promega). Mutants in which the 5′ flanking regions were serially deleted were constructed by fusing the pCAT-Basic vector which did not contain a promoter. The GGN4 gene fragment containing the region from −1938 to +65 was subcloned upstream of the CAT gene in vector pCAT-Basic to make pCAT/1.95. Five deletion mutants were constructed by digestion of pCAT/1.95 with various restriction enzymes and re-ligation. The 5′ flanking region of GGN4 was deleted up to −1050 (pCAT/1.05), −750 (pCAT/0.75), −450 (pCAT/0.45), −350 (pCAT/0.35) and the last mutant contains the DNA fragment from −1950 to −1051 (Fig. 4A). Supercoiled plasmid DNA encoding the wild type gene and each gene deletion was prepared by alkaline lysis and purified by CsCl density gradient centrifugation. The A6 cell line that originated from *Xenopus* kidney epithelial cells was maintained in Leivovitz L15 medium (Gibco BRL) supplemented with 10% fetal bovine serum, 100 μU/ml of penicillin and streptomycin at 25°C. The cells were plated at a density of approximately 2×10⁵ cells/35 mm dish and were transfected with 2 μg of plasmid DNA using 25 μl lipofectin (Gibco BRL) as described in the supplier’s manual [18]. 1 μg of pCAT-Basic vector was used as a control and each pCAT promoter deletion mutant was transfected into A6 cells. Forty-eight hours after adding plasmids to the medium, CAT activities were determined from cell lysates in 0.25 M Tris buffer (pH 7.8). For each construct, CAT activity was expressed as a relative value and the values were obtained by taking mean values of three separate experiments. The acetylated and non-acetylated forms were separated by thin layer chromatography (Merck) and then quantified using a phosphorimager (Fuji). As shown in Fig. 4B, the intact GGN4 promoter (pCAT/1.95) showed inhibited promoter activity. However, the truncated forms, pCAT/0.75 and pCAT/0.45, derepressed the CAT activity approximately 37.0 and 12.6%, respectively. These results suggest that the 5′ flanking region from −1938 and −750 contains a negatively regulated element. Since the promoter activity in pCAT/0.45 was reduced and virtually removed in pCAT/0.35, a positive regulatory element (promoter) may be located near −450.

The strong and rapid induction of antimicrobial peptide genes in *Drosophila* fat body cells after septic injury has served as a model system for the analysis of innate immunity [1,19]. *Drosophila* produces at least seven distinct antimicrobial peptides [1,7] and much of their promoter structure is known to have binding sites of Rel-like protein such as Dif or Dorsal [20,21]. Both molecular and bio-
chemical approaches have shown that inducibility of these genes is directly regulated by the Rel protein [22]. Rel proteins are maintained in the cytoplasm in association with IκB proteins and are activated by phosphorylation and degradation of IκB proteins. This sequence of events leads to the nuclear translocation of Rel proteins and the transcription activation of target genes [22]. As discussed above, the promoter region of the GGN4 gene contains a recognition site for nuclear factors of dl (Fig. 1B). The fact that pCAT/0.35, in which the dl binding site has been deleted, lost its promoter activity suggests that regulation by Rel protein plays an important role in GGN4 gene expression.

The transcriptional activities of deletion mutants were also measured in Xenopus oocytes after microinjecting them (Fig. 4B) by the standard microinjection technique [23]. In this in vivo assay system, the promoter activity of pCAT/0.45 was completely abolished indicating that transcription is more thoroughly controlled intracellularly. It is known from other studies that bacterial lipopolysaccharide can induce innate immunity related genes [24]. To test whether expression of the GGN4 gene is induced when the cell is challenged with bacteria, heat-killed Gram-negative Pseudomonas aeruginosa was added to cultured cells (see also Fig. 4B) [24]. CAT assays from cultured cells revealed that activity levels were elevated approximately two-fold in response to heat-killed bacteria, measured in experiments carried out in triplicate.

To obtain more direct evidence that the hypothetical promoter element of the dl binding site is involved in the transcription of GGN4, an electrophoretic mobility shift assay was performed with double-stranded oligonucleotides containing the GGN4 promoter derived putative dl binding site. The oligonucleotide ligand corresponding to the sequence from 3284 to 3364 contains a sequence

---

Fig. 4. Expression of the GGN4 gene. (A) Deletion mutants in the 5' flanking region of GGN4 fused to the CAT reporter gene that were used in transfection assays are shown. The pCAT-Basic (Promega) construct is used with a backbone vector. (B) CAT activities in A6 cells are shown. A6 cells were transiently transfected with the CAT constructs containing different lengths of the 5' flanking region of the GGN4 gene. Heat treated P. aeruginosa at 65°C was added to cultured cells 24 h after transfection (10^6 cells/ml). (C) CAT activities in Xenopus oocytes are shown. CAT activity was assayed in Xenopus oocytes following microinjection of the indicated constructs. The level of CAT activity shown above was determined by Imagequant program supported by a phosphorimager (Fuji).

---

Fig. 5. Electrophoretic mobility shift assays using purified GST-dl recombinant protein. The binding activity of the recombinant proteins (10 ng) to the DNA fragment amplified from −284 to −364 (80 bp including the dl binding site) of the GGN4 gene is shown. Unlabeled PCR amplified probe was used as a specific competitor (SC) in a dose dependent manner. λ DNA was used as a non-specific competitor (NC).

---
with homology to the dl binding site (5′-GGAGAA-ACCC-3′). Recombinant dl protein that was used for binding was produced in *Escherichia coli* as glutathione S-transferase (GST) fusion protein. pGST-dl vector (kindly donated by Michael Levine) was transformed into *E. coli* DH5α, the gene induced with 0.2 mM IPTG and the GST fused dl protein was isolated with glutathione S-4B resin (Pharmacia). As shown in Fig. 5, recombinant dl produced retarded bands and the binding was inhibited by increasing amounts of the specific inhibitor (see the legend to Fig. 5) in a quantity dependent manner. A non-specific competitor, λ DNA, however, did not affect the binding activity. These results strongly support the idea that the dl binding site plays a regulatory role in the transcription of the GGN4 gene.

Acknowledgements

This study was supported by a grant of the Korea R&D Project, Ministry of Health and Welfare, Republic of Korea (HMP-97-B-2-0021) to BJL. S.Y. Kwon was supported by BK21 Research Fellowship from the Korean Ministry of Education.

References