Inducible gene expression of moricin, a unique antibacterial peptide from the silkworm (*Bombyx mori*)

Seiichi FURUKAWA*, Hiromitsu TANAKA†, Hiroshi NAKAZAWA†, Jun ISHIBASHI†, Toshio SHONO* and Minoru YAMAKAWA†[1]

*Laboratory of Applied Zoology, Institute of Agriculture and Forestry, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan, and
†Laboratory of Biological Defense, National Institute of Sericultural and Entomological Science, Tsukuba, Ibaraki 305-8634, Japan

Molecular cloning of cDNAs encoding moricin, a novel antibacterial peptide from the silkworm (*Bombyx mori*), was performed using a fat-body cDNA library. A reverse-transcription PCR product encoding a partial nucleotide sequence of moricin was used as a probe. Nucleotide sequencing of four positive clones revealed two types of moricin cDNAs designated moricin 1 and 2. cDNAs for moricin 1 and 2 shared 97.2% identity in their nucleotide sequences. Although one amino acid residue (Phe^6) of moricin 1 in the putative signal peptide was replaced with Lys^6 in moricin 2, amino acid sequences of their mature portions were identical. Moricin gene expression in *B. mori* larvae injected with *Escherichia coli* was observed in fat-bodies, haemocytes and the Malpighian tubule, but not in other tissues like the midgut and silk glands. Accumulation of moricin gene transcripts induced by *E. coli* reached a maximum level 8 h after injection and persisted up to 48 h. It was confirmed that lipopolysaccharide (LPS) and lipid A, which are cell-wall components of *E. coli*, triggered moricin gene expression. Comparison of gene expression between moricin 1 and 2 by PCR using specific primers indicated that moricin 2 gene was more strongly expressed than moricin 1 gene. A genomic clone encoding moricin 2 was screened from a *B. mori* genomic library using a moricin cDNA as a probe. Regulatory motifs for gene expression such as nuclear-factor-κB-binding-site-like sequence (κB site) and nuclear-factor–interleukin-6-binding-site-like sequence (NF-IL-6 site) were found in the 5′-upstream regulatory region. An electrophoretic-mobility-shift assay revealed that there are bacterial LPS-inducible nuclear proteins that can bind to the κB site and other sites in the regulatory region.

Key words: cDNA and gene cloning, insect immunity, nuclear binding proteins, regulatory motifs, tissue-specific expression.

INTRODUCTION

Insects have an efficient innate immunity against microbial infection, although they have not acquired the adaptive immune systems that play an extremely important role in vertebrates [1–4]. Antibacterial peptides, one group of factors involved in insect humoral defence reactions, are rapidly secreted into the haemolymph upon bacterial infection [5–8]. To date, more than 150 antibacterial peptides and polypeptides have been isolated from different insect species. These antibacterial peptides and polypeptides are classified into five major groups: cecropins, insect defensins, attacin-like (glycine-rich) proteins, proline-rich peptides and lysozymes [3,7,8]. Moricin, which was isolated from the silkworm (*Bombyx mori*), consists of 42 amino acids. This peptide does not have a significant similarity in amino acid sequences with other known antibacterial peptides belonging to the classification described above [9]. Moricin shows antibacterial activity against both Gram-positive and Gram-negative bacteria and was demonstrated to form ion channels in the bacterial membrane [9]. In addition, an artificial moricin gene was synthesized and expressed in *Eschrichia coli*, and it was confirmed that the physical and biological characteristics of the recombinant moricin were identical with those of natural moricin [10]. The recombinant moricin strongly suppressed the growth of methicillin-resistant *Staphylococcus aureus* isolated from patients [11].

Although the characteristics of moricin have been extensively analysed, data on gene expression are still missing. Thus we cloned two moricin cDNAs and a gene, determined their nucleotide sequences and performed an electrophoretic-mobility-shift-assay (EMSA) to analyse the induction mechanisms of gene expression. Here we present data on the structure and expression of a gene encoding moricin, a unique antibacterial peptide from the silkworm.

EXPERIMENTAL

Biological materials

Silkworms (*Bombyx mori*; Tokai × Asahi and Chu 602 strains), were reared on artificial diet (Nihonnosanko) at 25 °C. *Escherichia coli* K12 strain JM 109 [12], grown in Luria–Bertani medium [13] at 37 °C with shaking, was used for injections into silkworms.

Preparation of a probe

Reverse-transcription PCR (RT-PCR) was performed using the following primers synthesized based on the amino acid sequence of moricin [9] to prepare a probe for the screening of moricin

Abbreviations used: RT-PCR, reverse-transcription PCR; κB site, nuclear-factor-κB-binding-site-like sequence; NF-IL-6 site, nuclear-factor–interleukin-6-binding-site-like sequence; EMSA, electrophoretic-mobility-shift assay; LPS, lipopolysaccharide; poly(A)$^+$, polyadenylated.

[1] To whom correspondence should be sent (e-mail yamakawa@nises.affrc.go.jp).

The nucleotide sequence data reported here have been submitted to the GSDB/DDBJ/EMBL/NCBI Nucleotide Sequence Databases and are available under accession numbers AB0003915 and AB014092 for moricin 1 cDNA and moricin 2 cDNA respectively and AB019538 for the moricin 2 gene.

cDNA: 5′-CGGAATTCGC(T/C/A/G)AA(A/G)AT(T/C/A)-CC(T/C/A/G)AT(T/C/A)AA(A/G)GC-3′ (forward primer derived from Ala[1] to Ala[7]) and 5′-CGGAATTCCA(A/G)TG-(T/C)TT(A/T/G/C)C(G/T)(T/C)TT(T/C)TT(A/T/G/C)GG-(T/C) TT-3′ (reverse primer derived from Lys-36 to His-42). The reaction mixture was kept at 94 °C for 1 min and 25 cycles of RT-PCR (94 °C for 30 s, 50 °C for 30 s, 72 °C for 3 min) were carried out. The resultant PCR fragment (130 bp) was subcloned into a T-vector (InVitrogen, San Diego, CA, U.S.A.) and sequencing was carried out according to dye-terminator cycle-sequencing methods using a DNA sequencer (ABI 373A; Applied Biosystems, Foster City, CA, U.S.A.). On the basis of the nucleotide sequence, the following new primers were synthesized: 5′-CGATAAAGGCCATTAAGACTG-3′ (forward primer derived from Pro-4 to Val-11) and 5′-TCGTTTCTTAGGCTT-CAAGAA-3′ (reverse primer derived from Phe[34] to Arg[40]). The sample was kept first at 94 °C for 1 min, then 25 cycles of RT-PCR (94 °C for 1 min, 58 °C for 2 min, 72 °C for 3 min) were performed. The resultant PCR fragment (110 bp) was subcloned and sequencing was carried out as described above.

Probes for EMSA were prepared based on the sequence of the 5′-upstream region of the moricin 2 gene. The nucleotide sequences of two probes containing the nuclear-factor-κB-binding-site-like sequence (κB site) or the nuclear-factor–interleukin-6-binding-site-like sequence (NF-IL-6 site) were as follows: 5′-TTCATTGGGATTCAGTAAAGTCCCGGAGTT-3′ (M2κB) and 5′-CTGCAGTGAGTGGAAATCTTACGAA-ATCCA-3′ (M2IL-6), where the underlined regions indicate the κB site and the NF-IL-6 site in that order. To prepare other probes, the 5′-upstream region of moricin 2 gene was digested with *Pst*I and *Sac*I and the resultant three DNA fragments (115, 283 and 203 bp) were separated by electrophoresis (see Figure 7 below). These three fragments, designated M2-a (for 115 bp), M2-b (for 283 bp) and M2-c (for 203 bp), were used as probes for the EMSA.

Screening of moricin cDNAs and genomic clones

The 110 bp DNA fragment was labelled with DIG-conjugated dUTP (Boehringer Mannheim). We screened 1×10^5 plaques of *B. mori* fat-body cDNA library using the PCR fragment as a probe. Ten positive clones were obtained in the first screening. Screening conditions were the same as those described previously [14]. These clones were further screened with the same probe and all the clones remained positive.

Moricin genomic clones were screened from 6×10^4 plaques of *B. mori* genomic library [15] with a probe (320 bp) synthesized by RT-PCR under the same conditions as those described above using the following primers; 5′-TTGTGGCAATGTCTCTGG-TG-3′ (forward primer) and 5′-TAGTACACAGTTGTCGTT-GC-3′ (reverse primer). The probe was labelled with DIG-conjugated dUTP. Two positive plaques were obtained in the first screening and remained positive in the second screening. Southern-blot analysis of these two clones revealed that they had the same restriction-enzyme cleavage sites (results not shown).

Nucleotide sequencing

cDNA inserts were excised by *Eco*RI from vectors and the length of the inserts was compared by 1.2 %-agarose gel electrophoresis. Four out of ten clones, which contained the longest inserts, were subcloned into pBluescript II (SK^+) vectors (Stratagene). The nucleotide sequence of the four subcloned DNA fragments was determined by a dye-terminator cycle-sequencing method using a DNA sequencer (ABI 373A).

The phage DNA was digested with *Hin*dIII and *Sal*I and subcloned into the same vector described above for determination of the nucleotide sequence of the moricin gene. The nucleotide sequencing was performed as described above.

Northern blotting

Total RNA was extracted from fat-bodies, haemocytes, midgut, the Malpighian tubule and silk glands from fifth-instar larvae (3 days) 8 h after immunization with *E. coli* (2×10^7 cells) using an Isogen kit (Wako, Tokyo, Japan) to examine tissue-specific expression. Extracted RNA samples (7 μg) were electrophoresed and transferred on to a GeneScreen Plus membrane (DuPont) and hybridized with the same probe used for screening of moricin genomic clones. Prehybridization and hybridization were carried out at 50 °C for 2 h and overnight respectively. Other conditions were the same as those described previously [14]. Silkworm larvae were injected with *E. coli* (2×10^7 cells), LPS (Difco, 20 μg dissolved in LPS-free physiological saline from Otsuka Pharmaceutical Corp., Tokyo, Japan) or lipid A (Daiichi Kagaku, Tokyo, Japan; 10 μg dissolved in 0.025 % triethylamine) to examine the effects of *E. coli* cell-wall components on the induction of moricin gene expression. Fat-bodies were excised 8 h after injection, total RNA was extracted and Northern blotting was performed as described above. For the analysis of the time course of moricin gene expression, *B. mori* larvae were immunized with *E. coli* and fat-bodies were excised at the indicated time intervals (see Figure 3 below). Other experimental conditions for Northern blotting were the same as those described above.

RT-PCR for the quantitative analysis of two moricin gene transcripts

The following specific primers were designed and synthesized to distinguish gene transcripts of the two moricin genes: 5′-TGGCAATGTCTCTGGTGT-3′ (common forward primers), 5′-GTAAGTACTACACAGGGT-3′ (reverse primer for moricin 1) and 5′-GTAAGTACTACAAAGGGG-3′ (reverse primer for moricin 2). PCR was performed for 8 min at 94 °C and for each 30 s at 94 °C, 55 °C and 72 °C. An aliquot (10 μl) of reaction mixture (total volume 55 μl) after 27, 31 and 35 PCR cycles was analysed by 1.2 %-agarose-gel electrophoresis. Other experimental conditions for injection of silkworm larvae with *E. coli*, RNA preparation and first-strand cDNA synthesis were the same as those described above.

EMSA

In order to prepare crude nuclear extracts from fat-bodies, *B. mori* larvae (Chu 602 strain) were injected with 20 μg of LPS (Difco) dissolved in physiological saline (Otsuka Pharmaceutical Corp.). As a control, non-injected larvae were used. Fat-bodies from the larvae were excised 4 h after LPS injection. All the following procedures were carried out at 0–4 °C. The first half of the crude nuclear extract preparation depended principally on the methods of Mine et al. [16]. Briefly, 10 g of fat-bodies were washed twice in PBS (137 mM NaCl/2.7 mM KCl/8.1 mM Na_2PO_4,$12H_2O$/1.5 mM KH_2PO_4, pH 7.4–7.6) and suspended in 10 ml of buffer A [10 mM Hepes/NaOH (pH 7.9)/10 mM KCl/1.5 mM $MgCl_2$/1 mM dithiothreitol]. The suspension was diluted with 10 ml of buffer A containing 0.5 % Nonidet P40, shaken softly for 10 min and centrifuged for 10 min at 6300 ***g***. The latter half of the procedure depended principally on the methods of Suzuki et al. [17]. Briefly, the resultant precipitate was suspended in 10 ml of extraction buffer [40 mM Tris/HCl (pH 7.9)/5 mM $MgCl_2$/12.5 % sucrose/25 % glycerol/2 mM dithiothreitol]. This

suspension was homogenized by two or three strokes with a Dounce homogenizer using a tight-fitting pestle and centrifuged for 10 min at 6300 ***g***. The precipitate was suspended in 10 ml of the extraction buffer and homogenized by 15 strokes with the same homogenizer. Satd. (4 M) $(NH_4)_2SO_4$ solution (0.1 vol.) was added to the homogenate within 15 min under stirring. The extract was centrifuged for 3–4 h at 160000 ***g*** and $\frac{2}{3}$ vol. of the supernatant was transferred. Proteins in the supernatant were precipitated by the addition of pulverized $(NH_4)_2SO_4$ to 0.33 g/ml under stirring. After additional stirring for 30 min, the proteins were recovered by centrifugation for 20 min at 7200 ***g***. The precipitate was dissolved or suspended in dialysis buffer [20 mM Hepes/NaOH (pH 7.9)/100 mM KCl/12.5 mM $MgCl_2$/0.1 mM EDTA/17 %(v/v) glycerol/2 mM dithiothreitol] and dialysed against the dialysis buffer with one change overnight. The dialysis residue was centrifuged for 60 min at 100000 ***g*** and the supernatant was stored at −80 °C until use. Protein concentration was determined with a Bio-Rad Protein Assay kit using a BSA as a standard.

The DNA probes M2κB and M2IL6 were labelled with [γ-^{32}P]ATP (ICN) using T_4 polynucleotide kinase (Nippon-gene, Tokyo, Japan). Other probes (M2-a, -b and -c) were labelled with [α-^{32}P]dCTP (ICN) using klenow fragment (Nippon Gene). These labelled probes were purified with a NICK column (Pharmacia). Preincubation for the binding reaction was carried out for 10 min at room temperature with 2.6 μg of nuclear extract in the reaction mixture (total 18 μl) containing 4 μl of 5× binding buffer [0.1 M Hepes/NaOH (pH 7.8)/0.5 M NaCl/5 mM EDTA/25 % glycerol/2.5 mM dithiothreitol], 2 μl of poly dI·dC (1 mg/ml) and 1 μl of BSA (2 mg/ml). For the competition experiment, 10–60-fold excess amounts of a non-labelled probe were added to the reaction mixture just before the preincubation (see Figure 7 below). A 2 μl volume of ^{32}P-labelled DNA probe (25 fmol/μl) was added to the reaction mixture and incubated for 20 min at room temperature.

The 5 % native polyacrylamide gel containing 5 % glycerol was prepared with 0.5× Tris/borate buffer [0.089 M Tris/borate (pH 8.0)/0.089 M boric acid/2 mM EDTA]. A pre-run was carried out at 18–25 mA for 30 min at 4 °C in 0.5× Tris/borate buffer. After gel electrophoresis of the sample at 18 to 25 mA for 150 min at 4 °C, the gel was dried and exposed overnight to X-ray film at −80 °C.

RESULTS

cDNA cloning and nucleotide sequencing

In order to screen moricin cDNA clones, we first prepared a probe by RT-PCR using primers based on the amino acid sequences of moricin [9]. A single DNA band (110 bp) appeared, and nucleotide sequencing of this PCR product indicated that it encodes moricin (results not shown). We screened a fat-body cDNA library to obtain a full-length cDNA using the PCR product as a probe. Ten clones from 1×10^5 plaques showed positive signals, and four out of ten positive clones, showing maximum-length cDNA inserts, were subjected to nucleotide sequencing. The results revealed the presence of two cDNA types that differed in nucleotide sequences but encoded the same mature portion of moricin (Figure 1). Thus we designated these two clones moricin 1 and moricin 2. Overall identity in nucleotide sequences between moricin 1 and moricin 2 was 97.2 %. Only one amino acid residue in the putative signal peptide sequence of these two moricin clones was different, namely, Phe^6 of moricin 1 was replaced by Lys^6 in moricin 2. Changes in nucleotide sequences were clustered in the 3′-non-coding region. A poly-adenylated [poly(A)$^+$] addition signal (ATAAA) was identified at the typical sites of both cDNAs.

Tissue specificity of moricin gene expression

A longer probe (322 bp) for Northern-blot analysis was newly prepared and used to examine the tissue specificity of moricin gene expression. For this purpose, silkworm larvae were injected with *E. coli* and various tissues were excised. Northern blotting was performed with RNA samples extracted from the tissues. The results indicated that the fat-body is the main site for moricin gene expression (Figure 2). Weak signals were also observed in RNA samples from haemocytes and the Malpighian tubules (Figure 2).

Time course of moricin gene expression

Kinetics of moricin mRNA accumulation were analysed by Northern blotting after injection of *E. coli* into fifth-instar larvae. As shown in Figure 3, a weak signal appeared 1 h after the injection, and the intensity gradually increased with time and reached a maximum level 8 h after injection. Strong gene expression persisted up to 48 h after injection of *E. coli*.

Effect of cell-wall components on the induction of moricin gene expression

LPS and lipid A, which are known to be effective for the induction of gene expression of *B. mori* antibacterial peptides [18], were examined by Northern-blot analysis to test their ability to induce the expression of the moricin genes. As controls, saline (solvent for LPS) and triethylamine (solvent for lipid A) were also injected into *B. mori* larvae. As shown in Figure 4, LPS and lipid A induced the accumulation of moricin gene transcripts as well as *E. coli* cells. On the other hand, RNA samples from untreated or control larvae injected with saline or triethylamine did not show positive signals, suggesting that the expression of the moricin genes can be induced by bacterial infection but not by wounding.

Comparison of the extent of gene expression between moricin 1 and 2

We quantitatively analysed the expression of the two moricin genes by RT-PCR using specific primers. For this purpose we designed specific reverse primers for moricin 1 and 2 based on the divergent nucleotide sequences in the 3′-non-coding regions. PCR products were analysed by agarose-gel electrophoresis after three different PCR cycles. Results of the quantitative analysis of both moricin gene transcripts revealed that moricin 2 gene was expressed more strongly than the moricin 1 gene (Figure 5).

Nucleotide sequence of moricin 2 gene

As two positive clones out of 6×10^4 plaques were shown to be the same by Southern blotting, the nucleotide sequence of one clone was determined. The clone contained a 6.5 kbp insert. The results revealed that moricin 2 gene consisted of 2452 bp and the deduced amino acid sequence was coincident with that of moricin 2 (Figure 6). An intron (883 bp) was detected between Val^{39} and

```
Bmmor 1   GCAAAAACAGTAAACCGCGCAGTTATTTAAAACATGAATATTTTAAAATTTTTCTTTGTT   60
Bmmor 2   .....................T..............C...........                49

Moricin 1                                     M  N  I  L  K  F  F  F  V    9
Moricin 2                                     .  .  .  .  .  L  .  .  .    9

Bmmor 1   TTTATTGTGGCAATGTCTCTGGTGTCATGTAGTACAGCCGCTCCAGCAAAAATACCTATC   120
Bmmor 2   ............................................................   109

Moricin 1  F  I  V  A  M  S  L  V  S  C  S  T  A  A  P  A  K  I  P  I   29
Moricin 2  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .   29

Bmmor 1   AAGGCCATTAAGACTGTAGGAAAGGCAGTCGGTAAAGGTCTAAGAGCCATCAATATCGCC   180
Bmmor 2   ............................................................   169

Moricin 1  K  A  I  K  T  V  G  K  A  V  G  K  G  L  R  A  I  N  I  A   49
Moricin 2  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .   49

Bmmor 1   AGTACAGCCAACGATGTTTTCAATTTCTTGAAACCGAAGAAAAGAAAGCATTAACAAAAG   240
Bmmor 2   ..........................................................G.....  229

Moricin 1  S  T  A  N  D  V  F  N  F  L  K  P  K  K  R  K  H  ***        66
Moricin 2  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .              66

Bmmor 1   AAATTGAGTGAATGGTATTAGATATATTACTAAAGGATCGATCACAATGATATATAGATA   300
Bmmor 2   ............................................................   289

Bmmor 1   GGTCATAAGATGTCAACGTGAATTTATGGATTTTTGTTTTACCCTGTGTAGTACTTACTT   360
Bmmor 2   ......-.....................................C....T.............  348

Bmmor 1   ATAGTCAGTAGGTACATATATTGATTGCAACGACAACTGTGTACTATTTTTTATATTTGG   420
Bmmor 2   ..........---.-.T..A.....................................      404

Bmmor 1   TTCGAAAAGTTGCATTATTAACGATTTTAGAAAATAAAACTACTTTACTTTTAC------   474
Bmmor 2   ..........................................................GAAAAA  466

Bmmor 1   ----                                                          474
Bmmor 2   AAAA                                                          470
```

Figure 1 Nucleotide and deduced amino acid sequences of cDNAs encoding moricin 1 and moricin 2

The nucleotide sequences of moricin 1 and moricin 2 along with the deduced amino acid sequences are shown. Numbers in the left-hand margin denote nucleotide and amino acid numbers. Mature peptide is indicated by **bold** letters. The translation stop codon, TAA, is depicted with three stars (***). The poly(A)$^+$ addition signal, ATAAA, is shown with a wavy line (∼∼∼∼). Deduced amino acid sequences of the two moricins are also indicated by a single letter. Different amino acids are boxed.

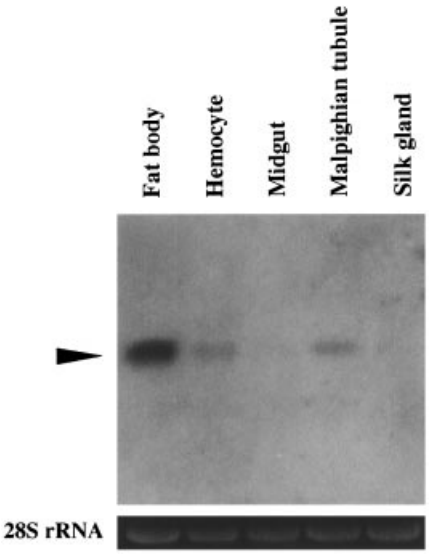

Figure 2 Tissue-specific expression of moricin genes

Total RNA samples (7 μg) from fat-bodies, haemocytes, midgut, Malpighian tubules and silk glands were analysed by Northern blotting. Details of experimental conditions are described in the Experimental section. An internal control (28 S rRNA) is also shown.

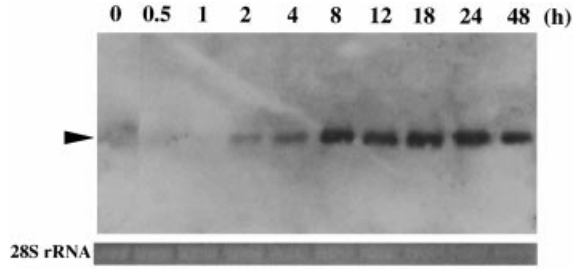

Figure 3 Time course of accumulation of moricin gene transcripts after immunization with *E. coli*

Silkworm larvae were immunized with *E. coli* and fat-bodies were excised at the indicated time intervals. Total RNA samples were analysed by Northern blotting under the same experimental conditions as those described in Figure 2.

Gly^{40}, and a typical intron consensus sequence [19] was conserved in this gene. Two κB sites and an NF-IL-6 site, but no GATA motif, were present in the 5′-upstream regulatory region. The transcription-initiation site (adenine), which is located in downstream of the TATA box (see Figure 6), was determined by the primer-extension method [20] (results not shown).

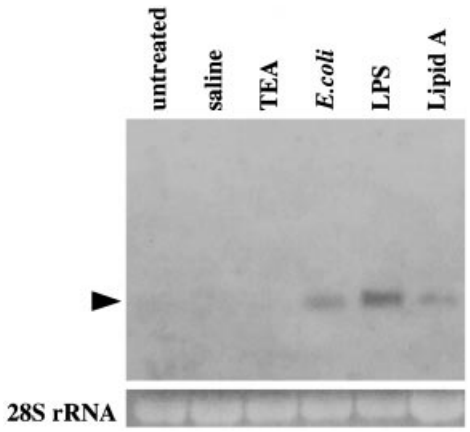

Figure 4 Effects of LPS and lipid A from *E. coli* on the induction of moricin gene expression

Silkworm larvae were injected with *E. coli* (2×10^7 cells), LPS (20 μg) or lipid A (10 μg). As controls, saline (solvent for LPS) or triethylamine (TEA, solvent for lipid A) were injected into larvae. Fat-bodies were excised 8 h after injection. Total RNA extracted from fat-bodies was analysed by Northern blotting. Other experimental conditions were the same as those described in Figure 2.

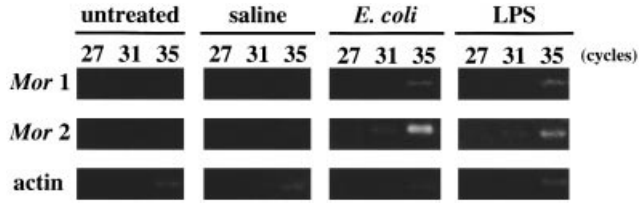

Figure 5 Comparison of the extent of expression of two moricin genes

RT-PCR was performed using specific primers to distinguish the two moricin gene transcripts. The quantity of PCR products after given reaction cycles was analysed by 1.2%-agarose-gel electrophoresis. Actin gene transcripts were also analysed as internal controls under the same experimental conditions. 'untreated', non-injected larvae; 'saline', larvae injected with LPS-free saline. '*E. coli*', larvae injected with *E. coli*; 'LPS', larvae injected with LPS.

```
AGTATTAAACTTTAATATATCTCGTCGATTACTACCGTGAAGATGCAATGCGTATCGATT  -849
CAGAATATCGGTTATATTTTAGTTAAAATTTAGATCTCGTATGCAAATATCGGTGGTCGC  -789
TCTCATGTCTGTGATTTTTATGATTTCACTTGTATCCACTTGTGGGTAACTGGTTGTTCA  -729
TGAGCTCGCTAGTGGGTCATGAGCTTTTCGACTCAGGAAATTAAACAAATAAACAATAGC  -669
AATCTATGTACGCTACGCCAACAATAACAACAAGTCTGCTCCTGTACTAACTTTAGATAG  -609
GGAGGCATTACAAAATAGAGATATATTATTATCATATAGGCTATTTCGTCCTAAAAGCAA  -549
ATATTCGATTGGCACACAAATTATTGTCTGTCGCACATGCTAAAATTCTGTTTAAATATA  -489
ACCTGCTAATTGTGCTCTATATTACTGATAAACATATTGTAAATAATTCTCAGTCTGCAG  -429
TGAGTGGAAATCTTACGAAATCCACTTAATAGCGATAACTGATATCACATCGTTGCGTTG  -369
TTCTAATGTTTACCCGGTATAATTTATTTCTTGATTTTAATGAATATAACCGATCTATAA  -309
GTATCAAATAAGCAAAAACAATTCTATTCATTATTTCTCGCAGTTCACAAACATCTAACC  -249
TGCACTTCATTTCACGTATAAACAATTATTCATTTAAGATTTTGCACTCGACCGATAACA  -189
TTATTTTTATGAGGATTAACCTCTTAAAAGATGAAGTACTTATCTGTTTTATCTTTTAC  -129
GATGTTATGACGTTCACAAGAATACACACATATTTAAACAAACTTGTGATGATGACCCAA   -69
CCGTTCATTGGGATTCAGTAAAGTCCCGGAGTTACATATAAAACGATTGGTAACGGCGGA    -9
ATCAATTCAATGTGTTCTAAAAATTTGCAAAAACAGTAAACCGCGCAGTTATTTAAAATA    52
                (+1)
TGAATATTTTAAAACTTTCTTTGTTTTATTGTGGCAATGTCTCTGGTGTCATGTAGTA    112
M  N  I  L  K  L  F  F  V  F  I  V  A  M  S  L  V  S  C  S
CAGCCGCTCCAGCAAAAATACCTATCAAGGCCATTAAGACTGTAGGAAAGGCAGTCGTAA   172
T  A  A  P  A  K  I  P  I  K  A  I  K  T  V  G  K  A  V
GTAAAAATAAATTTATCTAGAAATATATTTAATTCATTTTCTCAATAATATACTTTTTTC   232
TTTATTTCGGTTGATTTTCTGGAACAAAATATGCTTGCCTTAATACAGGCTATTTGTTAT   292
GATTTTTGTGTTTCAGTGAACAGTATCGAAATGCATATGTTATATATGTATAATAAAAAA   352
AAATTGTGCATAGTTGTATATGTAATTGTTGGTAACGTGAAAGCACTAATGGTTTGAAAA   412
ATATGAGAATATTTTAGATTACTTAATTGCGTAATGGTTAGCATCTCACCTGAGTCTTAA   472
TGGATGTAATGGAGGTTTAAATAAGGTGAACTACACCAGAAAAAAAAACGTATACTGTACT   532
TAAAATTCAAGTTTTGAGGAAATTTTCCAGTATAAAATTTTAATATGATACAGACTTTTT   592
AAAATTGCATTTGTAGGCAGACGAGCATACGGTCTACCTGATGGTAAGTGGTTACCGTCG   652
CTCATGGACGTCAGCAATGCCAGGGGCAGAGCCAAGCCGCTGCCTACAACTGTCTATTAC   712
TCTTCGACGTTTATAGTTTTCGTAAGCGTGGACAATTAATCATAAAGCCTCCTCCTCCTT   772
CCTTGCGTCTTATTCCTCACTGCTGAGGGTAGTGACCACCCCTTTGTATCACGTTCGCAC   832
GCACCATCTTTCTCCATCTATTCTTGTCTCTGGCGGTGTGGAGAGCGTTGTGAAATGTGG   892
AATCAAGAGCGGTGCGGATCTGGTCGGACTAACGTATTGGGCTGCGCCCCCGAGGACTTT   952
TTCCATCTACCTTACCAGTCATGATGAGCCTCTCGAGATTGCTACCATCCTTATATAGTC  1012
GTTCCAAGAAATTATATATTAATATGTTTTTTTTTCAGGGTAAAGGTCTAAGAGCCATC  1072
                                    G  K  G  L  R  A  I
AATATCGCCAGTACAGCCAACGATGTTTTCAATTTCTTGAAACCGAAGAAAAGAAAGCAT  1132
 N  I  A  S  T  A  N  D  V  F  N  F  L  K  P  K  K  R  K  H
TAAGAAAAGAAATTGAGTGAATGGTATTAGATATATTACTAAAGGATCGATCACAATGAT  1192
***
ATATAGATAGGTCATAGATGTCAACGTGAATTTATGGATTTTTGTTTTCCCCTTTGTAGT  1252
ACTTACTTATAGTCAGTTCTTAAATTGATTGCAACGACAACTGTGTACTATTTTTTATAT  1312
TTGGTTCGAAAAGTTGCATTATTAACGATTTTAGAAAATAAAACTACTTTACTTTTACAC  1372
GATGTGTTTGTTTTGAATTGAAATCTTAAATATGAAAGAATAATTTCGTAAATTTACATT  1432
TTTATGGATTCCTTGTCATTCGATTTTCTTGTTTACCTGTGAACTTTTATTGTAAGCTTT  1492
GTGCTCAGTAGTAGTGACATACCTTATATTTTATACTTTGTACCTACCCGTTTTACGAAA  1552
AT                                                             1554
```

Figure 6 Nucleotide and deduced amino acid sequences of the moricin 2 gene

The nucleotide number, which starts from the transcription initiation site (+1), is indicated in the right-hand margin of the panel. TATA box and the poly(A)$^+$ addition signal (ATAAA) are shown with ***bold italic*** letters. κB site and NF-IL-6 site are indicated by double and single underlines respectively. Stars (***) denote the translation stop codon. The deduced amino acid sequence is expressed by a single letter and mature peptides are shown with **bold** letters.

Nuclear protein binding to the 5′-upstream regulatory region

EMSA was performed with five different probes and the crude nuclear extracts from fat-bodies excised from LPS-injected silkworm larvae to analyse the induction mechanisms of moricin gene expression. As shown in Figure 7(A), LPS-inducible nuclear proteins were found to bind the κB probe. In one case the binding was observed with a nuclear protein from non-injected larvae and an enhanced signal was detected with the nuclear extract from LPS-injected larvae. In the other case, binding was not observed at all with the sample from non-injected larvae, but a strong signal was found with the sample from LPS-injected larvae. This binding was completely abolished with an excess of competitor, suggesting that binding is a specific interaction between the nuclear proteins and the probe. Of three longer probes (M2-a, M2-b and M2-c) no binding was found with M2-b probe under our experimental conditions (results not shown). By contrast, nuclear proteins were observed to bind to M2-a and M2-c probes (Figures 7B and 7C). In the case of M2-b probe, a protein present in the nuclear extract from non-injected larvae bound to the probe, and a stronger signal was obtained with the sample from LPS-injected larvae. The binding was not affected by the small amount of the non-labelled probe, but was strongly affected by a 50-fold excess amount of the competitor (Figure 7B). Specific binding was also detected with the M2-c probe containing two κB sites (Figure 7C). The binding was not affected at all by an M2κB probe, although a complete competition was observed with the M2κB probe when the same probe was used for EMSA (Figure 7A). The result suggests that a nuclear protein binds to the region where the κB site is not involved. The result also suggests that the M2κB-specific nuclear binding protein might not able to bind to its site when another protein binds to its binding site, excluding the M2κB region, because of a conformational change.

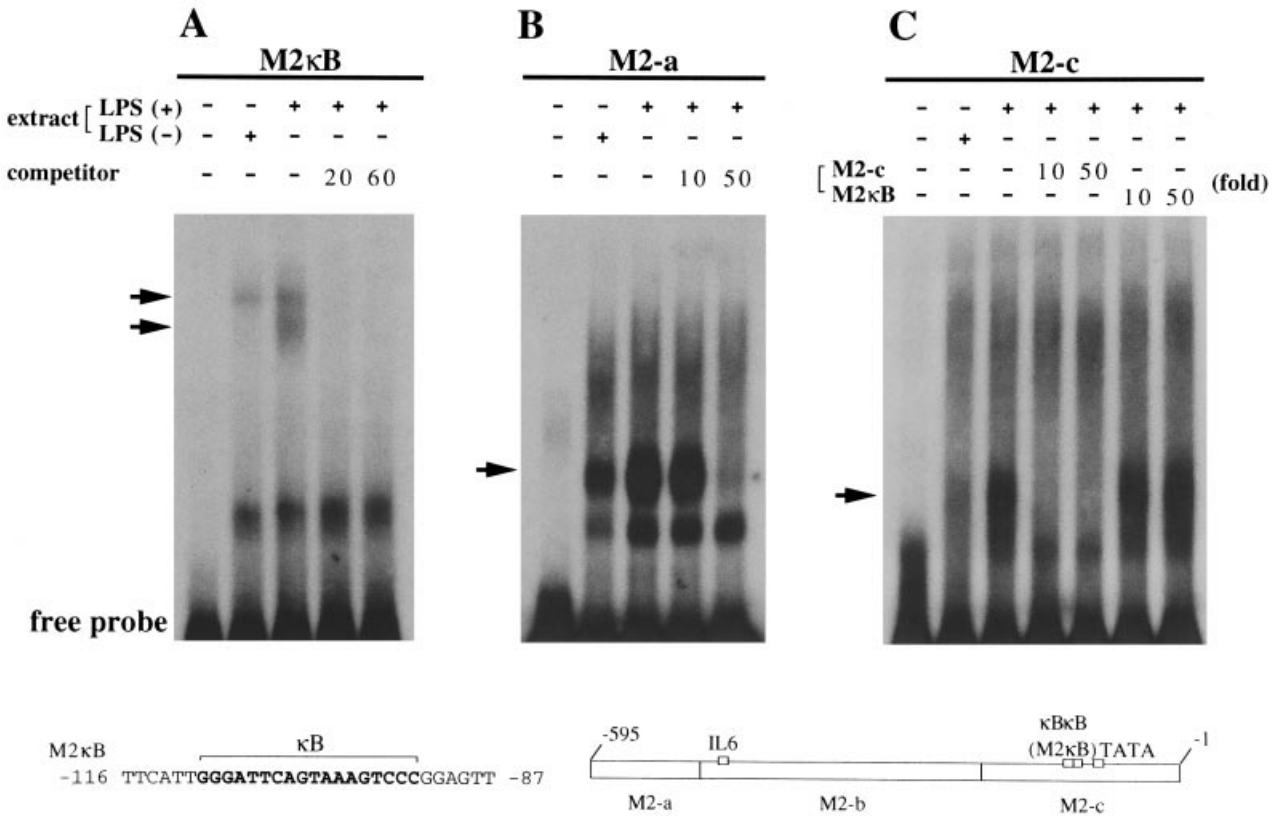

Figure 7 EMSA with nuclear extracts and labelled DNA probes corresponding to the 5′-upstream region of moricin 2 gene

Nuclear extracts from LPS-injected or non-injected larval fat-bodies were used. For the competition experiments, the indicated molar excess of the cold probe was added to the reaction mixture before the addition of the labelled probe. Arrows indicate the position of the specific DNA–protein complexes. (**A**) A probe (M2κB) containing two κBs including a complementary sequence was used. The nucleotide sequence is shown in the lower panel. **Bold** letters denote the κB sequence. (**B**) and (**C**) Longer probes corresponding to M2-a (**B**), M2-b and M2-c (**C**) were used. The position of the probes on the 5′-upstream region is indicated in the lower panel. Note that M2-c contained a longer sequence including M2κB. The result of EMSA using an M2-b probe is omitted, because there were no positive signals. Conditions are described in the Experimental section.

DISCUSSION

In the present study we cloned two types of moricin cDNAs whose nucleotide sequences differ in the coding and non-coding regions, resulting in one amino acid replacement in the putative signal peptide. The results suggest that moricin forms a multiple gene family, as is seen in other insect antibacterial peptides such as cecropins [16], or that two moricin cDNA clones may have arisen from alleles. Concerning these two possibilities, our preliminary genomic Southern-blot analysis using a moricin probe suggested that a multiple gene family may possibly be formed (results not shown). The screening of moricin genes resulted in the cloning of the moricin 2 gene. This gene contained a single intron, which has also been observed in cecropin genes [21].

It is noteworthy that gene expression of moricin 1 and 2 was simultaneously induced upon bacterial infection, suggesting that gene dosage is an effective strategy for insects to protect themselves against bacterial invasion. In addition, moricin is a very unique antibacterial peptide that shows strong antibacterial activity against both Gram-positive and Gram-negative bacteria [9]. Accordingly, it is suggested that defence mechanisms in *B. mori* against bacterial infection depend on a very efficient strategy including the combination of gene dosage and a wide spectrum of antibacterial peptides.

As mentioned above, the structure of moricin is unique, and it cannot be classified into any category of antibacterial peptides already identified [9]. Thus we performed a computer-aided homology search for amino acid sequences, including the putative signal peptide and prosegment, and the nucleotide sequences of the 3′-non-coding region. A lack of significant homology with registered data in the EMBL Nucleotide Sequence Database suggests that not only the mature peptide, but also other portions of moricin cDNAs, are different from those of other antibacterial peptides.

As moricin showed a novel structure and function, we analysed expression patterns of the moricin gene to clarify the induction mechanisms as compared with other antibacterial peptides. We observed that moricin has, in principle, a similar gene-expression pattern induced by bacteria or the cell-wall components to that of other antibacterial peptides such as cecropin [18,22,23], attacin [24] and lebocin [25]. On the other hand, moricin gene expression was observed in the Malpighian tubule as well as in fat-bodies and haemocytes. It is the first report to show that an antibacterial peptide gene from *B. mori* is expressed in the Malpighian tubule. It has previously been demonstrated that the larval integuments derived from the ectoderm of lepidopteran insects such as *B. mori* and *Hyalophola cecropia* (the giant silk moth) develop an immune response, when challenged with live bacteria or their cell-wall components [26]. In addition, cecropin B was isolated directly

from the cuticle of *B. mori* [27]. Moreover, results reported recently for transgenic fruitflies (*Drosophila melanogaster*) suggest that insects harbour a local immune response system against microbial infection [28]. It is interesting to speculate that moricin may play a role in the local immune response in the Malpighian tubule, which is derived from the ectoderm.

Results of EMSA indicated that LPS induces or enhances the binding of nuclear proteins to the specific sequences containing a κB site(s) and other regions in the 5′-upstream of the moricin 2 gene. The results paralleled the accumulation of moricin gene transcripts by the injection of LPS or its components. By contrast, binding was not observed with a probe containing NF-IL-6 site, suggesting that this motif does not contribute to the induction of moricin 2 gene expression. In addition, we could not find a GATA motif near the κB sites in this gene, which plays an important role in antibacterial-peptide gene expression of *D. melanogaster* [29]. This suggests the GATA motif is not essential for moricin gene expression. An efficient gene-expression system for the quantitative analysis of the expression of artificial gene constructs must be developed in *B. mori* to determine the promoter function of these regulatory motifs in antibacterial peptide genes of this insect.

This work was supported by Enhancement of Center of Excellence, Special Coordination Funds for Promoting Science and Technology, Science and Technology Agency, Japan.

REFERENCES

1 Boman, H. G. and Hultmark, D. (1987) Annu. Rev. Microbiol. **41**, 103–126
2 Dunn, P. E. (1986) Annu. Rev. Entomol. **31**, 321–329
3 Hultmark, D. (1993) Trends Genet. **9**, 178–183
4 Hoffmann, J. A. and Reichhart, J.-M. (1997) Trends Cell Biol. **7**, 309–316
5 Boman, H. G. (1991) Cell **65**, 205–207
6 Boman, H. G., Faye, I., Gudmundsson, G. H., Lee, J.-Y. and Lidholm, D.-A. (1991) Trends Genet. **9**, 178–183
7 Kimbrel, D. A. (1991) BioEssays **13**, 657–663
8 Wicker, C., Reichhart, J.-M., Hoffmann, D., Hultmark, D., Samakovlis, C. and Hoffmann, J. A. (1990) J. Biol. Chem. **265**, 22493–22498
9 Hara, S. and Yamakawa, M. (1995) J. Biol. Chem. **270**, 29923–29927
10 Hara, S. and Yamakawa, M. (1996) Biochem. Biophys. Res. Commun. **220**, 664–669
11 Hara, S., Asaoka, A. and Yamakawa, M. (1996) Appl. Entomol. Zool. **31**, 465–466
12 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene **33**, 103–119
13 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
14 Furukawa, S., Taniai, K., Ishibashi, J., Hara, S., Shono, T. and Yamakawa, M. (1997) Biochem. Biophys. Res. Commun. **238**, 769–747
15 Fujii, T., Sakurai, H., Izumi, S. and Tomino, S. (1989) J. Biol. Chem. **264**, 11020–11025
16 Mine, E., Sakurai, H., Izumi, S. and Tomino, S. (1995) Nucleic Acids Res. **23**, 2648–2653
17 Suzuki, Y., Obara, T., Takiya, S., Hui, C.-C., Matsuno, K., Suzuki, T., Suzuki, E., Ohkubo, M. and Tomino, T. (1990) Dev. Growth Differ. **32**, 179–187
18 Taniai, K., Furukawa, S., Shono, T. and Yamakawa, M. (1996) Biochem. Biophys. Res. Commun. **226**, 783–790
19 Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. and Sharp, P. A. (1986) Annu. Rev. Biochem. **55**, 1119–1150
20 Devine, J. M., Tsang, A. S. and Williams, J. G. (1982) Cell **28**, 793–800
21 Taniai, K., Kadono-Okuda, K., Kato, Y., Yamamoto, M., Shimabukuro, M., Chowdhury, S., Xu, J., Kotani, E., Tomino, S. and Yamakawa, M. (1995) Gene **163**, 215–219
22 Taniai, K., Kato, Y., Hirochika, H. and Yamakawa, M. (1992) Biochim. Biophys. Acta **1132**, 203–206
23 Kato, Y., Taniai, K., Hirochika, H. and Yamakawa, M. (1993) Insect Biochem. Mol. Biol. **23**, 285–290
24 Sugiyama, M., Kuniyoshi, H., Kotani, E., Taniai, K., Kadono-Okuda, K., Yamamoto, M., Shimabukuro, M., Chowdhury, S., Choi, S. K., Kataoka, H. et al. (1995) Insect Biochem. Mol. Biol. **25**, 385–392
25 Chowdhury, S., Taniai, K., Hara, S., Kadono-Okuda, K., Kato, Y., Yamamoto, M., Xu, J., Choi, S. K., Debnath, N. C., Choi, H. K. et al. (1995) Biochem. Biophys. Res. Comm. **214**, 271–278
26 Brey, P. T., Lee, W.-J., Yamakawa, M., Koizumi, Y., Perrot, S., François, M. and Ashida, M. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 6275–6279
27 Lee, W.-J. and Brey, P. T. (1994) Anal. Biochem. **217**, 231–235
28 Ferrandon, D., Jung, A. C., Crigui, M.-C., Lemaitre, B., Uttenweiler-Joseph, S., Michaut, L., Reichhart, J.-M. and Hoffmann, J. A. (1998) EMBO J. **17**, 1217–1227
29 Kadalayil, L., Petersen, U.-M. and Engström, Y. (1997) Nucleic Acids Res. **25**, 1233–1239

Received 8 January 1999/18 February 1999; accepted 10 March 1999