abf-1* and *abf-2*, ASABF-type antimicrobial peptide genes in *Caenorhabditis elegans

Yusuke KATO^{*1}, Tomoyasu AIZAWA[†], Hirokazu HOSHINO[‡], Keiichi KAWANO[§], Katsutoshi NITTA[‡] and Hong ZHANG^{*}

^{*}Department of Developmental Biology, National Institute of Agrobiological Sciences, Oowashi 1-2, Tsukuba, Ibaraki 305-8634, Japan, [†]Bio-oriented Technology Research Advancement Institution, Saitama 331-8537, Japan, [‡]Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan, and [§]Faculty of Pharmaceutical Sciences, Toyama Medical and Pharmaceutical University, Toyama 930-0194, Japan

Two genes encoding the ASABF (*Ascaris suum* antibacterial factor)-type antimicrobial peptide, *abf-1* and *abf-2*, were identified in *Caenorhabditis elegans*. Recombinant ABF-2 exhibited potent microbicidal activity against Gram-positive and Gram-negative bacteria, and yeasts. The tissue-specific distribution estimated by immunofluorescence staining and transgenic analysis of a *gfp* fusion gene (where GFP corresponds to green fluorescent protein) suggested that ABF-2 contributes to surface defence in the pharynx. *abf-1* contains a single intron at a conserved position, suggesting that *asabf* and *abf* originated from a common ancestor. Both transcripts for *abf-1* and *abf-2* were detected as two distinct forms, i.e. spliced leader (SL)1-*trans*-spliced with a long 5'-untranslated region (UTR) and SL-less with a short 5'-UTR. A polycistronic precursor RNA

encoding ABF-1 and ABF-2 was detected, suggesting that these genes form an operon. An 'opportunistic operon' model for regulation of *abf* genes, including the generation of short SL-less transcripts, is proposed. In conclusion, *C. elegans* should have an immune defence system due to the antimicrobial peptides. *C. elegans* can be a novel model for innate immunity. Furthermore, the combination of biochemical identification in *Ascaris suum* and homologue hunting in *C. elegans* should be a powerful method of finding rapidly evolved proteins, such as some immune-related molecules in *C. elegans*.

Key words: defensin, innate immunity, invertebrates, nematodes, operon.

INTRODUCTION

Immunity is often classified into adaptive and innate immunity. T- and B-cell-based adaptive immunity is thought to be specific to higher vertebrates. In contrast, innate immunity, such as phagocytosis or the production of antimicrobial peptides, is observed in a wide variety of animals. Such generality indicates that we can use invertebrate model animals to investigate innate immunity.

A good example is a series of studies on the immune regulatory system by the *Toll* signalling pathway (reviewed in [1]). *Toll* was originally identified as a factor regulating dorso-ventral patterning in *Drosophila melanogaster*. Later, the *Toll* signalling pathway was suggested to also regulate the gene expression of antimicrobial peptide genes in *D. melanogaster*. This report stimulated the search for *Toll* homologues in mammals. Finally, some of the *Toll* homologues were revealed to be actual regulators of mammalian innate immunity. This example suggests that invertebrate model animals can be powerful tools for identifying innate immune-related molecules by genetic study.

The nematode, *Caenorhabditis elegans*, is a model animal widely used in the laboratory. Forward and reverse genetic methods have been established, and the entire genomic sequence is now available for *C. elegans* [2]. *C. elegans* is thus one of the best candidates for an invertebrate model to study innate immunity. Recently, the mechanisms of bacterial virulence of *Pseudomonas aeruginosa* ([3–6] and reviewed in [7]) and *Salmonella typhimurium* [8,9] were investigated using the *C. elegans* pathogenesis model. The morphological change induced by a novel bacterial pathogen, *Microbacterium nematophilum*, was also investigated using the advantages of genetic analysis in *C.*

elegans [10]. These attractive approaches have established a novel model for host-pathogen interaction using *C. elegans*. Immunity should be one of the major host factors influencing host-pathogen interactions. However, the immune defence system of *C. elegans* is poorly understood.

In the present study, we report two of the ASABF (*Ascaris suum* antibacterial factor)-type antimicrobial peptide genes [11–13], *abf-1* and *abf-2*, toward the elucidation of *C. elegans* immunity. These genes were at least partially transcribed as a polycistronic precursor RNA, i.e. *abf-1* and *abf-2* form an operon. However, the characterization of these transcripts revealed some unusual properties, which are also described in the present paper.

EXPERIMENTAL

Worm strains

Wild-type Bristol N2 was cultured and maintained as previously described [14]. *glp-4(bn2)* was maintained at 15 °C and cultured at 25 °C to obtain a sterile phenotype [15].

Preparation of worm RNA

C. elegans in a mixed stage was frozen at –80 °C and ground using a Coolmil (Tokken, Chiba, Japan). The ground nematodes were subsequently extracted with extraction buffer (50 mM NaCl, 5 mM EDTA, 0.5 % SDS and 50 mM Tris/HCl, pH 7.5). After two phenol/chloroform (1:1, v/v) extractions, the total RNA was obtained by ethanol precipitation. Polyadenylated [poly(A)⁺] RNA was isolated using Oligotex-dT30 Super (Takara, Tokyo, Japan).

Abbreviations used: ASABF, *Ascaris suum* antibacterial factor; EIF, eukaryotic initiation factor 4A; GFP, green fluorescent protein; ORF, open reading frame; poly(A)⁺, polyadenylated; RACE, rapid amplification of cDNA ends; RT, reverse transcription; SL, spliced leader; UTR, untranslated region.

¹ To whom correspondence should be addressed (e-mail kato@affrc.go.jp).

PCR primers

The PCR primers used in the present study were as follows: *abf-10s*, 5'-CCTTGTTCTCCTACTTCCGAACA-3'; *abf-11s*, 5'-GTTCCGGTAATGCAGAGGA-3'; *abf-11as*, 5'-GTTGTGTGGGAGCTGAGA-3'; *abf-11asx*, 5'-GGTTAGTTGTGTGGGAGCTGAGA-3'; *abf-11as7*, 5'-GCATTAATACGACTCATATAGGGAGGTTAGTTGTGTGGGAGCTGAGA-3'; *abf-12s*, 5'-CGACAAGGACTTTGCACATCT-3'; *abf-12as*, 5'-GACACATCGTTGGCGTTTGA-3'; *abf-13s*, 5'-GCTGCCGACATCGACTTTAG-3'; *abf-13as*, 5'-CCTCTCTTAATAAGAGCACCA-3'; *abf-14s*, 5'-GCATTACCAGCTGCAGTATG-3'; *abf-14as*, 5'-CCTCCATTGTGCACACCGGTA-3'; *abf-15as*, 5'-CCTCCACAGAAGCATGTTGGAT-3'; *abf-s3R*, 5'-GGTCGGGTCAGGATATCTCGCTCA-3'; *abf-as7R*, 5'-GATCGATCCTGGATGCGAGTGCAGCTTTTTTGTCA-3'; *abf-as4R-BamHI*, 5'-CGACCCGGGATCCACCAAGTGAATATCTCCTCT-3'; *abf-s3r-NsiI*, 5'-CCACAATGCATGTTTGGTCGGGTCAGGATATCTCGCTCA-3'; *abf-20as-BamHI*, 5'-CCACAGGATCCGGTTCGAAATCTGAACGAACGGTTGGGA-3'; *T7-abf-13as*, 5'-GCATTAATACGACTCACTATAGGGAGACCTCTCTTAATAAGAGCACCA-3'; *BamHI-abf-13s*, 5'-GCAAGGATCCGCTGCCGACATCGACTTTAG-3'; *1a-5'f*, 5'-GATATTCAGCAAAAAAGACAAAACACTACTGTCGACC-3'; *anti 1a R53-C46*, 5'-TCGTCCACATCGATCGCAAAACAC-3'; *1a-A1-A8*, 5'-GCAGTCGACTTTTCATCATGCGCA-3'; *1a-as-A1-7C*, 5'-GCATGATGAAAAGTCGACTGC-3'; *1a-s-86s-92H*, 5'-AGTTCACGTGGACGAAGGCA-3'; *new inv-1a-s*, 5'-GAGCAAAGTGGCGCAAGG-3'; *aup-1as*, 5'-CCAAACTTCCACATGCTGCCT-3'; *aup-2as*, 5'-GACACCTCTTGCACGCTGTCT-3'; *aup-1s*, 5'-GTCTCGACTGCTCCTCTTCA-3'; *aup-2s*, 5'-GGTGGCTGATGAGAATGATGT-3'; *spliced leader (SL)1-20*, 5'-GGTTTAATTA-CCCAAGTTTG-3'; *SL2-20*, 5'-GGTTTTAACCCAGTTAC-TC-3'; *SL3*, 5'-GGTTTTAACCCAGTTAACCAAG-3'; *SL4*, 5'-GGTTTTAACCCATATAACCAAG-3'; *SL5*, 5'-GGTTTAACCCAGTTAACCAAG-3'; *HindIII-eif-1s* (where eif corresponds to eukaryotic initiation factor 4A), 5'-GCATAAGCTTCGACGACATGGAGCTCAAGGA-3'; *SP6-eif-1as*, 5'-GCATATTTAGGTGACACTATAGAAGAGCTGGATGTGTCAAGAGCATTTCT-3'; and *gpd-as1*, 5'-GGACAGTGGTCATAAGTCCCTCA-3'. (A,B) means a PCR using primers A and B, and (A,B) × (C,D) means a nested-PCR using primers A and B for the first round of PCR and primers C and D for the second round of PCR.

cDNA cloning

Mature transcripts for *abf-1* and *abf-2* were cloned using reverse transcription (RT)-PCR. The poly(A)⁺ RNA isolated from the mixed-stage *C. elegans* was reverse-transcribed using random 9-mer primers [RNA LA PCR kit (AMV) version 1.1; Takara], and a single-stranded cDNA mixture was obtained. First, the internal regions were amplified using the cDNA mixture as templates, i.e. (*abf-11s*,*abf-11as*) × (*abf-12s*,*abf-12as*) for *abf-1*, and (*abf-13s*,*abf-13as*) × (*abf-14s*,*abf-14as*) for *abf-2*. Next, the 5' terminal regions were amplified using the SL1 primer, i.e. (*SL1*,*abf-11as*) × (*SL1*,*abf-12as*) for *abf-1*, and (*SL1*,*abf-13as*) × (*SL1*,*abf-14as*) for *abf-2*. Finally, polyadenylation sites were determined by 3' rapid amplification of cDNA ends (RACE) using reverse-transcripts and an oligo dT-M13M4 primer as templates, i.e. (*abf-11s*,M13M4) × (*abf-12s*,M13M4) for *abf-1*, and (*abf-13s*,M13M4) × (*abf-14s*,M13M4) for *abf-2*. A polycistronic precursor RNA was also detected by RT-PCR, i.e. (*abf-11s*,*abf-13as*) × (*abf-12s*,*abf-14as*) for the *abf* operon.

Determination of 5' ends of transcripts for *abf-1* and *abf-2*

A RT-PCR was carried out using specific SL primers, i.e. (*SLX*,*abf-11as*) × (*SLX*,*abf-12as*) for *abf-1*, and (*SLY*,*abf-13as*) × (*SLY*,*abf-14as*) for *abf-2*, where 'SLX' is an SL1-20 or SL2-20 primer, and 'SLY' is one of the SL1-20, SL2-20, SL3, SL4 or SL5 primers. *gpd-2* and *gpd-3* were used as positive controls for the SL2-accepting genes [16], i.e. (*SL2-20*,*gpd-as1*) × (*SL2-20*,*GPD-58*). The *GPD-58* primer was originally used by Spieth et al. [16]. 5' RACE was performed using the 5' RACE system, version 2.0 (Gibco BRL).

Detection of a transcript for C50F2.4

Poly(A)⁺ RNA isolated from the mixed-stage *C. elegans* was reverse-transcribed using a random 9-mer primer. A transcript for C50F2.4 was detected by RT-PCR using the reverse-transcripts as templates, i.e. (*aup-1s*,*aup-1as*) × (*aup-2s*,*aup-2as*).

Preparation of genomic DNA from *Ascaris suum*

The ovaries and uteri were isolated from the adult female *A. suum*. These tissues were ground in liquid nitrogen. An aliquot of the tissue powder (7 ml) was suspended in 50 ml of an extraction buffer (0.1 M EDTA, 0.5 % SDS and 10 mM Tris/HCl, pH 8.0) containing 20 µg/ml (final concentration) RNase-A (Nippon Gene, Tokyo, Japan). The suspension was incubated at 37 °C for 1 h, in order to degrade the RNA. Subsequently, protease K (150 µg/ml, final concentration) was added, and the suspension was incubated at 37 °C for 18 h. After two extractions with water saturated phenol/chloroform (1:1, v/v; pH 8.0) and propan-2-ol precipitation, the genomic DNA was dissolved in 1 ml of 1 × TE [where 1 × TE corresponds to 10 mM Tris/HCl (pH 8.0)/1 mM EDTA] [17]. This DNA solution was dialysed against 1 × TE at 37 °C for 2 days. This partially purified DNA solution was turbid, due to contamination by polysaccharides. Further purification was achieved using a Qiagen Blood and Cell Culture DNA kit. After this step, the highly purified DNA solution became clear. A total of 140 µg of genomic DNA was obtained. Although other tissues were also subjected to the same procedure, these trials resulted in a lower yield.

Cloning of the *asabf* gene

The intron of *asabf* was identified by PCR using the genomic DNA as a template, i.e. (*1a-A1-A8*,*anti 1a R53-C46*). The flanking sequences of the *asabf* gene were cloned by inverse PCR. The genomic DNA of *A. suum* (0.5 µg) was digested with *NsiI* (for the 5' flanking region) or *KpnI* (for the 3' flanking region). The restriction endonucleases were inactivated by phenol/chloroform (1:1, v/v) extraction. Cyclic DNAs were generated by self-ligation of the fragments and were used as templates for inverse PCR, i.e. (*1a-as-A1-7C*,*new inv-1as*) for the 5' flanking region, and (*1a-s-86S-92H*,*anti 1a R53-C46*) for the 3' flanking region.

Immunofluorescence staining

A peptide from the C-terminus of ABF-2 (Cys⁶⁹-Ala⁸⁰; CANGGGDIPLGA) was synthesized and conjugated to keyhole limpet haemocyanin. Rabbits were injected with the conjugated peptide, followed by multiple boosts. Affinity-purified anti-(ABF-2 peptide) antibodies were prepared. The immunofluorescence

staining was carried out using whole-mount-fixed larvae and adults as described by Miller and Shakes [18]. Briefly, we fixed the worms in 1% (v/v) formaldehyde and then permeabilized them by freeze-thawing. The fixed worms were incubated with the affinity-purified anti-(ABF-2 peptide) antibody. A Texas Red-conjugated secondary antibody (Vector Laboratories, Burlingame, CA, U.S.A.) was used at 5 µg/ml for immunofluorescence staining.

***gfp* reporter constructs (where GFP corresponds to green fluorescent protein)**

The putative promoter region of the *abf* operon was amplified by high-fidelity-PCR using Pyrobest polymerase (Takara) and cosmid C50F2 as a template, i.e. (abf-s3R,abf-as7R) for *abf-1* and (abf-s3R,abf-as4R-*Bam*HI) for *abf-2*. The PCR products were digested with *Bam*HI. A *gfp* reporter plasmid, pPD95.77, was digested with *Hind*III, treated with the Klenow fragment and was subsequently digested with *Bam*HI. The PCR products were finally introduced into pPD95.77 using the smooth end and the *Bam*HI site. The plasmids were highly purified using the Qiagen plasmid Midi kit. Subsequently, RNase-A and protease K treatments were performed. Following ethanol precipitation, the plasmids were subjected to microinjection into *C. elegans*.

Generation of transgenic animals

Hermaphrodites were injected with a mixture of pRF4 (90 µg/ml) and a *gfp* reporter plasmid (75 µg/ml) in 1 × TE [19]. Transgenic animals were identified by a roller phenotype.

Northern-blot analysis

Total RNA or poly(A)⁺ RNA was isolated as described above from the embryos obtained using an alkaline hypochlorite treatment [14]. Poly(A)⁺ RNA (2–4 µg) was electrophoresed on a 1.2% (w/v) agarose slab gel containing 6.6% (v/v) formaldehyde and was capillary-transferred on to a Gene Screen Plus membrane (DuPont). Parts of cDNA for *abf-1* and *abf-2* were amplified by PCR, i.e. (abf-10s,abf-11asx) × (abf-11s,abf-11as7) for *abf-1* and (*Bam*HI-abf-13s,T7-abf-13as) for *abf-2*. The PCR product was digested with *Bam*HI and used as a template for the RNA probe for *abf-2*. A digoxigenin-labelled RNA probe was synthesized using a digoxigenin RNA labelling kit (Boehringer Mannheim GmbH). As a control for the loading calibration, the *C. elegans* homologue of *eif* was used [20]. The template of the RNA probe for *eif* was prepared by RT-PCR, i.e. (*Hind*III-eif-1s,SP6-eif-1as), and was used after *Hind*III digestion. Prehybridization, hybridization, detection using alkaline phosphatase and CSPD [3-(4-methoxyphosphoryl)-1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1^{3,7}]decan-4-yl] phenyl phosphate], and autoradiography were performed according to the recommendations of the manufacturer.

Computer-assisted sequence analysis

Blast database searches were performed via the following on-line servers: Washington University (<http://genome.wustl.edu/gsc/Blast/client.pl>), the National Center of Biotechnology Information (<http://www.ncbi.nlm.nih.gov/blast/>), and the National Institute of Genetics (<http://www.ddbj.nig.ac.jp/c-elegans/html/CE-BLAST.html>). Alignment and phylogenetic analysis were carried out using ClustalW (<http://www.ddbj.nig.ac.jp/htmls/E-mail/clustalw-j.html>) with compensation by eye-fitting. Secretory signal regions were predicted by SignalP (<http://www.cbs.dtu.dk/services/SignalP/>). The

theoretical pI and molecular mass were estimated by ExPASy (<http://www.expasy.ch/tools/pi-tool.html>). Other standard sequence analyses were performed using Genetyx-Mac version 7.3 (Software Development, Tokyo, Japan).

Expression and purification of ABF-2

The open reading frame (ORF) of a cDNA clone encoding ABF-2 was amplified by PCR using primers with engineered restriction sites. The PCR product was inserted into pPIC3 (Invitrogen) using the *Bam*HI/*Eco*RI site. After amplification using *Escherichia coli* and linearization, the ABF-2 expression construct was transformed into *Pichia pastoris* GS115 (his4). A transformed strain was grown in 2% (v/v) methanol/buffered glycerol-complex medium [BMGY; 1% (w/v) yeast extract/2% (w/v) peptone/100 mM potassium phosphate (pH 6.0)/1.34% (w/v) yeast nitrogen base/4 × 10⁻⁵% biotin/1% (v/v) glycerol] in flasks shaken at 30 °C for 7 days. The culture supernatant was diluted with water to a conductivity of 5 mS/cm and then adjusted to pH 6.0 with NaOH. The recombinant ABF-2 was partially purified using cation-exchange chromatography (SP-Sepharose FF). Further purification was achieved by cation-exchange HPLC (Hi Trap SP) using a linear gradient elution (0–300 mM NaCl). The fractions containing ABF-2 were identified using SDS/PAGE. These fractions were subjected to final purification by reversed-phase HPLC (Finepak SIL 300C18 T7) using a linear gradient elution [20–40% (v/v) acetonitrile containing 1% (v/v) trifluoroacetic acid]. The purified recombinant ABF-2 was freeze-dried and stored at 4 °C.

Micro-organisms

E. coli (JM109) was purchased from Takara. *Staphylococcus aureus* (A.T.C.C. 6538P) was a gift from Dr Masanori Yamamoto (Teijin Institute for Bio-Medical Research, Tokyo, Japan). *Bacillus subtilis* (IFO3134) was purchased from the Institute for Fermentation (IFO), Osaka, Japan. *Agrobacterium tumefaciens* (MAFF1001), *Kocuria varians* (MAFF118076), *Bdellovibrio bacteriovorus* (MAFF106101), *Klebsiella pneumoniae* (MAFF519002), *Candida krusei* (MAFF114085), *Kluyveromyces thermotolerans* (MAFF113848), *Pichia anomala* (MAFF113717), *Debaryomyces hansenii* (MAFF113836), *Torulaspora delbrueckii* (MAFF113811) and *Saccharomyces cerevisiae* (MAFF113011) were obtained from the National Institute of Agrobiological Sciences, Ibaraki, Japan.

Microbicidal assays

Each microbial strain in exponential phase was suspended in 10 µl of 10 mM Tris/HCl, pH 7.5, containing 3-fold serially increasing concentrations of purified recombinant ABF-2 (i.e. 1, 3, 9, 27 etc. µg/ml ASABF). The attenuation of the microbial suspension was adjusted to 0.02 at 650 nm. After a 2 h incubation, the test suspension (5 µl) was diluted 1000 times. The diluted sample (200 µl) was inoculated on to an optimum medium. The number of colonies was counted, and the 50% microbicidal concentration was determined.

Tricine/SDS/PAGE and Western-blot analysis

To prepare the *C. elegans* extracts, worms were suspended in 1 × SDS/PAGE loading buffer and heated at 100 °C for 20 min. The extract was centrifuged and the supernatant was subjected to electrophoresis. Tricine/SDS/PAGE was performed using 15% (w/v) polyacrylamide gels with the PAGEL system (ATTO, Tokyo, Japan). Following electrophoresis, the peptides were

Figure 3 Transcripts and organization of *abf* genes

Top left-hand panel: nucleotide sequence of *abf-1* and *abf-2*. Numbers indicate the nucleotide position on the cosmid C50F2 (GenBank® accession number, U80445). The deduced amino acid sequences are given below the nucleotide sequences. Exons are shown in capital letters. Splice donor (gt) and acceptor sites (ag) are indicated in boxes. Double underlining shows TATA-box-like sequences and CCAAT sequences. White letters on a black background indicate polyadenylation signals in *C. elegans*. Note that only 78 bp of the intergenic region was detected between the polyadenylation site of *abf-1* and the SL1-acceptor site of *abf-2*, which is typically observed in operons. Top right-hand panel: *trans*-splicing of transcripts. The type of SL was identified using RT-PCR. Note that the transcript of *abf-2* was exclusively *trans*-spliced to SL1, whereas *abf-1* is the downstream gene of the *abf* operon with an intergenic region. *gpd-2* and *-3* were used as positive controls for the SL2 acceptor gene. Bottom left-hand panel: the schematic structure of transcripts for *abf-1* and *abf-2*. The transcripts of *abf* genes were cloned by RT-PCR. Long SL-*trans*-spliced and short SL-less mRNAs were identified for both *abf-1* and *abf-2*. A polycistronic precursor RNA containing the ORFs of *abf-1* and *abf-2* was detected, suggesting that the *abf* genes form an operon. Bottom right-hand panel: schematic gene organization of the *abf* operon. The *abf* operon was located in an intron of a reverse-directional gene, C50F2.2. A single intron was found in *abf-1* at the conserved position identical with that of *asabf*. The organization of *gfp* fusion constructs are also indicated.

Figure 4 Immunofluorescent detection of ABF-2

(A and B) Animals were immunostained with an anti-[ABF-2 (Cys⁶⁰–Ala⁸⁰)] antibody. ABF-2 was detected only in the pharyngeal marginal zone. (C and D) Pre-immune IgG instead of the anti-(ABF-2) antibody was used for immunostaining. (E and F) Secondary antibody alone without the anti-(ABF-2) antibody. (A, C and E) Images taken using Nomarski differential interference contrast microscopy. (B, D and F) Fluorescent images.

In vitro antimicrobial activity of recombinant ABF-2

To confirm antimicrobial activity, recombinant ABF-2 was prepared using the yeast *P. pastoris* expression system. A total of 100 µg of the recombinant ABF-2 was obtained from 1 litre of culture supernatant. No significant contaminant was detected by SDS/PAGE (Figure 1C) or N-terminal sequencing (results not shown). The first ten amino acid residues at the N-terminus of the recombinant ABF-2 were identical with the predicted mature ABF-2 (DIDFSTXARM), indicating that the native secretory signal was cleaved as predicted by SignalP (Figure 2). We next submitted the recombinant ABF-2 to mass spectral analysis and obtained a molecular mass of 6537.3 Da. This is in good agreement with the mass of 6536.6 Da calculated for the predicted mature ABF-2 containing four intramolecular disulphide bridges without a four-residue peptide (IKRG) at the C-terminus, as observed in ASABF [12].

The microbicidal activity of recombinant ABF-2 was tested (Table 1). All tested Gram-positive bacteria were sensitive. Some Gram-negative bacteria and yeasts were sensitive although others were resistant. The Gram-positive bacteria tended to be more sensitive than the Gram-negative bacteria and yeasts. This tendency is identical with that of ASABF [12,13]. 50% microbicidal concentrations were less than 1 µM for the sensitive organisms, suggesting that ABF-2 is a microbicidal peptide.

Both *abf-1* and *abf-2* generate long SL-*trans*-spliced and short SL-less mRNAs

The cDNAs for *abf-1* and *abf-2* were cloned by RT-PCR (Figure 3, top left-hand panel). The structures of the transcripts for *abf-1* and *abf-2* are summarized in Figure 3 (bottom left-hand panel). Some of the mature mRNAs of *C. elegans* are known to be modified at their 5' termini by the addition of a 22 nt leader sequence designated SL [22]. To date, SL1–SL5 have been reported (see the Discussion section). The 5' terminal regions were determined by specific amplification using SL primers and 5' RACE. Both transcripts of *abf-1* and *abf-2* were amplified by the SL1 primer, but not the SL2–SL5 primers (Figure 3, top right-hand panel). It is noteworthy that the distance between the SL and the translation initiation site (i.e. the 5'-UTR) was long

(191 bp for *abf-1*, and 104 bp for *abf-2*; see Figure 3, top left-hand panel). On the other hand, short SL-less transcripts were mainly detected by 5' RACE (the lengths of the 5'-UTRs were 25 bp for *abf-1*, and 17 bp for *abf-2*). The *abf-2* transcripts in which SL2–SL5 were *trans*-spliced were not detected by 5' RACE.

abf-1 and *abf-2* form an operon

The cDNA sequences were compared with the genomic DNA sequence to determine the gene organization (Figure 3, top left-hand panel and bottom right-hand panel). Both *abf-1* and *abf-2* were found in the cosmid C50F2 from chromosome I. *abf-1* and *abf-2* were aligned in the same direction and formed a gene cluster. Although *abf-2* was an intron-less gene, *abf-1* was interrupted by a single intron.

In *C. elegans*, a number of genes are arranged in operons, transcribed as polycistronic precursor RNAs [22]. The polycistronic precursor RNAs are processed to generate monocistronic mature RNAs. An operon is a cluster of closely spaced genes whose intergenic space is 0–400 nt [23]. Because the distance between the polyadenylation site of *abf-1* and the SL1 acceptor site of *abf-2* was only 78 bp, *abf-1* and *abf-2* were predicted to form an operon [16]. A polycistronic precursor RNA encoding both the ORFs of *abf-1* and *abf-2* despite loss of the *abf-1* intron was detected by RT-PCR, suggesting that *abf-1* and *abf-2* actually form an operon and are at least partially transcribed as a polycistronic precursor (Figure 3, bottom left-hand panel). The *abf* operon was sufficiently compact to be involved in a 1.8 kb intron of the reverse-directed gene C50F2.2 (Figure 3, bottom right-hand panel).

Both *abf-1* and *asabf* have a single intron at a conserved position

To estimate the evolutionary relationship, the exon–intron junction of *asabf* was determined and compared with that of *abf-1*. The ORF region of *asabf* was cloned by PCR. The genomic sequence obtained was compared with that of cDNA, and the position of the exon–intron junction was determined. *asabf* was interrupted by a single intron at the same position as that of *abf-1*

Figure 5 Expression of *abf* genes

Left-hand panels: expression of *abf-gfp* fusion genes in transgenic animals. ABF-1-GFP (construct 1 in Figure 3, bottom right-hand panel) was observed in the pharyngeal lumen. ABF-2-GFP (construct 2 in Figure 3, bottom right-hand panel) was mainly detected in the pharyngeal tissue. Right-hand panel: Northern-blot analysis. Total RNA (20 µg) was loaded. As a control for loading calibration, the *C. elegans* homologue of *eif* was used. The *abf-1* transcripts were detected as five bands. The short SL-less and long SL-*trans*-spliced transcripts of *abf-1* could be the 0.3 and 0.5 kb bands. In contrast, the *abf-2* transcript was detected as a single band (0.4 kb, identical with the short SL-less form). The accumulation of both *abf-1* and *abf-2* transcripts was almost identical in adults of N2 and *glp-4(bn2)*, and the eggs.

(Figure 3, bottom right-hand panel). Furthermore, *asabf* was analysed by inverse-PCR to determine the entire gene structure, including the 5'- and 3'-UTR regions. No intron was found in *asabf* other than the intron identically positioned with that of *abf-1*. Thus the exon-intron organization of *asabf* is identical with that of *abf-1*. These results suggest that both the *asabf* and *abf* genes have evolved from a common ancestor.

Expression analysis of *abf* genes

To explore the tissue-specific distribution of *abf* gene products, an immunofluorescence study was performed using the anti-(ABF-2) antibody. Only the pharyngeal marginal zone was stained in all larval and adult stages, but not in pre-hatching embryos (Figure 4).

The expression of *abf-2* was also examined using *gfp* reporter constructs. The closest upstream gene of the *abf* operon was predicted to be the putative gene C50F2.4 (Figure 3, bottom right-hand panel). The transcript of C50F2.4 was detected by RT-PCR, suggesting that C50F2.4 is a functional gene (results

not shown). Based on this result, the *gfp* reporter plasmid, containing the 2 kb region between the farthest downstream exon of C50F2.4 and *abf-1* was used to monitor the expression of *abf-2* (Figure 3, bottom right-hand panel; construct 1). ABF-2-GFP was detected in the pharynx in all larval and adult stages, but not in pre-hatching embryos (Figure 5, left-hand panels). This result was in accordance with that obtained by the immunofluorescence study. ABF-2-GFP was also observed in the pharyngeal neurons and excretory cells. The expression of *abf-1* was also explored by *gfp* reporter study (Figure 3, bottom right-hand panel; construct 2). ABF-1-GFP was detected in the pharynx and a part of the body wall muscle in all larval and adult stages, but not in pre-hatching embryos (Figure 5, left-hand panels).

Because the *gfp* transgene experiment is not effective to test the gene expression in germ line cells [24], Northern-blot analyses were carried out in adults of N2, temperature-induced adult *glp-4(bn2)* with few eggs, and eggs of N2 (Figure 5, right-hand panel). We detected *abf-2* transcripts in the adults of N2 and *glp-4*, and eggs. No obvious difference in the accumulation of *abf-2*

transcripts was observed. The *abf-2* transcript was detected as a single band of 0.4 kb, which was thought to be the short SL-less *abf-2* mRNA. The accumulation of *abf-1* transcripts was also observed in the adults of N2 and *glp-4*, and the eggs. The *abf-1* transcripts were detected as at least five bands (0.3, 0.5, 1, 2 and 4.5 kb). Although the short SL-less and long SL-*trans*-spliced transcripts of *abf-1* could be the 0.3 and 0.5 kb bands respectively, the 1, 2 and 4.5 kb bands were unexpected. No genes whose sequence was significantly similar to *abf-1* or *abf-2* have been detected in the *C. elegans* genome database.

DISCUSSION

ASABF-type antimicrobial peptides in *C. elegans*

In the present study, we have reported the presence of ASABF-type antimicrobial peptide-like genes in *C. elegans*, *abf-1* and *abf-2*. The antimicrobial activity of the recombinant ABF-2 against various microbes was confirmed *in vitro*. *C. elegans* should at least possess an immune defence system against microbes, due to these antimicrobial peptides. This is believed to be the first report of immune effector molecules in *C. elegans*. In addition, the gene organization and expression of the ASABF-type antimicrobial peptide family is described. In our previous study, an expressed sequence tag-predicted peptide, yk150c7, and a gene predicted by the genomic sequence, T22H6.5, were described as ASABF homologues in *C. elegans* [12]. yk150c7 and T22H6.5 are still thought to be ASABF-related, because of the significant similarity in their predicted amino acid sequence and gene organization (H. Zhang and Y. Kato, unpublished work). However, *abf-1* and *abf-2* should be genes more closely related to ASABF because the sequence identity of these peptides to ASABF is much higher than that of yk150c7 and T22H6.5. yk150c7 and T22H6.5.

Advantage of combinational analysis with *A. suum*

In many invertebrates, antimicrobial peptides are produced for defence against the infection of microbes [25]. As the first step to reveal the immune defence system of *C. elegans*, we decided to identify the immune effector molecules that contribute to the defence against microbes, such as antimicrobial peptides, as previously described [11,12]. However, database searches could not find any clear-cut homologues for known antimicrobial peptides isolated from other animals, and this situation has not changed. The length of the antimicrobial peptides is not very considerable, e.g. cecropins and defensins are approx. 30–40 residues long [25]. In addition, the molecular evolution of antimicrobial peptides tends to be rapid, presumably due to co-evolution with invading microbes [25]. An extreme example is the mature region of mammalian defensins whose amino acid changes have occurred in a co-ordinated fashion, i.e. changed molecules were positively selected [26]. Thus the database searches using sequences from evolutionally distant animals should not be effective in identifying the immune-related peptides susceptible to variation, such as antimicrobial peptides, even though the *C. elegans* genome sequence project has already been completed. Next, we intended to directly purify antimicrobial peptides from *C. elegans*. However, *C. elegans* is a very small animal (approx. 1 mm in length). The whole bodies of *C. elegans* contain a wide variety of proteins/peptides from all tissues, and are not suitable for the purification of tissue-specific peptides. In contrast, the parasitic nematode *A. suum* is big enough (adult females are 25–35 cm in length) for isolation of the pseudocoelomic fluid. The pseudocoelomic fluid was much less complex. Actually, antimicrobial activity was detected in the pseudocoelomic fluid,

in addition to bacteriolytic and agglutinating activities [11]. Finally, we decided to purify the antimicrobial peptides in the body fluid of *A. suum* and subsequently search the *C. elegans* genome sequence database to identify their homologues. Consequently, ASABF was purified from the pseudocoelomic fluid of *A. suum*, and its primary structure was determined [12]. Currently, ABF-1 and ABF-2 are the only clear-cut homologues of ASABF, detected by database searches using the Blast algorithm. No homologues have been found in other evolutionally distant organisms. In addition, we could not find any *C. elegans* homologues of antimicrobial peptides isolated from higher animals, such as insects or vertebrates. This observation is consistent with the tendency of rapid molecular evolution of antimicrobial peptides [25], i.e. even if antimicrobial peptides in nematodes were evolutionally related to those in higher animals, it is difficult to detect their counterparts using sequence similarity. The recombinant ABF-2 was a potent antimicrobial peptide and is thought to be functionally identical to ASABF. In addition, the same origin of *abf-1* and *asabf* was suggested by the study of the exon–intron junction, suggesting that *A. suum* is proximal enough in evolution to *C. elegans* to predict true homologues of immune-related factors liable to vary, such as antimicrobial peptides. Furthermore, *A. suum* is big enough for purification and characterization of tissue-specific proteins/peptides. These results suggest that the combination of biochemical identification in *A. suum* and homologue hunting in *C. elegans* is a powerful method of finding rapidly evolved proteins, such as some immune-related molecules in *C. elegans*.

Physiological role of *abf* genes

Immunofluorescence and transgenic studies revealed that ABF-2 is mainly produced in the pharynx and may be secreted into the lumen. This observation is consistent with the prediction that ABF-2 is a secretory peptide. It is noteworthy that live bacteria are concentrated in the pharynx due to normal feeding, because *C. elegans* is a natural bacteria eater. Thus ABF-2 should defend the pharyngeal tissue against microbial infection from the lumen. The recombinant ABF-2 was a potent antimicrobial peptide effective against Gram-positive and Gram-negative bacteria, and yeasts, suggesting that ABF-2 should contribute to the defence against a wide variety of microbes. The surface defence in the digestive tract, respiratory apparatus, epidermis, or reproductive organs by antimicrobial peptides has also been observed in insects and vertebrates, suggesting that this defence strategy is of evolutionally ancient origin. Another possible role of ABF-2 is to help with the digestion of ingested bacteria. The possibility that these antimicrobial peptides also contribute to defence in other tissues, in which GFP fusion peptides were detected, is not yet excluded.

Both ABF-1 and ABF-2 were undetectable in pre-hatching embryos by immunofluorescence and transgenic studies. Worms could be exposed to more severe bacterial stress after hatching because of the loss of egg shells as a physical barrier and the start of ingestion. Thus the post-hatching production of these antimicrobial peptides is thought to be well adapted. However, the transcripts of *abf-1* and *abf-2* were also detected in eggs by Northern-blot analysis, suggesting some possibilities of contribution to the defence in eggs and/or embryos.

Model for *abf* gene regulation: a hypothesis

abf-1 and *abf-2* are located side-by-side, in the same direction. The gene organization predicted that these genes form an operon. A polycistronic precursor RNA encoding both ORFs of *abf-1* and *abf-2* was detected by RT-PCR, suggesting that *abf-1* and

Figure 6 Model for gene regulation of *abf* genes

abf-2 actually form an operon. Therefore *abf-1* and *abf-2* are at least partially synchronously transcribed as an operon.

However, some atypical properties were found in the *abf* operon. First, the downstream gene, *abf-2*, received only SL1, but not SL2 or SL3–SL5. The addition of SL1 and SL2 is thought to be regulated in different manners depending on gene organization [22]. Usually the most upstream genes in the operons or non-operon genes exclusively receive SL1. In contrast, the downstream genes in the operons are known to be *trans*-spliced to SL2 or ‘novel’ SLs, i.e. SL3, SL4 or SL5. In some cases, the downstream genes also receive a mixture of SL1 and SL2. Only the downstream genes with no intercistronic regions are exclusively *trans*-spliced to SL1 [23]. *abf-2* was exclusively *trans*-spliced to SL1, and there was an intercistronic region of 78 bp (Figure 3, top left-hand panel).

Secondly, short SL-less transcripts for both *abf-1* and *abf-2* were detected. The *trans*-splicing of SL is thought to be essential for the processing of polycistronic precursor RNAs. These short SL-less mRNAs are not predicted to be the transcripts generated from the operon. No similarity was detected near the 5' terminal region of these short SL-less transcripts, suggesting that these transcripts are not generated by a common machinery. In addition, TATA-box-like (–29 for *abf-1*, and –25 for *abf-2*) and CCAAT (–67 for *abf-1*) sequences were found in the

upstream region of the short SL-less transcripts, suggesting that these could be produced as the transcripts of single genes (Figure 3, top left-hand panel). We, thus, speculate that *abf-1* and *abf-2* could be expressed in two separate ways, i.e. (1) synchronous expression as an operon, and (2) independent expression as multiple single genes (Figure 6). This hypothesis does not conflict with the observation that the tissue distributions of ABF-1–GFP and ABF-2–GFP were not completely identical, although both GFP fusion proteins were detected in the pharynx, i.e. *abf-1* and *abf-2* could not always be synchronously expressed, and the accumulation of these transcripts might be independently regulated. Further study is needed to reveal the exact mechanism of the gene regulation and RNA processing for the *abf* genes, also including the unexpectedly large transcripts of *abf-1*. However, it is at least suggested that some operons predicted by genomic sequences can be independent genes. Twenty-five percent of genes in *C. elegans* were predicted to form operons [27]. This ‘opportunistic’ operon hypothesis should be examined in other predicted operons.

We thank the following researchers: Professor Thomas Blumenthal (University of Colorado Health Science Center) for his technical suggestions and helpful discussion on the regulation of *abf* genes, Professor Andrew Fire (Department of Embryology, Carnegie Institution of Washington, Baltimore, MD, U.S.A.) for pPD95.77, Professor James M. Kramer (Northwestern University Medical School, Chicago, IL, U.S.A.) for pRF4, Dr Alan Coulson (The Sanger Centre, Cambridge, U.K.) for C50F2, Dr Shigenobu Yoshida (National Institute for Agro-Environmental Sciences, Ibaraki, Japan) and Dr Ritsuko Murakami (National Institute of Agrobiological Sciences, Ibaraki, Japan) for help with the antimicrobial assay, and Professor Ronald H.A. Plasterk (the Netherlands Cancer Institute) and Professor Yasumi Ohshima (Kyushu University) for their helpful suggestions. This work was supported by the Bio Design Program (BDP-01-V-1-1), the Pioneer Research Project Fund of the Ministry of Agriculture, Forestry and Fisheries, and Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan.

REFERENCES

- Hoffmann, J. A., Kafatos, F. C., Janeway, C. A. and Ezekowitz, R. A. (1999) Phylogenetic perspectives in innate immunity. *Science* (Washington, D.C.) **284**, 1313–1318.
- The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* (Washington, D.C.) **282**, 2012–2018.
- Mahajan-Miklos, S., Tan, M.-W., Rahme, L. G. and Ausubel, F. M. (1999) Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa*–*Caenorhabditis elegans* pathogenesis model. *Cell* (Cambridge, Mass.) **96**, 47–56.
- Tan, M.-W., Mahajan-Miklos, S. and Ausubel, F. M. (1999) Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 715–720.
- Tan, M.-W., Rahme, L. G., Sternberg, J. A., Tompkins, R. G. and Ausubel, F. M. (1999) *Pseudomonas aeruginosa* killing of *Caenorhabditis elegans* used to identify *P. aeruginosa* virulence factors. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2408–2413.
- Darby, C., Cosma, C. L., Thomas, J. and Manoil, C. (1999) Lethal paralysis of *Caenorhabditis elegans* by *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15202–15207.
- Tan, M.-W. and Ausubel, F. M. (2000) *Caenorhabditis elegans*: a model genetic host to study *Pseudomonas aeruginosa* pathogenesis. *Curr. Opin. Microbiol.* **3**, 29–34.
- Aballay, A., Yorgey, P. and Ausubel, F. M. (2000) *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. *Curr. Biol.* **10**, 1539–1542.
- Labrousse, A., Chauvet, S., Couillault, C., Leopold-Kurz, C. and Ewbank, J. J. (2000) *Caenorhabditis elegans* is a model host for *Salmonella typhimurium*. *Curr. Biol.* **10**, 1543–1545.
- Hodgkin, J., Kuwabara, P. E. and Corneliusen, B. (2000) A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Curr. Biol.* **10**, 1615–1618.
- Kato, Y. (1995) Humoral defense of the nematode *Ascaris suum*: antibacterial, bacteriolytic and agglutinating activities in the body fluid. *Zool. Sci.* **12**, 225–230.
- Kato, Y. and Komatsu, S. (1996) ASABF, a novel cysteine-rich antibacterial peptide isolated from the nematode *Ascaris suum*. Purification, primary structure, and molecular cloning of cDNA. *J. Biol. Chem.* **271**, 30493–30498.

- 13 Zhang, H., Yoshida, S., Aizawa, T., Murakami, R., Suzuki, M., Koganezawa, N., Matsuura, A., Miyazawa, M., Kawano, K., Nitta, K. and Kato, Y. (2000) *In vitro* antimicrobial properties of recombinant ASABF, an antimicrobial peptide isolated from the nematode *Ascaris suum*. *Antimicrob. Agents Chemother.* **44**, 2701–2705
- 14 Lewis, J. A. and Fleming, J. T. (1995) Basic culture methods. In *Caenorhabditis elegans*, Modern Biological Analysis of an Organism (Epstein, H. F. and Shakes, D. C., eds.), pp. 3–29, Academic Press, San Diego, CA
- 15 Beanan, M. J. and Strome, S. (1992) Characterization of a germ-line proliferation mutation in *C. elegans*. *Development* (Cambridge, U.K.). **116**, 755–766
- 16 Spieth, J., Brooke, G., Kuersten, S., Lea, K. and Blumenthal, T. (1993) Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. *Cell* (Cambridge, Mass.) **73**, 521–532
- 17 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
- 18 Miller, D. and Shakes, D. C. (1995) Immunofluorescence microscopy. In *Caenorhabditis elegans*, Modern Biological Analysis of an Organism (Epstein, H. F. and Shakes, D. C., eds.), pp. 373–375, Academic Press, San Diego, CA
- 19 Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991) Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970
- 20 Roussel, D. L. and Bennett, K. (1992) *Caenorhabditis* cDNA encodes an eIF-4A-like protein. *Nucleic Acids Res.* **20**, 3783
- 21 Cociancich, S., Goyffon, M., Bontems, F., Bulet, P., Bouet, F., Menez, A. and Hoffmann, J. (1993) Purification and characterization of a scorpion defensin, a 4 kDa antibacterial peptide presenting structural similarities with insect defensins and scorpion toxins. *Biochem. Biophys. Res. Commun.* **194**, 17–22
- 22 Blumenthal, T. and Steward, K. (1997) RNA processing and gene structure. In *Caenorhabditis elegans*, Modern Biological Analysis of an Organism (Epstein, H. F. and Shakes, D. C., eds.), pp. 117–146, Academic Press, San Diego, CA
- 23 Williams, C. and Blumenthal, T. (1999) SL1 trans splicing and 3'-end formation in a novel class of *Caenorhabditis elegans* operon. *Mol. Cell. Biol.* **19**, 376–383
- 24 Kelly, W., Xu, S.-Q., Montgomery, M. K. and Fire, A. (1997) Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238
- 25 Boman, H. G. (1996) Gene-encoded peptide antibiotics and the concept of innate immunity: an update review. *Scand. J. Immunol.* **43**, 475–482
- 26 Hughes, A. L. and Yeager, M. (1997) Coordinated amino acid changes in the evolution of mammalian defensins. *J. Mol. Evol.* **44**, 675–682
- 27 Zorio, D. A., Cheng, N. N., Blumenthal, T. and Spieth, J. (1994) Operons as a common form of chromosomal organization in *C. elegans*. *Nature* (London) **372**, 270–272

Received 14 May 2001/27 September 2001; accepted 14 November 2001