Bombyxin F1 Gene: Structure and Expression of a New Bombyxin Family Gene That Forms a Pair with Bombyxin B10 Gene

Yoshida Ikuyo, Tsuzuki Seiji, Salah Eldin Abdel Salam, Ino Masaya, Korayem Ahme Mohamed, Sakurai Sho, Iwami Masafumi

<table>
<thead>
<tr>
<th>Journal or publication title</th>
<th>Zoological Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>14</td>
</tr>
<tr>
<td>Number</td>
<td>4</td>
</tr>
<tr>
<td>Page range</td>
<td>615-622</td>
</tr>
<tr>
<td>Year</td>
<td>1997-08-01</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2297/10546">http://hdl.handle.net/2297/10546</a></td>
</tr>
</tbody>
</table>
Bombyxin F1 Gene: Structure and Expression of a New Bombyxin Family Gene That Forms a Pair with Bombyxin B10 Gene

Ikuyo Yoshida¹,², Seiji Tsuzuki², Salah Eldin Abdel Salam¹,², Masaya Ino¹, Ahmed Mohamed Korayem¹,², Sho Sakurai¹,², and Masafumi Iwami¹,²*

¹Department of Biology, Faculty of Science and ²Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa 920-11, Japan

ABSTRACT—Bombyxin F1 gene, a new bombyxin family gene, has been identified. The F1 gene forms a pair with bombyxin B10 gene with an opposite transcriptional orientation and the gene pair F1/B10 is located between bombyxin gene pairs B9/C1 and A7/B7 in a bombyxin gene cluster. The nucleotide sequence of the F1 gene and its deduced amino acid sequence deviate moderately from those characterized previously for the family-A, family-B, family-C, family-D, and family-E bombyxin genes; the bombyxin F1 gene and preprobombyixin F1 share no more than 62% and 53% sequence identities with other bombyxin members, respectively. Harr-plot analysis indicated that the spacer of the F1/B10 gene pair has low sequence similarity with that of other bombyxin gene pairs characterized. The bombyxin F1 mRNA in Bombyx mori brain was shown to locate in four pairs of medial neurosecretory cells, which also produce other bombyxin family mRNAs. Genomic Southern hybridization indicated that the Bombyx haploid genome contains a single copy of the family-F bombyxin gene.

INTRODUCTION

The bombyxin gene, belonging to the insulin-related multigene family, comprises at least 30 gene copies per haploid genome of the silkmoth Bombyx mori (Kondo et al., 1996; Tsuzuki et al., 1997). The 30 bombyxin genes have been classified into five families, A, B, C, D, and E, according to their sequence similarity. The Bombyx genome includes 10 family-A genes, 12 family-B genes, six family-C genes, one family-D gene, and one family-E gene. The bombyxin genes except the family-E gene cluster in two genomic segments in unique distribution patterns (Kondo et al., 1996). The arrangement of these genes has been classified into three categories: gene pairs, gene triplets, and single genes.

Bombyxin genes share structural features with the vertebrate insulin gene. Bombyxin genes encode precursor molecules which consist of four domains in the order of a signal peptide, B chain, C-peptide, and A chain (Iwami et al., 1989). The C-peptide is flanked by the processing signals, indicating the excision of the C-peptide after translation (Iwami et al., 1989) and the A and B chains of the mature molecules are linked together by disulfide bonds (Nagasawa et al., 1988), like insulin (Steiner et al., 1985). The bombyxin genes, however, possess some characteristic features markedly different from the vertebrate insulin gene. The copy number of bombyxin gene is more than 30 per haploid genome (Kondo et al., 1996; Tsuzuki et al., 1997). All bombyxin genes so far characterized lack introns, indicating that they are presumably the functional processed genes or genes derived from the processed genes (Iwami, 1990; Iwami et al., 1989, 1990; Kawakami et al., 1989; Kondo et al., 1996; Tsuzuki et al., 1997). Bombyxin genes are expressed predominantly in the brain (Kawakami et al., 1989) and at low levels in a number of other tissues of Bombyx mori (Iwami et al., 1996b), in contrast to the vertebrate insulin gene which is expressed in the gastroenteronic organ and almost silent in the brain.

In the course of determination of nucleotide sequence of the upstream region of bombyxin B10 gene, we found a new bombyxin gene whose sequence deviates moderately from any of the previously identified bombyxin family genes. We therefore define this bombyxin gene group as family-F and named the presently isolated gene as F1 gene. We determined the sequences of two allelic forms of the gene, one from Japanese race, Kinshu, and the other from Chinese race, Showa. Genomic Southern hybridization indicated the presence of a single copy of the family-F bombyxin gene in the Bombyx haploid genome. We also localized the expression site to four pairs of the medial neurosecretory cells in the brain by whole-mount and section in situ hybridizations.

* Corresponding author: Tel. +81-76-264-5711; FAX. +81-76-264-5977 or 234-4140.
MATERIALS AND METHODS

Animals
Eggs of racial hybrid, Kinshu (Japanese race) x Showa (Chinese race), of the silkmoth Bombyx mori were obtained from Kanebo Silk Elegance (Kasugai, Aichi, Japan). Larvae were reared at 25 ± 1°C under a photoperiod of 12 h light and 12 h dark on artificial diet (Silkmate, Nihon Nosan Kogyo, Yokohama, Japan).

Clone isolation and sequencing of bombyxin F1 gene
Two phage clones (107 and 421) previously isolated (Kondo et al., 1996) were used (Fig. 1). Clone 107 contains a part of bombyxin gene cluster derived from the Bombyx Kinshu race and clone 421 contains the allelic component of the same locus derived from the Showa race. The upstream regions, which contain the newly identified bombyxin F1 gene copies, of bombyxin B10 genes in both clones were sequenced; the nucleotide sequences of both allelic forms of bombyxin F1 gene and the spacer between genes F1 and B10 were determined. Both DNA strands were sequenced by the A. L. F. DNA sequencer (Pharmacia, Sweden) using an AutoRead sequencing kit (Pharmacia) or a Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham, UK). Nucleotide sequence data were assembled and analyzed using computer program in DNASIS (Hitachi Software Engineering, Yokohama, Japan).

In situ hybridization
For whole-mount in situ hybridization, brains of the fifth-instar Bombyx larvae were fixed for 40 min in a fixative containing 85% ethanol, 4% formaldehyde, and 5% acetic acid at room temperature, and washed in 10 mM phosphate-buffered saline (PBS), pH 7.4, containing 15% (w/v) sucrose at 4°C for 15 to 20 hr (Bandtlow et al., 1987). After washing with PBT (PBS containing 0.05% Tween20), the brains were treated with protease K (0.05 mg/ml in PBT) at 37°C for 40 min, and fixed again with 3% (w/v) parafomaldehyde in PBS at room temperature for 20 min. The brains were washed three times with PBT and hybridized with 100 ng of a bombyxin F1 gene probe (5'-ATCAGCCGGAAGTTCCGCGAT-3') labeled with digoxigenin at the 5'-end at 37°C for 20-24 hr in 100 µl solution of 50% formamide, 5 x SSC (SSC: 0.15 M NaCl, 0.015 M sodium citrate), and 5% (w/v) dextran sulfate. The probe was specific for the C-peptide domain of bombyxin F1 gene. We also used a digoxigenin-labeled probe whose sequence (5'-CGGAAATCAACCAGCAGATAAC-3') was specific for the signal peptide domain of bombyxin F1 gene. The condition prevented cross-hybridizations among bombyxin family mRNAs. After washing with 50% formamide, 5 x SSC and then with PBT at room temperature, the brains were treated with 5% sheep serum (Capel, USA) at 4°C for 15-20 hr and with an alkaline phosphatase conjugated anti-digoxigenin antibody (Boehringer Mannheim, Germany) at room temperature for 2 hr. The color development was done with 4-nitroblue-tetrazolium chloride and 5-bromo-4-chloro-3-indoly-phosphate. After dehydration, the brains were clarified with methyl salicylate.

Section in situ hybridization was carried out essentially the same as whole-mount in situ hybridization as described previously (Iwami et al., 1996b). After fixation and serial dehydration, brains were embedded in paraaffin wax (Paraplast, Oxford Labware, USA) and sectioned at 10 µm. The sections were then dewaxed, hydrated, and treated with 0.2 N HCl for 20 min followed by protease K (1 ng/ml in 20 mM Tris-HCl, pH 7.4, 2 mM CaCl2) at 37°C for 25 min. The sections were fixed again with 4% paraformaldehyde for 15 min, washed with PBS, and hybridized with both of the F1 probes used for whole-mount in situ hybridization under the above condition. The color development was done in the presence of 1 mM levamisole, a potent inhibitor of lysosomal phosphatase.

Genomic Southern hybridization
Large molecular mass DNA was prepared from the first-instar larvae, according to the method of Weeks et al. (1986). A 10 µg aliquot of the Bombyx genomic DNA was digested with BamHI, EcoRI, or HindIII, separated by electrophoresis on 0.7% (w/v) agarose gel, and transferred to nylon membrane filter with 0.4 N NaOH, 0.6 M NaCl. Hybridization was done with the 32P-labeled bombyxin F1 gene fragment (position 421-682 in Fig. 2) at 65°C in the presence of 6 x SSC. The filter was washed at 65°C with 2 x SSC, 0.1% SDS and then with 0.2 x SSC, 0.1% SDS. The condition did not allow cross-hybridization among bombyxin families.

RESULTS
Structure of bombyxin F1 gene
In the upstream region of bombyxin B10 gene (Fig. 1), we identified a gene which forms a pair with bombyxin B10 gene. The gene had sequence similarity with the previously
characterized bobynxin family genes but was found to belong to none of the five bobynxin families (Iwami et al., 1989, 1990; Kawakami et al., 1989; Kondo et al., 1996; Tsuzuki et al., 1997) because its sequence similarity with any of other family genes was 62% at most. The bobynxin genes belonging to the same family have more than 85% similarity in nucleotide sequence with each other with no exceptions (Kondo et al., 1996). We thus establish a new bobynxin gene group as family-F and designate the isolated gene from the segment A as gene F1 and that from the segment B as gene F1', since we previously designated the bobynxin genes carried by the segment B, which is allelic to segment A, with prime (') (Kondo et al., 1996). Genes F1 and F1' therefore represent the allelic forms from Chinese race and Japanese race, respectively (see DISCUSSION).

Figure 2 shows the nucleotide sequences of genes F1 and F1'. The nucleotide sequence of gene F1' is identical with that of gene F1, except for the position 625 of the sequence

---

**Fig. 2.** Nucleotide and deduced amino acid sequences of bobynxin F1 gene and the spacer between bobynxin F1 and B10 genes. F1 and F1' represent the genes carried by the Showa and Kinshu genome, respectively. Signal peptide, B chain, C-peptide, and A chain are indicated. Slash indicates the termination codon. Dashes (--) in the F1' sequence represent the nucleotides identical to those of the F1 sequence. Amino acid sequences are shown in single-letter code and asterisks (*) in the preprobobynxin F1' sequence represent the amino acids identical to those of the preprobobynxin F1 sequence. One amino acid residue in the C-peptide domain is substituted between preprobobynxin F1 and F1'. TATA box-like sequence and polyadenylation signal for the genes F1 and F1' are indicated by thick underlines. TATA box-like sequence for the B10 gene is located at position 82-88 and indicated by thin underline.
where G is replaced by A. The nucleotide sequences have been deposited in the DDBJ/GenBank/EMBL DNA data base under accession numbers AB001049 (gene F1 and the spacer between genes F1 and B10) and AB001050 (gene F1' and the spacer between genes F1' and B10). The transcriptional orientation of bombyxin F1 gene is opposite to that of the B10 gene. The open reading frame of the F1 gene apparently encodes prepombbyxin F1 with four domains in the order of signal peptide/B chain/C-peptide/A chain, as in other prepombbynkinins (Iwami et al., 1989, 1990; Kawakami et al., 1989; Kondo et al., 1996; Tsuzuki et al., 1997), prepombbyxin-related peptides (Kimura-Kawakami et al., 1992; Iwami et al., 1996a), and preproinsulins (Steiner et al., 1985). A TATA box-like sequence (TATTTAAA) and a polyadenylation signal sequence (AATAAAA) are observed 99 bp upstream from the translation initiation site and 41 bp downstream from the termination site, respectively. The coding region of the bombyxin F1 gene lacks introns and no splicing signal is found in the untranslated regions as in other bombyxin genes (Iwami, 1990; Iwami et al., 1989, 1990; Kawakami et al., 1989; Kondo et al., 1996; Tsuzuki et al., 1997).

**Brain neurosecretory cells that produce bombyxin F1 mRNA**

Figure 3 shows the cellular localization of bombyxin F1 mRNA in the brain of *Bombyx* larvae. The F1 mRNA was localized to four pairs of medial large neurosecretory cells which are classified as M1 subgroup-a cells (Ichikawa, 1991). The M1 subgroup-a cells also produce the family-A (Iwami, 1990), family-B (Iwami, 1990), family-C (unpublished), family-D (unpublished), and family-E (Tsuzuki et al., 1997) bombyxin mRNAs. We previously localized the expression site of the bombyxin-related genes of *Samia* (Inoue et al., unpublished) and *Agrius* (Iwami et al., 1996a) to the medial neurosecretory cells in the brain. The expression of bombyxin family genes in the medial neurosecretory cells is therefore an intriguing feature, in contrast to the vertebrate insulin gene which is expressed in the gastroenteric organ and is almost silent in the brain.

**The Bombyx genome contains a single copy of the family-F bombyxin gene**

To estimate the number of the family-F gene copy in the *Bombyx* genome, the genomic DNA prepared from the racial hybrid, Kinshu × Showa, was subjected to Southern hybridization analysis (Fig. 4). Two bands (13.5 kbp and 4.5 kbp)

---

**Fig. 3.** (a) Schematic drawing of a *Bombyx* brain. The portions boxed with solid and broken lines represent the areas shown by photographs in (b) and (c), respectively. (b) Whole-mount in situ hybridization of a *Bombyx* brain. The bombyxin F1 mRNA was detected using the C-peptide specific oligonucleotide as a probe. Arrows indicate the neurosecretory cells which produce the bombyxin F1 mRNA. The same cells were also positive when the signal peptide specific probe was used. The photograph shows the medial portion of the brain. Scale bar = 50 μm. (c) In situ hybridization. A section of a brain was hybridized with the F1 probes. Four medial neurosecretory cells (arrows) are positive here, but serial sections (not shown) demonstrate eight (four in each brain hemisphere) positive cells. Scale bar = 50 μm.

**Fig. 4.** Genomic Southern hybridization analysis of the family-F bombyxin gene. Genomic DNA from the *Bombyx* racial hybrid (Kinshu × Showa) digested with BamHI (lane B), EcoRI (lane E), HindIII (lane H), or PstI (lane P) was analyzed. The sizes of DNA fragments containing the family-F bombyxin genes are shown at the right in kbp.
were observed when the DNA was digested with BamHI. The 13.5 kbp and 4.5 kbp bands correspond in size to a BamHI fragment bearing bombyxin gene F1 and another BamHI fragment bearing gene F1', respectively (Fig. 1). The 8.0 kbp and 6.2 kbp bands in the EcoRI-digested DNA correspond in size to EcoRI fragments bearing genes F1' and F1, respectively. The size of the HindIII fragment bearing gene F1 or F1' is 2.7 kbp when calculated from the restriction maps. The 2.7 kbp band were observed with about double intensity. When the DNA was digested with PstI, the 6.9 kbp band with about double intensity was observed. The 6.9 kbp band contained two DNA fragments, one for gene F1 and the other for gene F1'. The sizes of the hybridized fragments digested with BamHI, EcoRI, HindIII, and PstI therefore exactly matched those of the expected fragments from the restriction maps and no extra fragment bearing another copy of the family-F gene could be detected. We thus concluded that the Bombyx haploid genome contains a single copy of the family-F gene.

**DISCUSSION**

**Organization of bombyxin F1 gene**

We identified the bombyxin F1 gene which forms a pair with the bombyxin B10 gene. The gene pair F1/B10 is located between bombyxin gene pairs B9/C1 and A7/B7. We previously characterized the allelic components of the genomic DNA segment which includes multiple copies of bombyxin gene (Kondo et al., 1996). One genomic segment (map A in Fig. 1) derived from Chinese race, Showa, and the other (map B in Fig. 1) derived from Japanese race, Kinshu. The right-hand part of segment A was very similar to the majority of segment B with respect to the distribution patterns of bombyxin genes and restriction sites. These two segments were thus concluded to represent the heterozygous components of the same chromosomal segment (Kondo et al., 1996). Segment C was presumed to have no allelic variation between the two races. The haploid genome therefore contains segments A (or B) and C. Now the segment A contains 24 bombyxin gene copies and the segment C contains six bombyxin gene copies (Fig. 1). Bombyxin E1 gene, a unique family-E gene, has not been localized on these segments (Kondo et al., 1996). Total number of the bombyxin gene copy is thus revealed to be 31 in the haploid Bombyx genome. The family-A, family-B, and family-C bombyxin genes have 10, 12, and six copies in the genome, respectively. The family-D, family-E, and family-F bombyxin genes exist in single copy in the genome.

The arrangement of the bombyxin gene copies on these segments can be classified into three categories: gene pairs, gene triplets, and single genes, as mentioned previously (Kondo et al., 1996). Sixteen out of the 31 bombyxin genes form pairs, 12 form triplets, and two lie singly. (Bombyxin E1 gene has not been known to form a pair or triplet or to exist as single. See Tsuzuki et al., 1997). In the pairs, two genes, each of which belongs to different families, are opposed with an opposite transcriptional orientation. Most gene pairs are composed of family-A and family-B members, while gene pairs B/C and F/B exist singly. All triplets are composed of the family-A, family-B, and family-C genes and the three genes are opposed to each other. The transcriptional direction of the family-A and family-C genes is opposite to that of the family-B gene. All the family-C genes in the triplets are pseudogenes. The triplets have probably been generated by an unequal crossing-over between two gene pairs, B/A and B/C (Kondo et al., 1996). Bombyxin C2 and D1 genes are present singly, without forming either a pair or triplet but there is the possibility that they forms pairs or triplets with undetected bombyxin genes because of the presence of molecules whose primary structures resemble, but deviate appreciably from, the six bombyxin families characterized so far (Nagasawa et al., 1984; Jhoti et al., 1987). Further studies expect to disclose more bombyxin and bombyxin-related gene copies in Bombyx.

**Nucleotide sequence of the spacer between bombyxin F1 and B10 genes**

The bombyxin gene pair F1/B10 is located between the gene pairs B9/C1 and A7/B7 (Fig. 1). The distance between genes F1 and B10 is 0.4 kbp, which is shorter than that between the spacer of any other bombyxin gene pairs (0.5 to 2.3 kbp; Kondo et al., 1996). The spacer regions of a similar gene pair organization have been extensively studied for the silkmoth chorion genes (Goldsmith and Kafatos, 1984). The expression of various chorion gene pairs are coordinately regulated in a developmental-stage-specific manner by the spacer sequences (Spoor et al., 1986, 1993). It is thus probable that the spacers of bombyxin gene pairs play important role(s) in regulating gene expression.

Figure 5 compares the nucleotide sequence of the spacer of gene pair F1/B10 with that of other gene pairs, B1/A2 (map C in Fig. 1), A7/B7 (map A in Fig. 1), and B9/C1' (map B in Fig. 1), by Harr-plot analysis. The spacers of these three gene pairs have different nucleotide sequence each other and therefore are presumed to contain different regulatory units. The upstream region of the gene B10 contains a TATA box-like sequence (Fig. 2) and the sequence identity extends to about 150 bp from the translation initiation site. The upstream region of gene F1 except for a TATA box-like sequence, however, does not show sequence similarity with that of genes A2, A7, and C1'. In addition, the central region of the spacer of F1/B10 also has no significant sequence similarity with that of gene pairs B1/A2, A7/B7, and B9/C1'. We previously suggested that two gene pairs B1/A2 and B2/A3 (map C in Fig. 1) must have been generated by duplication of an original pair and may share a transcriptional regulation mechanism because of their high sequence similarity (Kawakami et al., 1989). On the other hand, the central region of spacer of A7/B7 had no significant sequence similarity with that of gene pairs B1/A2 and B2/A3 (Kondo et al., 1996), and therefore gene pair A7/B7 could be under the control of a regulation mechanism different from that of B1/A2 and B2/A3. Similarly, the regulatory mechanism of gene pair F1/B10 may be also different from that of any of bombyxin gene pairs characterized previously.
Amino acid sequence of preprobombbyxin F1

Figure 6 compares the deduced amino acid sequence of preprobombbyxin F1 with that of preprobombbyxins (Iwami et al., 1989, 1990; Kawakami et al., 1989; Kondo et al., 1996; Tsuzuki et al., 1997), prepro-bombbyxin-related peptides of the saturnid moth Samia cynthia ricini (Kimura-Kawakami et al., 1992), prepro-bombbyxin-related peptides of the hornworm Agrius convolvuli (Iwami et al., 1996a), and human proinsulin (Bell et al., 1980). The amino acid sequence of prepro-bombbyxin F1 deviates considerably from that of any of other preprobombbyxins. Preprobombbyxins within the same family have at least 73% identical sequence with each other, whereas preprobombbyxin F1 has no more than 53% identical sequence with other preprobombbyxins and 31% with proinsulin.

Similarities were found between residues of bombbyxin F1 and other insulin family members that contribute to the tertiary structure of insulin (Blundell et al., 1970). All cysteins (positions 6, 7, 11, 20 of A chain and positions 7 and 19 of B chain; taking the amino termini of the A and B chains of insulin as position 1) are conserved. All residues responsible for the hydrophobic core formation are conserved or substituted by hydrophobic residues (positions 2, 3, 16, 19 of A chain and positions 2, 6, 11, 15, and 18 of B chain). Glycines (position 1 of A chain and position 8 of B chain), which contribute to the main chain conformation, are also conserved. It therefore seems highly probable that bombbyxin F1 has an insulin-like globular structure (see Jhoti et al., 1987 and Nagata et al., 1995). Interestingly, the sequence of the central part of B chain (position 6-20 of B chain) is conserved or conservatively substituted between bombbyxin F1 and other bombbyxin family
Fig. 6. Amino acid sequence comparison among bombixins, bombixin-related peptides, and insulin. Amino acids are numbered from the amino terminus of each domain of insulin.

ACKNOWLEDGMENTS

This study was supported in part by grants from the Ministry of Education, Science, Sports and Culture of Japan.

REFERENCES

Iwami M, Tanaka A, Hano N, Sakurai S (1996b) Bombyxin gene expression in tissues other than brain detected by reverse transcription-polymerase chain reaction (RT-PCR) and in situ hybridization. EXPERIMENTA 52: 882–887
search Foundation, Osaka, pp 123–126

(Received March 19, 1997 / Accepted April 21, 1997)