

# Fibroin-modulator-binding protein-1 (FMBP-1) contains a novel DNA-binding domain, repeats of the score and three amino acid peptide (STP), conserved from *Caenorhabditis elegans* to humans

Shigeharu Takiya<sup>1,2,\*</sup>, Tetsuro Ishikawa<sup>2</sup>, Katsuya Ohtsuka<sup>2</sup>, Yoshinori Nishita<sup>1</sup> and Yoshiaki Suzuki<sup>3</sup>

<sup>1</sup>Center for Advanced Science and Technology and <sup>2</sup>Graduate School of Environmental Earth Science, Hokkaido University, North 10, West 8, Kita-ku, Sapporo 060-0810, Japan and <sup>3</sup>National Institute for Basic Biology, Nishigo-naka 38, Okazaki 444-8585, Japan

Received October 14, 2004; Revised and Accepted January 18, 2004

DDBJ/EMBL/GenBank accession no. AB163434

## ABSTRACT

The predicted transcriptional regulatory factor for the fibroin gene of the silkworm *Bombyx mori*, fibroin-modulator-binding protein-1 (FMBP-1), was purified by sequential DNA affinity column chromatography, and cDNA clones corresponding to FMBP-1 were isolated from a library. The N-terminal half of FMBP-1 has a weak similarity to the DNA-binding domain of several transcriptional regulatory factors in higher plants. The C-terminal half contains four tandem repeats of a novel 23 amino acid motif, which we named the score and three amino acid peptide (STP). Other genes containing STP repeats were found in *Drosophila*, *Caenorhabditis elegans*, mouse and human. Mutational analysis of FMBP-1 showed that the STP repeats form a novel DNA-binding domain. Sequences flanking STP repeats modulated DNA-binding activity. The FMBP-1 gene was expressed during the fourth to fifth instar. FMBP-1 activity appeared to be regulated at the transcriptional level and by the post-transcriptional modification.

## INTRODUCTION

The larval silk gland of *Bombyx mori* is distinctly divided into anterior, middle and posterior regions, based on structural and functional criteria. Several silk protein genes are expressed in a spatial- and temporal-specific manner in the silk gland (1,2).

Fibroin-H-chain (usually fibroin), fibroin-L-chain and fibrohexamerin (previously P25) genes are expressed in the posterior region (3–8), while the sericin-1 and -2 genes are expressed in the middle region with sublocal specificity (6,9–13). These silk genes are expressed at every feeding stage, but not at every molting stage, during larval development (4–7,11,13,14).

Several protein factors that bind the transcriptional regulatory elements of silk genes have been detected and characterized in *B.mori* (15–21). Fibroin-modulator-binding proteins (FMBP)-1, FMBP-2 and FMBP-3 were first detected as factors bound to the intronic elements of the fibroin gene, but it was later found that all these FMBPs also bind to the upstream elements (21). cDNAs for the FMBP-2 and FMBP-3 were isolated and identified as Bm fork head (Bm Fkh) and POU-M1, respectively, the transcription factors having a winged-helix type and a POU-homeodomain type DNA-binding domain (22,23). Bm Fkh was expressed in both PSG and MSG, and bound to the regulatory elements of both the fibroin and sericin-1 genes (15,16,21). Bm Fkh is thought to be a common regulatory factor for the fibroin and sericin-1 genes. Expression of the fibrohexamerin gene is also regulated by Bm Fkh (20,24). POU-M1 is expressed in many organs during the organogenesis of the *Bombyx* embryo, and the silk gland is one of the organs that expresses POU-M1 (13,25). Though POU-M1 binds to the transcriptional activation element of the upstream region of the sericin-1 gene, and was first thought to be an activator for the sericin-1 gene (26), POU-M1 might function as a repressor or, in coordination with another factor, as an activator. POU-M1 is highly expressed in ASG, and the expression is complementary to that of the sericin-1 gene in MSG (13).

\*To whom correspondence should be addressed. Tel: +81 11 706 3590; Fax: +81 11 737 0536; Email: takiya@rmg.hokudai.ac.jp  
Present address:

Yoshiaki Suzuki, 5651-451 Matsukawa Village, North Azumi County, Nagano Prefecture 399-8501, Japan

The FMBP-1 activity detected by a DNA-binding assay closely correlated with the specificity of the fibroin gene expression (21). FMBP-1 was abundant in PSG at the fourth and fifth feeding stages, but barely detectable in MSG and PSG at the fourth molting stage. FMBP-1 might play an important role in the regulation of the tissue- and stage-specificity of the fibroin gene. In the present study, we purified FMBP-1 protein, isolated its cDNA and analyzed the structure and expression. We found that FMBP-1 contains a novel DNA-binding domain composed of tandem repeats of the score and a three amino acid long peptide (STP), conserved from *Caenorhabditis elegans* to humans.

## MATERIALS AND METHODS

### Animals

*B.mori* eggs, hybrids of Kinshu x Showa, were purchased from Ueda Sanshu (Ueda, Japan). Resulting larvae were reared on an artificial diet from Nippon Nohsankoh (Yokohama, Japan) or mulberry leaves, and developed to the appropriate developmental stages. Larvae were staged as described previously (5,27,28).

### Preparation of poly(dA-dT)-Cellulofine

Poly(dA-dT)-Cellulofine was prepared as recommended in the instruction manual (Seikagaku-kogyo). Poly(dA-dT) was mixed with FMP-activated Cellulofine in 50 mM phosphate buffer (pH 8.0), and the active FMP residues were blocked by 100 mM monoethanolamine in 10 mM Tris-HCl, pH 8.0. After washing with 10 mM sodium acetate, pH 4.5 and 0.5 M NaCl, and then with 10 mM Tris-HCl, pH 8.0 and 0.5 M NaCl, the poly(dA-dT)-Cellulofine was stored at 4°C.

### Purification of FMBP-1

Silk gland extracts were prepared as described previously (29,30). For the large-scale preparation of FMBP-1, PSG extracts were prepared with a modified method of Takiya *et al.* (31), dialyzed against K500 buffer [20 mM HEPES, pH 7.9, 500 mM KCl, 12.5 mM EDTA, pH 8.0, 15% (v/v) glycerol and 2 mM DTT] and loaded on a poly(dA-dT)-Cellulofine column. The column was washed, and bound proteins were fractionated sequentially with N500 (500 mM NaCl instead of 500 mM KCl in the K500 buffer), N700, N900, N1200 and N2000 buffer. Each fraction was analyzed for protein components by SDS-PAGE and FMBP-1 activity by an electrophoresis mobility-shift assay (EMSA). The FMBP-1 eluted in the N1200 fractions was pooled, diluted with N0 buffer below 0.9 M NaCl and reloaded on the poly(dA-dT)-Cellulofine. The proteins were again fractionated, and the eluates containing FMBP-1 activity were pooled and stored at -80°C.

### Screening of the cDNA library

A cDNA library was constructed with poly(A)<sup>+</sup> RNA from PSG 12 h after the fourth ecdysis, as described previously (23,31). A nitrocellulose filter merged with phage plaques was prepared as the standard method for screening. Before hybridization, the filter was washed in washing solution

(6× SSC, 0.5% SDS and 1 mM EDTA) at 46°C for 2 h, and soaked in prehybridization solution (6× SSC, 5× Denhardt's reagent, 0.5% SDS and 100 µg/ml ssDNA) at 60°C for 2 h. The prehybridization solution was replaced with new solution, and an oligonucleotide probe (5'-GTATTGTATGCATTCATCACTTAATCTAAAGTGTGTAGTGTACTAG-3') was added and incubated at 60°C for 16-24 h. After hybridization, the filter was washed with 2× SSC and 0.1% SDS at room temperature, and washed three times with 0.5× SSC and 0.1% SDS at 46°C for 20 min. The filter was air-dried, and exposed to X-ray film at -80°C overnight with an intensifying screen.

### Recovery of proteins from SDS-PAGE gels

SDS-PAGE was performed with 2.5% stacking and 12% separation gels. When the dilution of the protein sample was below detectability, proteins were adsorbed to StrataClean resin (Stratagene) and loaded onto gels following treatment with a gel loading solution, as described previously (31). Once the FMBP-1 protein was recovered from the polyacrylamide gel, the gel was negatively stained with zinc (32), and the protein was prepared as described previously (31,33). For the peptide sequencing of FMBP-1, the gel was stained with CBB, and the FMBP-1 band was cut out and treated for sequencing.

### EMSA

EMSAs were carried out as described by Takiya *et al.* (21). The oligonucleotide +290 probe (0.4-0.5 ng) (Figure 4) was incubated with 12 µg of PSG extract and an appropriate amount of purified FMBP-1 or recombinant FMBP-1 (rFMBP-1) in 10 µl of reaction mixture [10 mM Tris-HCl, pH 7.9, 7.5 mM MgCl<sub>2</sub>, 60 mM NaCl, 1 µg poly(dI-dC), 20% glycerol and 0.6% NP40] for 20 min on ice. The protein-probe complexes were separated on 7% polyacrylamide gels. In the competition assay, 1 µl (0-50 ng) of the unlabeled double-stranded competitor (10 mM Tris-HCl, pH 7.9, 1 mM EDTA and 200 mM NaCl) was added to the EMSA reaction mixture. In the hypershift assay, 1 µl of antiserum against the region corresponding to amino acids 1-98 of FMBP1 was added.

### Northern blot hybridization

Total RNA of PSG and MSG was prepared with phenol/chloroform treatment from the staged larvae, and 20 µg of RNA was electrophoresed per lane after denaturation, as below. Poly(A)<sup>+</sup> RNA was prepared from PSG and MSG on day 1 of the fifth instar with a Quick Prep mRNA isolation kit (Amersham Bioscience). Poly(A)<sup>+</sup> RNA (2 µg) was denatured in glyoxal/dimethyl sulfoxide solution, fractionated on 1% agarose/formaldehyde gels and transferred to a nylon membrane using standard methods. The entire coding region of FMBP-1 was labeled with the megaprime DNA-labeling system (Amersham Bioscience) and used for hybridization. Hybridization was performed in 6× SSC, 5× Denhardt's solution, 0.1% SDS and 100 µg/ml ssDNA at 50°C, and then the filter was washed with 0.2× SSC at 65°C. An oligonucleotide probe from *Bombyx* actin A3, 5'-TCCTGTACAATGGAGGGACCAG-3' (34), and a cDNA probe of BmHNF4 were used as controls.

### Western blotting

Silk gland extracts (50 µg protein) were incubated for 10 min at 65°C in a sample buffer (50 mM Tris-HCl, pH 6.8, 4% SDS, 100 mM DTT, 0.2% BPB and 10% glycerol), and separated by SDS-PAGE (12%) according to the standard methods. Proteins were transferred to a PVDF membrane (BioRad) by electroblotting at 50 mA for 90 min in a transfer buffer (20 mM glycine, 25 mM Tris and 20% methanol) using a semi-dry blotting apparatus. Blots were blocked by treatment with 5% (W/V) skim milk (Difco) in 1× TBS-T (2 mM Tris-HCl, pH 7.6, 13.7 mM NaCl and 1% Tween20) for 1–2 h at room temperature or overnight at 4°C, and then incubated with a 1:1000 dilution of rabbit anti-FMBP-1 antibody for 1 h at room temperature. After three rinses with TBS-T, the blots were treated with a second antibody (goat anti-rabbit immunoglobulin G conjugated to horseradish peroxidase) for 1 h. The signal was detected with an enhanced chemiluminescence (ECL) western blotting kit (Amersham Bioscience).

### Preparation of GST-FMBP-1 fusion proteins

Appropriate combinations of the oligonucleotide primers were selected and the regions of the cDNA corresponding to the parts of the FMBP-1 protein (Figure 8) were amplified by PCR. Products of the appropriate size were purified, and the sequences attached to the primers were digested by the restriction enzymes EcoRI and Sall or XhoI, and subcloned into the expression vector pGEX-4T-1 or pGEX-6P-1 (Amersham Bioscience).

*Escherichia coli* transformed with pGEX-FMBP1 DNA was cultured at 20 or 30°C, and the expression of the GST-FMBP-1 fusion protein was induced by the addition of 1 mM (20°C) or 0.1 mM (30°C) isopropyl-β-D-thiogalactopyranoside. After incubation for a further 4 h, cells were harvested, washed and resuspended in phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na<sub>2</sub>HPO<sub>4</sub> and 3.2 mM KH<sub>2</sub>PO<sub>4</sub>). Cells were broken by sonication and the addition of 10% Triton X-100 to a final concentration of 1%. The supernatant obtained by centrifugation was mixed with glutathione-Sepharose at 4°C overnight. Sepharose beads bound with the fusion protein were washed with PBS and then with 50 mM Tris-HCl, pH 8.0. Finally the bound protein was eluted with elution buffer (50 mM Tris-HCl, pH 9.5 and 5 mM glutathione), dialyzed against NP40 buffer (50 mM Tris-HCl, pH 7.9, 12.5 mM MgCl<sub>2</sub>, 100 mM NaCl, 0.1 mM EDTA, 20% glycerol, 0.1% NP40 and 1 mM DTT) and stored at -80°C.

## RESULTS

### Purification of FMBP-1 with poly(dA-dT)-affinity-resin

To search for an appropriate carrier for FMBP-1, we conducted an EMSA using the double-stranded synthetic polynucleotides, poly(dA-dT), poly(dA)poly(dT) or poly(dG-dC), and heat-denatured salmon sperm DNA or tRNA. High concentrations of any of these polynucleotides with ssDNA or tRNA did not compete efficiently in binding of FMBP-1 to the +290 region probe (21) (see also Figure 4), but poly(dA-dT) did compete at least as effectively as the specific probe (data not shown). Accordingly, we decided to use

poly(dA-dT) for the purification instead of the FMBP-1-binding elements of the fibroin gene, and prepared poly(dA-dT)-Cellulofine.

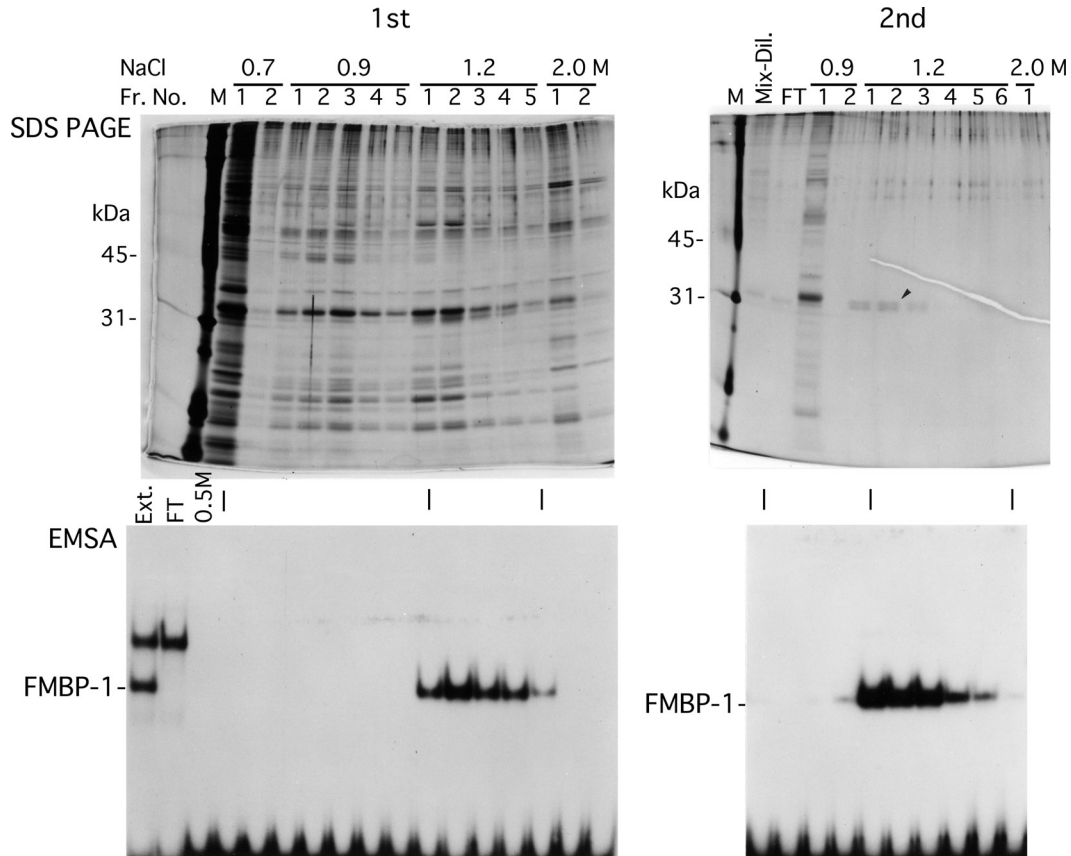
PSG extracts were loaded on the poly(dA-dT)-Cellulofine column, and fractionated with increasing concentrations of NaCl. The eluates were analyzed by SDS-PAGE, for protein composition, and an EMSA for FMBP-1 activity (Figure 1). The majority of the FMBP-1 activity was eluted at 1.2 M NaCl in two cycles of DNA affinity chromatography. In these fractions, two bands at ~32 kDa were detected in almost equal amounts by SDS-PAGE (Figure 1). We repeated the affinity purification on a large scale, and using these purification processes only one band of 32 kDa was obtained (see Discussion). The 32 kDa band was recovered from a preparative gel, and the binding specificity was examined in an EMSA with oligonucleotides of the FMBP-1-binding elements and their point mutants (21). The recovered 32 kDa protein showed the same binding specificity as FMBP-1 in crude PSG extracts (data not shown).

### Isolation and sequence analysis of FMBP-1 cDNA

The internal amino acid sequence of the 32 kDa protein was analyzed, and a total of 65 amino acids of four peptides were determined (Figure 2A). However, a search of various databases using the BLAST program found no amino acid sequence homologous to the peptides of FMBP-1. Therefore, we isolated a cDNA clone using PCR to determine the nucleotide sequence excluding degeneracy. The PCR was performed using primers containing degeneracy based on the amino acid sequence of peptide 2 (Figure 2A). The products were separated on a polyacrylamide gel, and those having the appropriate size were cloned. The inserted DNA was sequenced and a unique sequence of the internal region surrounded by the degenerated primers was used as a probe for library screening. As a result, we isolated several clones from the cDNA library of PSG at 12 h after the fourth ecdysis. The longest cDNA was 1481 nt, and included an open reading frame (ORF) that was 218 amino acids long with a predicted molecular mass of 25 kDa. The starting ATG codon was predicted at nucleotide position 111, and a poly(A) signal was found at position 1449 (AB163434). Southern blotting using the entire cDNA probe with genomic DNA demonstrated that FMBP-1 is encoded by a single copy gene in *Bombyx* (data not shown).

Figure 2A shows the deduced amino acid sequence of the putative protein encoded by these cDNAs. The protein contains all the peptides (P1-P4) whose amino acid sequences had been determined. The region from amino acid 55 to 84 is rich in acidic residues, while the region from 85 to 192 is rich in basic residues. The N-terminal region of the protein showed weak similarity to the AP2 domain, which is thought to be a DNA-binding domain of transcription factors in higher plants, such as APETALA2 (35), EREBP-1 (36) and RAV1 and RAV2 (37). The similarity to RAV1 and RAV2 extends to the C-terminus outside the AP2 domain (Figure 2B). At the N-terminal end of the basic region, there is a stretch rich in basic amino acids. In the C-terminal half of the predicted protein, there is a unique structure composed of four tandem repeats of 23 amino acids (Figure 2A). We again searched for sequences homologous to this unique repeat in various





**Figure 1.** Affinity purification of FMBP-1. FMBP-1 was fractionated and purified by two sequential rounds of affinity column chromatography with poly(dA-dT)-Cellulofine. The protein composition and FMBP-1 activity of each fraction was analyzed by SDS-PAGE (upper panels) and an EMSA (lower panels), respectively. Mix.-Dil. shows pooled and diluted fractions of the 1.2 M eluate of the first round fractionation. Fr. No., fraction number; M, protein marker; Ext., PSG extract; FT, flow through.

databases, and found similar repeated structures of the score and three amino acids in humans, mouse, *Drosophila* and *C.elegans* (Figure 2C). We named this repeated structure the STPR (score and three amino acid peptide repeat) domain.

### FMBP-1 is the STPR protein

The predicted molecular mass of the STPR protein is ~25 kDa, but FMBP-1 purified from PSG extracts was found to be 32 kDa on SDS-PAGE. We examined whether the STPR protein is FMBP-1. The ORF region was amplified by PCR and subcloned into an expression vector, and expressed as a fusion protein with glutathione *S*-transferase (GST). The rabbit antibody anti-FMBP-1 was prepared against the deletion protein del. 12 (see Figure 8), and added to the EMSA reaction. The bands shifted by FMBP-1 in crude PSG extract, the purified fraction and the fusion construct, the recombinant GST-STPR protein, were further shifted to the top of the gel (Figure 3), showing that the STPR protein is FMBP-1.

We compared nucleotide sequences recognized by FMBP-1 and the STPR protein, using various single point mutants of the +290 probe as competitors (Figure 4). FMBP-1 and the STPR protein competed for binding to the +290 probe in completely the same manner. On the other hand, Bm Fkh in crude extract, which bound to the same probe, competed differently from FMBP-1 with the same mutant oligonucleotides.

From these experiments, the recognition sequence of FMBP-1 was determined as ATNTWTNTA, whereas Bm Fkh is known to recognize overlapping but different nucleotides, TNT-GTAAATA (the complementary sequence is usually described), in the same DNA element (38).

### Tissue-specificity and developmental changes of the activity and expression of FMBP-1 from the fourth to fifth instar

FMBP-1 activity was detectable by an EMSA in PSG at the fourth and fifth feeding stages but rarely in MSG and the fourth molt in PSG (21). Using extracts from PSG at the fourth feeding, fourth molting and fifth feeding stages, and from MSG at the fourth molting and fifth feeding stages, western blotting was performed with anti-FMBP-1 antibody (Figure 5). Though the antibody reacted with some protein bands, the 32 kDa band was clearly detected in the extracts from PSG at the fourth and fifth feeding stages, and a very faint signal at the same position was observed in the extracts from MSG and PSG at the fourth molting stage.

To trace the activity of FMBP-1 at the early stages of the fifth instar precisely, PSG extracts were prepared at 2, 6, 12, 18, 24, 36 and 48 h after the fourth ecdysis, and used for the EMSA (Figure 6). Though the FMBP-1 activity was scarcely detectable at the fourth molting stage, it suddenly became

detectable at only 2 h after the fourth ecdysis, reached a peak at ~6 h, and then decreased gradually.

Northern blotting was performed to clarify whether the expression of the FMBP-1 gene corresponds with the DNA-binding activity. Total RNA was extracted from the staged PSG, and poly(A)<sup>+</sup> RNA was extracted from PSG and MSG at day 1 of the fifth instar, and used for northern blotting. As shown in Figure 7, long bands of 7, 5, 2.4, 1.6 and 1.35 kb were detected in poly(A)<sup>+</sup> RNA from both PSG and MSG. FMBP-1 RNA was also detectable in PSG during the fourth molting stage. While it was difficult to detect RNAs of 2.4 kb and longer in the total RNA, the amount of 1.6 and 1.35 kb FMBP-1 RNA consistently increased from 48 h after the third ecdysis to 6 h after the fourth ecdysis and then decreased from 12 h after the fourth ecdysis.

### STPR is a novel DNA-binding domain

To map the DNA-binding domain of FMBP-1, various deletion proteins were prepared, as shown in Figure 8A and B, and used for the EMSA. Clearly, the deletion proteins still completely covering the STPR domain bound to the +290 probe, while the proteins lacking even part of the STPR domain lost most of their DNA-binding activity (Figure 8C). The protein that covered only the STPR domain (del. 13) still had ~50% of the DNA-binding activity of the complete FMBP-1. Accordingly, the STPR domain conserved from *C.elegans* to humans is a novel DNA-binding domain.

### DISCUSSION

Competition of the binding of FMBP-1 with poly(dA-dT) was consistent with the recognition sequence of FMBP-1 as ATNTWTNTA. FMBP-1 binding sites of the fibroin gene and poly(dA-dT) contain the FMBP-1-recognizing sequence (21). Thus, poly(dA-dT) was used for DNA affinity chromatography of FMBP-1, and a 32 kDa protein was purified. Though two bands of ~32 kDa were detected in a preliminary purification, only one major band was recovered in the large-scale purification. Western blotting detected the band at ~32 kDa, and so FMBP-1 was probably modified or decomposed by proteolysis during the first purification.

The predicted molecular mass encoded by the cDNAs for FMBP-1 was ~25 kDa, but recombinant protein produced from the cDNA showed the same DNA-binding specificity as that of native FMBP-1 in crude PSG extracts. An antibody against the recombinant protein reacted with protein bands, including 32 kDa band in western blotting. These proteins might be produced by differential splicing of mRNA and/or modification of the protein; however, the 32 kDa protein alone was purified with DNA affinity chromatography.

FMBP-1 possesses an acidic region similar to a region in the plant proteins RAV1 and RAV2, and a basic region including a unique repeated structure. Acidic regions often function as activation domains in transcription factors (39–41). On the N-terminal side of the basic region, at about the center of FMBP-1, there is a span that is especially rich in basic

A

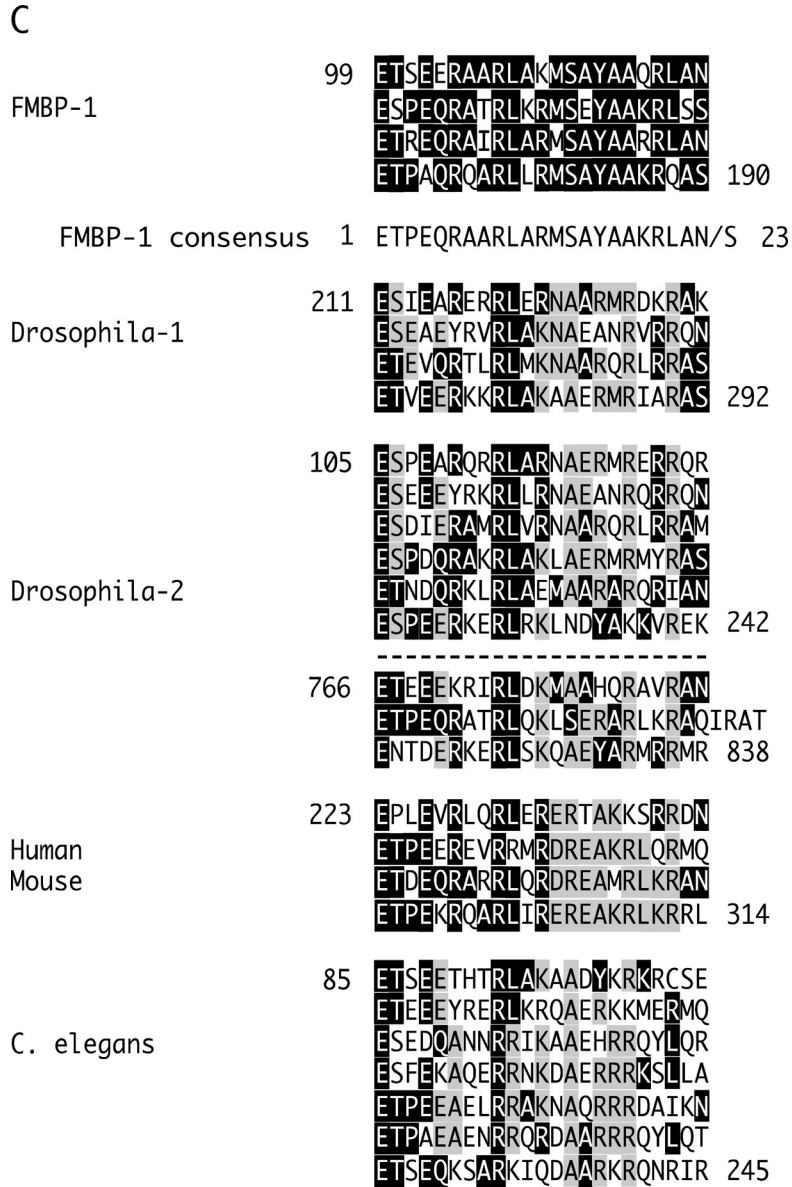
```

1  MGDLSQADSS WKKMVIAKTN NEWKSSVGED ANAASNTTHF RLSDECIQYD 50
      P1
51  NIKKEIEEID EQETLNTVEP VMIQEMDPL SLLEPKARRR RKGSGPKSET 100
      P2
101 SEERAARLAK MSAYAAQRLA NESPEQRATR LKRMSEYAAK RLSSETREQR 150
      P3
151 AIRLARMSAY AARRLANETP AQRQARLLRM SAYAAKRQAS KKSLSLVNDS 200
      P4
201 LNYSIMPNQS RATNHPFT 218

```

### B

FMBP-1	1	MGDL	SQADS	SWKKMV	--IA	KTN-NEWKSS	VGEDANAASN	---TTH-FRL	SDE	45
RAV1	63	KGVV	POPNG	RWGAQI	--YE	KHQ-RVWLGT	FNEEDEAARA	YDVAVHRFR	RDA	111
RAV2	66	KGVV	POPNG	RWGAQI	--YE	KHQ-RVWLGT	FNEQEEAARS	YDIAACRFRG	RDA	114
AP2-1	133	RGVT	FYRRTG	RWESHI	--WD	CGK-QVYLG	FDTAHAARS	YDRAAIKFRG	VEA	182
AP2-2	225	RGVT	LHKCG	RWEARMGQFL		GKK-YVYLG	FDTEVEAARA	YDKAAIKCNG	KDA	275
EREBP1	100	RGVR	RRPWG	KFAAEIRDPA		KNGARVWLGT	YETDEEAIA	YDKAAYRMG	SKA	151
FMBP-1	46	CIQYDNI	KKEIEEID	Q	ETLNTVEPVD	MIQEMDPLSL	LEPKARRRRK	GSGPKS		98
RAV1	112	VTNFKDV	KMDEDEVD	L	NSHSKSEIVD	MLRKHTYNEE	LE-QSKRRPN	GNGNMT		163
RAV2	115	VVNFKNV	LEDGD-LA	L	EAHSKAEIVD	MLRKHTYADE	LE-QNNKRQL	-FLSVD		164
AP2-1	173	DINFNID	DY							181
AP2-2	276	VTNFDPS	IY							284
EREBP1	152	HLNFPHR	IG							160

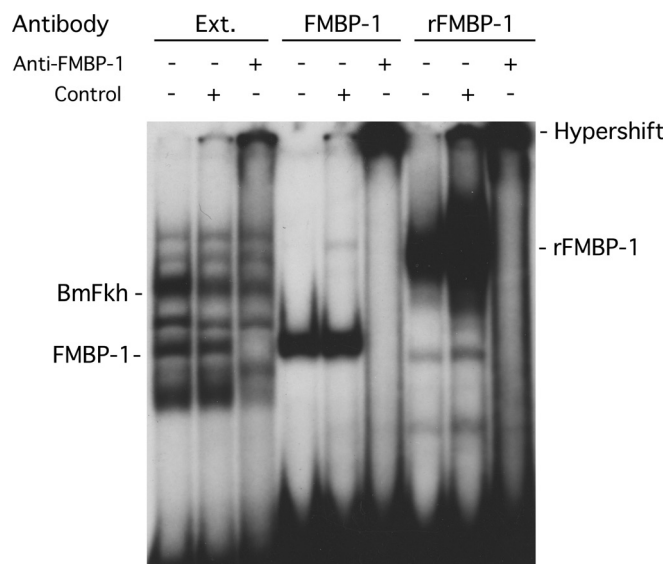


**Figure 2.** Structure of FMBP-1 protein. (A) Amino acid sequence deduced from the FMBP-1 cDNA. Repeated sequences of the 23 amino acid residues are indicated with thick bars. Peptide sequences determined from purified FMBP-1 protein are underlined (P1–P4), the acidic region is indicated with a dotted line, and the hyper-basic stretch is indicated by dots above the letters. (B) Comparison of the N-terminal half of FMBP-1 with plant proteins. Identical amino acids are indicated with white letters on black, and similar amino acids are shaded. RAV1 and RAV2, Proteins-1 and -2 related to ABI3/VP1 (*Arabidopsis*); AP2-1 and AP2-2, APETALA2 domains R1 and R2 (*Arabidopsis*); EREBP1, Ethylene-responsive element binding protein-1 (Tobacco). (C) Comparison of STPR domains of various organisms. *Drosophila* has two genes for STPR domain proteins. One (CG14440) of them has a tetrad of the STP, and the other (CG14442) has two separated STPR domains composed of six repeats and three repeats. The gene of *C.elegans* (Q11191) encodes seven repeats of the 23 amino acids. The amino acid sequences of the STPR domains of mouse (BC023180-1) and humans (BC012116) are identical. Amino acids identical to the consensus sequence of FMBP-1 are shown with white letters on black, while amino acids repeated  $\geq 50\%$  in the STPR domain of each protein are shaded.

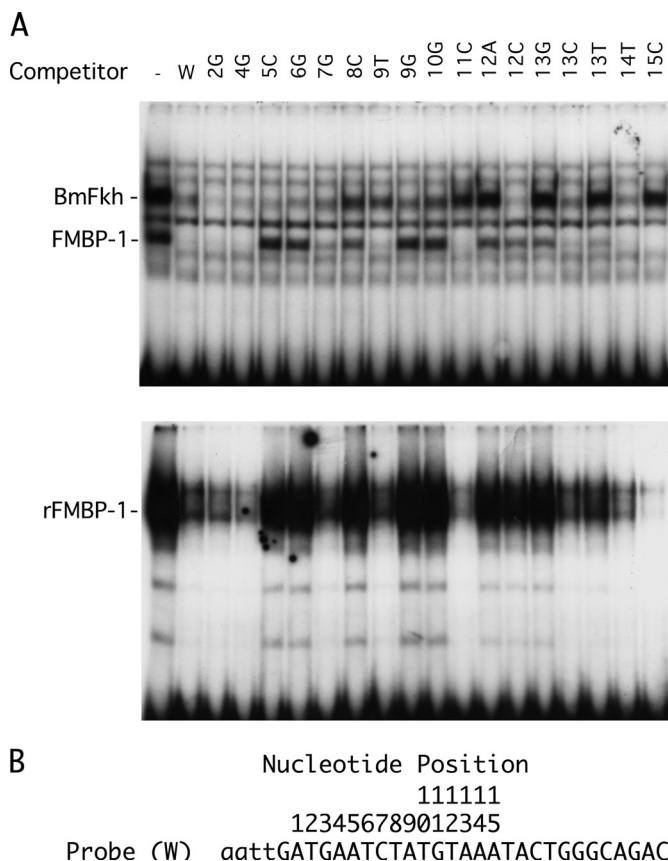
amino acids. Such hyper-basic regions can function as nuclear localization signals (42). The N-terminal region of FMBP-1 did not show any significant homology to other animal proteins, but we found that the region has a weak similarity to the AP2 domain observed in several transcription factors of higher plants (35–37). The AP2 domain is thought to be a part of the DNA-binding domain of the transcription factors, APETALA2, EREBPs, and RAV1 and RAV2. APETALA2 is a factor involved in the regulation of flower development in *Arabidopsis* (35), and EREBPs are known to be involved

in the regulation of ethylene response genes in tobacco (36). The function of RAVs has not yet been elucidated, but these proteins are partly homologous to factors ABI3 and VP1, which control the expression of abscisic acid-regulated genes (43,44). Though the significance of this weak similarity between FMBP-1 and the plant transcription factors is unclear at present, the possibility exists that some of the chemical components in plants ingested by insects regulate the expression of insect genes. FMBP-1 might be involved in such a process.

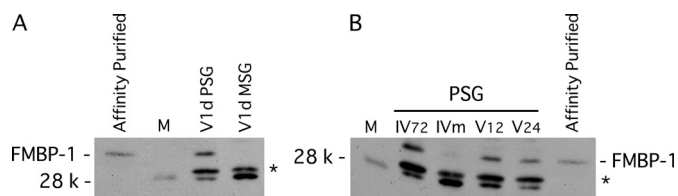




**Figure 3.** Effects of anti-FMBP-1 specific antiserum on an EMSA. Crude PSG extract (Ext.), purified FMBP-1 (FMBP-1) or recombinant GST-fused FMBP-1 (rFMBP-1) was incubated with or without (–) anti-FMBP-1 antiserum (Anti-FMBP-1) or control serum (Control).



**Figure 4.** Nucleotides recognized by FMBP-1. (A) PSG extract (upper panel) or recombinant GST-fused FMBP-1 (lower panel) was incubated with a +290 probe in the presence or absence (–) of various cold competitors with a single nucleotide mutation in the probe sequence, as shown at the top of the panel. (B) Nucleotide sequence of the top strand of the +290 probe and nucleotide position referring to the point mutants.

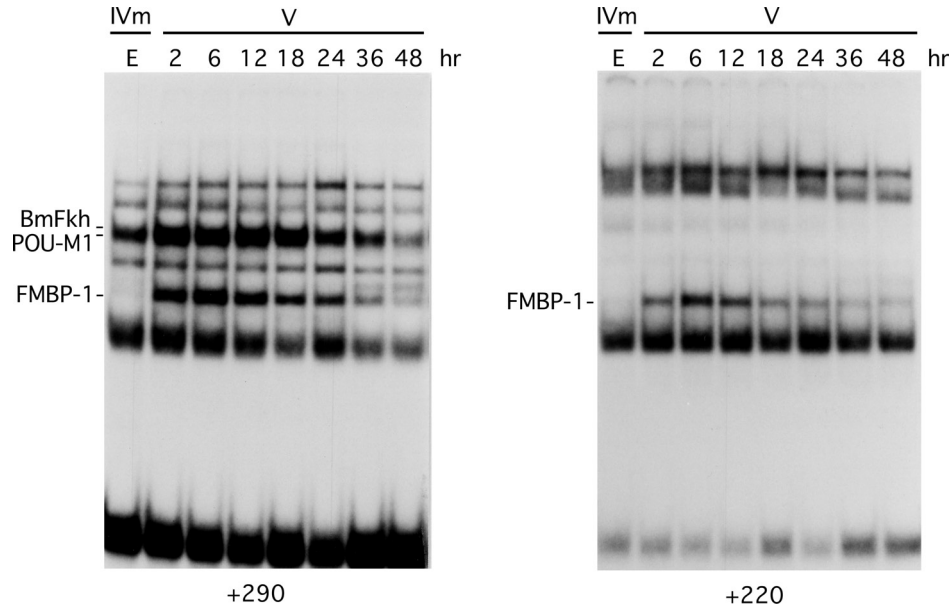


**Figure 5.** Western blot of FMBP-1 protein. (A) PSG or MSG extract (50 µg of protein) from day 1 of the fifth instar was separated by SDS-PAGE and blotted onto a PVDF membrane. The blot was treated with rabbit anti-FMBP-1 antiserum, and detected with the ECL detection system. M, 28 kDa marker protein; asterisk, a commonly detected band. (B) PSG extracts (50 µg) prepared at 72 h after the third ecdysis, at the fourth molt, and at 12, 24 h and 2 days after the fourth ecdysis, were separated by SDS-PAGE, blotted and detected with anti-FMBP-1 antiserum. The 28 kDa band increased and decreased inversely with 32 kDa FMBP-1.

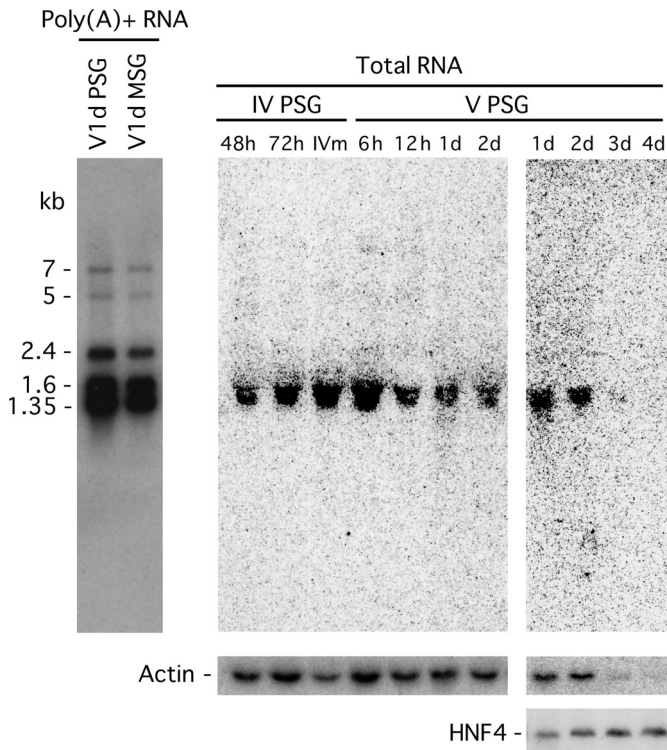
Genes encoding homologous repeats, such as the STPR of FMBP-1, have been found in humans, mouse, *Drosophila* and *C.elegans*, but the functions of these genes have not been reported. In *Drosophila*, there are two genes encoding STPR. One of the *Drosophila*, mouse and human genes contain a tetrad of the STP, as does FMBP-1. However, in the other *Drosophila* gene, the STP is repeated six times at the N-terminus and three times at the C-terminus. Furthermore, the unit is repeated seven times in the gene of *C.elegans*. STP is likely conserved as a basal motif constructing a larger protein domain. The homology of the 23 amino acids is greater between each unit of the same protein than between STPR proteins of different species. Functional and evolutionary pressures may select to maintain each unit of STPRs of the same protein.

We prepared a series of deletion constructs of FMBP-1 as GST-fusion proteins and examined the DNA-binding activity of these mutants. The STPR domain alone showed sufficient DNA-binding activity, whereas the proteins lacking the STPR domain lost almost all binding activity. Accordingly, the unique structure of FMBP-1, STPR, is a novel DNA-binding domain conserved in animals. Though the N-terminal region has similarity to the DNA-binding domain of transcription factors in higher plants, DNA-binding activity by the specific probe for FMBP-1 was never observed. However, the N-terminal region (amino acid 51–95) stimulated the DNA-binding activity of STPR. This effect was remarkable when the C-terminal end of FMBP-1 beyond the STPR domain was present (del. 3 versus the entire protein and del. 13 versus del. 7). The C-terminal region seems to prevent the binding activity of the STPR domain, and the N-terminal region seems to release the effect of the C-terminal region. The physical and functional interaction between these two regions surrounding the STPR should be elucidated.

In the transcription regulatory elements of the fibroin gene, there are multi-elements that bind BmFkh, POU-M1 and SGF2 (16,21). The binding sites of FMBP-1 overlap with the sites of those transcription factors (16,21). The coordination of FMBP-1 with these factors possibly raises its specificity as well as that of the fibroin gene. Besides the fibroin gene, sequences corresponding to ATNTWTNTA are observed in the promoter-proximal regions of other PSG specific genes, such as in the –33 and –55 regions of the fibroin light



**Figure 6.** Developmental profile of FMBP-1 activity at the fourth molt and just after the fourth ecdysis. PSG extracts (12 µg) prepared from each stage were incubated with the +290 or +220 probes, and the DNA-binding activity of FMBP-1 was analyzed by an EMSA. Bm fork head and POU-M1 were detectable with the +290 probe, but not with the +220 probe. Time (h) after the fourth ecdysis is shown at the top of the panel. E, stage E of the fourth molt.



**Figure 7.** Northern blot of FMBP-1 RNA. Poly(A)<sup>+</sup> RNA (2 µg) prepared from PSG or MSG 1 day after the fourth ecdysis was separated on a 1% agarose gel, transferred to a nylon membrane and detected by standard methods. The sizes of detected RNAs are shown on the left of the panel. Total RNA (20 µg) prepared from PSGs at 48 and 72 h after the third ecdysis, at the fourth molt (IVm), and was analyzed at 6 and 12 h and 1, 2, 3 and 4 days after the fourth ecdysis. Actin, RNA for actin A3 was detected with the same filter using an oligonucleotide probe. HNF4, RNA for a nuclear receptor HNF4 homolog was detected with a cDNA probe using the same samples but not the same filter.

chain gene and -70 of the fibrohexamerin gene. Whereas the corresponding sequence was not observed in such promoter-proximal regions in the MSG-specific sericin-1 and -2 genes.

The DNA-binding activity of FMBP-1 was clearly detectable at 2 h, peaked at 6–12 h after the fourth ecdysis, and then decreased gradually. Northern blot analyses showed that the expression of the FMBP-1 gene continued during the fourth instar and increased up to 6 h after the fourth ecdysis, and then decreased gradually. The decrease in FMBP-1 mRNA is consistent with the decrease in FMBP-1 activity from 12 h after the fourth ecdysis. This time point also coincides with when the expression of the fibroin gene reached its highest level, although the level is maintained during the fifth instar (5,11,14). On the other hand, the continuous expression of the FMBP-1 gene during the fourth ecdysis was inconsistent with the absence of FMBP-1 activity. mRNA for FMBP-1 was also present in MSG in similar amounts and with the same sizes as the mRNA from PSG. We examined the FMBP-1 activity of four different batches of MSG extracts in an EMSA. In these extracts, the FMBP-1 activity (judged by the mobility on the EMSA) was detectable, but the activity per unit protein was consistently very low compared with the activity in PSG extracts, although the activity in PSG extracts varied from batch to batch, probably reflecting the delicate differences of larval developmental stages (data not shown). Taken together, the DNA-binding activity of FMBP-1 seems to be regulated through post-transcriptional and/or post-translational modification as well as transcription. In this regard, it is of interest that the change in the amount of the 28 kDa band complemented that in the amount of 32 kDa FMBP-1 band, and that the antibody against FMBP-1 detected some other bands commonly found in PSG and MSG.





2. Suzuki, Y., Takiya, S., Suzuki, T., Hui, C.-C., Matsuno, K., Fukuta, M., Nagata, T. and Ueno, K. (1990) Developmental regulation of silk gene expression in *Bombyx mori*. In Hagedorn, H.H., Hildebrand, J.G., Kindwell, M.G. and Law, J.H. (eds), *Molecular Insect Science*. Plenum Press, NY, pp. 83–89.
3. Suzuki, Y. and Brown, D. (1972) Isolation and identification of the messenger RNA for silk fibroin from *Bombyx mori*. *J. Mol. Biol.*, **63**, 409–429.
4. Suzuki, Y. and Suzuki, E. (1974) Quantitative measurements of fibroin messenger RNA synthesis in the posterior silk gland of normal and mutant *Bombyx mori*. *J. Mol. Biol.*, **88**, 393–407.
5. Maekawa, H. and Suzuki, Y. (1980) Repeated turn-off and turn-on of fibroin gene transcription during silk gland development of *Bombyx mori*. *Dev. Biol.*, **78**, 394–406.
6. Ishikawa, E. and Suzuki, Y. (1985) Tissue- and stage-specific expression of sericin genes in the middle silk gland of *Bombyx mori*. *Dev. Growth Differ.*, **27**, 73–82.
7. Couble, P., Moine, A., Garel, A. and Prudhomme, J.-C. (1983) Developmental variations of a nonfibroin mRNA of *Bombyx mori* silk gland, encoding for a low-molecular-weight silk protein. *Dev. Biol.*, **97**, 398–407.
8. Kimura, K., Oyama, F., Ueda, H., Mizuno, S. and Shimura, K. (1985) Molecular cloning of the fibroin light chain complementary DNA and its use in the study of the expression of the light chain gene in the posterior silk gland of *Bombyx mori*. *Experientia*, **41**, 1167–1171.
9. Okamoto, H., Ishikawa, E. and Suzuki, Y. (1982) Structural analysis of sericin genes. Homologies with fibroin gene in the 5' flanking nucleotide sequences. *J. Biol. Chem.*, **257**, 15192–15199.
10. Couble, P., Michaille, J.-J., Garel, A., Couble, M.-L. and Prudhomme, J.-J. (1987) Developmental switches of sericin mRNA splicing in individual cells of *Bombyx mori* silkgland. *Dev. Biol.*, **124**, 431–440.
11. Obara, T. and Suzuki, Y. (1988) Temporal and spatial control of silk gene transcription analyzed by nuclear run-on assays. *Dev. Biol.*, **127**, 384–391.
12. Michaille, J.-J., Garel, A. and Couble, P. (1990) Cloning and characterization of the highly polymorphic *Ser2* gene of *Bombyx mori*. *Gene*, **86**, 177–184.
13. Matsunami, K., Kokubo, H., Ohno, K. and Suzuki, Y. (1998) Expression pattern analysis of SGF-3/POU-M1 in relation to sericin-1 gene expression in the silk gland. *Dev. Growth Differ.*, **40**, 591–597.
14. Suzuki, Y. and Giza, P.E. (1976) Accentuated expression of silk fibroin gene *in vivo* and *in vitro*. *J. Mol. Biol.*, **107**, 183–206.
15. Matsuno, K., Hui, C.-C., Takiya, S., Suzuki, T., Ueno, K. and Suzuki, Y. (1989) Transcriptional signals and protein binding sites for sericin gene transcription *in vitro*. *J. Biol. Chem.*, **264**, 18707–18713.
16. Hui, C.-C., Matsuno, K. and Suzuki, Y. (1990) Fibroin gene promoter contains a cluster of homeodomain binding sites that interact with three silk gland factors. *J. Mol. Biol.*, **213**, 651–670.
17. Suzuki, T., Matsuno, K., Takiya, S., Ohno, K., Ueno, K. and Suzuki, Y. (1991) Purification and characterization of an enhancer-binding protein of the fibroin gene: I. Complete purification of fibroin factor 1. *J. Biol. Chem.*, **266**, 16935–16941.
18. Suzuki, T., Takiya, S., Matsuno, K., Ohno, K., Ueno, K. and Suzuki, Y. (1991) Purification and characterization of an enhancer-binding protein of the fibroin gene: II. Functional analyses of fibroin factor 1. *J. Biol. Chem.*, **266**, 16942–16947.
19. Durand, B., Drevet, J. and Couble, P. (1992) P25 gene regulation in *Bombyx mori* silk gland: two promoter-binding factors have distinct tissue and developmental specificities. *Mol. Cell. Biol.*, **12**, 5768–5777.
20. Horard, B., Julien, E., Nony, P., Garel, A. and Couble, P. (1997) Differential binding of the *Bombyx* silk gland-specific factors SGFB to its target DNA sequence drives posterior-cell-restricted expression. *Mol. Cell. Biol.*, **17**, 1572–1579.
21. Takiya, S., Kokubo, H. and Suzuki, Y. (1997) Transcriptional regulatory elements in the upstream and intron of the fibroin gene bind three specific factors, POU-M1, Bm Fkh and FMBP-1. *Biochem. J.*, **321**, 645–653.
22. Fukuta, M., Matsuno, K., Hui, C.-C., Nagata, T., Takiya, S., Xu, P.-X., Ueno, K. and Suzuki, Y. (1993) Molecular cloning of a POU domain-containing factor involved in the regulation of the *Bombyx* sericin-1 gene. *J. Biol. Chem.*, **268**, 19471–19475.
23. Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T. and Suzuki, Y. (1995) Silk gland factor-1 involved in the regulation of *Bombyx* sericin-1 gene contains fork head motif. *J. Biol. Chem.*, **270**, 9340–9346.
24. Julien, E., Bordeaux, M.-C., Garel, A. and Couble, P. (2002) Fork head alternative binding drives stage-specific gene expression in the silk gland of *Bombyx mori*. *Insect Biochem. Mol. Biol.*, **32**, 377–387.
25. Kokubo, H., Takiya, S., Mach, V. and Suzuki, Y. (1996) Spatial and temporal expression pattern of *Bombyx fork head/SGF-1* gene in embryogenesis. *Dev. Genes Evol.*, **206**, 80–85.
26. Matsuno, K., Takiya, S., Hui, C.-C., Suzuki, T., Fukuta, M., Ueno, K. and Suzuki, Y. (1990) Transcriptional stimulation via SC site of *Bombyx* sericin-1 gene through an interaction with a DNA binding protein SGF-3. *Nucleic Acids Res.*, **18**, 1853–1858.
27. Kiguchi, K. and Agui, N. (1981) Ecdysteroid levels and developmental events during larval moulting in the silkworm, *Bombyx mori*. *J. Insect Physiol.*, **27**, 805–812.
28. Suzuki, Y., Obara, T., Takiya, S., Hui, C.-C., Matsuno, K., Suzuki, T., Suzukui, E., Ohkubo, M. and Tamura, T. (1990) Differential transcription of the fibroin and sericin-1 genes in cell-free extracts. *Dev. Growth Differ.*, **32**, 179–187.
29. Tsuda, M. and Suzuki, Y. (1981) Faithful transcription initiation of fibroin gene in a homologous cell-free system reveals an enhancing effect of 5' flanking sequence far upstream. *Cell*, **27**, 175–182.
30. Takiya, S., Hui, C.-C. and Suzuki, Y. (1990) A contribution of the core-promoter and its surrounding regions to the preferential transcription of the fibroin gene in posterior silk gland extracts. *EMBO J.*, **9**, 489–496.
31. Takiya, S., Nishita, Y., Ishikawa, S., Ohno, K., Tamura, T. and Suzuki, Y. (2004) *Bombyx* Y-box protein BYB facilitates specific DNA interaction of various DNA binding proteins independently of the cold shock domain. *J. Biochem.*, **135**, 683–693.
32. Oritiz, M.L., Calero, M., Parton, C.F., Castellanos, L. and Mendez, E. (1992) Imidazole-SDS-Zn reverse staining of proteins in gels containing or not SDS and microsequence of individual unmodified electroblotted proteins. *FEBS Lett.*, **296**, 300–304.
33. Inomata, Y., Kawaguchi, H., Hiramoto, M., Wada, T. and Handa, H. (1992) Direct purification of multiple ATF/E4TF3 polypeptides from HeLa cell crude nuclear extracts using DNA affinity latex particles. *Anal. Biochem.*, **206**, 109–114.
34. Mounier, N. and Prudhomme, J.-C. (1986) Isolation of actin genes in *Bombyx mori*: the coding sequence of cytoplasmic actin gene expressed in the silk gland is interrupted by a single intron in an unusual position. *Biochimie*, **68**, 1058–1061.
35. Jofuku, K.D., den Boer, B.G.W., Montagu, M.V. and Okamoto, J.K. (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. *Plant Cell*, **6**, 1211–1225.
36. Ohme-Takagi, M. and Shinsi, H. (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell*, **7**, 173–182.
37. Kagaya, Y., Ohmiya, K. and Hattori, T. (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.*, **27**, 470–478.
38. Takiya, S., Gazi, M. and Mach, V. (2003) The DNA binding of insect Fork head factor is strongly influenced by the negative cooperation of neighbouring bases. *Insect Biochem. Mol. Biol.*, **33**, 1145–1154.
39. Manfredi, J.-J. and Prives, C. (1994) The transforming activity of simian virus 40 large tumor antigen. *Biochem. Biophys. Acta*, **1198**, 65–83.
40. Hope, I.A. and Struhl, K. (1986) Functional dissection of a eukaryotic transcription factor protein, GCN4 of yeast. *Cell*, **46**, 885–894.
41. Sigler, P.B. (1988) Transcriptional activation. Acid blobs and negative noodles. *Nature*, **333**, 210–212.
42. Gill, G. and Ptashne, M. (1987) Mutants of GAL4 protein altered in an activation function. *Cell*, **51**, 121–126.
43. McCarty, D.R., Hattori, T., Carson, C.B., Vasil, V., Lazar, M. and Vasil, I.K. (1991) Viviparous-1 developmental gene of maize encodes a novel transcription activator. *Cell*, **6**, 895–905.
44. Hill, A., Nantel, A., Rock, C.D. and Quatrano, R.S. (1996) A conserved domain of the viviparous-1 gene product enhances the DNA binding activity of the bZIP protein EmBP-1 and other transcription factors. *J. Biol. Chem.*, **271**, 3366–3374.