Studies on Chitin Recognition Mechanism and Biosynthesis of Cuticle Proteins in Silkworm, *Bombyx mori*

Toru Togawa

Studies on Chitin Recognition Mechanism and Biosynthesis of Cuticle Proteins in Silkworm, *Bombyx mori*

by

Toru Togawa

A dissertation submitted to the Faculty of Graduate School,
Tokyo Metropolitan University
in partial fulfillment of the requirements
for the degree of Doctor of Science

March 2001

東京都立大学　博士（理学）学位論文（課程博士）

論　文　名　カイコ表皮タンパク質のキチン認識機構および生合成調節機構の解析（英文）

著　者　外川　徹

審査担当者

主　査　泉　進

委　員　駒野徹訓

委　員　八杉貞雄

委　員

上記の論文を合格と判定する

平成 13年 3月 6日

東京都立大学大学院理学研究科教授会

研究科長　尋澤清明

DISSERTATION FOR A DEGREE OF DOCTOR OF SCIENCE TOKYO METROPOLITAN UNIVERSITY

TITLE: Studies on Chitin Recognition Mechanism and Biosynthesis of Cuticle Proteins in Silkworm, *Bombyx mori*

AUTHOR: Toru Togawa

EXAMINED BY

Examiner in chief S. Izumi

Examiner T. Kamano

Examiner Sadao Yasugi

QUALIFIED BY THE GRADUATE SCHOOL OF SCIENCE
TOKYO METROPOLITAN UNIVERSITY

Dean Kiyoaki Kuwasawa

Date March 6, 2001

CONTENTS

ABBREVIATIONS

bp : base pair(s)

BSA : bovine serum albumin

DEAE : diethylaminoethyl

DEPC : diethylpyrocarbonate

DTT : dithiothreitol

EDTA : ethylenediaminetetraacetic acid

20HE : 20-hydroxyecdysone

IPTG : isopropyl thiogalactoside

JH : juvenile hormone

kb : kilobases

NaPB : sodium phosphate buffer

nt : nucleotide

PAGE : polyacrylamide gel electrophoresis

PBS : phosphate buffered saline

PCR : polymerase chain reaction

SDS : sodium dodecyl sulfate

Tris : tris(hydroxymethyl)aminomethane

ABSTRACT

Insect cuticle is composed mainly of chitin, a polymer of *N*-acetylglucosamine (GlcNAc), and chitin-binding cuticle proteins. The physical and mechanical properties of cuticle are highly diverse depending on the species, developmental stages, sex, body regions, and so on. These various properties are caused by cuticle protein composition. Therefore, to understand the mechanism by which cuticle maintains stable structure and obtains various mechanical properties, it is important to analyze chitin-recognition mechanism of cuticle proteins. On the other hand, the majority of cuticle protein genes is expressed in stage dependent manner. Moreover, it has been reported that genes for two of major cuticle proteins from silkworm, *Bombyx mori*, are transactivated by juvenile hormone (JH), whose mode of action has been unknown yet. These indicate that cuticle protein genes provide an excellent model system to study control of gene expression in post-embryonic development including JH signaling. I approached these two subjects, the chitin-recognition mechanism and the regulation mechanism of gene expression using cuticle proteins of *B. mori*.

(1) Chitin recognition mechanism of cuticle proteins from *Bombyx mori*

In larval cuticle of *B. mori*, four cuticle proteins, BMCP30, 22, 18, and 17 have been identified. To analyze chitin-binding activity of cuticle proteins, I attempted purification of these proteins in native form. Soluble proteins were extracted with guanidine-HCl solution from fifth instar larval cuticle, followed by dialysis against a guanidine-free buffer. Cuticle protein fraction was separated by ammonium sulfate precipitation (20% saturated ammonium sulfate), and each cuticle protein was purified by two cycles of ion-exchange chromatography. All purified cuticle proteins kept high affinity to chitin. The chitin-binding activity of purified BMCP30 was analyzed by chitin-affinity chromatography. It was revealed that pH optimum for binding to chitin of BMCP30 is pH 6.4, which corresponds to hemolymphal pH. Competition experiments using chitooligosaccharides showed that binding to chitin was not prevented by trimer of GlcNAc, but prevented by hexamer, indicating that BMCP30

recognizes 4-6 mer of GlcNAc as a unit for binding.

BMCPs possess about 70 amino acid region conserved widely among cuticle proteins from different insect species. To determine the chitin binding domain of BMCP30, I concerned my attention to the conserved region. The chitin-binding activities of BMCP30 and BMCP18 which have the conserved region in common were compared. They showed similar high chitin-binding activities, implying that the conserved region functions as a chitin binding domain. On the other hand, chitin-binding activity of BMCP30 showed resistance stronger than to urea solution than that of BMCP18. This result suggests that an additional chitin affinity region is present in another region of BMCP30.

To verify that the conserved region functions actually as chitin binding domain, a series of BMCP30 deletion mutants fused with glutathione S-transferase (GST) was synthesized in *E. coli*, and their chitin-binding activities were analyzed. The recombinant proteins bearing full length or only the conserved region of BMCP30 bound to chitin well. In contrast, the recombinant protein with only N-terminal region of BMCP30 and lacking the conserved region showed reduced chitin-binding activity. It indicated that the conserved region functions as the primary chitin binding domain. Some chitin-binding activity of the recombinant protein with only N-terminal of BMCP30 suggested an additional chitin-affinity region consistently with result from comparison of BMCP30 and BMCP18. The aromatic amino acid residues were highly conserved in the conserved region of cuticle proteins. The knowledge from chitin binding protein of other organisms suggested these aromatic residues in cuticle protein serve important role for binding to chitin.

(2) Expression control of cuticle protein gene in *Bombyx mori*

Although genomic clone for BMCP18 has been isolated previously, its structural analysis has been unsatisfactory. To study the control of cuticle protein gene expression, I analyzed the structure of BMCP18 genomic clone in detail. It was revealed that BMCP18 gene consists of three exons interspersed by two introns. The transcription initiation site was determined by primer extension analysis. A structural comparison of the BMCP18 gene and

related cuticle protein genes of other lepidopteran species (MSCP14.6 and HCCP12) showed that the 5' upstream region of the BMCP18, MSCP14.6, and HCCP12 genes has a 12-bp identical sequence matching the recognition sequence for transcription factors COUP-TF and HNF-4. This implies that molecular mechanisms regulating expression of these cuticle protein genes are also conserved. mRNAs coding for Bmsvp, the *B. mori* homolog of *Drosophila* Seven-up, which is known as a homolog of vertebrate COUP-TF, and BmHNF-4, a homolog of vertebrate HNF-4, were detected in the larval epidermis. Bmsvp bound to the 12-bp sequence *in vitro*, suggesting that Bmsvp regulates the BMCP18 gene expression.

Since JH is a small lipophilic hormone, it probably acts with a transcription factor which is a member of nuclear receptor superfamily as a ligand. Bmsvp is a member of nuclear receptors. The possibility that Bmsvp provides BMCP18 gene with JH responsiveness was discussed.

GENERAL INTRODUCTION

Cuticle is composite multilayered extracellular matrix, secreted by underlying monolayered epidermal cells in Arthropoda. It serves as skin and skeleton, so called exoskeleton, and determines the shape of bodies, and also functions as a barrier to environment surrounding the individuals (Locke, 1974). Although many properties of cuticle are convenient for animals, its extensibility is restricted. Therefore, arthropod animals must shed off the old cuticle and synthesize new one at each molt to grow up. During the molting process, most part of inner cuticle is taken up and re-used. Foregut, hindgut, dermal gland ducts and tracheae are lined with cuticle, known as the intima, which is shed at each molt in the same way as body surface. The arthropod cuticle is composed mainly of chitin and various proteins. Chitin is a linear chain of β–(1,4)-linked *N*-acetylglucosamine, organized into microfibrils, and embedded in protein matrix. A model of the three-dimensional structure of a chitin-protein complex in cuticle based on the X-ray fiber diagrams has been reported (Blackwell and Weih, 1980). In this model, it was shown that chitin microfibril is surrounded by six globular proteins per one round in a spiral. Crustacean cuticle also contains many calcium salts, involved in sclerotization, namely the calcification. In contrast, cuticles of almost all insects hardly contain calcium, but they contain a large quantity of phenol derivatives, which are involved in melanization and sclerotization (Hackman, 1984). Cuticle also contains lipids and pigments.

Cuticle is divided basically into two parts, outer thin epicuticle and inner thick procuticle (Hackman, 1984; Neville, 1984). The epicuticle is usually 1-3 μm thick, and contains no chitin. It is composed of lipoproteins, phenolic materials, and wax, and served several functions such as waterproofing and protection from fungi. The procuticle is a lamellate structure with ten to several hundreds μm thick, and composed of essentially chitin and proteins (Neville, 1984). The mechanical property of cuticle is mainly determined by procuticle. Procuticle is divided into further two subparts. The outside is the exocuticle, which is more hardened than the other part and secreted before each molt (pre-ecdysial procuticle). The inner layer, secreted after the molt, is the endocuticle (post-ecdysial procuticle).

The physical and mechanical properties of cuticle are diverse enormously depending on the species. Even in an individual, epidermal cells at different developmental stages synthesize cuticle with different properties suited for their function. In holometabolous insects such as Lepidoptera and Diptera, soft and flexible cuticle of larva, which is convenient for locomotion, changes into stiff and hard pupal cuticle to resist environment during metamorphosis. For diverse properties of cuticle, chitin can be varied somewhat in length and degree of acetylation. It is appeared that, however, the mechanical properties of cuticle depend mostly on the protein composition and cross-linkage of proteins. There are many reports showing that the alterations of cuticle mechanical properties during metamorphosis include the changes of cuticular protein compositions (Cox and Willis, 1985; Binger and Willis, 1990; Nakato *et al.*, 1990; Bærnholdt and Andersen, 1998; Mathelin *et al.*, 1998). At puparium formation of *Drosophila*, novel proteins are secreted into third instar larval cuticle, and the property of cuticle changes to hard without molting (Chihara *et al.*, 1982). Cuticle properties are also varied with sex and body regions. Sexually mature female locusts, when they oviposit into the soil, extend their abdomen more than twice in length. It is achieved by much more visco-elastic property of intersegmental cuticle of abdomen than other regions. Such extension of abdomen does not occur in male. It has been reported that abdominal cuticle of sexually mature female of locust, *Locusta migratoria*, has unique protein composition (Andersen *et al.*, 1986; Andersen and Højrup, 1987). On the other hand, cuticle sclerotization occurs as results of cross-linkage between amino groups and between histidyl residues (Christensen *et al.*, 1991) of protein through the intermediation of aromatic components such as catecholamine, DOPA, and quinones (quinone tanning) derived from hemolymphal tyrosine and phenylalanine or aromatic amino acid residues of proteins itself (Hackman, 1984). Hence, cuticle sclerotization seems to depend not only on sclerotizing agents and enzymes, but also on protein compositions. Therefore, it is very interesting things that how each set of cuticular proteins provides various properties of cuticle, and how epidermal cells control gene expression of cuticular proteins.

Cuticle Protein

Among proteins constructing cuticle, structural proteins binding to chitin and/or other proteins are especially called cuticle proteins. Cuticle proteins are synthesized in epidermal cells and secreted into the cuticle. After secretion, some of cuticle proteins are covalently linked with chitin or with one another. So far, huge numbers of cuticle proteins have been identified and isolated from various insect orders, including Diptera (Snyder *et al.*, 1982; Charles *et al.*, 1997; Dotson *et al.*, 1998), Lepidoptera (Horodyski and Riddiford, 1989; Binger and Willis, 1994; Nakato *et al.*, 1994), Coleoptera (Rondot *et al.*, 1996; Andersen *et al.*, 1997; Missios *et al.*, 2000), and Orthoptera (Højrup *et al.*, 1986; Andersen, 1998), and they are structurally analyzed. In addition to insects, cuticle proteins are also identified in other arthropod animals, including crustaceans, the American lobster, *Homarus americanus* (Nousiainen *et al.*, 1998) and the prawn, *Penaeus japonicus* (Watanabe *et al.*, 2000); and spider, *Araneus diadematus* (Norup *et al.*, 1996).

Accumulation of information on primary structures which were sequenced directly or predicted from their cDNA sequences revealed that cuticle proteins share some obvious homologous regions and/or motifs in common (Andersen *et al.*, 1995). They are listed below.

R & R Consensus: More than half of primary structures of cuticle proteins, whose structural information are available, are aligned well at their central or C-terminal region with 35 amino acid residues, and provide the consensus sequence. This conserved region was first mentioned by Rebers and Riddiford (1988), so called "the R & R Consensus". Increase in structural data showed that the conserved region is extended to N-terminal beyond the R & R Consensus with about 70 amino acids in total (Lampe and Willis, 1994; Andersen *et al.*, 1995), and this region was suggested to function in binding with other components in cuticle such as chitin or proteins. Iconomidou *et al.* (1999) predicted that this region of cuticle proteins consists of several β-strands, and proposed this conformation is essential for binding to chitin fibrils and defining cuticle's helicoidal architecture. Andersen (1998, 2000) proposed that, according to structure of the conserved region, this group of cuticle proteins could be divided

to RR-1, RR-2, and RR-3 subgroups, which include cuticle proteins mainly derived from soft and flexible, stiff and sclerotized, and post-ecdysial cuticle, respectively.

Proline/Alanine rich region: Some of R & R group cuticle proteins possess a region whose amino acid content is unusually abundant in proline and alanine, which is just following the R & R Consensus. Several proteins from *Locusta*, *Schistocerca*, and *Drosophila* bear the octapeptide sequence, xPTPPPxP (x represents hydrophobic residue), within the Pro/Ala rich region. This sequence resembles to the recognition sequence for SH3 domain of transducing tyrosine kinase family (Ren *et al.*, 1993). Therefore, this motif has been discussed to be involved in protein/protein interaction (Andersen *et al.*, 1995).

Alanine rich repeat: Andersen *et al.* (1995) and Andersen (2000) mentioned a kind of repeats, whose consensus appears to be (P/V)xDTPEVAAA(K/R)AA(H/F)xAA(H/Y), reiterated in some cuticle proteins. This repeat was first discovered by Nakato *et al.* (1990) in the primary structure of a pupal cuticle protein from *Bombyx mori*, suggested involvement in protein/chitin interaction.

Ala-Ala-Pro-Ala/Pro motif: The AAP(A/P) motif also occurs over and over in some proteins (Andersen *et al.*, 1995). It is known that the peptides with proline tend to bend, therefore cuticle proteins bearing this motif presumably is apt to fold.

Glycine rich motif: Cuticle proteins possessing glycine rich stretch, usually GGx tripeptide repeats, at N-terminal or all over the sequence have been identified in *Drosophila* (Apple and Fristrom, 1991), *Locusta* (Andersen *et al.*, 1995; Andersen, 2000), *Tenebrio* (Bouhin *et al.*, 1992a; Mathelin *et al.*, 1995), *Manduca* (Okot-Kotber *et al.*, 1996), and *Bombyx* (Suzuki *et al.*, 2000). Glycine rich stretches are also present in many extracellular structural proteins, such as insect chorion proteins (Tsitilou *et al.*, 1983) and collagen, in both invertebrates and vertebrates.

Hydrophobic cuticle protein: Cuticle proteins in this group have hardly any obvious conserved sequence or common motif. However, they have unique amino acid content, which is abundant in valine, alanine, and proline, and hydrophobic in total. Proteins of this type have been identified in *Bombyx* (Shiomi *et al.*, 1998; Sawada *et al.*, personal

communication) and *Manduca* (Okot-Kotber *et al.*, 1996).

Hormonal Regulation of Post-embryonic Development in Insects

Post-embryonic development of insects is roughly classified into three steps, growth, molt, and metamorphosis. These processes are mainly controlled by two peripheral hormones; those are 20-hydroxyecdysone (20HE), the active metabolite of ecdysone, and juvenile hormone (JH). In larval stages of lepidopteran species, a pulse of 20HE in presence of JH induces the larval-larval molt and makes epidermal cells synthesize the larval cuticle. In contrast, after the titer of JH in hemolymph declines in final instar larva, raise of 20HE causes larval-pupal molt (metamorphosis) and allows epidermal cells to produce the pupal cuticle (Riddiford, 1976).

20HE demonstrates its actions mainly as transcriptional regulation. Studying the puffing patterns of the salivary gland chromosome in *Drosophila*, Ashburner *et al.* (1974) proposed a model for mode of 20HE action. In the model 20HE is accepted by a receptor and activates directly small set of genes (early genes) at first. The proteins encoded by these early genes are mostly transcription factors, and activate a second set of genes (late genes). The products of early genes also repress expression of themselves. 20HE is a small lipophilic hormone similar to vertebrate steroids, retinoids, and thyroid hormones. These molecules are accepted by members of structurally and functionally related nuclear receptor superfamily, which act as a ligand dependent transcription factor composed of homo- or heterodimer, and regulate gene expression in a target cell (Mangelsdorf *et al.*, 1995). In many cases the responses of these hormones take place in two steps; the direct induction of the transcription from a small number of specific genes known as the "primary response", and subsequent activation of other genes by the products from the primary response as the "secondary response" (Yamamoto, 1985). 20HE, in the same way, is accepted by the ecdysone receptor, a heterodimer of two nuclear receptors, EcR and USP (Koelle *et al.*, 1991; Yao *et al.*, 1992), and stimulates sequential ecdysone signal cascade (Thummel, 1995; Lam *et al.*, 1997). This cascade is basically conserved in other insect species including *Manduca*

sexta (Fujiwara *et al.*, 1995; Hiruma *et al.*, 1997), *Bombyx mori* (Swevers *et al.*, 1995; Kamimura *et al.*, 1996), *Galleria mellonella* (Jindra *et al.*, 1994a, b), and *Aedes aegypti* (Li *et al.*, 2000).

JH is an insect sesquiterpenoid hormone, which is synthesized and secreted by the corpora allata. Seven structurally related molecules have been identified in insects. JH-I and JH-II appear to be restricted in Lepidoptera, and JH-0 and 4-methyl JH-I have been isolated only from lepidopteran eggs. JH-III is the most popular molecule and has been found in all insect orders investigated (Nijhout, 1994). JH-III-bisepoxide and methyl farnesoate are found only in higher Diptera and a cockroach, respectively (Brüning *et al.*, 1985; Richard *et al.*, 1989).

JH has pleiotropic actions with two major roles widespread into probably all insect species. One is to maintain larval or nymphal characters, so-called "status quo" action. JH modifies ecdysone action and its cascade at each molt, causing prevention of metamorphosis (Riddiford, 1976; Nijhout, 1994). The other is regulation of reproductive maturation with two steps, induction of yolk protein biosynthesis (vitellogenesis) and stimulation of follicle to uptake yolk proteins (Nijhout, 1994). JH also serves other roles including destruction of adult diapause of Hemiptera (Morita *et al.*, 1999), caste determination of Isoptera (Nijhout, 1994), control of polymorphism (Emlen and Nijhout, 1999), and maturation of sex pheromone glands in Lepidoptera (Fan *et al.*, 1999).

Because of various roles, JH is supposed to have some different modes of action at molecular and cellular levels. It appears to act at the membrane level in ovarian follicle cells to uptake yolk proteins (Ilenchuk and Davey, 1987; Sevala *et al.*, 1995), and in contrast, it seems to act at inner cell or nuclear levels in the epidermal cells and fat body (Osir and Riddiford, 1988; Engelmann, 1995). It has been also reported that JH regulates expression of several genes involved in both developmental and reproductive process at transcriptional level (Zhang *et al.*, 1993; Nakato *et al.*, 1994; Hirai *et al.*, 1998; Comas *et al.*, 1999; Dubrovsky *et al.*, 2000). Compared with ecdysone, however, the molecular mechanism of JH action have been still dark, nevertheless big efforts have been made to understand it and to identify

specific receptor of JH. Some nuclear proteins with affinity to JH have been reported (Palli *et al.*, 1994; Jones and Sharp, 1997). And, since JH is a small lipophilic hormone, the possibility that JH correlates with members of nuclear receptor superfamily in gene expression has been examined (Forman *et al.*, 1995; Harmon *et al.*, 1995). Although these experiments seemed to have brought some positive data, it has become clear after all that they were artificial or that candidate of receptor requires high concentration of JH over the physical condition in binding reaction. The possibility that JH acts in the novel molecular system for modulating ecdysone action has been also discussed (Riddiford, 1996).

Hormonal Regulation of Cuticle Protein Genes

The majority of insect cuticle proteins is biosynthesized in stage dependent manner. It has been shown that expression of several genes for cuticle proteins is regulated by 20HE and/or JH. In *Drosophila melanogaster*, it was reported that transcripts of pupal cuticle protein genes (EDG78, 84, and 91) accumulate in isolated imaginal discs after exposure to a pulse of 20HE, but not in discs cultured in the continuous presence or absence of the hormone (Apple and Fristrom, 1991).

In epidermal culture system of tobacco hornworm, *Manduca sexta*, it was shown that expression of gene coding for larval-cuticle specific protein, LCP14, is rapidly suppressed in response to 20HE. This suppression by 20HE could be caused by both cessation of the gene transcription and destabilization of the mRNA (Hiruma *et al.*, 1991).

Expression of a pupal cuticle protein gene (GmPCP52) of the wax moth, *Galleria mellonella*, reaches a peak within several hours after pupation, and declines rapidly. 20HE accelerates the decrease of GmPCP52 expression both transcriptionally and post-transcriptionally, but JH delays it (Krämer and Wolbert, 1998).

In silkworm, *Bombyx mori*, mRNAs for two major larval cuticle proteins, BMCP18 and 30, decrease a few days after last larval molt correspondingly to disappearance of JH in hemolymph. Nakato *et al.* (1994) and Shofuda *et al.* (1999) reported that topical application of JH to last instar larvae prolonged the expression of these mRNAs for more than six days.

On the other hand, mRNA for one of pupal cuticle proteins (PCP) is transactivated in isolated abdomen with an interval of 96h after 20HE administration (Nakato *et al.*, 1992). VAP-peptide (BmACP6.7) is another *Bombyx* cuticle protein constructing adult cuticle (Shiomi *et al.*, 1998). BmFTZ-F1, a member of nuclear receptors incorporated into the ecdysone signal cascade (Ueda and Hirose, 1990), recognized the promoter region of the gene for VAP-peptide, suggesting possibility that this gene is controlled by ecdysone signal (Shiomi *et al.*, 2000).

In contrast to other ecdysone stimulating cuticle protein genes, expression of an adult-specific cuticle protein gene of yellow mealworm, *Tenebrio molitor* (TMACP20) requires continuous presence of 20HE but not pulse stimulation of it in cultured epidermis (Braquart *et al.*, 1996), while JH analog application to newly ecdysis pupae prevents its expression (Charles *et al.*, 1992). Expression of mRNA for another adult-specific cuticle protein, TMACP22, is also prevented by JH analog application (Bouhin *et al.*, 1992b). Although adult-specific cuticle proteins were shown to require the absence of JH for their expression, genes for larval/pupal cuticle protein of *Tenebrio* (TMLPCP22 and 23) are activated by JH (Rondot *et al.*, 1996; Rondot *et al.*, 1998).

In the experiments with cultured epidermis of *Manduca* and *Tenebrio*, the effects of 20HE on the transcription of cuticle protein genes were prevented by protein synthesis inhibitors, indicating that 20HE does not act directly on these genes. Therefore, these cuticle protein genes could be late genes in secondary response of 20HE.

Scope of Project

Insect cuticle is exoskeleton, whose mechanical property is varied suitable for physiological state of the individual. This difference is thought to depend on cuticle protein composition in cuticle. Therefore, to understand the mechanism by which cuticle maintains stable structure and expresses various mechanical properties, it is important to analyze how cuticle proteins recognize chitin molecule and bind to chitin. Although enormous structural information is accumulating, little is known about chitin-binding property of cuticle proteins. On the other hand, cuticle protein genes provide an excellent model system for studying the

mechanism of gene regulation during metamorphosis, since their expression is stage dependent and controlled by 20HE and/or JH.

To approach these two subjects, I adopted cuticle proteins from silkworm, *Bombyx mori*, as model proteins because of following advantages. First, *B. mori* is a large insect, so it is easy to isolate interesting tissues and to obtain large amounts of samples useful to biochemical analyses. Second, endocrinology of *B. mori* is known well conveniently to study regulation of gene expression during post-embryonic development. Finally, expressions of genes for cuticle proteins abundant in larval cuticle are up-regulated by JH at transcriptional level. The mode of JH action is still unknown as described above. Until now, many studies to make it clear have been concentrated to identify the JH receptor. Analysis for mechanism of expression control of JH target genes such as BMCP30 and BMCP18 possibly provides a novel light to understand the JH actions.

In Section I of this thesis, I described the chitin-binding activity of BMCP30 particularly. To analyze the molecular property of cuticle proteins, I purified some cuticle proteins from larvae of *B. mori* in native form, and their chitin-binding activities were analyzed by chitin-affinity chromatography. The experiment using a series of deletion mutant proteins showed that the region conserved among cuticle proteins functions as the chitin binding domain. In Section II, I made reference to the regulation mechanism and the JH responsiveness of BMCP18 gene expression. Comparison of promoter region of BMCP18 gene with those of other lepidopteran cuticle protein genes (MSCP14.6 and HCCP12) revealed a conserved sequence of 12-bp. It was shown that Bmsvp, a member of nuclear hormone receptor superfamily, recognizes this sequence specifically *in vitro*.

§ Cuticle proteins identified in larval cuticle of silkworm, *Bombyx mori*, were previously referred as LCP30, 22, 18, and 17 by my collaborators (Nakato *et al.*, 1994, 1997; Shofuda *et al.*, 1999). Since the expression of these genes are not restricted to larval stages, I and colleague have changed the nomenclature to BMCP30, 22, 18, and 17, standing for *Bombyx mori* cuticle protein with indicated molecular weight.

REFERENCES

Andersen, S. O. (1998). Amino acid sequence studies on endocuticular proteins from the desert locust, *Schistocerca gregaria*. *Insect Biochem. Mol. Biol.* **28,** 421-434.

Andersen, S. O. (2000). Studies on proteins in post-ecdysial nymphal cuticle of locust, *Locusta migratoria*, and cockroach, *Blaberus craniifer*. *Insect Biochem. Mol. Biol.* **30,** 569-577.

Andersen, S. O. and Højrup, P. (1987). Extractable Proteins from abdominal cuticle of sexually mature locusts, *Locusta migratoria*. *Insect Biochem.* **17,** 45-51.

Andersen, S. O., Højrup, P., and Roepstorff, P. (1986). Characterization of cuticular proteins from the migratory locust, *Locusta migratoria*. *Insect Biochem.* **16,** 441-447.

Andersen, S. O., Højrup, P., and Roepstorff, P. (1995). Insect cuticular proteins. *Insect Biochem. Mol. Biol.* **25,** 153-176.

Andersen, S. O., Rafn, K., and Roepstorff, P. (1997). Sequence studies of proteins from larval and pupal cuticle of the yellow meal worm, *Tenebrio molitor*. *Insect Biochem. Mol. Biol.* **27,** 121-131.

Apple, R. T. and Fristrom, J. W. (1991). 20-Hydroxyecdysone is required for, and negatively regulates, transcription of *Drosophila* pupal cuticle protein genes. *Dev. Biol.* **146,** 569-582.

Ashburner, M., Chihara, C., Meltzer, P., and Richards, G. (1974). Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* **38,** 655-662.

Bærnholdt, D. and Andersen, S. O. (1998). Sequence studies on post-ecdysial cuticular proteins from pupae of the yellow mealworm, *Tenebrio molitor*. *Insect Biochem. Mol. Biol.* **28,** 517-526.

Binger, L. C. and Willis, J. H. (1990). *In vitro* translation of epidermal RNA from different anatomical regions and metamorphic stages of *Hyalophora cecropia*. *Insect Biochem.* **20,** 573-583.

Binger, L. C. and Willis, J. H. (1994). Identification of the cDNA, gene and promoter for a major protein from flexible cuticles of the giant silkmoth *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* **24,** 989-1000.

Blackwell, J. and Weih, M. A. (1980). Structure of chitin-protein complexes: ovipositor of the ichneumon fly *Megarhyssa*. *J. Mol. Biol.* **137,** 49-60.

Bouhin, H., Charles, J. P., Quennedey, B., Courrent, A., and Delachambre, J. (1992a). Characterization of a cDNA clone encoding a glycine-rich cuticular protein of *Tenebrio molitor*: developmental expression and effect of a juvenile hormone analogue. *Insect Mol Biol* **1,** 53-62Insect Molecular Biology.

Bouhin, H., Charles, J. P., Quennedey, B., and Delachambre, J. (1992b). Developmental profiles of epidermal mRNAs during the pupal-adult molt of *Tenebrio molitor* and isolation of a cDNA clone encoding an adult cuticular protein: effects of a juvenile hormone analogue. *Dev. Biol.* **149,** 112-122.

Braquart, C., Bouhin, H., Quennedey, A., and Delachambre, J. (1996). Up-regulation of an adult cuticular gene by 20-hydroxyecdysone in insect metamorphosing epidermis cultured *in vitro*. *Eur. J. Biochem.* **240,** 336-341.

Brüning, E., Saxer, A., and Lanzrein, B. (1985). Methyl farnesoate and juvenile hormone III in normal and precocene treated embryos and the ovoviviparous cockroach, *Nauphoeta cinerea*. *Int. J. Invertebr. Reprod. Dev.* **8,** 269-278.

Charles, J. P., Bouhin, H., Quennedey, B., Courrent, A., and Delachambre, J. (1992). cDNA cloning and deduced amino acid sequence of a major, glycine-rich cuticular protein from the coleopteran *Tenebrio molitor*. Temporal and spatial distribution of the transcript during metamorphosis. *Eur. J. Biochem.* **206,** 813-819.

Charles, J. P., Chihara, C., Nejad, S., and Riddiford, L. M. (1997). A cluster of cuticle protein genes of *Drosophila* melanogaster at 65A: sequence, structure and evolution. *Genetics* **147,** 1213-1224 [published erratum appears in *Genetics* 1998 ,**148**, 921].

Chihara, C. J., Silvert, D. J., and Fristrom, J. W. (1982). The cuticle proteins of *Drosophila melanogaster*: stage specificity. *Dev. Biol.* **89,** 379-388.

Christensen, A. M., Schaefer, J., Kramer, K. J., Morgan, T. D., and Hopkins, T. L. (1991). Detection of cross-links in insect cuticle by REDOR NMR spectroscopy. *J. Am. Chem. Soc.* **113,** 6799-6802.

Comas, D., Piulachs, M. D., and Belles, X. (1999). Fast induction of vitellogenin gene expression by juvenile hormone III in the cockroach *Blattella germanica* (L.) (Dictyoptera, Blattellidae). *Insect Biochem. Mol. Biol.* **29,** 821-827.

Cox, D. L. and Willis, J. H. (1985). The cuticular proteins of *Hyalophora cecropia* from different anatomical regions and metamorphic stages. *Insect Biochem.* **15,** 349-362.

Dotson, E. M., Cornel, A. J., Willis, J. H., and Collins, F. H. (1998). A family of pupal-specific cuticular protein genes in the mosquito *Anopheles gambiae*. *Insect Biochem.*

Mol. Biol. **28,** 459-472.

Dubrovsky, E. B., Dubrovskaya, V. A., Bilderback, A. L., and Berger, E. M. (2000). The isolation of two juvenile hormone-inducible genes in *Drosophila melanogaster*. *Dev. Biol.* **224,** 486-495.

Emlen, D. J. and Nijhout, H. F. (1999). Hormonal control of male horn length dimorphism in the dung beetle *Onthophagus taurus* (Coleoptera: Scarabaeidae). *J. Insect Physiol.* **45,** 45-53.

Engelmann, F. (1995). The juvenile hormone receptor of the cockroach *Leucophaea maderae*. *Insect Biochem. Mol. Biol.* **25,** 721-726.

Fan, Y., Rafaeli, A., Gileadi, C., and Applebaum, S. W. (1999). Juvenile hormone induction of pheromone gland PBAN-responsiveness in *Helicoverpa armidera* females. *Insect Biochem. Mol. Biol.* **29,** 635-641.

Forman, B. M., Goode, E., Chen, J., Oro, A. E., Bradley, D. J., Perlmann, T., Noonan, D. J., Burka, L. T., McMorris, T., Lamph, W. W., Evans, R. M., and Weinberger, C. (1995). Identification of a nuclear receptor that is activated by farnesol metabolites. *Cell* **81,** 687-693.

Fujiwara, H., Jindra, M., Newitt, R., Palli, S. R., Hiruma, K., and Riddiford, L. M. (1995). Cloning of an ecdysone receptor homolog from *Manduca sexta* and the developmental profile of its mRNA in wings. *Insect Biochem. Mol. Biol.* **25,** 845-856.

Hackman, R. H. (1984). Cuticle: Biochemistry. *In* "Biology of the Integument, Vol. 1, Invertebrates" (Bereiter-Hahn, J., Matoltsy, A. G., and Richards, K. S., Eds.). Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

Harmon, M. A., Boehm, M. F., Heyman, R. A., and Mangelsdorf, D. J. (1995). Activation of mammalian retinoid X receptors by the insect growth regulator methoprene. *Proc. Natl. Acad. Sci. USA* **92,** 6157-6160.

Hirai, M., Yuda, M., Shinoda, T., and Chinzei, Y. (1998). Identification and cDNA cloning of novel juvenile hormone responsive genes from fat body of the bean bug, *Riptortus clavatus* by mRNA differential display. *Insect Biochem. Mol. Biol.* **28,** 181-189.

Hiruma, K., Böcking, D., Lafont, R., and Riddiford, L. M. (1997). Action of different ecdysteroids on the regulation of mRNAs for the ecdysone receptor, MHR3, dopa decarboxylase, and a larval cuticle protein in the larval epidermis of the tobacco hornworm, *Manduca sexta*. *Gen. Comp. Endocrinol.* **107,** 84-97.

Hiruma, K., Hardie, J., and Riddiford, L. M. (1991). Hormonal regulation of epidermal

metamorphosis *in vitro*: control of expression of a larval-specific cuticle gene. *Dev. Biol.* **144,** 369-378.

Højrup, P., Andersen, S. O., and Roepstorff, P. (1986). Primary structure of a structural protein from the cuticle of the migratory locust, *Locusta migratoria*. *Biochem. J.* **236,** 713-720.

Horodyski, F. M. and Riddiford, L. M. (1989). Expression and hormonal control of a new larval cuticular multigene family at the onset of metamorphosis of the tobacco hornworm. *Dev. Biol.* **132,** 292-303.

Iconomidou, V. A., Willis, J. H., and Hamodrakas, S. J. (1999). Is β-pleated sheet the molecular conformation which dictates formation of helicoidal cuticle? *Insect Biochem. Mol. Biol.* **29,** 285-292.

Ilenchuk, T. T. and Davey, K. G. (1987). Effects of various compounds on Na/K-ATPase activity, JH I binding capacity and patency response in follicles of *Rhodnius prolixus*. *Insect Biochem.* **17,** 1085-1088.

Jindra, M., Sehnal, F., and Riddiford, L. M. (1994a). Isolation and developmental expression of the ecdysteroid-induced *GHR3* gene of the wax moth *Galleria mellonella*. *Insect Biochem. Mol. Biol.* **24,** 763-773.

Jindra, M., Sehnal, F., and Riddiford, L. M. (1994b). Isolation, characterization and developmental expression of the ecdysteroid-induced *E75* gene of the wax moth *Galleria mellonella*. *Eur. J. Biochem.* **221,** 665-675.

Jones, G. and Sharp, P. A. (1997). Ultraspiracle: an invertebrate nuclear receptor for juvenile hormones. *Proc. Natl. Acad. Sci. USA* **94,** 13499-13503.

Kamimura, M., Tomita, S., and Fujiwara, H. (1996). Molecular cloning of an ecdysone receptor (B1 isoform) homologue from the silkworm, *Bombyx mori*, and its mRNA expression during wing disc development. *Comp. Biochem. Physiol. B* **113,** 341-347.

Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P., and Hogness, D. S. (1991). The *Drosophila EcR* gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67,** 59-77.

Krämer, B. and Wolbert, P. (1998). Down-regulation of expression of a pupal cuticle protein gene by transcriptional and posttranscriptional control mechanisms during metamorphosis in *Galleria*. *Dev. Genes Evol.* **208,** 205-212.

Lam, G. T., Jiang, C., and Thummel, C. S. (1997). Coordination of larval and prepupal gene expression by the DHR3 orphan receptor during *Drosophila* metamorphosis.

Development **124,** 1757-1769.

Lampe, D. J. and Willis, J. H. (1994). Characterization of a cDNA and gene encoding a cuticular protein from rigid cuticles of the giant silkmoth, *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* **24,** 419-435.

Li, C., Kapitskaya, M. Z., Zhu, J., Miura, K., Segraves, W., and Raikhel, A. S. (2000). Conserved molecular mechanism for the stage specificity of the mosquito vitellogenic response to ecdysone. *Dev. Biol.* **224,** 96-110.

Locke, M. (1974). The structure and formation of the integument in insects. *In* "The physiology of Insecta" (Rockstein, M., Ed.), Vol. 6, pp. 123-213. Academic Press, New York.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83,** 835-839.

Mathelin, J., Bouhin, H., Quennedey, B., Courrent, A., and Delachambre, J. (1995). Identification, sequence and mRNA expression pattern during metamorphosis of a cDNA encoding a glycine-rich cuticular protein in *Tenebrio molitor*. *Gene* **156,** 259-264.

Mathelin, J., Quennedey, B., Bouhin, H., and Delachambre, J. (1998). Characterization of two new cuticular genes specifically expressed during the post-ecdysial molting period in *Tenebrio molitor*. *Gene* **211,** 351-359.

Missios, S., Davidson, H. C., Linder, D., Mortimer, L., Okobi, A. O., and Doctor, J. S. (2000). Characterization of cuticular proteins in the red flour beetle, *Tribolium castaneum*. *Insect Biochem. Mol. Biol.* **30,** 47-56.

Morita, A., Soga, K., Hoson, T., Kamisaka, S., and Numata, H. (1999). Changes in mechanical properties of the cuticle and lipid accumulation in relation to adult diapause in the bean bug, *Riptortus clavatus*. *J. Insect Physiol.* **45,** 241-247.

Nakato, H., Izumi, S., and Tomino, S. (1992). Structure and expression of gene coding for a pupal cuticle protein of *Bombyx mori*. *Biochim. Biophys. Acta* **1132,** 161-167.

Nakato, H., Shofuda, K., Izumi, S., and Tomino, S. (1994). Structure and developmental expression of a larval cuticle protein gene of the silkworm, *Bombyx mori*. *Biochim. Biophys. Acta* **1218,** 64-74.

Nakato, H., Takekoshi, M., Togawa, T., Izumi, S., and Tomino, S. (1997). Purification and cDNA cloning of evolutionally conserved larval cuticle proteins of the silkworm,

Bombyx mori. *Insect Biochem. Mol. Biol.* **27,** 701-709.

Nakato, H., Toriyama, M., Izumi, S., and Tomino, S. (1990). Structure and expression of mRNA for a pupal cuticle protein of the silkworm, *Bombyx mori*. *Insect Biochem.* **20,** 667-678.

Neville, A. C. (1984). Cuticle: Organization. *In* "Biology of the Integument, Vol. 1, Invertebrates" (Bereiter-Hahn, J., Matoltsy, A. G., and Richards, K. S., Eds.). Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

Nijhout, H. F. (1994). "Insect Hormones." Princeton University Press, Princeton.

Norup, T., Berg, T., Stenholm, H., Andersen, S. O., and Højrup, P. (1996). Purification and characterization of five cuticular proteins from the spider *Araneus diadematus*. *Insect Biochem. Mol. Biol.* **26,** 907-915.

Nousiainen, M., Rafn, K., Skou, L., Roepstorff, P., and Andersen, S. O. (1998). Characterization of exoskeletal proteins from the American lobster, *Homarus americanus*. *Comp. Biochem. Physiol. B* **119,** 189-199.

Okot-Kotber, B. M., Morgan, T. D., Hopkins, T. L., and Kramer, K. J. (1996). Catecholamine-containing proteins from the pharate pupal cuticle of the tobacco hornworm, *Manduca sexta*. *Insect Biochem. Mol. Biol.* **26,** 475-484.

Osir, E. O. and Riddiford, L. M. (1988). Nuclear binding sites for juvenile hormone and its analogs in the epidermis of the tobacco hornworm. *J. Biol. Chem.* **263,** 13812-13818.

Palli, S. R., Touhara, K., Charles, J. P., Bonning, B. C., Atkinson, J. K., Trowell, S. C., Hiruma, K., Goodman, W. G., Kyriakides, T., Prestwich, G. D., Hammock, B. D., and Riddiford, L. M. (1994). A nuclear juvenile hormone-binding protein from larvae of *Manduca sexta*: a putative receptor for the metamorphic action of juvenile hormone. *Proc. Natl. Acad. Sci. USA* **91,** 6191-6195.

Rebers, J. E. and Riddiford, L. M. (1988). Structure and expression of a *Manduca sexta* larval cuticle gene homologous to *Drosophila* cuticle genes. *J. Mol. Biol.* **203,** 411-423.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259,** 1157-1161.

Richard, D. S., Applebaum, S. W., Sliter, T. J., Baker, F. C., Schooley, D. A., Reuter, C. C., Henrich, V. C., and Gilbert, L. I. (1989). Juvenile hormone bisepoxide biosynthesis *in vitro* by the ring gland of *Drosophila melanogaster*: a putative juvenile hormone in the higher Diptera. *Proc. Natl. Acad. Sci. USA* **86,** 1421-1425.

Riddiford, L. M. (1976). Hormonal control of insect epidermal cell commitment *in vitro*.

Nature **259,** 115-117.

Riddiford, L. M. (1996). Juvenile hormone: the status of its "status quo" action. *Arch. Insect Biochem. Physiol.* **32,** 271-286.

Rondot, I., Quennedey, B., Courrent, A., Lemoine, A., and Delachambre, J. (1996). Cloning and sequencing of a cDNA encoding a larval-pupal-specific cuticular protein in *Tenebrio molitor* (Insecta, Coleoptera). Developmental expression and effect of a juvenile hormone analogue. *Eur. J. Biochem.* **235,** 138-143.

Rondot, I., Quennedey, B., and Delachambre, J. (1998). Structure, organization and expression of two clustered cuticle protein genes during the metamorphosis of an insect, *Tenebrio molitor*. *Eur. J. Biochem.* **254,** 304-312.

Sevala, V. L., Davey, K. G., and Prestwich, G. D. (1995). Photoaffinity labeling and characterization of a juvenile hormone binding protein in the membranes of follicle cells of *Locusta migratoria*. *Insect Biochem. Mol. Biol.* **25,** 267-273.

Shiomi, K., Niimi, T., Imai, K., and Yamashita, O. (2000). Structure of the VAP-peptide (BmACP-6.7) gene in the silkworm, *Bombyx mori* and a possible regulation of its expression by BmFTZ-F1. *Insect Biochem. Mol. Biol.* **30,** 119-125.

Shiomi, K., Niimi, T., Sato, Y., Imai, K., and Yamashita, O. (1998). A hydrophobic peptide (VAP-peptide) of the silkworm, *Bombyx mori*: a unique role for adult activity proposed from gene expression and production at the terminal phase of metamorphosis. *Insect Biochem. Mol. Biol.* **28,** 671-676.

Shofuda, K., Togawa, T., Nakato, H., Tomino, S., and Izumi, S. (1999). Molecular cloning and characterization of a cDNA encoding a larval cuticle protein of *Bombyx mori*. *Comp. Biochem. Physiol. B* **122,** 105-109.

Snyder, M., Hunkapiller, M., Yuen, D., Silvert, D., Fristrom, J., and Davidson, N. (1982). Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. *Cell* **29,** 1027-1040.

Suzuki, Y., Matsuoka, T., and Fujiwara, H. "Analysis for hormone response of *GCP1*, a novel integument specific gene in silkworm, *Bombyx mori*." *The 70th Annual Meeting of The Japanese Society of Sericultural Science*, Tokyo.

Swevers, L., Drevet, J. R., Lunke, M. D., and Iatrou, K. (1995). The silkmoth homolog of the *Drosophila* ecdysone receptor (B1 isoform): cloning and analysis of expression during follicular cell differentiation. *Insect Biochem. Mol. Biol.* **25,** 857-866.

Thummel, C. S. (1995). From embryogenesis to metamorphosis: the regulation and function of

Drosophila nuclear receptor superfamily members. *Cell* **83,** 871-877.

Tsitilou, S. G., Rodakis, G. C., Alexopoulou, M., Kafatos, F. C., Ito, K., and Iatrou, K. (1983). Structural features of B family chorion sequences in the silkmoth *Bombyx mori*, and their evolutionary implications. *ENBO J* **2,** 1845-1852.

Ueda, H. and Hirose, S. (1990). Identification and purification of a *Bombyx mori* homologue of FTZ-F1. *Nucleic Acids Res.* **18,** 7229-7234.

Watanabe, T., Persson, P., Endo, H., and Kono, M. (2000). Molecular analysis of two genes, DD9A and B, which are expressed during the postmolt stage in the decapod crustacean *Penaeus japonicus*. *Comp. Biochem. Physiol. B* **125,** 127-136.

Yamamoto, K. R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.* **19,** 209-252.

Yao, T. P., Segraves, W. A., Oro, A. E., McKeown, M., and Evans, R. M. (1992). *Drosophila ultraspiracle* modulates ecdysone receptor function via heterodimer formation. *Cell* **71,** 63-72.

Zhang, J., McCracken, A., and Wyatt, G. R. (1993). Properties and sequence of a female-specific, juvenile hormone-induced protein from locust hemolymph. *J. Biol. Chem.* **268,** 3282-3288.

Section I

Chitin Recognition Mechanism of Cuticle Proteins from *Bombyx mori*

INTRODUCTION

Insect surface is covered by exoskeleton, cuticle, whose mechanical property is highly diverse suitable for physiological state of the individual according as developmental stage, sex, body region, and so on. Cuticle is mainly composed of chitin, a polymer of *N*-acetylglucosamine (GlcNAc), and cuticle proteins, which are structural proteins binding to chitin and/or other proteins. Compositions of cuticle proteins are varied among cuticle possessing different mechanical properties (Chihara *et al.*, 1982; Cox and Willis, 1985; Andersen and Højrup, 1987; Nakato *et al.*, 1990). Therefore, it is believed that cuticle proteins have very important role in expression of cuticle mechanical properties. A number of cuticle proteins have been identified and characterized in various insect species (Snyder *et al.*, 1982; Rebers and Riddiford, 1988; Binger and Willis, 1994; Nakato *et al.*, 1994; Andersen *et al.*, 1995; Mathelin *et al.*, 1995; Charles *et al.*, 1997; Andersen, 1998; Dotson *et al.*, 1998). A comparison of the primary structures of cuticle proteins from different species showed that many of these proteins contain a consensus sequence at the central or C-terminal region, first mentioned by Rebers and Riddiford (1988), so called "R & R Consensus". The R & R Consensus exists not only in insect cuticle proteins but also in those from other arthropod species, including the American lobster, *Homarus americanus* (Nousiainen *et al.*, 1998) and the spider, *Araneus diadematus* (Norup *et al.*, 1996). Increase in the available information about the structures revealed that the conserved region is extended to N-terminal beyond the R & R Consensus (Lampe and Willis, 1994; Andersen *et al.*, 1995). Andersen (1998, 2000) proposed that, according to structure of the conserved region, cuticle proteins could be divided to RR-1, RR-2, and RR-3 subgroups, which include cuticle proteins mainly derived from soft and flexible, stiff and sclerotized, and post-ecdysial cuticle, respectively. Although structural information of cuticle proteins is accumulating, little is known about their chitin-binding property.

In larval cuticle of *Bombyx mori*, three major components of cuticle protein, BMCP30, BMCP22, and BMCP17 have been identified (Nakato *et al.*, 1990, 1994). BMCP30 has been

purified from the urea extract of the integument using chitin-affinity chromatography followed by a preparative SDS-PAGE (Nakato *et al.*, 1994). However, the proteins purified by this procedure were denatured, and were not suitable for analyses of native protein properties. In addition, other cuticle proteins have been left to be elucidated. In this study, to analyze the chitin-binding activity, I developed the purification method of cuticle proteins from *B. mori* larval integument in the native form, and the properties of chitin-binding activity of BMCP30 were analyzed by the chitin-affinity chromatography method. The analysis using deletion mutant of BMCP30 showed that the conserved region among cuticle proteins should function as a chitin binding domain.

MATERIALS AND METHODS

Animals

A commercial strain of the silkworm, *Bombyx mori* (*Kinshu* × *Showa*) was obtained from Ueda Sanshu, Ueda. Larvae were reared at 25°C on an artificial diet (Katakura Kogyo Corp.).

Chemicals and biological materials

The cDNA clones of *B. mori* cuticle proteins were kindly supplied by Dr. Hiroshi Nakato.

DEAE-cellulose (DE52) was obtained from Whatmann Paper Limited. Polyvinyl-difluoride (PVDF) membrane was from ATTO, Inc. Glutathione Sepharose 4B and 1-chloro-2,4-dinitrobenzene (CDNB) were purchased from Pharmacia Biotech. Chitin powder was from Wako Pure Chemicals. *N*-acetylglucosamine and chitooligosaccharides, tri-*N*-acetylchitotriose, $(GlcNAc)_3$; and hexa-*N*-acetylchitohexaose, $(GlcNAc)_6$ were purchased from Seikagaku Corp.

Restriction endonucleases were obtained from New England Biolabs, Inc. and Roche Molecular Biochemicals. DNA polymerases for PCR were purchased from TaKaRa Shuzo Corp. Synthetic oligonucleotides were from Sawady Technology.

Other chemicals and enzymes were obtained from commercial sources.

Purification and N-terminal sequencing of cuticle proteins

Fifth instar larvae were decapitated, and silkglands, gut, and other internal organs were removed by squeezing the body with fingers. Integument was cut open, and fat body, muscles, and epidermal cells were torn off from the cuticle carefully with a spatula. Cuticle was washed with 10 volumes (w/v) of a low-salt buffer (20 mM Tris-HCl, pH 7.5, a small amount of 1-phenyl-2-thiourea) and a high-salt buffer (20 mM Tris-HCl, pH 7.5, 1 M NaCl, and a small amount of 1-phenyl-2-thiourea) on ice. Cuticle proteins were extracted from well-washed cuticle with 6 M guanidine-HCl in 20 mM NaPB, pH 7.0. Extracted proteins

were renatured by dialysis against 20 mM NaPB, pH 7.0 containing 3 mM phenylmethylsulfonyl fluoride (PMSF), and insoluble proteins were removed by centrifugation. After cuticle proteins were collected by ammonium sulfate precipitation (20% saturated ammonium sulfate) and dialyzed as above, DEAE-cellulose (DE52) column chromatography in 20 mM NaPB, pH 7.0 was performed to fractionate cuticle proteins. After the first peak eluted, a linear gradient of 0 – 0.1 M NaCl in the same buffer was loaded to elute other proteins. The fractions in each peak were collected, and subjected to an additional ion-exchange chromatography to purify proteins to homogeneity.

To determine the amino-terminal sequences of interest proteins, each 150 – 300 pmol of purified protein was subjected to SDS-PAGE, and blotted to a polyvinyl-difluoride (PVDF) membrane in a transfer buffer (25 mM Tris, 0.2 M glycine, and 20% methanol) using a semi-dry blotting apparatus (ATTO, Inc). The membrane was stained with Coomassie Brilliant Blue (CBB), and stained spot was excised. The amino acid sequence of the protein on a PVDF membrane was determined by Edman degradation using an Applied Biosystems 477A Protein Sequencer.

SDS polyacrylamide gel electrophoresis

SDS polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to the method of Laemmli (1970), and proteins were stained with CBB.

DNA sequencing

Nucleotide sequences were determined by dideoxynucleotide chain termination method (Sanger, 1981) using ABI PRISM BigDye Terminator Cycle Sequencing Kit and ABI PRISM 310 Genetic Analyzer (PE Biosystems).

Chitin affinity chromatography

Powdered chitin was washed well with 0.1 N HCl, then with 0.1 N NaOH, and relatively small particles of chitin were collected. Each 80 μg of cuticle proteins in 400 μl of

a binding buffer (20 mM NaPB, pH 6.4, 0.1% Triton X-100) was loaded onto a 150 μl chitin column equilibrated with the binding buffer. After the column was washed with binding buffer, the proteins binding to chitin were eluted with 8 M urea in binding buffer.

The binding activities were analyzed by SDS-PAGE followed by densitometry. The definite amount of sample proteins using in the chitin-binding assay on chitin-affinity chromatography and that of proteins eluted from chitin column were subjected to SDS-PAGE. The electrophoresis pattern was taken in a personal computer as a digital file, and densitometry was performed using NIH image software. The binding activities were calculated as ratio of the amount of proteins eluted from chitin column to that of the sample proteins.

Construction of expression plasmids

The DNA regions of BMCP30 cDNA were amplified by PCR with pBMCP30c2 (Nakato *et al.*, 1994) as a template and a set of following two synthetic primers: LCP30RV70-E (5'-GCCGAATTCGCCGAAACTGGAAAGTAC-3') and LCP30FW750-X (5'-GCCCTCGAGGACTAGTATTTACTTAGG-3') for pGSTBMCP30; LCP30RV526-E (5'-GCCGAATTCGACGTCGCTCCCGAAGGC-3') and LCP30FW750-X for pGSTBMCP30CBD; LCP30RV70-E and LCP30FW525-X (5'-GCCCTCGAGTTAATTGTCGTAACGGATG-3') for pGSTBMCP30ΔCBD. DNA regions of BMCP18 cDNA were amplified with pBMCP18c1 (Shofuda *et al.*, 1999) as the template and following primers: LCP18RV93-E (5'-GCCGAATTCGTCGGCCCTGATGCCGAC-3') and LCP18FW414-X (5'-GCCCTCGAGACTTTTCAAATGAATGAG-3') for pGSTBMCP18; LCP18RV138-E (5'-GCCGAATTCAACATTGGTGTCGAAGG-3') and LCP18FW414-X for pGSTBMCP18CBD. Each DNA fragment was digested with *Eco*RI and *Xho*I, and inserted between those sites on pGEX-4T-1 (Pharmacia Biotech). The sequences of all inserts in expression vectors were checked by DNA sequencing.

Expression and purification of GST-fusion protein

Synthesis of GST-fused cuticle proteins was performed according to the outline

mentioned by manufacture. *E. coli* BL21 cells carrying one of pGSTBMCP expression plasmids were cultured in 200 ml of LB containing ampicillin at 30°C. When the optical density at 600 nm reached 0.5, IPTG was added to a final concentration of 0.5 mM to the culture. After incubation for a further 2hrs, cells were collected by centrifugation and stocked at –80°C until use.

The bacterial pellet was resuspended in 10 ml of PBS, and cells were disrupted by sonication. Triton X-100 (final concentration 1%) was added to the sonicates, and the sonicates was incubated at room temperature for 30min with gentle shaking. The soluble fraction was separated by a vigorous centrifugation, and then applied to a 500 μl of Glutathione-Sepharose column equilibrated with PBS. After the column was washed well with PBS, proteins binding to column were eluted by some aliquots of elution buffer containing 10 mM reduced glutathione. Each eluted fraction was analyzed by SDS-PAGE, and the fractions containing GST-fusion proteins were combined and dialyzed against 20 mM NaPB, pH 7.0. GST-fusion proteins were purified further to homogeneity using FPLC system with Mono Q column.

Chitin-binding assay of GST-fusion proteins

Two μg of purified GST-fusion protein in 1 ml of a binding buffer (20 mM NaPB, pH 6.4, 0.1% Triton X-100) were incubated with 100 μl of acid-alkaline-washed chitin powder equilibrated with the binding buffer at room temperature for 30 min with a gentle shaking. The supernatant was removed, and chitin was washed well with the binding buffer. The 100 μl of protein-bound chitin was incubated in 1 ml of a GST reaction solution (0.5 mM 1-chloro-2,4-dinitrobenzene (CDNB), 0.5 mM reduced glutathione, 100 mM K-phosphate buffer, pH 6.5) at room temperature for 5 min with a gentle shaking. After the reaction was stopped with 10 μl of 6N HCl, the absorbance of the supernatant at 340nm was measured.

RESULTS

Purification of cuticle proteins from larval integument

To analyze the chitin-binding activity of cuticle proteins, I attempted to purify cuticle proteins in native form. Proteins were extracted from well-washed larval cuticle with 6 M guanidine-HCl solution, and renatured by dialysis against guanidine-free buffer. Although the *B. mori* cuticle proteins of which I take notice are soluble in low ionic strength solution, in a high-salt solution they become insoluble by salting out. Therefore, I was able to collect them by ammonium sulfate precipitation at 20% saturation from the solution. Cuticle proteins collected by the ammonium sulfate precipitation were applied onto a DEAE-cellulose column equilibrated with 20 mM NaPB, pH 7.0. After a first peak was eluted, a linear 0 – 0.1 M gradient of NaCl in the same buffer was applied on the column to elute other proteins (Fig. 1). SDS-PAGE analysis showed that fractions in first three peaks, indicated as I, II, and III in Fig. 1, contained BMCP22, BMCP17, and BMCP30, respectively (data not shown). A peak indicated as IV contained a novel protein with molecular weight of 18 k. Since this protein showed a high chitin-binding activity (Fig. 6), I termed it BMCP18.

To proceed with the purification of each cuticle protein, additional ion-exchange chromatography was performed. BMCP30 and BMCP22 were purified to homogeneity using DE52 columns equilibrated with 20 mM Tris-HCl, pH 8.0, and BMCP18 and BMCP17 were also purified with DE52 columns equilibrated with 20 mM NaPB, pH7.0 (Fig. 2). Their N-terminal sequences were determined by automatic Edman degradation. All purified cuticle proteins investigated possessed high chitin-binding activity (data not shown).

Primary structure of *B. mori* cuticle proteins

BMCP30 has been isolated, and its cDNA and genomic sequences were also characterized (Nakato *et al.*, 1994). The putative amino acid sequence has shown no obvious homology with other known proteins including insect cuticle proteins. However, I found two mistakes on cDNA sequencing, one is an oversight of 30 nucleotides repeat, and the other is

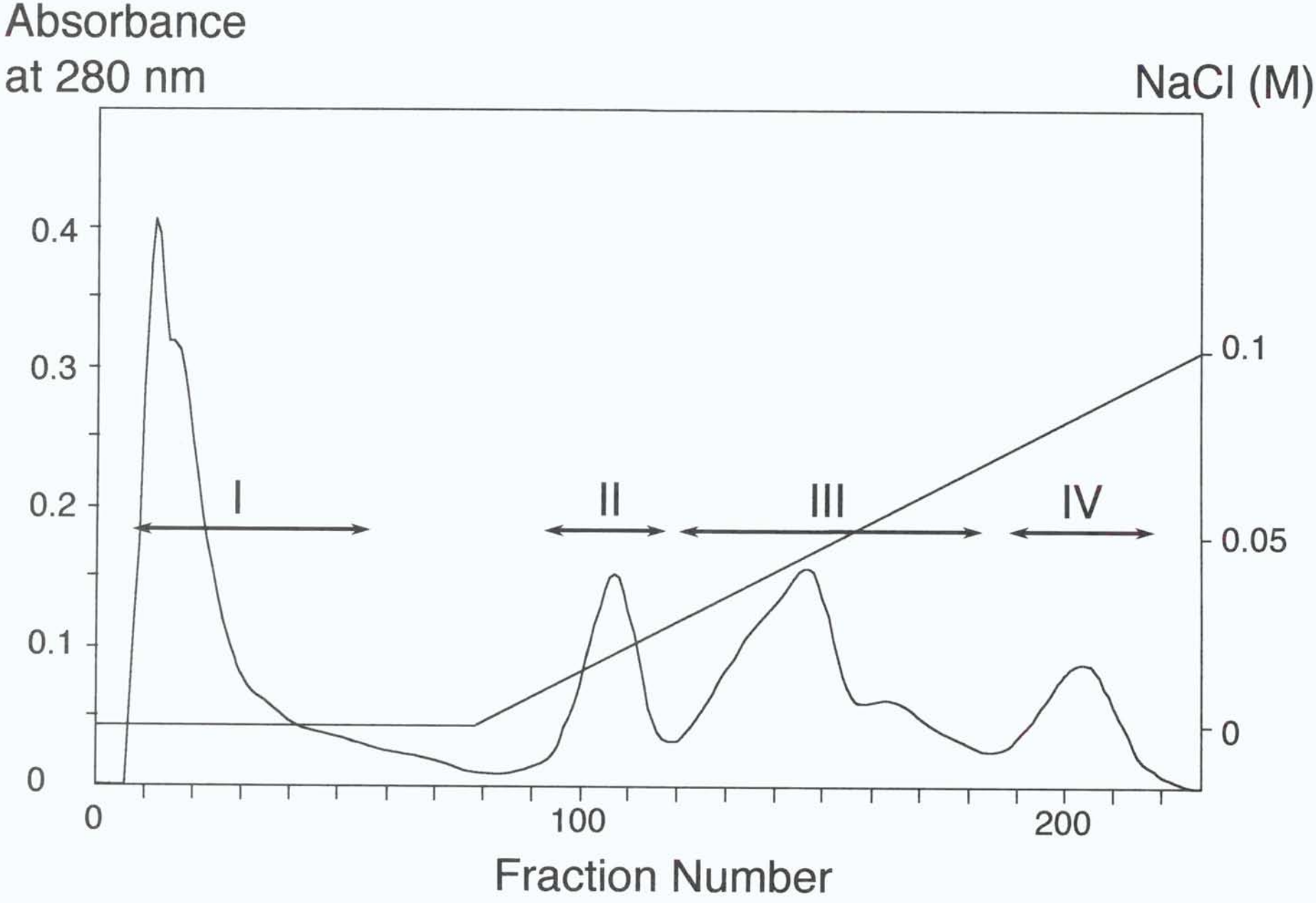

Absorbance
at 280 nm
NaCl (M)
0.4
0.3
0.2
0.1
0
0.1
0.05
0
I
II
III
IV
0
100
200
Fraction Number

Fig. 1. DE52 column chromatogram of *B. mori* cuticle proteins.

Cuticle proteins were extracted from well-washed larval cuticle by 6 M guanidine-HCl as described in Materials and Methods. After removal of guanidine-HCl by dialysis cuticle proteins were recovered by ammonium sulfate precipitation (20% saturation), and re-dialyzed against 20 mM NaPB, pH 7.0. This fraction was applied onto a DEAE-cellulose column equilibrated with 20 mM NaPB, pH 7.0. After the first peak was eluted, 0 – 0.1 M NaCl gradient in same buffer was loaded to the column. The *bars* with the *Roman numbers* indicate the pooling of fractions.

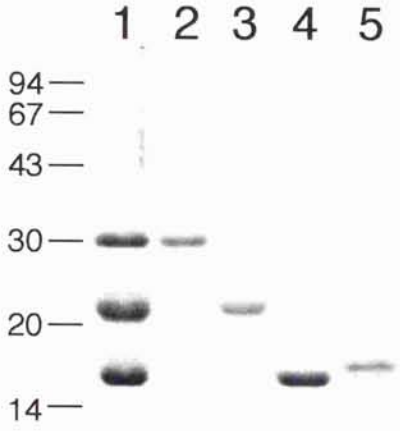


Fig. 2. SDS-PAGE of purified cuticle proteins.

Cuticle proteins precipitated by 20% ammonium sulfate (*lane 1*) were fractionated by cycles of ion-exchange chromatography, resulting purification of BMCP30 (*lane 2*), BMCP22 (*lane 3*), BMCP17 (*lane 4*), and BMCP18 (*lane 5*). *Numbers* to the left represent molecular weights of marker proteins.

lack of a guanine residue between nucleotide position 509 and 510 of original cDNA sequence causing translational frame shift at C-terminal region. Now, corrected cDNA sequence and primary structure of BMCP30 are available (Fig. 3). This new primary structure of BMCP30 possesses some noteworthy characteristics below. The molecular weight of pre-BMCP30 is 25.7 k, and that of mature BMCP30 without the signal peptide is 24.0 k with pI of 5.2. BMCP30 bears an Arg-Gly-Asp (RGD) sequence, which is known as the key sequence of cell-attachment domain identified in many extra-cellular matrix proteins (Pierschbacher and Ruoslahti, 1984; Yamada, 1991), as previously reported. There is a complete repeat of 10 amino acid residues derived from 30 nucleotides repeat mentioned above at central region. Moreover, C-terminal region of BMCP30 showed high similarity with the sequence conserved among many other insect cuticle proteins, called R&R Consensus (Andersen *et al.*, 1995).

mRNA sequences coding for BMCP22, BMCP17 (Nakato *et al.*, 1997), and BMCP18 (Shofuda *et al.*, 1999) have been isolated from *B. mori* larval epidermal cDNA library. Structural analysis of BMCP22 and BMCP17 cDNA sequences revealed that putative primary structures of these proteins show homology with cuticle proteins from other insect species, and possess the R & R Consensus in the central region. C-terminal region, which is abundant in proline and alanine residues, following R & R Consensus was also conserved among some cuticle proteins, such as SgAbd-4 from the desert locust, *Schistocerca gregaria* (Andersen, 1998), and LCP1 and LCP2 from *Drosophila melanogaster* (Snyder *et al.*, 1982) (Fig. 4). The deduced amino acid sequence of BMCP18 was also homologous to other cuticle proteins, and bears the R & R Consensus. Both BMCP30 and BMCP18 lack the Pro/Ala rich region following R & R Consensus.

The molecular weight of all of the four *B. mori* cuticle proteins determined by mobility in SDS-PAGE was 21 to 96% higher than those calculated from deduced amino acid sequences, in spite of suggestion that BMCPs received no modification such as glycosylation. This is a typical property of cuticle proteins because of possibly binding with SDS (Andersen *et al.*, 1995).

The schematic diagram representing putative primary structure of BMCPs was

illustrated in Fig. 4. Comparison of the primary structures revealed that they are well aligned around R & R Consensus and this conservation is extended 35 more residues to N-termini (represented as "the conserved region" in Fig. 4). According to the consensus sequence preceding the R&R Consensus, BMCPs listed here appear to belong to the RR1 subgroup of cuticle proteins (Andersen, 1998).

Characteristics of cuticle protein binding activity to chitin

All of the purified *B. mori* cuticle proteins showed high binding activity to chitin (data not shown). In addition to the chitin-binding activity, BMCP30 possesses cell-attachment signal (RGD tripeptide) in its primary structure (Fig. 3). Actually, BMCP30 purified from larval cuticle showed the cell-adhesive activity (Nakato, 1993). Therefore, BMCP30 seems to play an important role in stabilization of cuticle-epidermal cells connection. Hence, I examined the molecular mechanism underlying the formation of protein-chitin complex using BMCP30 as a model protein.

First of all to understand the chitin recognition mechanism of cuticle proteins, I analyzed the chitin-binding activity of BMCP30 using a chitin-affinity column chromatography, followed by SDS-PAGE and densitometry. To investigate the effect of pH on BMCP30 binding to chitin, binding activities under different pH conditions were analyzed. As shown in Fig. 5, BMCP30 bound to chitin over relatively wide pH range, and it showed maximal binding activity around pH 6.4 which corresponds to the pH value of silkworm hemolymph.

Binding stability of BMCP30 was compared with that of another cuticle protein, BMCP18. Both of cuticle proteins showed almost same binding activity at pH 6.4 (Fig. 6A) implying that the conserved region, which is shared by these proteins, functions as the chitin binding domain. In contrast to similar binding activities of both proteins in standard binding buffer, these proteins showed different resistance to urea concentration in binding buffer. Although almost all BMCP18 binding to chitin was eluted by 4 M urea, only 60% of BMCP30 was eluted from chitin column (Fig. 6B), suggesting that BMCP30 possesses additional chitin

affinity region in the structure, probably its N-terminal region.

Conserved region among cuticle proteins functions as chitin binding domain

To investigate whether the conserved region among cuticle proteins actually functions as a chitin binding domain, a series of deletion mutants of recombinant BMCP30 was synthesized in *E. coli*. To synthesize the recombinant proteins, I used the Glutathione S-transferase (GST) Gene Fusion System (Pharmacia Biotech), in which the resultant proteins were fused with GST. The various DNA regions of BMCP30 cDNA were prepared using PCR method and inserted into the *E. coli* expression vector pGEX-4T-1. DNA sequencing revealed that pGSTBMCP30ΔCBD carries one nucleotide substitution at nt position 408 of BMCP30 cDNA (adenine to guanine). This mutation is a silent mutation, which cause no amino acid substitutions. Each GST-fused cuticle proteins (Fig. 8A) was synthesized and purified as described in Materials and Methods, and their chitin-binding activity was analyzed.

Prior to analysis with deletion mutants, I established a chitin-binding assay utilizing GST activity of GST-fusion proteins. The GST activity was analyzed as described in Materials and Methods using CDNB and reduced glutathione as substrates. I confirmed that GST activity of GST-fusion protein is in proportion to the amount of the fusion protein in this assay. As shown in Fig. 7, the GST activities of GSTBMCP30 in sample solution (*a* in Fig. 7) and mixture of the protein solution and chitin (*b* in Fig. 7) were same, indicating GST-fusion proteins do not change their own GST activities when they bind to chitin. Therefore, the GST activity in the GST-fusion protein binding to chitin particle corresponds to the amount of GST-fusion proteins binding to chitin. The chitin-binding activity of GSTBMCP30 was calculated as the relative ratio of GST activity in chitin washed out non-binding proteins (*c* in Fig. 7) to that of GSTBMCP30 in sample solution (*a* in Fig. 7). Using this system, chitin-binding activities of recombinant cuticle proteins can be analyzed quickly and quantitatively.

Subsequently, the chitin-binding activities of some BMCP30 deletion mutants were analyzed. For negative control, GST derived from pGEX-4T-1 with no inserts was used. As shown in Fig. 8, GSTBMCP30 showed high chitin-binding activity. GSTBMCP30CBD

bearing only the conserved region also revealed high binding activity of the same level as it was observed using GSTBMCP30. In contrast, GSTBMCP30ΔCBD lacking the conserved region also possessed binding activity, but the level of it was quite low. These results indicated that the conserved region functions as the principal chitin binding domain of BMCP30. In addition, some binding activity of GSTBMCP30ΔCBD suggests that additional chitin-affinity region(s) exist in N-terminal structure. This is consistent with the result of chitin-affinity chromatography comparing BMCP30 and BMCP18 (Fig. 6B).

To confirm that the conserved region also functions as chitin binding domain in other cuticle proteins, experiment with a BMCP18 deletion mutant was done. Binding property of GSTBMCP18 to chitin was well similar to GSTBMCP30 (Fig. 8). GSTBMCP18CBD bearing the conserved region of BMCP18 also showed high binding activity as GSTBMCP18 and GSTBMCP30, suggesting that other cuticle proteins bind to chitin mediated by the conserved region.

Interaction with chitooligosaccharides

To determine that what size of GlcNAc oligomer BMCP30 can recognize, competition analysis using chitooligosaccharides was performed. Binding activities of recombinant BMCP30 (GSTBMCP30) to powdered chitin in presence of 1% GlcNAc, $(GlcNAc)_3$, or $(GlcNAc)_6$ were analyzed, and compared with those in absence of oligosaccharides (Fig. 9). Although GlcNAc and $(GlcNAc)_3$ did not show any effects on the binding of BMCP30, $(GlcNAc)_6$ diminished binding activity of BMCP30 to the chitin. These results suggest that BMCP30 recognizes 4-6 mer of GlcNAc as a unit for binding.

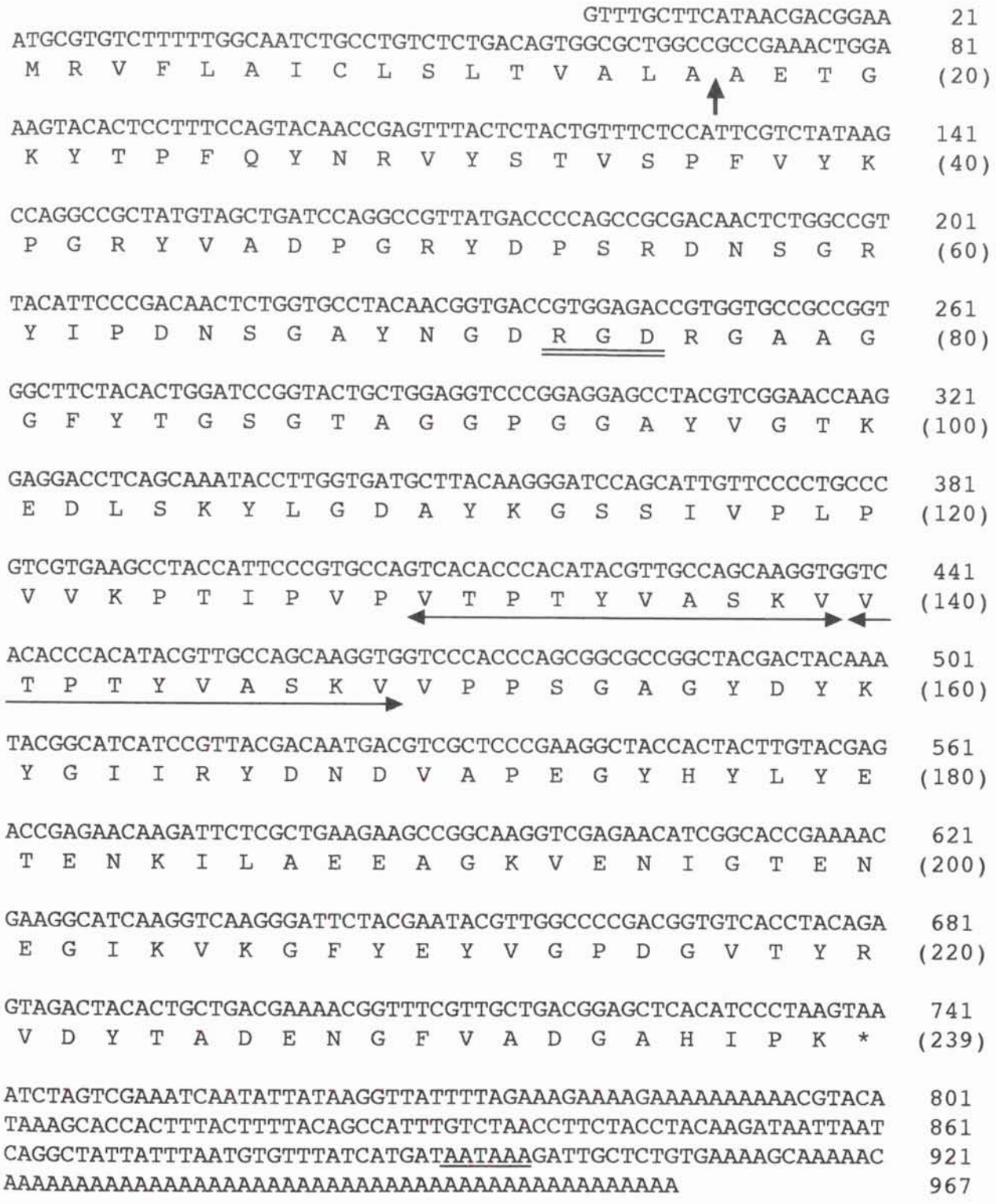

GTTTGCTTCATAACGACGGAA 21
ATGCGTGTCTTTTTGGCAATCTGCCTGTCTCTGACAGTGGCGCTGGCCGCCGAAACTGGA 81
M R V F L A I C L S L T V A L A A E T G (20)
AAGTACACTCCTTTCCAGTACAACCGAGTTTACTCTACTGTTTCTCCATTCGTCTATAAG 141
K Y T P F Q Y N R V Y S T V S P F V Y K (40)
CCAGGCCGCTATGTAGCTGATCCAGGCCGTTATGACCCCAGCCGCGACAACTCTGGCCGT 201
P G R Y V A D P G R Y D P S R D N S G R (60)
TACATTCCCGACAACTCTGGTGCCTACAACGGTGACCGTGGAGACCGTGGTGCCGCCGGT 261
Y I P D N S G A Y N G D R G D R G A A G (80)
GGCTTCTACACTGGATCCGGTACTGCTGGAGGTCCCGGAGGAGCCTACGTCGGAACCAAG 321
G F Y T G S G T A G G P G G A Y V G T K (100)
GAGGACCTCAGCAAATACCTTGGTGATGCTTACAAGGGATCCAGCATTGTTCCCCTGCCC 381
E D L S K Y L G D A Y K G S S I V P L P (120)
GTCGTGAAGCCTACCATTCCCGTGCCAGTCACACCCACATACGTTGCCAGCAAGGTGGTC 441
V V K P T I P V P V T P T Y V A S K V V (140)
ACACCCACATACGTTGCCAGCAAGGTGGTCCCACCCAGCGGCGCCGGCTACGACTACAAA 501
T P T Y V A S K V V P P S G A G Y D Y K (160)
TACGGCATCATCCGTTACGACAATGACGTCGCTCCCGAAGGCTACCACTACTTGTACGAG 561
Y G I I R Y D N D V A P E G Y H Y L Y E (180)
ACCGAGAACAAGATTCTCGCTGAAGAAGCCGGCAAGGTCGAGAACATCGGCACCGAAAAC 621
T E N K I L A E E A G K V E N I G T E N (200)
GAAGGCATCAAGGTCAAGGGATTCTACGAATACGTTGGCCCCGACGGTGTCACCTACAGA 681
E G I K V K G F Y E Y V G P D G V T Y R (220)
GTAGACTACACTGCTGACGAAAACGGTTTCGTTGCTGACGGAGCTCACATCCCTAAGTAA 741
V D Y T A D E N G F V A D G A H I P K * (239)
ATCTAGTCGAAATCAATATTATAAGGTTATTTTAGAAAGAAAAGAAAAAAAAAAACGTACA 801
TAAAGCACCACTTTACTTTTACAGCCATTTGTCTAACCTTCTACCTACAAGATAATTAAT 861
CAGGCTATTATTTAATGTGTTTATCATGATAATAAAGATTGCTCTGTGAAAAGCAAAAAC 921
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 967

Fig. 3. Corrected nucleotide sequence of BMCP30 cDNA and deduced primary structure of BMCP30.

Nucleotides are numbered from the 5' end of the BMCP30 cDNA clone, and amino acid residues are numbered in parentheses from translation initiation methionine residue. The signal peptide cleavage site is indicated by a *vertical arrow*. A cell adhesive RGD sequence and a 10 amino acid repeat are marked with *double underline* and *horizontal arrows*, respectively. A putative polyadenylation signal is *underlined*.

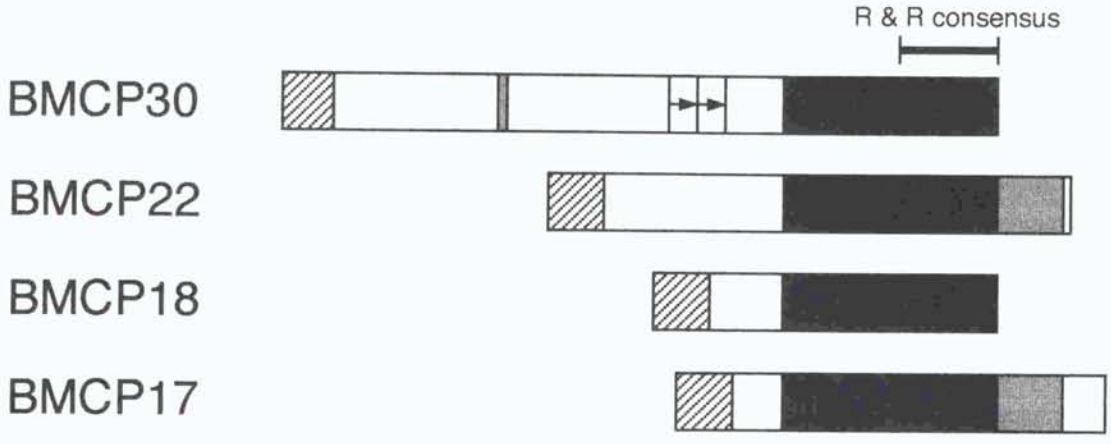


Fig. 4. Schematic representation of primary structures of *B. mori* cuticle proteins predicted from cDNA sequences.

Comparison of putative primary structures of *B. mori* cuticle proteins revealed the conserved region, represented by *solid boxes*, extending beyond the R & R Consensus to N-termini. Another conserved sequence, Pro/Ala rich region indicated by *shadowed boxes*, was found in C-termini of BMCP22 and BMCP17. Unique structures of BMCP30, RGD sequence and 10 amino acid repeat, are indicated by *shadowed box* and *boxes with horizontal arrows*. Signal sequences are *hatched*.

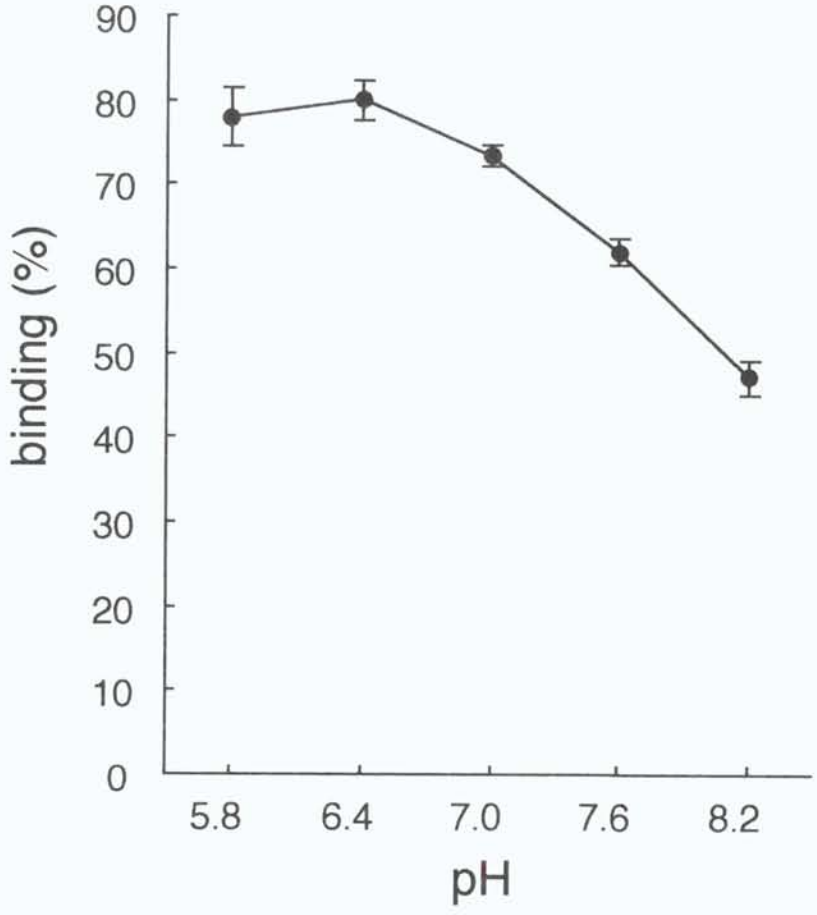
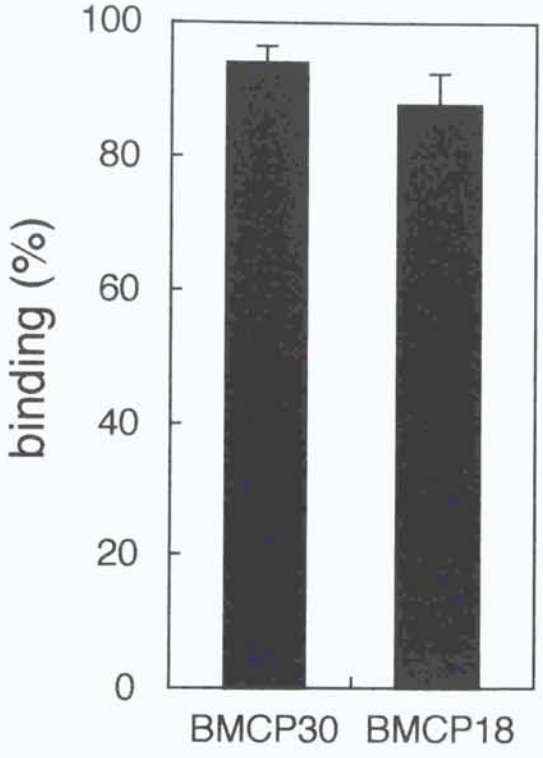


Fig. 5. Effect of pH on chitin binding of BMCP30.

Each 80 μg of purified BMCP30 in 400 μl of the binding buffer with indicated pH was loaded onto a 150 μl chitin column equilibrated with same buffer. After the column was washed well, the binding proteins were eluted with 8 M urea. The binding activity was calculated as described in Materials and Methods, and presented as the mean ± SE of more than three independent experiments.

A


100
80
60
40
20
0
binding (%)
BMCP30
BMCP18

B

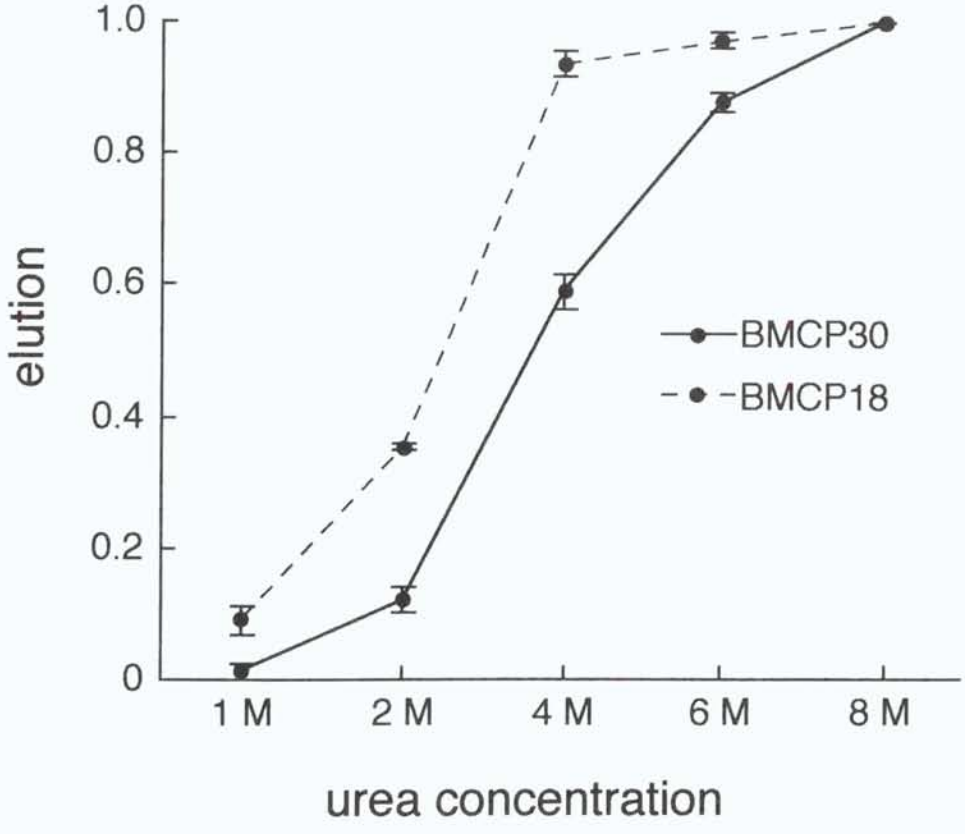
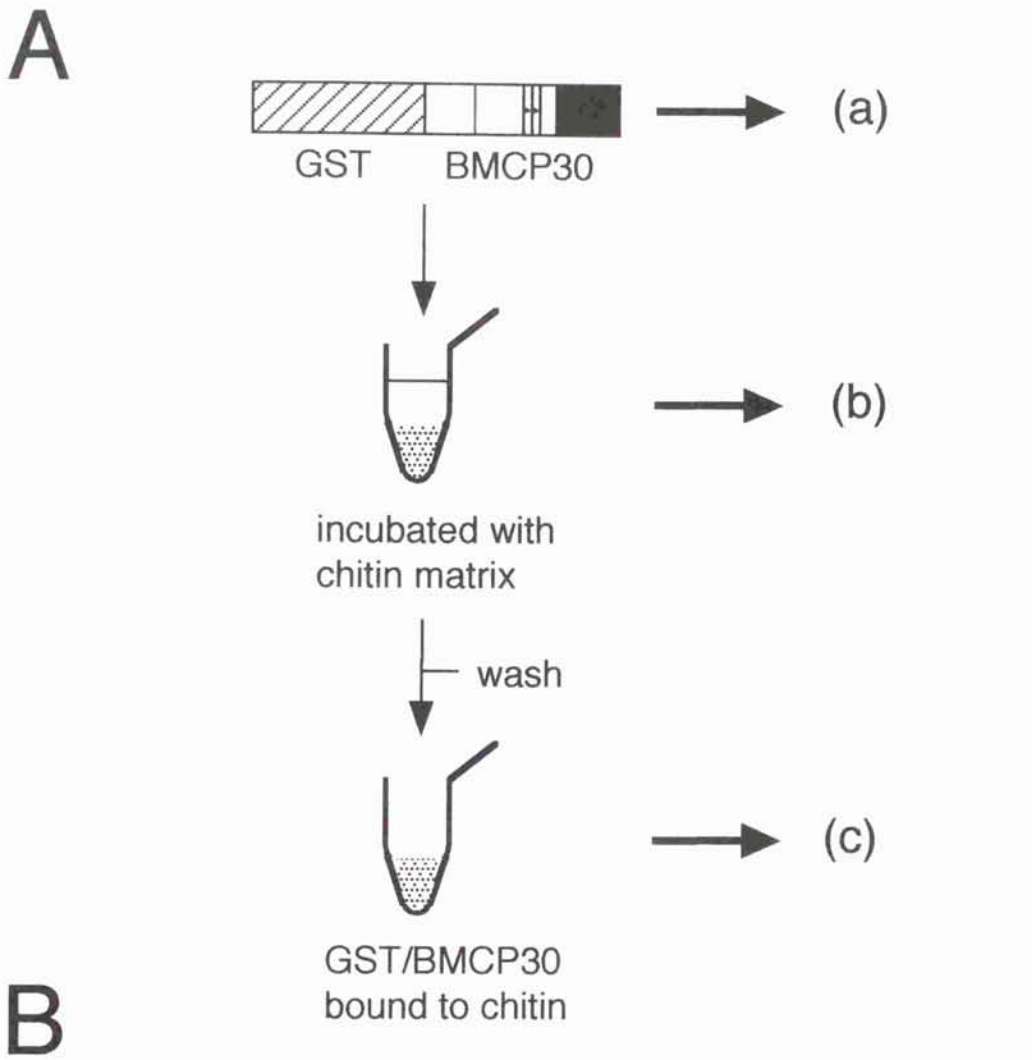

1.0
0.8
0.6
0.4
0.2
0
elution
BMCP30
BMCP18
1 M
2 M
4 M
6 M
8 M
urea concentration

Fig. 6. Comparison of chitin-binding activities of BMCP30 and BMCP18.

(A) Chitin-binding activities of BMCP30 and BMCP18 were analyzed by chitin-affinity chromatography as described in Materials and Methods. *Error bars* represent standard error (SE) of more than three experiments. (B) Resistance of chitin-binding activity to urea. BMCP30 or BMCP18 was loaded onto a chitin column in the same way with other chitin-affinity chromatography, and proteins bound to chitin were eluted by indicated concentration of urea. The elution efficiency was represented as the ratio of protein amount eluted by each concentration of urea to that by 8 M urea elution, with SE of more than three experiments.


A
GST
BMCP30
(a)
(b)
incubated with
chitin matrix
wash
(c)
GST/BMCP30
bound to chitin

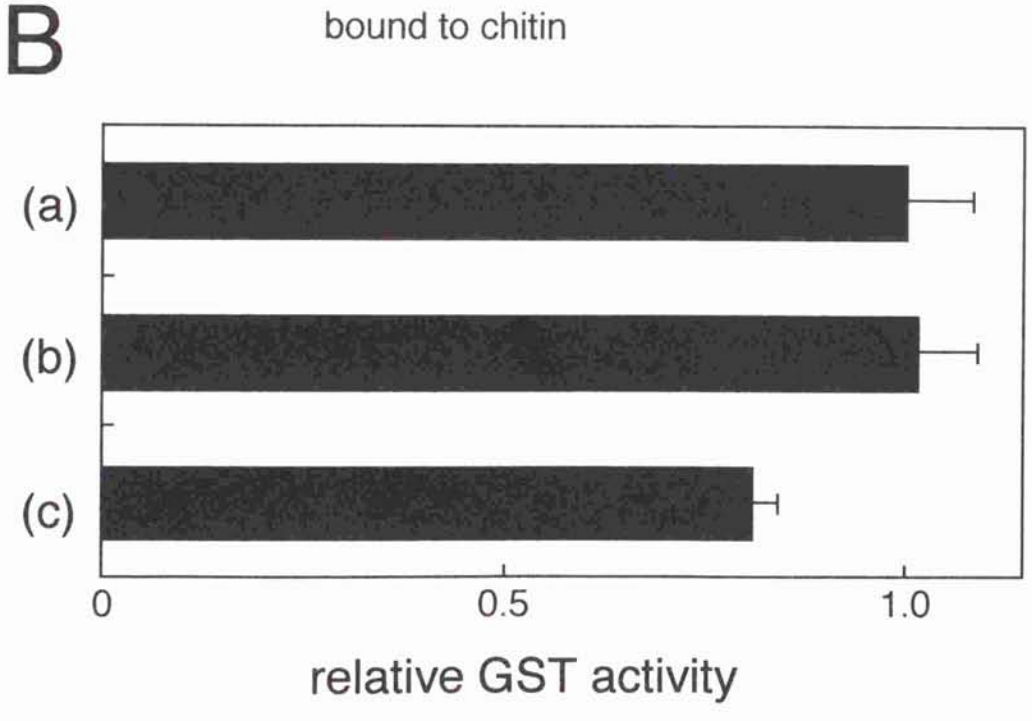

B
(a)
(b)
(c)
0
0.5
1.0
relative GST activity

Fig. 7. Chitin-binding assay of GST-fusion proteins.

(A) Experimental flow of chitin-binding assay utilizing GST activity of GST-fusion protein. Two μg of purified GST-fusion protein (GSTBMCP30) in 1 ml of the binding buffer (fraction *a*) were incubated with 100 μl of chitin slurry equilibrated with the binding buffer (fraction *b*). After incubation, the supernatant was removed and chitin was washed well with the binding buffer (fraction *c*). (B) Each fraction shown in (A) was subjected to GST-assay described in Materials and Methods. The GST activities of GSTBMCP30 sample solution (*a*), non-binding and binding protein to chitin (*b*), and only chitin binding protein (*c*) were presented as the mean ± SE of three independent experiments.

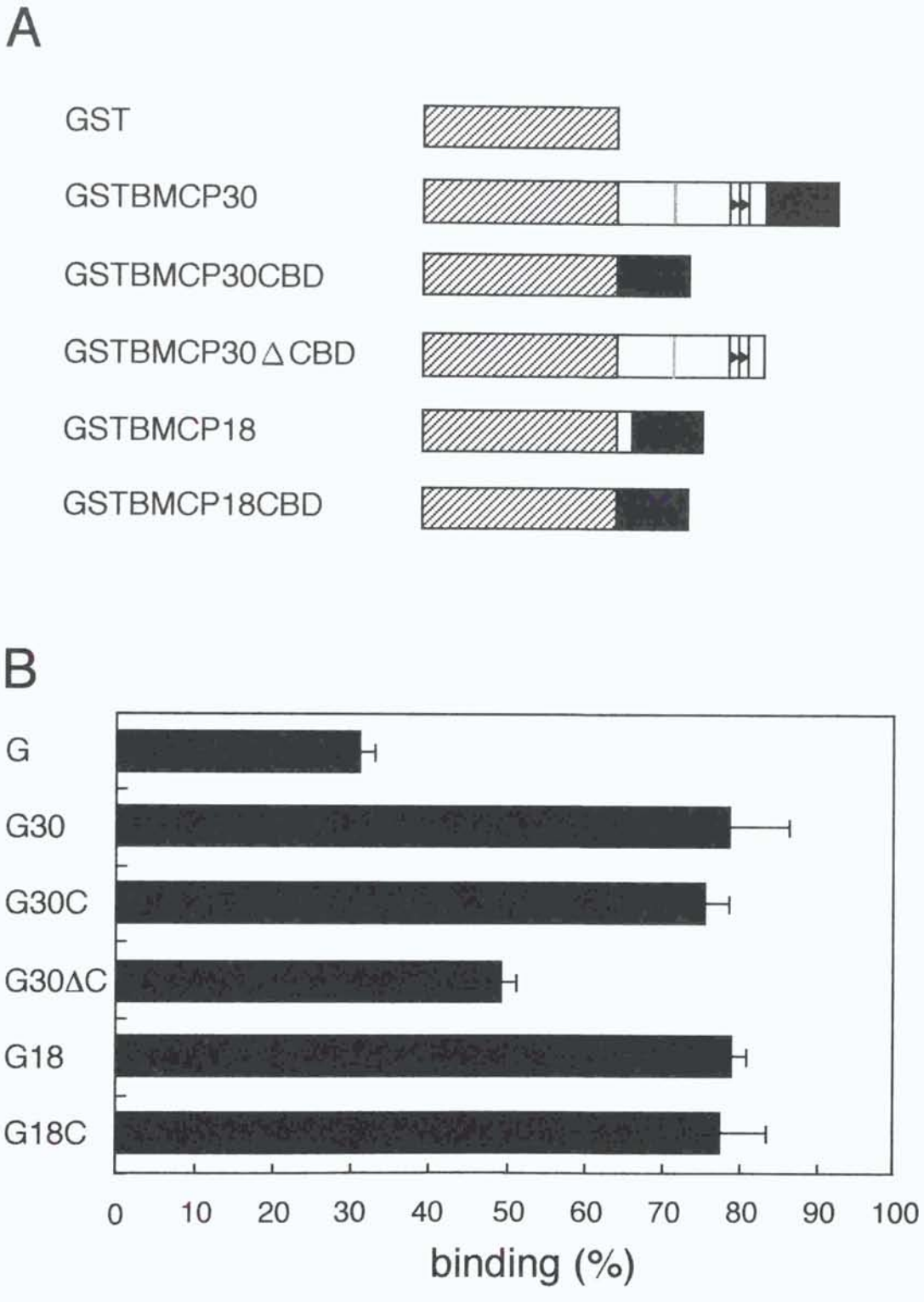

A
GST
GSTBMCP30
GSTBMCP30CBD
GSTBMCP30ΔCBD
GSTBMCP18
GSTBMCP18CBD
B
G
G30
G30C
G30ΔC
G18
G18C
0
10
20
30
40
50
60
70
80
90
100
binding (%)

Fig. 8. Chitin-binding activities of GST-fused cuticle proteins.

(A) Schematic representation of GST/cuticle protein fusion proteins. All GST-fused cuticle proteins were synthesized in *E. coli* BL21 cells, and purified by glutathione-affinity chromatography and FPLC with Mono Q column. The conserved regions of cuticle proteins are represented by *solid boxes*. The RGD sequence and 10 amino acid repeat are indicated by *shadowed bar* and *boxes with vertical arrows*, respectively. GSTs are shown with *hatched boxes*.

(B) Chitin-binding activities of GST/cuticle protein fusion proteins. Chitin-binding activities of GSTBMCP30, G30; GSTBMCP30CBD, G30C; GSTBMCP30ΔCBD, G30ΔC; GSTBMCP18, G18; GSTBMCP18CBD, G18C; and GST, G (as a negative control) were analyzed using their GST activities as described in Materials and Methods. Binding activities were represented as the ratio of GST-activity in protein-bound chitin matrix to that in protein sample solution with SE of three experiments.

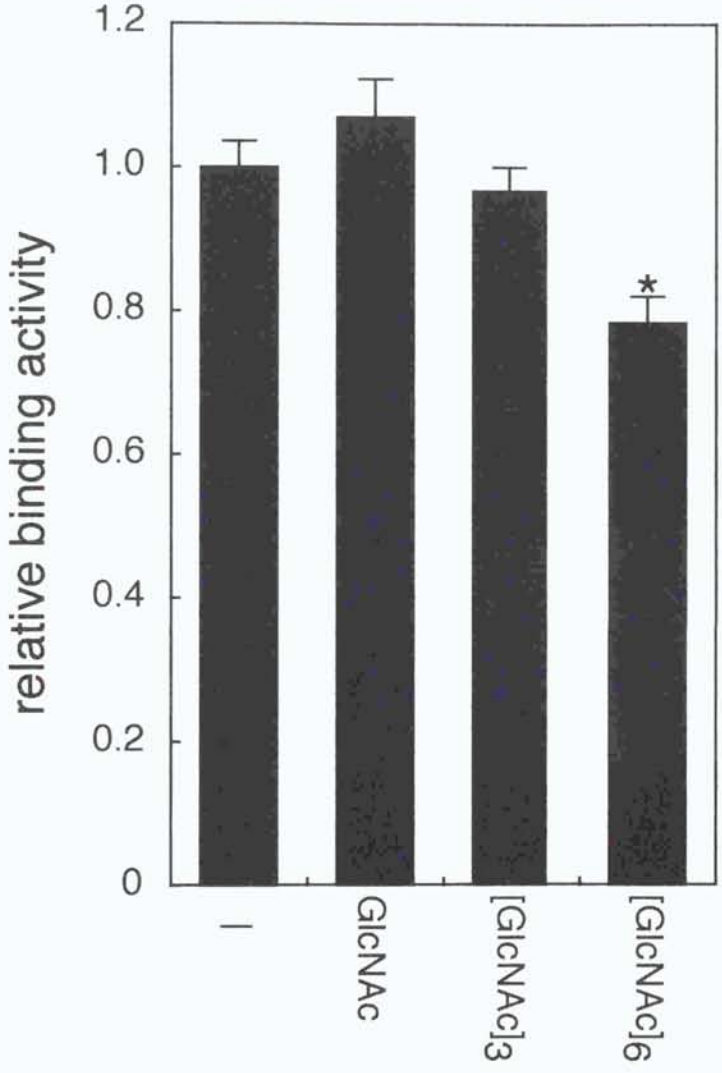

relative binding activity
1.2
1.0
0.8
0.6
0.4
0.2
0
*
–
GlcNAc
[GlcNAc]3
[GlcNAc]6

Fig. 9. Chitooligosaccharide competition in binding to chitin.

For chitooligosaccharide competition experiments, 0.2 μg of GSTBMCP30 was pre-incubated in 100 μl of binding buffer (20 mM NaPB, pH7.0, 0.1% Triton X-100) containing 1% of indicated sugar at room temperature for 15 min. Then, protein solution was mixed with 10 μl of chitin slurry equilibrated with sugar-containing binding buffer, and incubated for further 15 min. After removal of supernatant, chitin was washed well with sugar-containing binding buffer, and GST activity in chitin matrix was analyzed. The binding activities in presence of indicated sugars were represented as relative activity to that in absence of sugars (-) with SE of three experiments. Asterisk indicates $P < 0.05$ compared with no sugar addition.

DISCUSSION

To analyze molecular property of cuticle proteins, I purified four major cuticle proteins, BMCP30, 22, 18, and 17, in native form from larval cuticle of *Bombyx mori*. Molecular weight of each purified BMCP estimated by the mobility on SDS-PAGE is greater than that calculated from deduced amino acid sequence. It is known that BMCPs characterized until now do not receive any modification such as glycosylation. Similar examples were observed in cuticle proteins of other insect species (Andersen *et al.*, 1995). I think that this is a marked property of insect cuticle proteins.

All purified cuticle proteins examined kept the chitin-binding activity. I analyzed fundamental property underlying the formation of protein-chitin complex by use of BMCP30 as a model protein on behalf of BMCPs. BMCP30 showed the highest binding activity to chitin at pH 6.4, which is the hemolymphal pH value of *B. mori* larva (Fig. 5). The binding was competed by the hexamer but not by monomer and trimer of GlcNAc (Fig. 9), suggesting BMCP30 recognizes 4-6 mer of GlcNAc as a unit for binding.

The primary structures of BMCPs predicted from cDNA sequences resemble to each other, possessing the R & R Consensus at C-terminal or central region. Alignment of primary structure of *B. mori* cuticle proteins revealed that the conserved region is extended to N-termini beyond the R & R Consensus (Fig. 4). BMCP22 and BMCP17 possess other conserved motif, Pro/Ala rich region at C-terminal end. This motif occurs just after the R & R Consensus in several cuticle proteins from different insect species including *Manduca sexta* (Rebers and Riddiford, 1988), *Drosophila melanogaster* (Snyder *et al.*, 1982), and *Locusta migratoria* (Talbo *et al.*, 1991). BMCP30 bears unique motifs, the cell adhesive RGD sequence at the middle of N-terminal domain and the repeat of ten amino acid residues roughly located in the center of the primary structure (Fig. 3). The molecular property of BMCP30 which possesses both chitin-binding and cell-adhesive activities suggests that this protein functions likely as a glue binding epidermal cells and chitin filaments and makes cuticle stable. Actually, the *B. mori* cell line, Bm5, adhered to and proliferated on a chitin sheet coated by

BMCP30, but not with BSA in insect medium (Izumi, unpublished data). In vertebrates, collagen, which is an extracellular matrix protein, contributes greatly to the conservation of morphology of individuals. In the case of insect, proteins such as BMCP30 in *B. mori* appear to firmly bind the epidermal cells with chitin, forming a strong exoskeleton and hence conserving the morphology of individuals.

The extended conserved region including the R & R Consensus has been mentioned (Lampe and Willis, 1994; Andersen *et al.*, 1995). Iconomidou *et al.* (1999) predicted that this region of cuticle proteins consists of several β-strands, and proposed this conformation is essential for binding to chitin filaments and defining cuticle's helicoidal architecture. The results of experiments using mutant proteins of BMCP30 and BMCP18 strongly indicated that the conserved region functions as primary chitin binding domain of cuticle proteins. Fig. 10 shows alignment of the conserved region in cuticle proteins of some species from different orders. Based on the alignment, I proposed the following new consensus sequence;

rxrxraoxNGIxxxxxG – linker sequence – GxrxrxxxaGxxxxhxrxADExGrxxxxxxhP

where r, h, a, and o represent aromatic, hydrophobic, acidic residues, and residue with OH-group, respectively, and x stands for any amino acid residues. In this consensus, aromatic and glycine residues are highly conserved and found with correct spacing. Chitin binding domains were identified in plant chitin-binding lectins and chitinase (Chrispeels and Raikhel, 1991; Iseli *et al.*, 1993), and chitinase from bacteria (Watanabe *et al.*, 1994; Hashimoto *et al.*, 2000). Moreover, NMR experiments showed that conserved aromatic residues in the chitin binding domain of plant lectins was involved directly in binding to the ligands (Martins *et al.*, 1996; Weaver and Prestegard, 1998). And the aromatic residues in the chitin binding domain of bacterial chitinases is also known to be important for binding to the substrate (Ikegami *et al.*, 2000). The conserved region of cuticle proteins differs to these chitin binding domains in regard to the primary structure. In the conserved region of cuticle protein, however, aromatic residues were conserved well likely to plant and bacterial chitin binding domains. Therefore, in cuticle proteins, aromatic residues may also take an important role for binding to chitin. On the other hand, tyrosine and phenylalanine residues in cuticle proteins contribute to

quinone tanning in sclerotization of cuticle (Hackman, 1984). The conserved aromatic residues may also contribute to cross-linkage of cuticle proteins with chitin and/or other proteins.

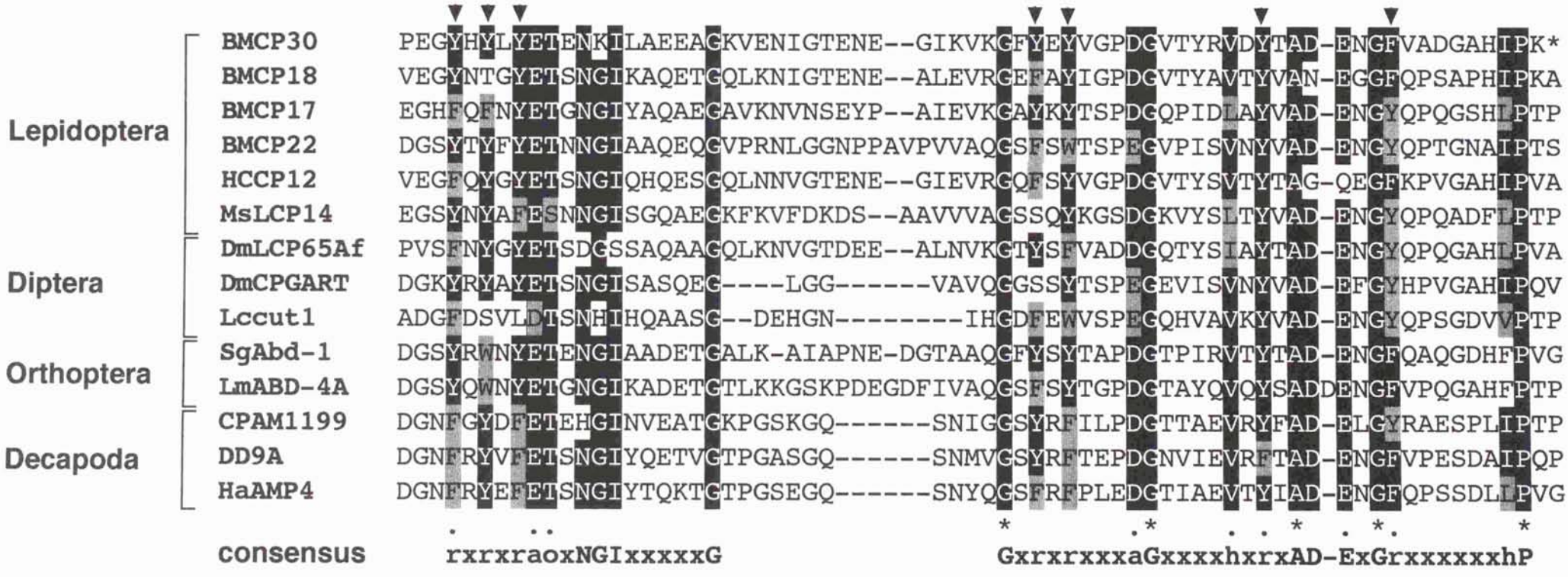

Lepidoptera
Diptera
Orthoptera
Decapoda
BMCP30 PEGYHYLYETENKILAEEAGKVENIGTENE--GIKVKGFYEYYVGPDGVTYRVDYTAD-ENGFVADGAHIPK*
BMCP18 VEGYNTGYETSNGIKAQETGQLKNIGTENE--ALEVRGEFAYIGPDGVTYAVTYVAN-EGGFQPSAPHIPKA
BMCP17 EGHFQFNYETGNGIYAQAEGAVKNVNSEYP--AIEVKGAYKYTSPDGQPIDLAYVAD-ENGYQPQGSHLPTP
BMCP22 DGSYTYFYETNNGIAAQEQGVPRNLGGNPPAVPVVAQGSFSWTSPEGVPISVNYVAD-ENGYQPTGNAIPTS
HCCP12 VEGFQYGYETSNGIQHQESGQLNNVGTENE--GIEVRGQFSYVGPDGVTYSVTYTAG-QEGFKPVGAHIPVA
MsLCP14 EGSYNYAFESNNGISGQAEGKFKVFDKDS--AAVVVAGSSQYKGSDGKVYSLTYVAD-ENGYQPQADFLPTP
DmLCP65Af PVSFNYGYETSDGSSAQAAGQLKNVGTDEE--ALNVKGTYSFVADDGQTYSIAYTAD-ENGYQPQGAHLPVA
DmCPGART DGKYRYAYETSNGISASQEG----LGG------VAVQGGSSYTSPEGEVISVNYVAD-EFGYHPVGAHIPQV
Lccut1 ADGFDSVLDTSNHIHQAASG--DEHGN--------IHGDFEWVSPEGQHVAVKYVAD-ENGYQPSGDVVPTP
SgAbd-1 DGSYRWNYETENGIAADETGALK-AIAPNE-DGTAAQGFYSYTAPDGTPIRVTYTAD-ENGFQAQGDHFPVG
LmABD-4A DGSYQWNYETGNGIKADETGTLKKGSKPDEGDFIVAQGSFSYTGPDGTAYQVQYSADDENGFVPQGAHFPTP
CPAM1199 DGNFGYDFETEHGINVEATGKPGSKGQ------SNIGGSYRFILPDGTTAEVRYFAD-ELGYRAESPLIPTP
DD9A DGNFRYVFETSNGIYQETVGTPGASGQ------SNMVGSYRFTEPDGNVIEVRFTAD-ENGFVPESDAIPQP
HaAMP4 DGNFRYEFETSNGIYTQKTGTPGSEGQ------SNYQGSFRFPLEDGTIAEVTYIAD-ENGFQPSSDLLPVG
consensus rxrxraoxNGIxxxxxG Gxrxrxxxa Gxxxxhxrx AD-ExGrxxxxxxhP

Fig. 10. Alignment of the conserved region of several cuticle proteins.

Deduced or sequenced amino acid sequences in the conserved region of BMCP30, 18, 17, and 22 from *B. mori*, HCCP12 (U08026) from *Hyalophora cecropia*, MsLCP14 (X13279) from *Manduca sexta*, DmLCP65Af (U84752) and DmCPGART (X06286) from *Drosophila melanogaster*, Lccut1 (AF026265) from *Lucilia cuprina*, SgAbd-1 (pir:S78091) from *Schistocerca gregaria*, LmABD-4A (swissprot:P21799) from *Locusta migratoria*, CPAM1199 (swissprot:P81577) from *Cancer pagurus*, DD9A (AB031223) from *Penaeus japonicus*, and HaAMP4 (swissprot:P81388) from *Homarus americanus* were aligned using the ClustalW Multiple Sequence Alignment Program (Thompson *et al.*, 1994). Gaps (-) were introduced to obtain optimal alignment. The identical residues were shown in *inverted type*, and functionally conserved residues were indicated with *shadowed background*. Functional classification of amino acid residues was according to Andersen *et al.* (1995), that is valine, leucine, and isoleucine as the hydrophobic residues; tyrosine, phenylalanine, and tryptophan as the aromatic residues; aspartic acid and glutamic acid as the acidic residues; lysine and arginine as the basic residues; asparagine and glutamine as the amides; serine and threonine as the residues with an OH-group. The consensus sequence deduced by this alignment was shown on the bottom. *r*, *h*, *a*, and *o* represent aromatic, hydrophobic, acidic residues, and residue with OH-group, respectively. The conserved aromatic residues were indicated with *arrow heads*.

REFERENCE

Andersen, S. O. (1998). Amino acid sequence studies on endocuticular proteins from the desert locust, *Schistocerca gregaria*. *Insect Biochem. Mol. Biol.* **28,** 421-434.

Andersen, S. O. (2000). Studies on proteins in post-ecdysial nymphal cuticle of locust, *Locusta migratoria*, and cockroach, *Blaberus craniifer*. *Insect Biochem. Mol. Biol.* **30,** 569-577.

Andersen, S. O. and Højrup, P. (1987). Extractable Proteins from abdominal cuticle of sexually mature locusts, *Locusta migratoria*. *Insect Biochem.* **17,** 45-51.

Andersen, S. O., Højrup, P., and Roepstorff, P. (1995). Insect cuticular proteins. *Insect Biochem. Mol. Biol.* **25,** 153-176.

Binger, L. C. and Willis, J. H. (1994). Identification of the cDNA, gene and promoter for a major protein from flexible cuticles of the giant silkmoth *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* **24,** 989-1000.

Charles, J. P., Chihara, C., Nejad, S., and Riddiford, L. M. (1997). A cluster of cuticle protein genes of *Drosophila* melanogaster at 65A: sequence, structure and evolution [published erratum appears in *Genetics* 1998 ,**148**, 921]. *Genetics* **147,** 1213-1224.

Chihara, C. J., Silvert, D. J., and Fristrom, J. W. (1982). The cuticle proteins of *Drosophila melanogaster*: stage specificity. *Dev. Biol.* **89,** 379-388.

Chrispeels, M. J. and Raikhel, N. V. (1991). Lectins, lectin genes, and their role in plant defense. *Plant Cell* **3,** 1-9.

Cox, D. L. and Willis, J. H. (1985). The cuticular proteins of *Hyalophora cecropia* from different anatomical regions and metamorphic stages. *Insect Biochem.* **15,** 349-362.

Dotson, E. M., Cornel, A. J., Willis, J. H., and Collins, F. H. (1998). A family of pupal-specific cuticular protein genes in the mosquito *Anopheles gambiae*. *Insect Biochem. Mol. Biol.* **28,** 459-472.

Hackman, R. H. (1984). Cuticle: Biochemistry. *In* "Biology of the Integument, Vol. 1, Invertebrates" (Bereiter-Hahn, J., Matoltsy, A. G., and Richards, K. S., Eds.). Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

Hashimoto, M., Ikegami, T., Seino, S., Ohuchi, N., Fukada, H., Sugiyama, J., Shirakawa, M., and Watanabe, T. (2000). Expression and characterization of the chitin-binding domain of chitinase A1 from *Bacillus circulans* WL-12. *J. Bacteriol.* **182,** 3045-3054.

Iconomidou, V. A., Willis, J. H., and Hamodrakas, S. J. (1999). Is β-pleated sheet the

molecular conformation which dictates formation of helicoidal cuticle? *Insect Biochem. Mol. Biol.* **29,** 285-292.

Ikegami, T., Okada, T., Hashimoto, M., Seino, S., Watanabe, T., and Shirakawa, M. (2000). Solution structure of the chitin-binding domain of *Bacillus circulans* WL-12 chitinase A1. *J. Biol. Chem.* **275,** 13654-13661.

Iseli, B., Boller, T., and Neuhaus, J. M. (1993). The N-terminal cysteine-rich domain of tobacco class I chitinase is essential for chitin binding but not for catalytic or antifungal activity. *Plant Physiol.* **103,** 221-226.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680-685.

Lampe, D. J. and Willis, J. H. (1994). Characterization of a cDNA and gene encoding a cuticular protein from rigid cuticles of the giant silkmoth, *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* **24,** 419-435.

Martins, J. C., Maes, D., Loris, R., Pepermans, H. A., Wyns, L., Willem, R., and Verheyden, P. (1996). 1H NMR study of the solution structure of Ac-AMP2, a sugar binding antimicrobial protein isolated from *Amaranthus caudatus*. *J. Mol. Biol.* **258,** 322-333.

Mathelin, J., Bouhin, H., Quennedey, B., Courrent, A., and Delachambre, J. (1995). Identification, sequence and mRNA expression pattern during metamorphosis of a cDNA encoding a glycine-rich cuticular protein in *Tenebrio molitor*. *Gene* **156,** 259-264.

Nakato, H. (1993). Structure and expression of cuticle protein genes in silkworm, *Bombyx mori*. Doctor Sci thesis, Tokyo Metropolitan University.

Nakato, H., Shofuda, K., Izumi, S., and Tomino, S. (1994). Structure and developmental expression of a larval cuticle protein gene of the silkworm, *Bombyx mori*. *Biochim. Biophys. Acta* **1218,** 64-74.

Nakato, H., Takekoshi, M., Togawa, T., Izumi, S., and Tomino, S. (1997). Purification and cDNA cloning of evolutionally conserved larval cuticle proteins of the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **27,** 701-709.

Nakato, H., Toriyama, M., Izumi, S., and Tomino, S. (1990). Structure and expression of mRNA for a pupal cuticle protein of the silkworm, *Bombyx mori*. *Insect Biochem.* **20,** 667-678.

Norup, T., Berg, T., Stenholm, H., Andersen, S. O., and Højrup, P. (1996). Purification and characterization of five cuticular proteins from the spider *Araneus diadematus*. *Insect*

Biochem. Mol. Biol. **26,** 907-915.

Nousiainen, M., Rafn, K., Skou, L., Roepstorff, P., and Andersen, S. O. (1998). Characterization of exoskeletal proteins from the American lobster, *Homarus americanus*. *Comp. Biochem. Physiol. B* **119,** 189-199.

Pierschbacher, M. D. and Ruoslahti, E. (1984). Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **309,** 30-33.

Rebers, J. E. and Riddiford, L. M. (1988). Structure and expression of a *Manduca sexta* larval cuticle gene homologous to *Drosophila* cuticle genes. *J. Mol. Biol.* **203,** 411-423.

Sanger, F. (1981). Determination of nucleotide sequences in DNA. *Science* **214,** 1205-1210.

Shofuda, K., Togawa, T., Nakato, H., Tomino, S., and Izumi, S. (1999). Molecular cloning and characterization of a cDNA encoding a larval cuticle protein of *Bombyx mori*. *Comp. Biochem. Physiol. B* **122,** 105-109.

Snyder, M., Hunkapiller, M., Yuen, D., Silvert, D., Fristrom, J., and Davidson, N. (1982). Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. *Cell* **29,** 1027-1040.

Talbo, G., Højrup, P., Rahbek-Nielsen, H., Andersen, S. O., and Roepstorff, P. (1991). Determination of the covalent structure of an N- and C-terminally blocked glycoprotein from endocuticle of *Locusta migratoria*. Combined use of plasma desorption mass spectrometry and Edman degradation to study post-translationally modified proteins. *Eur. J. Biochem.* **195,** 495-504.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673-4680.

Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S., and Tanaka, H. (1994). The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *J. Bacteriol.* **176,** 4465-4472.

Weaver, J. L. and Prestegard, J. H. (1998). Nuclear magnetic resonance structural and ligand binding studies of BLBC, a two-domain fragment of barley lectin. *Biochemistry* **37,** 116-28.

Yamada, K. M. (1991). Adhesive recognition sequences. *J. Biol. Chem.* **266,** 12809-12812.

Section II

Expression Control of Cuticle Protein Gene in *Bombyx mori*

INTRODUCTION

Juvenile hormone (JH) is an insect sesquiterpenoid hormone, which has pleiotropic actions including the two major roles, the regulation of metamorphosis and reproductive maturation. In post-embryonic development, JH modifies the effect of the molting hormone, ecdysone. In the presence of JH, a rise of hemolymphal ecdysone titer leads to larval molt, and epidermal cells synthesize larval cuticle. In contrast, ecdysone without JH causes metamorphosis. It is known that JH regulates expression of several genes involved in both developmental and reproductive processes (Zhang *et al.*, 1993; Nakato *et al.*, 1994; Hirai *et al.*, 1998; Comas *et al.*, 1999; Dubrovsky *et al.*, 2000). As JH is a small lipophilic molecule, it is thought to act with mediation of some nuclear receptors on gene expression like as vertebrate steroid and thyroid hormones. So far, however, the modes of JH action are still unknown in contrast to ecdysone, nevertheless enormous efforts have been made to clarify them. 20-hydroxyecdysone (20HE), the active metabolite of ecdysone, is a steroid hormone, and whose nuclear receptor has been identified and lower cascade has been revealed in detail (Koelle *et al.*, 1991; Yao *et al.*, 1992; Thummel, 1995; Lam *et al.*, 1997). Although some nuclear proteins which interact with JH were reported from different insect species (Palli *et al.*, 1994; Jones and Sharp, 1997), the functional JH receptor has not been identified yet.

BMCP18 is one of the major cuticle proteins in larval cuticle of silkworm, *Bombyx mori*. Biosynthesis of this protein in epidermal cells is regulated in stage-dependent fashion at the mRNA level (Shofuda *et al.*, 1999). BMCP18 mRNA is abundant in epidermal cells of larvae throughout feeding stage, and the mRNA level declines around four days after the fourth molt possibly correlated with JH disappearance from hemolymph. A topical application of JH analog, methoprene, to the early stage of fifth instar larvae sustains the expression of BMCP18 gene. This indicates that expression of BMCP18 gene is up-regulated by JH, and the regulation mechanisms of BMCP18 gene expression provide a model system for investigation of JH action mode.

In this section, I analyzed the structure of BMCP18 gene and compared with those of

other lepidopteran cuticle proteins (MSCP14.6 from *Manduca sexta* and HCCP12 from *Hyalophora cecropia*) which are highly structurally related with BMCP18. I showed that these cuticle protein genes have a common 12-bp identical sequence at the 5' upstream region, which matches to the recognition sequence of vertebrate transcription factors, COUP-TF and HNF-4. I also discuss the possible participation of Bmsvp, a *Bombyx* homolog of vertebrate COUP-TF and *Drosophila* Seven-up, in the expression control and JH responsiveness of BMCP18 gene.

MATERIALS AND METHODS

Animals

A commercial strain of the silkworm, *Bombyx mori* (*Kinshu* × *Showa*) was obtained from Ueda Sanshu, Ueda. Larvae were reared at 25°C on an artificial diet (Katakura Kogyo Corp.).

Chemicals and biological materials

The BMCP18 genomic clone, pBMCP18G1Hd1, was kindly supplied by Dr. Hiroshi Nakato.

[α-^{32}P] dCTP (specific activity; 3 kCi/mmol) and [γ-^{32}P] ATP (specific activity; 3 kCi/mmol) were purchased from the Hungarian Academy of Sciences. T4 polynucleotide kinase was obtained from Nippon Gene Corp. Rous associated virus 2 (RAV-2) reverse transcriptase, restriction endonucleases, and other DNA modifying enzymes were from Takara Shuzo Corp. and New England Biolabs, Inc. Nylon membrane, oligo(dT)$_{12\text{-}18}$, and poly(dI)poly(dC) were obtained from Amersham Pharmacia Biotech. Synthetic oligonucleotides were from Sawady Technology.

Other chemicals and enzymes were obtained from commercial sources.

DNA sequencing

Nucleotide sequences were determined by dideoxynucleotide chain termination method (Sanger, 1981) using ABI PRISM BigDye Terminator Cycle Sequencing Kit and ABI PRISM 310 Genetic Analyzer (PE Biosystems). To generate sequencing ladders of BMCP18 gene corresponding to the primer extension products, dideoxynucleotide chain termination experiment was performed using *Bca*BEST Dideoxy Sequencing Kit (TaKaRa), BMCP18 genomic clone (pBMCP18G1Hd1) as a template, and the 25 base synthetic primer used in primer extension analysis.

Primer extension analysis

The transcription initiation site was determined by primer extension analysis with a 25 base synthetic primer (5'- CGATTTCATTTTGATTGATTGTTTG-3', nt position +71 to +47 of the BMCP18 gene). The primer was radiolabeled at the 5' end with T4 polynucleotide kinase and [γ-^{32}P] ATP. The 5'-labeled primer was hybridized with 5 μg of total RNA isolated from the integument of fifth instar day-2 larvae, and cDNA was synthesized by RAV-2 reverse transcriptase (Ausubel *et al.*, 1987). The resultant cDNA was electrophoresed on 8% acrylamide / 7M urea gel and detected by autoradiography.

Southern blotting

For genomic Southern blot analysis, 20 μg of DNA isolated from silk glands was digested with restriction enzymes and electrophoresed in 0.7% agarose gel. To hydrolyze high molecular weight DNAs, the gel was submerged in 0.25 N HCl. After wash, the gel was soaked in a denaturing solution (1.5 N NaOH, 0.5 M NaCl) for 30 min. DNA fragments were transferred onto a nylon membrane with an alkaline solution (0.4 N NaOH) and hybridized with a radiolabeled DNA probe at 65°C (Sambrook *et al.*, 1989). After hybridization, the membrane was washed and subjected to autoradiography.

Total RNA isolation

The dorsal integument was dissected from larvae, and the fat body, tracheae, Malpighian tubules, and muscle were carefully removed from the integument of the first abdominal segment (A1) in cold PBS under microscopic observation, then the integument of A1 was cut off. Several pieces of the integument were homogenized in 400 μl of solution D (4 M guanidine thiocyanate, 25 mM sodium citrate, pH 7.0, 0.5% sarcosyl, and 0.1 M 2-mercaptoethanol). Subsequent procedures followed the method of Chomczynski and Sacchi (1987). If total RNA was not prepared by this way, I adopted the method for RNA extraction from adults of fruit flies (Ashburner, 1989).

RT-PCR analysis

Total RNA (0.5 μg) was reverse-transcribed by RAV-2 reverse transcriptase using an oligo(dT)$_{12-18}$ primer in 10 μl of reaction solution. PCR experiments were performed using 0.5 μl aliquot of 1st stranded cDNA solution (derived from 25 ng of total RNA) as a template, and primer sets following: BmsvpE5'FW-E (5'- GGCGAATTCTCGCTGCTACTGCGGGC-3') and BmsvpE3'RV-E (5'- GGCGAATTCCAGCATGTCGCGGATTAG-3') for Bmsvp detection; BmHNF-4E5'FW-E (5'-GGCGAATTCAGCGACGCCGAAATAAACA-3') and BmHNF-4E3'RV-E (5'-GGCGAATTCCCCGCCGAGCAGCATCT-3') for BmHNF-4 detection; BMCP18RV93-E (5'- GCCGAATTCGTCGGCCCTGATGCCGAC-3') and BMCP18FW414-X (5'- GCCCTCGAGACTTTTCAAATGAATGAG-3') for BMCP18 detection; 30kc65'FW-E (5'- GGCGAATTCACAGTGTTGTGACTGCTT-3') and 30kc63'RV-E (5'- GGCGAATTCTCGTTTTCAGCTTCAGCT-3') for detection of one isoform of 30K protein mRNAs (30Kc6, Sakai *et al.*, 1988).

***In vitro* transcription and translation**

Bmsvp-α cDNA, pBmsvpα, was kindly provided by Dr. Yaginuma, Nagoya University. The translation initiation consensus, AAAAATG (Cavener, 1987), was introduced at the upstream region of Bmsvp cDNA by PCR using pBmsvpα as a template and primers of BmsvpATG-E (5'-GCCGAATTCAAAAATGCGTCGCGCGTCTTCTCGC-3') and M13 forward primer. PCR products were digested by *Eco*RI, and inserted at the *Eco*RI site of the pBluescript KS$^+$ under the control of T7 promoter. The resulting plasmid was termed pBmsvpATG. BmHNF-4b cDNA clone, subcloned in pBluescript SK$^+$, was a gift from Dr. Iatrou (Swevers and Iatrou, 1998). These constructs were expressed *in vitro* using the TNT® T7 Quick Coupled Transcription/Translation System (Promega) according to the manufacturer's instruction.

Gel retardation assay

Gel retardation assay was conducted as described (Kapitskaya *et al.*, 1998) with some

alterations. Two μl of *in vitro* translation reactions was added to 10 μl of reaction buffer (10 mM TrisHCl, pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 4% glycerol, 0.5 mM DTT, 0.5 mM EDTA, 0.25 μg poly(dI)poly(dC)) and incubated at 25°C for 15 min. Subsequently, 25 fmol of 5'-labeled DNA probe was added, and the reaction solution was incubated for 15 more min at 25°C. For competition analysis, 2.5 pmol (100-fold molar excess) of competitor DNA was added to the reaction. The reaction mixture was electrophoresed in 5% nondenaturing acrylamide gel at room temperature, and the gel was subjected to autoradiography.

RESULTS

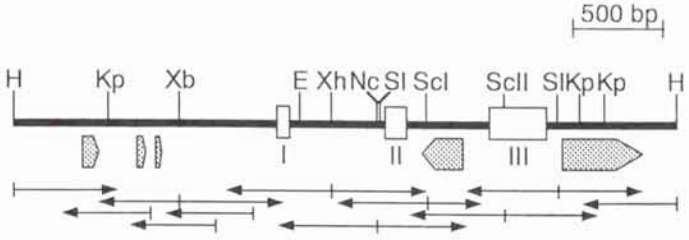
Structure of the BMCP18 gene

Structure of the BMCP18 genomic clone, pBMCP18G1Hd1, bearing the 3.7 kb *Hin*dIII fragment of genomic sequence was analyzed. Figure 1 shows the restriction map and the complete nucleotide sequence of pBMCP18G1Hd1 together with the deduced amino acid sequence of BMCP18. A comparison of the genomic sequence with cDNA showed the BMCP18 gene to consist of three exons interspersed by two introns.

The transcription initiation site of the BMCP18 gene was determined by primer extension analysis. A 25 base synthetic primer complementary to the first exon of the BMCP18 gene was hybridized with total epidermal RNA prepared from the integument of fifth instar day-2 larvae (V2), the primer was extended with reverse transcriptase, and the synthesized DNA fragment was analyzed by gel electrophoresis. Dideoxy-sequencing ladders extended from the same primer were electrophoresed in parallel with the synthesized cDNA. Almost all primer DNA was elongated 46 nucleotides and terminated at the thymine residue (Fig. 2, arrowhead), therefore I determined the corresponding adenine residue as the transcription initiation site of the BMCP18 gene (+1 in Fig. 1B). This is in agreement with the report that cap sites of most eukaryotic mRNAs are adenine (Busslinger *et al.*, 1980). The sequence ATCAGTC (underlining indicates the +1 site) in the BMCP18 mRNA matches well with the consensus sequence ATCA$^{G}/_{T}$T$^{C}/_{T}$ at the cap site of insect mRNAs (Hultmark *et al.*, 1986).

A typical TATA box (TATAAAA) occurs 30 bp upstream from the transcription initiation site (Fig. 1). Structures around splice sites agree with the consensus sequence for the general splice junction. Sequences closely similar to the Bm1 element, a highly reiterated retroposon-like element dispersed throughout the *B. mori* genome (Adams *et al.*, 1986), exist in the second intron and the 3' downstream region in mutually inverted orientation. In addition, three truncated sequences of the Bm1 element were also identified in the 5' upstream region of the BMCP18 gene.

A



B

```
-1444 AAGCTTTTGCTGCCTATATTAACCCTCTTTATTATCAGACGATTCGGGAATACTATCCGATTTTCCATTCGTAATACCTAGTTTATTTATTTAAAAAATC
-1344 TGTCAACAAAGCGAAAACCAATATCGTCTGTTCACCTTTAGACTTAATAGAAGTATTGAACGCGATTCCGATCCGGTAGTAGATTCATTCGCAAAGCAGC
-1244 TGCTCTTGAGTTGTTAGGGTCTTCTTCGGAGGCGCTCGGGCAGCTGTTAGCAAATCCCACCCTTTGAGCCTTTGCTCGTCCACCTGTCCTGATAAAACTG
-1144 GAAAGGCCTCCGGGCTACCAGTAATCAGCTTGGTAATCAGTTTGGTAGGCAGCGGCTTGGCTCTGCCCCTGGCATTGCTGACGTCCATGAGCGACGGTGA
-1044 CCACTTACCATCAGGTGGGCCGTATGCTCGTCTGCCTACAAAGGCAATAAAAAAAAAAATAAAAAAAAAATCCCTCATTCATTTTAATTAACGGGTAAAAC
 -944 ACCCGTTAGCAGGTACCCTTATTTTTCAGTTTTGTACACTAAGTTCCTATGTAACAAGTGCACCTAAGACTATAATCCCCATAAGCCCTACGTAGGAACT
 -844 CTACGTAACACGCATTTACATCTAAGTCCAAAATAAACTCATCATATCTGCTCGTACGATGTACCATGAATACGAAAGGGTAGTACTACCGCTCTGCCTA
 -744 TTTCTGCCGCGAAGCAGTAATGCGGTAGACGGCTCCCACACTTCGAACTGGAGAAAGCCGTTATAATAATTCACGTTGCTACATCTATGGACTTCGGTTA
 -644 CCACTTAACACTAGGTGCACCGTGAACTATTTCTGTTGGAATAGTGAACTATCCGAAAATTACTGTGTAAATCTATAACTTACTCGTTATATACATAGTT
 -544 CACTCATTGTCTAGAAAGTTAATGGAAGGAAGTCTATCGAATTTTTAAAGAAAGCACTTGGTTTGTCGTTTTTAAATCATATTAAAATTATTATATAATA
 -444 AAATAATATTAAGCTCTTTTATTATAATTTATTTTAAGGACAAAACCAGATTTAGCAAATAGTTTGAAAGCAAAAGTAACTCATTCTTAAAATGCTGAAT
 -344 CGAAAATCGATTCCACTGACATATTATGTTACAATTCATCTAAGTTTAAACAGAGTTATTTTTATAGTAAATCGAGTTTTAATAAACACGTGCACACATA
 -244 TCGAAAAACACGTAGCTTATGTCGCTGTGCATGAATAACTAATGCTCTCAAAACTCAGCTGTGACGTGTGAGAATCTCAGGGAAATATTTTAAGTTTATA
 -144 ATCTTGTAATTTATTACTTCGAAAAGCGCGCGAATCCAACTCAGGCTCAGTGGACATCTTTGAAGTTCAAATGTCAAAATGCACGTCCTCGTCTGCGGCG
  -44 ATGGCCCGACCGAGTATAAAAACCGCTGCGCGCATCTTCGTAGC +1 ATCAGTCAATCGTTCACTGTCCTCAAACAACAGTCTTAGGACAAAACAAACAATC
  +56 AATCAAAATGAAATCG gtacgtgataactgtctacttttttattactattccaaattattattttatgtgaaatgtgaattcataattaacaccagtata
                M K S
 +155 atttgctaaatgtattcaaatgtattgtctgaccattgatccgagtaagccaacattggaaaatagttctaactcatcggcattagctatccttcacacg
 +255 taattcagagttgccaatctatatatatataaatgaattgctgttcgttagtctcgttaaaactcgagagcggctggaccgatttggctataattagcca
 +355 taaaatggtcttgaattattcgtgaaaaatatgaaaatgctcggaattacataaaaacaaacaatttgtttttccttgatgtgtccccgtcggacgga
 +455 ttcctttttgttttttttaaagtttatttatacaaaagtttaggtcttttatttatcgattgaggcattcgaagtctgccgggtcagctagtctgatat
 +555 aaaacaagtgagtccatggatatgtcgactcggtcctgattgttcgatttttttctag TTCTTCGTTGTCGCTCTCTTCGTCGCCGTAGCTGTCGCTG
                                                                F F V V A L F V A V A V A
 +654 CCCCAGTCGGCCCTGATGCCGACGCTGTCGTCATCAGACAAGACAACGACAACATTGGTGTCGAAGGATACAACACCGG gtaagatttctttaaataaa
      A P V G P D A D A V V I R Q D N D N I G V E G Y N T G
 +753 atgatcactaatacaatgtttgaattactgtagcgttcgttctctgggtacagtttaggcgcacatttttttattgcttagatgggtggacgagctcaca
 +853 gcccaccggtgttaagtggttactggagcccatagacacctacgacgttaatgcgccacccactttgagatataagttctaaggtctcaagtataattac
 +953 aacggctgccccacccttcaaaccgaaacgcattactgcttcacggcagaaataggcagggtgactcacaagaggtcctaccactagtaatacaagtcag
+1053 ctttcttataatttatgtcaatggccaataagctgatcctccgcaagcatgacacgaaatatatcgttcagcagtgtctctaatccgagtcgttactcaa
+1153 taaaatagggaaaataatcggtgattttctcccggcag CTATGAAACCTCCAACGGAATCAAGGCTCAGGAAACCGGTCAACTGAAGAACATTGGCACTG
                                              Y E T S N G I K A Q E T G Q L K N I G T
+1252 AAAACGAAGCTCTTGAGGTCCGCGGAGAATTTGCGTACATCGGCCCTGACGGAGTTACTTACGCTGTGACCTACGTCGCTAACGAGGGTGGATTCCAGCC
     E N E A L E V R G E F A Y I G P D G V T Y A V T Y V A N E G G F Q P
+1352 TTCGGCACCTCACATTCCCAAGGCTTAAGTCAACCGAACAGGCCCCAGCATCTAGTCCAAAAATACTCATTCATTTGAAAAGTTTTCGTTTGTTATTAGA
       S A P H I P K A *
+1452 ATCCAAGCGATCTGTGATACGATCCTGATGACAATACATACGTTTAAGTAA TCTGTATTGTTTAATTTACTCCTACTGAGGCGGCACATTTTGAACAGA
+1551 ATAGTTTGGCGGCGTTTGCCTTAAGTCGACGAGTGCCCATACCTAATACATACATACTTAAAGTCGTCGTGGCCTAACGGATAAGACGTCCGGTGCATTC
+1651 GTGTTGAGCGATGCACCGGTGTTCGAATCCCGCAGGCGGGTACCAATTTTTCTAATGAAATACGTACTCAATAAATGTTCACGATTGACTTCCACGGTGA
+1751 AGGAATGACATCGTGTTATAAAATGAAACCCGCAAAATTATAATTTGCGTAATTACTGGTGGTAGGACCTCTTGTGAGTCCGCGCGGCTGGGTACCACCA
+1851 CCCTGCCTATTTCTGCCGTGAAGCAGTAATGCGTTTCGGTTTGAAGGGTGGGGCAGCCGTCGTTACTATACTTGAGACCTTAGAACTTGTATCTCAAGGT
+1951 GGGTGGCGCCTTTACGTTGTGGATGTCTATGGGCTCCAGTAACCACTTAACATCAGGTGGGCTGTGATCCCGTCCACACATCTAAGCAATAAAAAAAAAC
+2051 TTAAGTAAAATTACAAACGAAATTCAGAAATATAATTACAAGAGTAGCTTAGTAGTTACCACGTAACAAAGTTAAAAACGAAAAGAAATTGTGCAAAAGG
+2151 TTAGTGAAAAGTTAATAAATACTATCGATATGCGCACAATCTTAATGAGTTTTTGAGTCACTGGATACTTGAATAGAAGCTT
```

Fig. 1. BMCP18 gene structure.

(A) Restriction map of pBMCP18G1Hd1. *Open bars* with Roman numbers represent exons. *Thin arrows* under the restriction map indicate the extent and direction of sequencing. *Shadowed bold arrows* show sequences homologous to the Bm1 element. Symbols for restriction enzyme cleavage sites: E, *Eco*RI; H, *Hin*dIII; Kp, *Kpn*I; Nc, *Nco*I; ScI, *Sac*I; ScII, *Sac*II; Sl, *Sal*I; Xb, *Xba*I; Xh, *Xho*I. (B) Complete BMCP18 gene nucleotide sequence. Nucleotides are numbered from the transcription initiation site (position +1). Three exons are *boxed*. The deduced amino acid sequence is shown under the nucleotide sequence. The TATA box is *double-underlined*, and the putative polyadenylation signal is *underlined*. *Arrows* indicate Bm1-like sequences. The *arrowhead* shows the signal peptide cleavage site. The *underline* following arrowhead indicates N-terminal amino acid sequence of mature BMCP18 determined by automatic Edman degradation described in Section I.

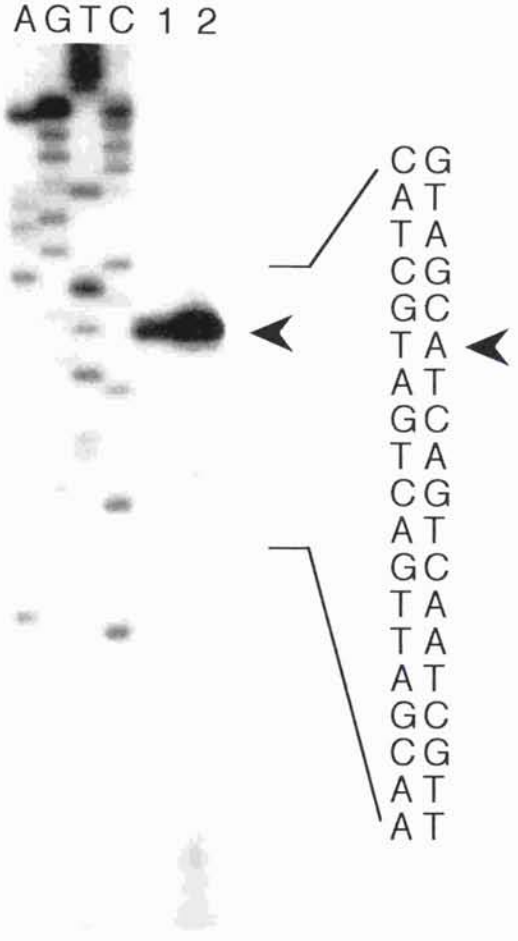


Fig. 2. BMCP18 gene transcription initiation site.

A synthetic 25 base single-stranded primer complementary to the 5' proximal region of BMCP18 cDNA was hybridized with epidermal RNA and elongated by reverse transcription. Extended cDNA was analyzed by gel electrophoresis followed by autoradiography. The amounts of loaded reaction were 1/10 (*lane1*) and 1/5 (*lane2*) of the total primer extension reaction. *A*, *G*, *T*, and *C* at the top indicate dideoxynucleotides added to sequencing reaction of the BMCP18 genomic clone. The genomic sequence corresponding to sequencing ladders is shown on the right. The *arrowhead* indicates the putative transcription initiation site.

Copy number of LCP18 gene

Genomic Southern blot analysis was conducted to analyze the genomic organization of the BMCP18 gene (Fig. 3). Each 20 μg of genomic DNA digested with *Hin*dIII (H), *Kpn*I (K), *Xba*I (X), and *Pst*I (P) was electrophoresed and transferred onto a nylon membrane. The membrane was subjected to Southern blot analysis using BMCP18 cDNA as a probe. As expected from the restriction map (Fig. 1A), *Hin*dIII excised a 3.7 kb fragment and *Kpn*I a 2.6 kb fragment. Both *Xba*I and *Pst*I digestions provided a single fragment, suggesting that the BMCP18 gene is a single copy gene in *B. mori*.

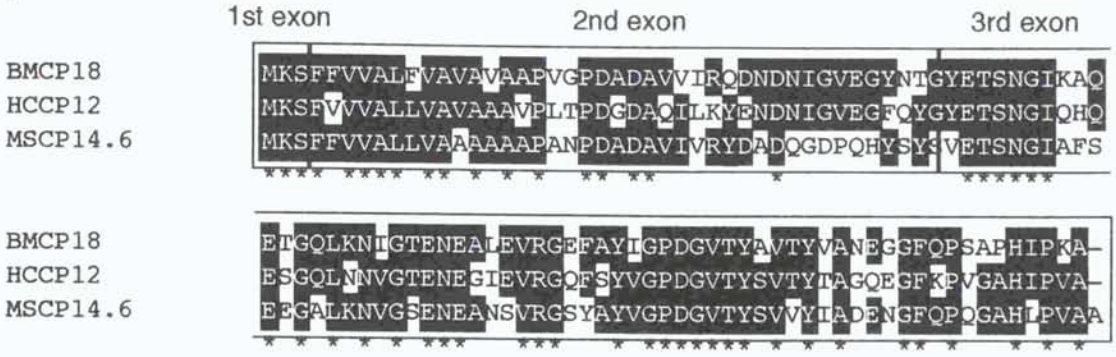
Comparison of the 5' upstream sequence of the BMCP18 gene with those of two other lepidopteran cuticle protein genes

The primary structure of BMCP18 is similar to those of cuticle proteins from different insect species (Shofuda *et al.*, 1999). This protein especially resembles MSCP14.6 of *Manduca sexta* (Rebers *et al.*, 1997) and HCCP12 of *Hyalophora cecropia* (Binger and Willis, 1994) (Fig.4A). As shown in Fig. 4B, exon/intron structures of genes encoding these three proteins are also similar to each other. Moreover, the mode of codon-interruption at exon/intron boundary is identical in all three genes (Fig.4A). The nucleotide sequences of the 5' untranslated region of mRNAs coding these three proteins are also similar (Fig. 4C). Alignment of the upstream sequence of BMCP18 with those of MSCP14.6 and HCCP12 revealed a conserved sequence of 12 bp (5'-GTTCAAATGTCA-3'), located at nt position –80 of the BMCP18 gene (Fig. 4C). The structure around the 12-bp sequence greatly resembles the consensus sequence ($^{A}/_{G}{}^{A}/_{G}G^{T}/_{G}TCA^{A}/_{G}{}^{A}/_{G}G^{T}/_{G}TCA$) for the binding of vertebrate transcription factors COUP-TF and HNF-4 (Kimura *et al.*, 1993). Some sequences resembling the recognition sequences for the Broad-Complex (BR-C) (von Kalm *et al.*, 1994), which is a group of transcription factors mediating ecdysone signals (Karim and Thummel, 1992), were also found in the 5' upstream region of each gene at –155 of the BMCP18 gene for BR-C Z4, at –278 and –222 of the MSCP14.6 gene for BR-C Z4 and Z1, respectively, and at –317 of the HCCP12 gene for BR-C Z3 (Fig. 4C).

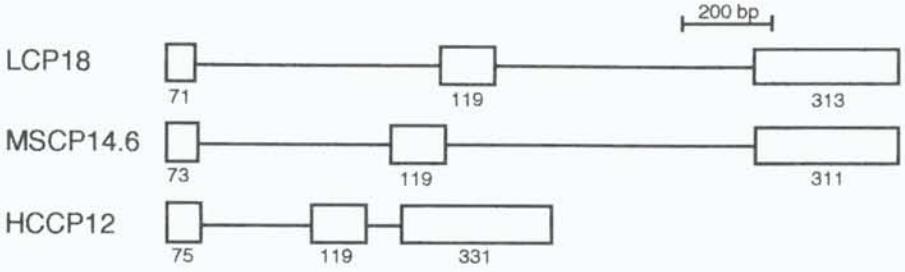
Fig. 3. Southern blot analysis for *B. mori* genomic DNA.

Each 20 μg of *B. mori* genomic DNA was digested with *Hin*dIII (*H*), *Kpn*I (*K*), *Xba*I (*X*), and *Pst*I (*P*). DNA fragments were electrophoresed, hydrolyzed with HCl, and transferred onto a nylon membrane. The membrane was washed and hybridized with ^{32}P-labeled BMCP18 cDNA. The *numbers* at left are size markers.

A



B



C

```
BMCP18      ATTCCACTGACATATTATGTTACAATTCATCTAAGTTTAAACAGAGTTATTTTTATAGTA----A   -275
MSCP14.6    TAAGTAGAGACATAATATACTAGACTCGACTACTTTCTAAACAGAGTTTTAAACAAAACTTCGCA   -259
HCCP12      TGTGATTATTTCTATAATTTTTAGTTT-ATTA-TAGATACCTAGATGACTGAGATAACTACCATT   -271
                  **  **  *   *  *        **   ***    *     *

BMCP18      ATCGAGTTTTAATAAACACGTGCACACATATCGAAAAACACGTAGCTTATGTCGCTGT-GCATGA   -211
MSCP14.6    ATTAAATATCCAACGTATTTTATTACCACATTTAAAATTATAAAATACAAAAATACATCGCATTA   -194
HCCP12      ATAGTGCTTTCATTAT-TGTCTTTTTCATATTTTAAGTGATAAACGTTATGTGTAGGT---ACTA   -210
            **      *  *             ** **   **   *   *    *        *   *  *

BMCP18      ATAACTAATGCTCTCAAAACTCAGCTGTGACGTGTGAGAATCTCAGGG--AAATATTTTAAGTTT   -148
MSCP14.6    ATTTCCGTGACTCCAAAAGGAAATGTAGG-CAAAAAAGGACCTAAACAGCATAAATATCACATAT   -130
HCCP12      ATTGAAAAAAAGTTAATAAAGTTTACGTA-GATACAATATTATAAAGA-CGTTTATAATAGCAAG   -147
            **             * *                 *     * *         **   *

BMCP18      ATAATCTT---GTAATTTATTACTTCGAAAAGCGCGCGAATCCAAC-TCAGGCTCAGTGGACATC   -87
MSCP14.6    CTAATGCAGAACTAACCGGACATGTCAAGATGAACCTTTA----------GCGTCCATAGACTCA   -75
HCCP12      ATAAAACA-AGCTAACCGGACATGTCAAAAAGAACCTTTACCTAAGATCGGTAACTGAAGACGCA   -83
             ***        ***      *  ** * * *  *    *          *   *    ***

BMCP18      TTTGAAGTTCAAATGTCAAAATGCACGTCCTCGTCTGCGGCGATGGCCCGACCGAGTATAAAAAC   -22
MSCP14.6    TCT-CAGTTCAAATGTCAGA----GTATCCTTGCCT-CGCCATC-CACGCGCTGG-TATAAATTG   -18
HCCP12      TCT-AGGTTCAAATGTCAGATA--GTGTCCTTGGCA-CGACGACGCGCCCGCTTACTATAAATGC   -22
            * *   ************ *       **** * *  ** *      *   *    ******
                                +1
BMCP18      CGCTGCGCGCATCTTCGTAGCATCAGTCAATCGTTCACTGTCCTCAAAC--AACAGTCTTAGGAC   +42
MSCP14.6    CCGCCGCGACAGCACTCAGGCATCAGTCGTTCGATCGTCATCAACATATCAAACCGTCTTAGGAC   +48
HCCP12      TTGACCCGACGACACTCTGGCATCAGTCGCTCGATTGCCTTCT-CATCACAAACAGTCTTAGGAC   +43
                  *  *      *********  *** *     **  **     *** **********

BMCP18      AAAACAAACA---ATCAATCAAAATGAAATCGGTACGTGA-TAACTGTCTACTTTTTATTACTAT   +103
MSCP14.6    TTTACAAAAA-------TCAAAAATGAAATCGGTACGTTAATTTCGAAACATCTGTGAACGTAAT   +106
HCCP12      AAAAAGAAAACAAATCATCAAAAATGAAATCGGTACGTTG--TTTGATAGATTTATAAAACGTAT   +106
               *  ** *          ******************          *  * * *     **
```

Fig. 4. Comparison of BMCP18, MSCP14.6, and HCCP12.

(A) Comparison of primary structure. The deduced amino acid sequences of BMCP18, MSCP14.6 (Rebers *et al.*, 1997), and HCCP12 (Binger and Willis, 1994) are aligned using the ClustalW Multiple Sequence Alignment Program (Thompson *et al.*, 1994). The identical residues are shown in *inverted type*, and functionally conserved residues are indicated with *shadowed background*. Amino acid sequence coded by each three exon was showed with *box*. (B) Exon/intron structures. *Boxes* and *bars* represent exons and introns, respectively. The *numbers* below the boxes indicate nucleotide numbers contained each exon. (C) Comparison of 5' upstream sequences of the genes. Nucleotide sequences of the promoter region and the first exon of three cuticle protein genes are aligned using the ClustalW. Gaps (-) are introduced to obtain optimal alignment. *Asterisks* show identical nucleotides. The first exons of three genes are *boxed*, and transcription initiation sites (+1) are shown. Translation initiation codons (ATG) are *dotted*. The conserved 12-bp sequence and putative TATA box are shown in *inverted type*. *Arrows* indicate sequences similar to Broad-Complex recognition sequences.

Bombyx homologs of COUP-TF and HNF-4 in larval epidermis

I focused on analyzing transcription factors that recognize the 12-bp sequence because this sequence is highly conserved among three similar cuticle protein genes of different species and may function as a transcriptional *cis*-element.

It is known that Seven-up (Svp), a *Drosophila* transcription factor, is a homolog of COUP-TF (Mlodzik *et al.*, 1990), and a *Bombyx* homolog of Svp, Bmsvp, were identified recently (Yaginuma, personal communication). Swevers and Iatrou (1998) reported that BmHNF-4, a *Bombyx* homolog of vertebrate HNF-4, recognizes the above consensus sequence. To determine whether Bmsvp and BmHNF-4 are expressed in larval epidermal cells, I conducted an RT-PCR experiment using total epidermal RNA. Preliminary experiments showed that Bmsvp and BmHNF-4 mRNA are abundant in the fat body (Fig. 5A, lane FB), so fat body cells were carefully removed from the larval integument. I examined the contamination of fat body cells in the integument preparation by RT-PCR using primers specific for a 30K protein mRNA (Sakai *et al.*, 1988), expressed only in fat body cells. Bmsvp and BmHNF-4 mRNAs existed in the RNA fraction isolated from integument cells of fifth instar day-1 larvae, but mRNA coding for a 30K protein was hardly detected (Fig. 5A, lane Epi). I have already confirmed that these RT-PCR products were not derived from genomic sequences by negative control experiment in which RT-PCR was performed without reverse transcriptase (data not shown). These results clearly indicated that Bmsvp and BmHNF-4 are expressed in epidermal cells at this stage (Fig. 5A).

Subsequently, I analyzed developmental changes of mRNA levels coding for Bmsvp and BmHNF-4 in epidermal cells (Fig. 5B). Preceding it, I verified the correlation between amounts of Bmsvp mRNA and PCR products. As shown in Fig. 5C, it is clear that the products resulted from RT-PCR reflect relative amounts of Bmsvp mRNA sequences. Developmental changes of mRNA levels for Bmsvp and BmHNF-4 are shown in Fig. 5B. Bmsvp was expressed almost evenly throughout fourth larval stages and feeding stages of fifth instar larvae. In contrast, the expression of BmHNF-4 increased greatly after day-2 of the fifth instar.

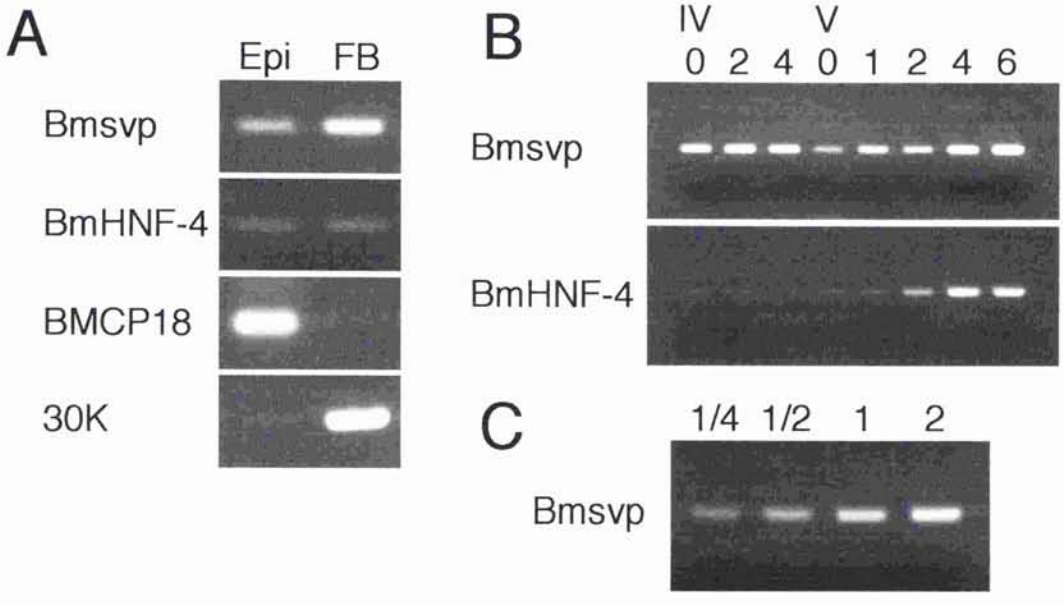

A
Epi
FB
Bmsvp
BmHNF-4
BMCP18
30K
B
IV
V
0 2 4 0 1 2 4 6
Bmsvp
BmHNF-4
C
1/4 1/2 1 2
Bmsvp

Fig. 5. RT-PCR to detect Bmsvp and BmHNF-4 expression in epidermal cells.

(A) Total RNA was isolated from the integument (*Epi*) and the fat body (*FB*) of larvae one day after the fourth molt (V1). Total cDNA was synthesized using a RAV-2 reverse transcriptase and oligo(dT) primer. Bmsvp cDNA, BmHNF-4 cDNA, BMCP18 cDNA (for a positive control), and an isoform of cDNA for 30K protein (30Kc6, Sakai *et al.*, 1988) were amplified by PCR using the primers described in Materials and Methods. (B) Developmental profiles of Bmsvp and BmHNF-4 mRNA in the epidermis. Total RNA was prepared from the integument during larval development and Bmsvp and BmHNF-4 cDNAs were detected by RT-PCR. Developmental stages are *IV0*, the day of the third molt, and *V0*, the day of the fourth molt. *Numbers* indicate days after each molt. (C) RT-PCR for Bmsvp mRNA was carried out using various amounts of reverse transcripts from the V1 epidermal RNA to confirm quantitative amplification. Relative amounts of templates used were indicated at the top. The same amount of reverse transcripts as in the reaction labeled lane *1* was used for other RT-PCR experiments shown in (A) and (B).

Specific interaction of Bmsvp with the 5' upstream sequence of the BMCP18 gene *in vitro*

I analyzed the binding activity of Bmsvp and BmHNF-4 to the 12-bp sequence upstream of the BMCP18 gene by gel retardation assay, using Bmsvp and BmHNF-4 expressed *in vitro*. Production of Bmsvp and BmHNF-4 *in vitro* was confirmed by [^{35}S] methionine incorporation followed by SDS-PAGE (data not shown). The ^{32}P-labeled 5' upstream sequence of the BMCP18 gene containing the 12-bp conserved sequence was used as a probe (Fig. 6A). Unlabeled DNA bearing the sequence identical to the probe was used as a specific competitor, and the 5' upstream sequence of the *B. mori* SP1 gene (5'-CGAGAGCTGATAATGGTTGGATAGT-3', Sakurai *et al.*, 1988) was used as a nonspecific competitor. A complex was detected when Bmsvp was incubated with the probe DNA (Fig. 6B). The formation of the complex was competed by a specific, but not by a nonspecific, competitor. No binding activity of BmHNF-4 to the probe was observed. These results indicated that Bmsvp binds specifically to the sequence *in vitro*. The presence of Bmsvp mRNA in the epidermis and the binding activity of Bmsvp to the 5' upstream sequence of the BMCP18 gene suggest the involvement of Bmsvp in BMCP18 gene expression control.

A

CTTTGAAGTTCAAATGTCAAAATGC

B

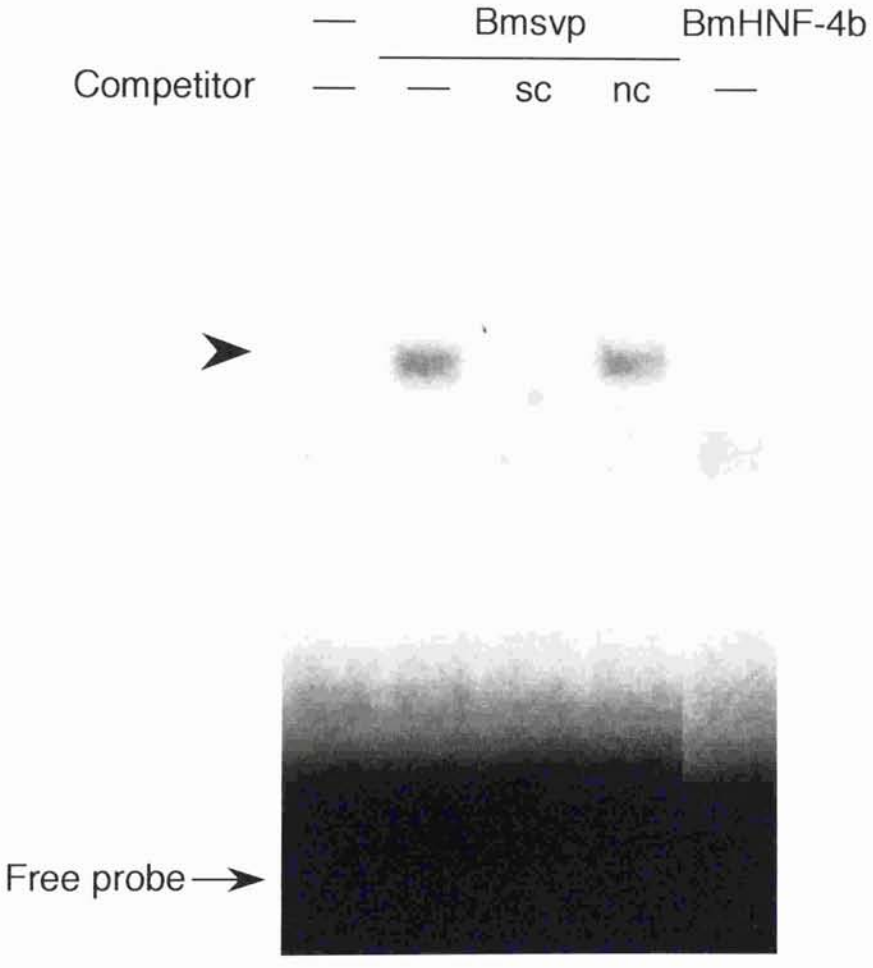


Fig. 6. Gel retardation assay of Bmsvp and BmHNF-4.

(A) The 5' upstream sequence of BMCP18 gene used as a probe in the gel retardation assay. *Arrows* indicate the degenerate direct repeat of the nuclear hormone receptor recognition half-site, AGGTCA, separated by one nucleotide. (B) Bmsvp and BmHNF-4b nuclear receptors were synthesized *in vitro* using a TNT system. Competitor DNA was added at a 100-fold molar excess to the labeled probe. An unlabeled probe was used as a specific competitor (*sc*), and the 5' upstream sequence of the storage protein SP1 gene was used as a nonspecific competitor (*nc*). Shifted bands are indicated by *arrowhead*.

DISCUSSION

To obtain the fundamental information for analyzing the regulation mechanisms of BMCP18 gene expression, structure of genomic clone coding for BMCP18 was analyzed. A comparison of the genomic sequence with that of cDNA showed that the BMCP18 gene consists of three exons separated by two introns. Sequences homologous to the Bm1 element, a repetitive sequence dispersed throughout the *B. mori* genome, were observed in the 5' upstream region, the second intron, and downstream of the third exon. Transposons and other repetitive sequences have been implicated in gene duplication during evolution of multigene clusters in many species (Maeda and Smithies, 1986). Moreover, Bm1-mediated gene duplication has been suggested in *B. mori* (Rodakis *et al.*, 1984; Mori *et al.*, 1991). It has been proposed that gene clusters of several insect cuticle proteins have arisen due to gene duplication (Snyder *et al.*, 1982; Charles *et al.*, 1997; Rondot *et al.*, 1998). The primary structure of BMCP18 resembles closely those of other cuticle proteins of *B. mori*, BMCP17, BMCP22, and BMCP30 as described in Section I. Although there are no data showing that *B. mori* cuticle protein genes are clustered near one another on the chromosome, the distribution of multiple copies of the Bm1 element around the BMCP18 gene suggests that Bm1-mediated gene duplication may have occurred as these genes evolved.

A group of insect cuticle proteins is structurally related to one another, possessing a conserved region termed the R&R Consensus (Andersen *et al.*, 1995; Iconomidou *et al.*, 1999). BMCP18 is a member of this group, exhibiting homology of other cuticle proteins. BMCP18 shows a striking similarity to MSCP14.6 and HCCP12 from other lepidopteran species (Shofuda *et al.*, 1999). A comparison of genes for these three proteins showed that primary structures and exon/intron structures were conserved. Expression patterns of genes for BMCP18, MSCP14.6, and HCCP12 are also comparable. The BMCP18 gene is expressed in larval epidermal cells until the mid-fifth instar and disappears before the spinning stage. A temporal rise of BMCP18 mRNA in the epidermis is seen on the day of pupation and at the pharate adult stage (Shofuda *et al.*, 1999). The expression of MSCP14.6 mRNA also occurs

throughout larval feeding stages and disappears before the wandering stage. In pharate pupae and pharate adults, MSCP14.6 mRNA is detectable again at restricted regions of the integument (Riddiford *et al.*, 1986; Rebers *et al.*, 1997). HCCP12 accumulates in the flexible cuticle from the larval to adult stages. In contrast, the abdominal rigid region of the integument at larval and adult, but not pupal, stages synthesizes HCCP12 (Binger and Willis, 1990). In summary, genes for BMCP18, MSCP14.6, and HCCP12 appear to be expressed throughout the larval epidermis at feeding stages, and the expression is spatially and temporally restricted at pupal and adult stages, although the spatial distribution of BMCP18 mRNA at later stages remains unclear. Similarity in developmental patterns suggests that the expression of these cuticle protein genes is controlled by conserved mechanisms.

In many studies, to identify potential *cis*-acting elements, 5' upstream sequences of orthologous genes have been compared among different species (Fujioka *et al.*, 1996; Lessing and Nusse, 1998) based on the assumption that functionally important sequences have been conserved during evolution. Following this strategy, I compared the 5' upstream sequences of the three lepidopteran cuticle protein genes. Rebers *et al.* (1997) reported a striking similarity between 5' upstream sequences of MSCP14.6 and HCCP12 genes. A sequence highly homologous to the HCCP12 gene (27/29 bp identical) occurs at -118 from the transcription initiation site of the MSCP14.6 gene. This sequence, however, was not identified in the 5' upstream of the BMCP18 gene. Instead, a comparison of the upstream sequences of these three genes showed an additional conserved sequence of 12-bp (5'-GTTCAAATGTCA-3') near the core promoter region (Fig. 4C). The 13-bp region including the above sequence (5'-AGTTCAAATGTCA-3' in the BMCP18 gene) consists of an incomplete direct repeat of the nuclear hormone receptor recognition half-site, AGGTCA, with separation by one nucleotide. The sequence matches the recognition sequence of vertebrate nuclear receptors COUP-TF and HNF-4 (Kimura *et al.*, 1993). Occurrence of the identical sequence in comparable regions of the three cuticle protein genes implies that the sequence is responsible for temporal and/or spatial activation of these genes with intermediation of transcription factors such as COUP-TF and HNF-4.

Svp is a *Drosophila* homolog of COUP-TF that was first identified as a gene required for normal eye development. The *svp* mutant shows transformation of photoreceptor cells, the R3/R4 pair and the R1/R6 pair toward R7 cell fate (Mlodzik *et al.*, 1990). Ubiquitous expression of *svp* causes some disorders of cell fate determination in ommatidia including transformation of non-neuronal cone cells to R7 cells. These results suggest that *svp* controls decisions among photoreceptor subtypes as well as between neuronal and non-neuronal fates (Hiromi *et al.*, 1993). Svp is also involved in cell proliferation of the Malpighian tubules under EGF receptor signals (Kerber *et al.*, 1998). Biochemical studies on ecdysone signaling pathways showed that Svp can heterodimerize with EcR, a subunit of a functional ecdysone receptor, and inhibit ecdysone signals (Zelhof *et al.*, 1995). It is still unclear, however, what factors actually interact with Svp in developing tissues, including eye ommatidia and the Malpighian tubules. Furthermore, a ligand for Svp is also unknown.

I showed that mRNA for Bmsvp, the *B. mori* homolog of Svp, existed in larval epidermal cells where the BMCP18 gene is expressed (Fig. 5). The Bmsvp expression pattern during the fifth instar, however, is inconsistent with that of BMCP18. Although BMCP18 expression declines about four days after the fourth molt (Shofuda *et al.*, 1999), Bmsvp mRNA remains high for six days in the final larval stage (Fig. 5B). Since Bmsvp is a member of the ligand-dependent nuclear receptor superfamily, the capacity of Bmsvp to activate transcription may also be ligand-dependent. I clearly showed that Bmsvp expressed *in vitro* recognized the conserved 12-bp sequence in a sequence-specific manner (Fig. 6). This suggests the possible participation of Bmsvp in the regulation of BMCP18 gene expression and supports the idea that the conserved 12-bp sequence is involved in a common regulatory mechanism of cuticle protein gene expression in the lepidopteran species.

It is also possible that Bmsvp contributes to JH responsiveness of BMCP18 gene. Since Bmsvp is a member of nuclear receptor superfamily, it may act as a JH receptor. Another possibility is that Bmsvp functions as a mediator of JH signals. JH signaling possibly involves the primary and the secondary responses, such as those of steroids and thyroid hormones. Developmental pattern of Bmsvp in larval epidermal cells is inconsistent

to hemolymphal JH titer, which declines rapidly one day after the fourth molt (Fig. 5B). Therefore, Bmsvp is not expectable as a primary response product of JH signal. Many nuclear hormone receptors usually act as homo- or heterodimer, so heterodimeric partner of Bmsvp may be produced at the primary response and mediate JH signaling with Bmsvp.

REFERENCE

Adams, D. S., Eickbush, T. H., Herrera, R. J., and Lizardi, P. M. (1986). A highly reiterated family of transcribed oligo(A)-terminated, interspersed DNA elements in the genome of *Bombyx mori*. *J. Mol. Biol.* **187,** 465-478.

Andersen, S. O., Højrup, P., and Roepstorff, P. (1995). Insect cuticular proteins. *Insect Biochem. Mol. Biol.* **25,** 153-176.

Ashburner, M. (1989). Preparing RNA from adult flies. *In* "*Drosophila*: A Laboratory Manual" (Ashburner, M., Ed.), pp. 110-111. Cold Spring Harbor Laboratory, New York.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987). "Current protocols in molecular biology." John Wiley & Sons, New York.

Binger, L. C. and Willis, J. H. (1990). *In vitro* translation of epidermal RNA from different anatomical regions and metamorphic stages of *Hyalophora cecropia*. *Insect Biochem.* **20,** 573-583.

Binger, L. C. and Willis, J. H. (1994). Identification of the cDNA, gene and promoter for a major protein from flexible cuticles of the giant silkmoth *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* **24,** 989-1000.

Busslinger, M., Portmann, R., Irminger, J. C., and Birnstiel, M. L. (1980). Ubiquitous and gene-specific regulatory 5' sequences in a sea urchin histone DNA clone coding for histone protein variants. *Nucleic Acids Res.* **8,** 957-977.

Cavener, D. R. (1987). Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res.* **15,** 1353-1361.

Charles, J. P., Chihara, C., Nejad, S., and Riddiford, L. M. (1997). A cluster of cuticle protein genes of *Drosophila* melanogaster at 65A: sequence, structure and evolution [published erratum appears in *Genetics* 1998 ,**148**, 921]. *Genetics* **147,** 1213-1224.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal. Biochem.* **162,** 156-159.

Comas, D., Piulachs, M. D., and Belles, X. (1999). Fast induction of vitellogenin gene expression by juvenile hormone III in the cockroach *Blattella germanica* (L.) (Dictyoptera, Blattellidae). *Insect Biochem. Mol. Biol.* **29,** 821-827.

Dubrovsky, E. B., Dubrovskaya, V. A., Bilderback, A. L., and Berger, E. M. (2000). The

isolation of two juvenile hormone-inducible genes in *Drosophila melanogaster*. *Dev. Biol.* **224,** 486-495.

Fujioka, M., Miskiewicz, P., Raj, L., Gulledge, A. A., Weir, M., and Goto, T. (1996). *Drosophila* Paired regulates late *even-skipped* expression through a composite binding site for the paired domain and the homeodomain. *Development* **122,** 2697-2707.

Hirai, M., Yuda, M., Shinoda, T., and Chinzei, Y. (1998). Identification and cDNA cloning of novel juvenile hormone responsive genes from fat body of the bean bug, *Riptortus clavatus* by mRNA differential display. *Insect Biochem. Mol. Biol.* **28,** 181-189.

Hiromi, Y., Mlodzik, M., West, S. R., Rubin, G. M., and Goodman, C. S. (1993). Ectopic expression of *seven-up* causes cell fate changes during ommatidial assembly. *Development* **118,** 1123-1135.

Hultmark, D., Klemenz, R., and Gehring, W. J. (1986). Translational and transcriptional control elements in the untranslated leader of the heat-shock gene *hsp22*. *Cell* **44,** 429-438.

Iconomidou, V. A., Willis, J. H., and Hamodrakas, S. J. (1999). Is β-pleated sheet the molecular conformation which dictates formation of helicoidal cuticle? *Insect Biochem. Mol. Biol.* **29,** 285-292.

Jones, G. and Sharp, P. A. (1997). Ultraspiracle: an invertebrate nuclear receptor for juvenile hormones. *Proc. Natl. Acad. Sci. USA* **94,** 13499-13503.

Kapitskaya, M. Z., Dittmer, N. T., Deitsch, K. W., Cho, W. L., Taylor, D. G., Leff, T., and Raikhel, A. S. (1998). Three isoforms of a hepatocyte nuclear factor-4 transcription factor with tissue- and stage-specific expression in the adult mosquito. *J. Biol. Chem.* **273,** 29801-29810.

Karim, F. D. and Thummel, C. S. (1992). Temporal coordination of regulatory gene expression by the steroid hormone ecdysone. *ENBO J* **11,** 4083-4093.

Kerber, B., Fellert, S., and Hoch, M. (1998). Seven-up, the *Drosophila* homolog of the COUP-TF orphan receptors, controls cell proliferation in the insect kidney. *Genes Dev.* **12,** 1781-1786.

Kimura, A., Nishiyori, A., Murakami, T., Tsukamoto, T., Hata, S., Osumi, T., Okamura, R., Mori, M., and Takiguchi, M. (1993). Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-

4 (HNF-4). *J. Biol. Chem.* **268,** 11125-11133.

Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P., and Hogness, D. S. (1991). The *Drosophila EcR* gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67,** 59-77.

Lam, G. T., Jiang, C., and Thummel, C. S. (1997). Coordination of larval and prepupal gene expression by the DHR3 orphan receptor during *Drosophila* metamorphosis. *Development* **124,** 1757-1769.

Lessing, D. and Nusse, R. (1998). Expression of *wingless* in the *Drosophila* embryo: a conserved *cis*-acting element lacking conserved Ci-binding sites is required for *patched*- mediated repression. *Development* **125,** 1469-1476.

Maeda, N. and Smithies, O. (1986). The evolution of multigene families: human haptoglobin genes. *Annu. Rev. Genet.* **20,** 81-108.

Mlodzik, M., Hiromi, Y., Weber, U., Goodman, C. S., and Rubin, G. M. (1990). The *Drosophila seven-up* gene, a member of the steroid receptor gene superfamily, controls photoreceptor cell fates. *Cell* **60,** 211-224.

Mori, S., Izumi, S., and Tomino, S. (1991). Structures and organization of major plasma protein genes of the silkworm *Bombyx mori*. *J. Mol. Biol.* **218,** 7-12.

Nakato, H., Shofuda, K., Izumi, S., and Tomino, S. (1994). Structure and developmental expression of a larval cuticle protein gene of the silkworm, *Bombyx mori*. *Biochim. Biophys. Acta* **1218,** 64-74.

Palli, S. R., Touhara, K., Charles, J. P., Bonning, B. C., Atkinson, J. K., Trowell, S. C., Hiruma, K., Goodman, W. G., Kyriakides, T., Prestwich, G. D., Hammock, B. D., and Riddiford, L. M. (1994). A nuclear juvenile hormone-binding protein from larvae of *Manduca sexta*: a putative receptor for the metamorphic action of juvenile hormone. *Proc. Natl. Acad. Sci. USA* **91,** 6191-6195.

Rebers, J. E., Niu, J., and Riddiford, L. M. (1997). Structure and spatial expression of the *Manduca sexta* MSCP14.6 cuticle gene. *Insect Biochem. Mol. Biol.* **27,** 229-240.

Riddiford, L. M., Baeckmann, A., Hice, R. H., and Rebers, J. (1986). Developmental expression of three genes for larval cuticular proteins of the tobacco hornworm, *Manduca sexta*. *Dev. Biol.* **118,** 82-94.

Rodakis, G. C., Lecanidou, R., and Eickbush, T. H. (1984). Diversity in a chorion multigene family created by tandem duplications and a putative gene-conversion event. *J. Mol. Evol.* **20,** 265-273.

Rondot, I., Quennedey, B., and Delachambre, J. (1998). Structure, organization and expression of two clustered cuticle protein genes during the metamorphosis of an insect, *Tenebrio molitor*. *Eur. J. Biochem.* **254,** 304-312.

Sakai, N., Mori, S., Izumi, S., Haino-Fukushima, K., Ogura, T., Maekawa, H., and Tomino, S. (1988). Structures and expression of mRNAs coding for major plasma proteins of *Bombyx mori*. *Biochim. Biophys. Acta* **949,** 224-232.

Sakurai, H., Fujii, T., Izumi, S., and Tomino, S. (1988). Structure and expression of gene coding for sex-specific storage protein of *Bombyx mori*. *J. Biol. Chem.* **263,** 7876-7880.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual; Second Edition." Cold Spring Harbor Laboratory, New York.

Sanger, F. (1981). Determination of nucleotide sequences in DNA. *Science* **214,** 1205-1210.

Shofuda, K., Togawa, T., Nakato, H., Tomino, S., and Izumi, S. (1999). Molecular cloning and characterization of a cDNA encoding a larval cuticle protein of *Bombyx mori*. *Comp. Biochem. Physiol. B* **122,** 105-109.

Snyder, M., Hunkapiller, M., Yuen, D., Silvert, D., Fristrom, J., and Davidson, N. (1982). Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. *Cell* **29,** 1027-1040.

Swevers, L. and Iatrou, K. (1998). The orphan receptor BmHNF-4 of the silkmoth *Bombyx mori*: ovarian and zygotic expression of two mRNA isoforms encoding polypeptides with different activating domains. *Mech. Dev.* **72,** 3-13.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673-4680.

Thummel, C. S. (1995). From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor superfamily members. *Cell* **83,** 871-877.

von Kalm, L., Crossgrove, K., Von Seggern, D., Guild, G. M., and Beckendorf, S. K. (1994). The *Broad-Complex* directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of *Drosophila* metamorphosis. *ENBO J* **13,** 3505-3516.

Yao, T. P., Segraves, W. A., Oro, A. E., McKeown, M., and Evans, R. M. (1992). *Drosophila ultraspiracle* modulates ecdysone receptor function via heterodimer formation. *Cell*

71, 63-72.

Zelhof, A. C., Yao, T. P., Chen, J. D., Evans, R. M., and McKeown, M. (1995). Seven-up inhibits ultraspiracle-based signaling pathways *in vitro* and *in vivo*. *Mol. Cell. Biol.* **15,** 6736-6745.

Zhang, J., McCracken, A., and Wyatt, G. R. (1993). Properties and sequence of a female-specific, juvenile hormone-induced protein from locust hemolymph. *J. Biol. Chem.* **268,** 3282-3288.

Acknowledgments

I am very grateful to Dr. Susumu Izumi, Tokyo Metropolitan University, for his instruction and encouragement throughout this work and for his critical review of this manuscript. I wish to express my gratitude to Dr. Hiroshi Nakato, Tokyo Metropolitan University, for his encouragement and variable suggestions throughout this work. I would like to thank Prof. Shiro Tomino, Okayama University of Science, for his instruction and encouragement. I acknowledge providing the Bmsvp cDNA with Dr. Toshinobu Yaginuma, Nagoya University. I am indebted to a number of my colleagues at Department of Biology, Tokyo Metropolitan University, for providing biological materials and helpful discussion.