

# Bombyxin: An Insect Brain Peptide that Belongs to the Insulin Family

Author: Iwami, Masafumi

Source: Zoological Science, 17(8): 1035-1044

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.17.1035

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at <u>www.bioone.org/terms-of-use</u>.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

### Bombyxin: An Insect Brain Peptide that Belongs to the Insulin Family

Masafumi Iwami\*

Department of Biology, Faculty of Science and Division of Life Science, Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa 920-1192, Japan

ABSTRACT—Bombyxin is a 5 kDa secretory brain peptide that belongs to the insulin family. Bombyxin of the silkmoth Bombyx mori can induce adult development when injected into brain-removed dormant pupae of the saturniid moth Samia cynthia ricini by activating the prothoracic glands to synthesize and release ecdysone. Bombyx bombyxin has been shown to lower the concentration of the major haemolymph sugar, trehalose, and to elevate the trehalase activity in the midgut and muscles in Bombyx, but the doses required to be effective are higher than the amounts in the feeding larvae. The exact physiological function of bombyxin in Bombyx itself is therefore still obscure, but its insulin-like structure suggests it has important roles. Bombyxin comprises a mixture of highly heterogeneous molecular forms whose amino acid sequences have 40% identity with human insulin. The Bombyx bombyxin gene encodes a precursor consisting of the signal peptide, B chain, C peptide, and A chain, in that order from the N terminus. So far, 32 bombyxin genes have been identified in Bombyx, and they are classified into 7 families, A to G, according to their sequence similarity. The bombyxin genes have no introns and cluster in unique distribution patterns. The gene arrangement in the cluster has been classified into three categories: gene pairs, gene triplets, and single genes. Nucleotide sequence analysis indicates that equal and unequal crossings-over and duplications may have generated these unique distribution patterns. The Bombyx bombyxin genes are expressed predominantly in the brain and at low levels in a number of other tissues. Genes of all 7 families are expressed in four pairs of the medial neurosecretory cells of the brain. Detailed examination indicated that only a limited number of genes in the A, B and C family members are expressed and that their expression shows a gene-arrangement-dependent pattern.

### INTRODUCTION

Insect molting and metamorphosis have remained interesting and important subjects throughout the history of developmental biology and endocrinology. In 1922, Kopeć found that a humoral factor from the gypsy moth *Lymantria dispar* was responsible for the induction of larval to pupal molting (Kopeć, 1922). This finding was the first demonstration of the presence of hormone in invertebrates. He named that hormone the brain hormone. His finding was unfortunately long ignored, until 1940 when Wigglesworth threw light on the brain hormone by confirming Kopećs discovery (Wigglesworth, 1940). Shortly after, Fukuda demonstrated that the prothoracic glands in the thorax secrete a hormone essential for

\* Corresponding author: Tel. +81-76-264-5711;

FAX. +81-76-264-5977. E-mail: masafumi@kenroku.kanazawa-u.ac.jp insect development (Fukuda, 1944). The hormone secreted from the prothoracic gland is now known as ecdysone, a hormone directly responsible for the development of peripheral tissues. Fukuda's work on the silkmoth *Bombyx mori* was one of the outstanding scientific achievements in Japan. His work was immediately confirmed by Williams (1947). Williams then proved that the brain hormone, now known as the prothoracicotropic hormone (PTTH) stimulates the prothoracic glands to secrete ecdysone. By those pioneering works, an important concept, known as the classical scheme, was established (for review, Bollenbacher and Granger, 1985).

In the early 20th century, Japan's economy was highly dependent on silk production so a large amount of research was done on *Bombyx* (for review, Okada, 1994). The abundance of the silkmoth larvae was a great advantage for biologists at that time. In the course of the purification of PTTH, a peptide was identified in an extract from *Bombyx* brains (Ichikawa and Ishizaki, 1963) and, after many years, the peptide was purified almost homogeneously from millions of Bombyx heads (Suzuki et al., 1982). The peptide has a molecular weight of 5 kDa and the ability to stimulate the prothoracic glands of the saturniid moth Samia cynthia ricini to synthesize and release ecdysone. Soon after, the 5 kDa peptide was however revealed to be inactive in Bombyx at a physiological dose, but instead a 30 kDa peptide is active in Bombyx itself (Kawakami et al., 1990). Thus, the 5 kDa peptide has not been regarded as a "pure" PTTH. Unexpectedly, the 5 kDa peptide was shown to be homologous to insulin, a peptide hormone that plays crucial roles in the energy metabolism and growth of vertebrates (Nagasawa et al., 1984, 1986). Although the existence of insulin-like peptides in invertebrates was predicted in 1960s through biological and/or immunological assays, little was known about their structure (for review, Kramer, 1985). The elucidation of the structure of the 5 kDa peptide was therefore the first demonstration of the presence of insulin or insulin-related peptides in invertebrates at the molecular level. The silkmoth insulin-related peptide is now called bombyxin and has been proved to exist widely in insects. In this review, molecular aspects of bombyxin, especially the structure and expression of its gene, are summarized.

### STRUCTURE OF BOMBYXIN AND THE BOMBYXIN GENE

Bombyxin comprises highly heterogeneous molecular forms. Five molecular forms, I, II, III, IV and V, have so far been identified from *Bombyx* heads (Nagasawa *et al.*, 1984, 1986, 1988; Jhoti *et al.*, 1987; Maruyama *et al.*, 1988). The primary structures have been determined completely for

bombyxins II and IV and partially for bombyxins I, III and V. They are heterodimers of the A and B chains whose amino acid sequences show about 50% and 30% identity to the A and B chains of human insulin (Fig. 1). The A and B chains of bombyxin are connected by two inter- and one intra-chain disulfide bonds in exactly the same manner as insulin. Bombyxins II and IV have been chemically synthesized and proved to have the same prothoracicotropic activity in Samia as natural bombyxins (Nagasawa et al., 1988; Maruyama et al., 1990). Molecular modeling for the three-dimensional structure showed that bombyxin resembles insulin in adopting a globular-like core structure (Jhoti et al., 1987). Solution structure analysis of bombyxin by nuclear magnetic resonance further demonstrated that the overall main-chain fold of bombyxin is similar to those of insulin in solution, insulin in the crystalline T-state, and relaxin in the crystalline form (Nagata et al., 1995). In fact, a hybrid molecule consisting of the A chain of Bombyx bombyxin and the B chain of human insulin stimulates 2deoxyglucose uptake and DNA synthesis in CHO cells. Bombyxin resembles relaxin in having a pyroglutamic acid residue at the B chain N terminus (Nagasawa et al., 1986).

Thirty-two genes that encode bombyxin have been cloned from the *Bombyx* genome (Iwami *et al.*, 1989, 1990; Kawakami *et al.*, 1989; Kondo *et al.*, 1996; Tsuzuki *et al.*, 1997; Yoshida *et al.*, 1997, 1998). These genes have been classified into 7 families, A, B, C, D, E, F and G, according to their sequence similarity. The family B members can be further divided into three subfamilies BI, BII and BIII (Kondo *et al.*, 1996). The family A bombyxin consists of 10 gene copies, the family B 12 gene copies, the family C 6 gene copies. The bombyxin fami-





lies D to G stand as single genes. The amino acid sequences deduced from the bombyxin genes show that bombyxin II is the product of genes A6 and/or A7 (Kondo *et al.*, 1996) and bombyxin IV is that of gene E1 (Tsuzuki *et al.*, 1997). On the other hand, bombyxins III and V do not coincide with any bombyxins deduced from the genes, indicating that more bombyxin genes remain undetected. The deduced amino acid sequences of the bombyxin genes show 41% to 56% identity with each of the other families and 28% to 35% identity with

 Table 1. Overall sequence identity among the prepropeptides of

 Bombyx bombyxins and human insulin

			Bombyxin							Humar
			A1	B1	C1	D1	E1	F1	G1	Insulin
Bombyxin	A1	(89)	100							
	B1	(73)	55	100						
	C1	(91)	51	51	100					
	D1		55	51	52	100				
	E1		48	49	47	47	100			
	F1		48	47	51	53	48	100		
	G1		50	54	47	56	41	50	100	
Human Insulin			35	33	34	33	34	31	28	100

Amino acid sequence identity was calculated using the computer program Maximum Matching in DNASIS (Hitachi Software). Numbers in parentheses indicate the minimum sequence identity of preprobombyxins within the same family. Amino acid sequences are from A1, Iwami *et al.*, 1989; B1, Kawakami *et al.*, 1989; C1, Iwami *et al.*, 1990; D1, Kondo *et al.*, 1996; E1, Tsuzuki *et al.*, 1997; F1, Yoshida *et al.*, 1997; G1, Yoshida *et al.*, 1998; human insulin, Bell *et al.*, 1980. human preproinsulin, as shown in Table 1. Preprobombyxins within the same family have at least 73% identical sequences with each other. Compared with the limited structural variation of vertebrate insulins (Steiner *et al.*, 1985), *Bombyx* bombyxin has a large diversity in structure. In vertebrate insulins, there may be little room for mutational divergence to occur because of the low copy number of their genes (see Discussion in Kondo *et al.*, 1996). It is reported that even point mutations resulting in abnormal human insulins can cause diabetes mellitus (Steiner *et al.*, 1985).

The amino acid sequences of the protein products deduced from the genes are listed in Fig. 2. Except for C3 to C6, all the products have the same basic structure as that of preproinsulin (Kondo *et al.*, 1996). They consist of the signal peptide, B chain, C peptide, and A chain, in that order from the N terminus. Bombyxin genes C3 to C6 have an in-frame stop codon so as to encode only the signal peptide and an N terminal portion of the B chain. The bombyxin A10 gene also has an in-frame stop codon at about the center of the region coding for the A chain (Kondo *et al.*, 1996). These genes are thus presumably pseudogenes.

The amino acid sequences of preprobombyxins show a high similarity in the A and B chains throughout the 7 families, and the level of conservation is remarkably high within each family (Fig. 2). On the other hand, the conservation of amino acid sequences is relatively low in the C peptide and even lower in the signal peptide, as has been shown for preproinsulins (Steiner *et al.*, 1985). The sequences of the C

1         10         20         1         10         20         30         1         10         20         30         1         10         20         30         1         10         20         30         1         10         20         30         1         10         20			Signal Peptide	B Chain		С	Peptide		A Chain
Bondbyx         A1         ZNI ILLAI ALM/STVMWYST         OOPORVHT/SECTURATION         Stable Press         Stable Pre			1 10 20	1 . 10 .	20 . 30	1.	10 20	. <sup>30</sup>	1 10 20
Bondbyx In Family A         A2         TitLA IALMESTVMMVST         OOPGEVHTTCGETURATE/A         DIMEEGVD         SEDAGRAS         IS         SAULHEYS         AG         IV         IV         AG         IV         IV         A3         TitLA IALMESTVMMVST         OOPGEVHTTCGETURATE/A         DIMEEGVD         SEDAGRAS         IS         SAULHEYS         AG         IV         IV         A4         TitLA IALMESTVMMVST         OOPGEVHTCGETURATE/A         DIMEAGVD         SEDAGRAS         IS         SAULHYS         EG3DT         IV         IV <td>Bombyx</td> <td>E A1</td> <td>MILLALALMI STVMWVST</td> <td>OOPORVHTXCCHI ART</td> <td></td> <td><b>KE</b>SGAOFAS</td> <td>🛱 GSA 🗒 MPYS</td> <td>FGRIEKE</td> <td></td>	Bombyx	E A1	MILLALALMI STVMWVST	OOPORVHTXCCHI ART		<b>KE</b> SGAOFAS	🛱 GSA 🗒 MPYS	FGRIEKE	
Family A         A3         TillLaiaLLISTVMWST         Opeoprint/CERTURED         ViewEAGVD         SDADIVS         IC         EG320         IC	Bombyxin	A2	MULLALALALSTVMWVST	OOPOEVHT CGRHLARINA	DOWEEGVD	KRSDAQFAS	G SA LMPYS	AG	
A4         TRULLAIALMESTVANVST         OOPOGVHT CCRTUARUE         DOMEAGVD         CSGADFAS         G SALLMPYS         EGTEXT         EI DEGG LPP SVD/L SYG           A5         TILLAIALMUTTVWWAST         OOPOGVHT CCRTUARUE         DOWEAGVD         CSGADFAS         G SALLMPYS         EGTEXT         EI DEGG LPP SVD/L SYG           A6         TILLAIALMUTTVWWAST         OOPOAHTI CCRTUARUE         DOWEAGVD         CSGADFAS         G SALLMPYS         EGTEXT         EGTEXT         C DEGG LPP SVD/L SYG           A7         TILLIAIALMITTVWWAST         OOPOAHTI CCRTUARUE         DOWEAGVD         CSGADFAS         G SALLMPYS         EGTEXT         C DEGG LPP SVD/L SYG           A9         TILLAIALMITTVWWAST         OOPOAHTI CCRTUARUE         DOWEAGVD         CSGADFAS         G SALLMPYS         EGTEXT         C DEGG LPP SVD/L SYG           A10         CILLAIALMITTVWWAST         OOPOAHTI CCRTUARUE         DOWEAGVD         CSGADFAS         G SALLMPYS         EGTEXT         C DEGG LPP SVD/L SYG           Subfamil V         B1         B1         MITSWHMULVYISUMCSSEA         OEVARTICCRTUARUE         DOPGAVCY         CSGADFAS         G SALLMPYS         EGTEXT         C DEGG CP SVD/L SYG           B1         B2         MITSWHMULVYISUMCSSEA         OEVARTICCRTUARUE         DOPGAVCY         CSGADFAS         G SALLMPYS </td <td>Family A</td> <td>A3</td> <td>MAILLAIALMESTVMWVST</td> <td>QOPOGVHTYCGRHLART</td> <td>NUWEAGVD</td> <td><b>KR</b>SDAQYVS</td> <td>G SAULMPYS</td> <td>EGROKE</td> <td>CINE COLRPOSED LSYC</td>	Family A	A3	MAILLAIALMESTVMWVST	QOPOGVHTYCGRHLART	NUWEAGVD	<b>KR</b> SDAQYVS	G SAULMPYS	EGROKE	CINE COLRPOSED LSYC
A5       TXLLAIALMUTTVMMAST       OOPOTVHTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         A6       TXLLAIALMUTTVMMAST       OOPOANTHTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         A7       TXLLAIALMUTTVMMVST       OOPOANTHTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         A8       TXLLAIALMUTTVMMVST       OOPOANTHTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         A9       TXLLAIALMUTTVMMAST       OOPOANTHTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         A10       TXLLAIALMUTTVMMAST       OOPOANTHTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         B1       MIXTSVMMUVISLICSSEA       OEVARTTRCGHTARTIL       DIMEAGVD       TXSDAGFAS       IG SAUMPYS       EGDOT       INTEGELPT SVDUL STO         B2       MIXTSVMMUVISLICSSEA       OEVARTTRCGHTARTIL       DIMEGVD       TXSDAGAVD       IF       TTROYL       CSTART       VMEGOFPFTLDUL STOG         B1       B4       MIXTSVMUNUVISLICSGEA       OEVARTTRCGHTARTIL       DIG	,	A4	MAILLA I ALMLSTVMWVST	OOPOGVHT CGRHLART A	DIGWEAGVD	SGAOFAS	G SATLMPYS	EGROKE	CINECCLRPCSVDVILSYC
A6       TKILLAIALMLSTVMWYST       OUPOAVHTYGGETLATUL       DIGMEAGVD       TSGAGAFAS       IG       SAUMPYS       EGTSG       IVEGGE PPSVDULSTO         A7       TXLLLAIALMLTIVMWYST       OUPOAVHTYGGETLATUL       DIGMEAGVD       TSGAGAFAS       IG       SAUMPYS       EGTSG       IVEGGE PPSVDULSTO         A8       TXLLLAIALMLTIVMWST       OUPOAVHTYGGETLATUL       DIGMEAGVD       TSGAGAFAS       IG       SAUMPYS       EGTSG       IVEGGE PPSVDULSTO         A9       TXLLLAIALMLTIVMMAST       OUPOAVHTYGGETLATUL       DIGMEAGVD       TSGAGAFAS       IG       SAUMPYS       EGTDG       IVEGGE PPSVDULSTO         A10       MALLAIALMLTIVMMAST       OUPOAVHTYGGETLATUL       DIGMEAGVD       TSGAGAFAS       IG       SAUMPYS       EGTDG       IVEGG PPSVDULSTO         B1       B4       MKTSVMFMLVIVISLINCSSEA       OEVARTYGGETLATUL       DIGMEAVD       TROVL       GSTSR       VMEGGPPFTLUVLSTO         B6       MKTSVMFMLVVISLINCSSEA       OEVARTYGGETLATULT       DIGFGVE       TROAU       GSTSR       VMEGGPPFTLUVLSTO       STGG         B7       MKTSVMFMLVVISLINCSSEA       OEVARTYGETLATUT       DIGFGVE       TROAU       GSTSR       VMEGGPPFTLUVLSTO       STGG         B8       MKTSVMFMLVVISLINCSSEA       OEVARTYGETLATUT       DIGFGVE		A5	MALLLA   ALMLTTVMWAST	<b>OOPOTVHTYCGRHLARTLA</b>	DEWEAGVD	<b>KR</b> SDAQFAS	G SATLMPYS	EGEDOR	GIVDECCLRPCSVDVLLSYC
A7       WILLLA IALMUT IVMINYST       00P0AVHTYCGETHARTER       DISMEAGVD       RISGAOFAS       VG       SAULMPYS       EGIGYT		A6	I ILLA I ALMLSTVMWVST	00P0AVHTYCGRHLARTLA	DEWEAGVD	KRSGAOFAS	G SAULMPYS	EGRGKR	GIVDECCLRPCSVDVLLSYC
A8       MALLTIALMESTYMMYST       OOPOEVHT/CGENLARIM DUGWEAGVD       SISDAOFAS       GLSATUPYS       AGE       ENDEGO PPSVD/LSG         A10       MALLAIALMALTTYMMAST       OOPOAVHT/CGENLARIM DUGWEAGVD       SISDAOFAS       GLSATUPYS       EGBOG       ENDEGO PPSVD/LSG         FamilyB       B1       MALLAIALMALTTYMMAST       OOPOAVHT/CGENLARIM DUGWEAGVD       SISDAOFAS       GLSATUPYS       EGBOG       ENDEGO PPSVD/LSG         B1       B2       MATSWEMLVVISLMCSSEA       OEVART/CGENLADIA       DUGFGVE       SIGGAOYAP       GLSATUPYS       EGBOG       ENDEGO PPSTLD/LSG         B1       B2       MATSWEMLVVISLMCSSEA       OEVART/CGENLADIA       DUGFGVE       SIGGAOYAP       GLTAROYL       SISSAGE       SVDECOFRPCTLD/LSGS         B1       B4       MATSWEMLVVISLMCSSEA       OEVART/CGENLADIA       DUGFGVE       SIGGAOYAP       GLTAROYL       SISSAGE       SVDECOFRPCTLD/LSGS         B6       MATSWEMLVVISLMCSSEA       OEVART/CGENLADIA       DUGFGVE       SIGGAOYAP       GLTAROYL       SISSAGE       SVDECOFRPCTLD/LSGS         B7       MATSWEMLVVISLMCSSEA       OEVART/CGENLADIA       DUGFGVE       SIGGAOYAP       GLTAROYL       SISSAGE       SVDECOFRPCTLD/LSGS         B8       MATSVIEMLVVISLMCSGEA       OEVART/CGENLADIA       DUGFGVE       SIGGAOYAP </td <td></td> <td>A7</td> <td>M LLLA I ALMLT I VMWVST</td> <td>QOPOAVHTYCGRHLARTLA</td> <td>DEWEAGVD</td> <td>KRSGAQFAS</td> <td>G SAULMPYS</td> <td>EGRGKR</td> <td>CIVDECCLRPCSVDVLLSYC</td>		A7	M LLLA I ALMLT I VMWVST	QOPOAVHTYCGRHLARTLA	DEWEAGVD	KRSGAQFAS	G SAULMPYS	EGRGKR	CIVDECCLRPCSVDVLLSYC
A9       WillLalialMLTTVINNAST       OOPANHTGGERUTATUR DUGNEAGYD       GSADAFAS       GG SAUMPYS       EGDOG       SINDEGLPPSSVDULSTG         Family B       Family B       Family B       MixTSVMFILVIVISLMSGEA       OOPANHTGGERUTATUR DUGNEAGYD       GSADAFAS       GG SAUMPYS       EGDOG       SINDEGRPSSVDULSTG         B1       B2       MixTSVMFILVIVISLMSGEA       OEVARTGGERUTATUR DUGNEAGYD       GSAGAYAP       F       TROYL       GSGSG       SVDEGGRPSTLDULSTG         B1       B4       MixTSVMFILVIVISLMSGEA       OEVARTGGERUTATUR DUGNEAGYD       GSGAYAP       F       TROYL       GSGSG       SVDEGGRPSTLDULSTG         B1       B4       MixTSVMFILVIVISLMSGEA       OEVARTGGERUTATUR DUGNEAGYD       GSGAYAP       F       TROYL       GSGSG       SVDEGGRPSTLDULSTG         B6       MixTSVMFILVIVISLMSGEA       OEVARTGGERUTATUR DUGNEAV       GGGAYAP       F       TROHL       GNEKK       SVDEGGRPSTLDULSTG         B10       B10       MixTSVMFILVIVISLMSGEA       OEVARTGGERUTATUR DUGNEAV       GGGAYAP       F       TROHL       GNEKK       SVDEGGRPSTLDULSTG         B11       B12       MixTSVMFILVISLMSGEA       OEVARTGGERUTATUR DUGNEAV       GSGAYAP       F       TROHL       GNEKK       SVDEGGRPSTLDULSTG         Subfamily       B3       Mi		A8	LLLTI ALMESTVMWVST	QOPOEVHTYCGRHLARIMA	DEWEAGVD	KRSDAOFAS	G SAULMPYS	AGE	CIVDECCLRPCSVDVLLSYC
A10       MALLAIALMLTTVMMAST       00P0AVHTVGGRTUPTLA DUGWEAGVD       KENA0YAS       KG SAVLMPYS       EGROF       SUBECOMPTONE         Family B       B1       MARTSVMFMLVIVISLMCSGEA       0EVARTVGGRTUPTLA       DUGFOVE       KGGA0YAP       VF       TIROYL       GSRKK       SVDECOMPTONE       SVDECOMPTON		A9	MALLLA I ALMLTTVMWAST	00P0AVHTYCGRHLAFTLA	DEWEAGVD	KRSDAOFAS	G SAULMPYS	EG DO	CIVECCLRPCSVDVLLSYC
Family B       B1       MittsWrFMLVIVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B1       B2       MittsWrFMLVFVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B1       B4       MittsWrFMLVFVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B6       MittsWrFMLVVVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B7       MittsWrFMLVVVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B8       MittsWrFMLVVVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B10       MittsWrFMLVVVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B10       MittsWrFMLVVVISLMCSGEA       GEVARTYGERHLOTUL DUGFOVE       KrGGAQYAP       VF       TROYL       GS36XR       WDECGFRPCTLDULSTGE         B10       MittsWrFMLVVVISLMCSGEA       GEVARTYGERHLOTUL		LA10	M LLLAI ALMLTTVMWAST	<b>OOPOAVHTYCGRHLARTLA</b>	DCWEAGVD	KESNAQYAS	🔓 SAÜLMPYS	EGIGOE	GIVDECCLRP
Subfamily       B2       MXTSWHRLVFVISLICSSEA       OEVARTXGERHADILY       DIGFOVE       GSGA0YAP       VF       TROYL       GSGKF       SWDEGERPETLDUL SYGG         B1       B4       MXTSWHRLVTVISLICSGEA       OEVARTYGERHADILY       DIGFOVE       GGA0YAP       VF       TROYL       GSGKF       SWDEGERPETLDUL SYGG         B6       MXTSWHRLVTVISLICSGEA       OEVARTYGERHADILY       DIGFOVE       GGA0YAP       VF       TROYL       GSGKF       SWDEGERPETLDUL SYGG         B6       MXTSWHRLVTVISLICSGEA       OEVARTYGERHADILY       DIGFOVE       GGA0YAP       VF       TROYL       GSGKF       SWDEGERPETLDUL SYGG         B7       MXTSWHRLVTVISLICSGEA       OEVARTYGERHADILY       DIGFOVE       GGA0YAP       VF       TROYL       GSRKF       SWDEGERPETLDUL SYGG         B8       MXTTSWHRLVTVISLINYSGEA       OEVARTYGERHADILY       DIGFOVE       GGA0YAP       VF       TROYL       GSRKF       SWDEGERPETLDUL SYGG         Subfamily       B1       B10       MXTTSWHRLVVISLITYSSEE       OEVARTYGERHANILY       DIGFOVE       GGA0YAP       VF       TROYL       GSRKF       SWDEGERPETLDUL SYGG         Subfamily       B5       MXTTSWHRLVVISLTYSSEE       OEVARTYGERHANILY       DIGFOVE       GGA0YAP      VF       TROYL       GSRKF </td <td>Family B</td> <td><b>F</b>B1</td> <td>MMTSVMFMLVIVISLMCSGEA</td> <td>0EVARTY<u>CGRHLADTLA</u></td> <td>DEFGVE</td> <td>KRGGAQYAP</td> <td>F TROYL</td> <td>GSRGKE</td> <td>GVVDECCFRPCTLDVLLSYCG</td>	Family B	<b>F</b> B1	MMTSVMFMLVIVISLMCSGEA	0EVARTY <u>CGRHLADTLA</u>	DEFGVE	KRGGAQYAP	F TROYL	GSRGKE	GVVDECCFRPCTLDVLLSYCG
BI       B4       MXTSVMFMLV11SLMCSGEA       QEVART/CGRILADILA       DUCFOVE       KTGGAQYAP       VF       TROYL       GSBKR       SV/DECCFRPGTLD/LLSYG         B6       MXTSVMEML/VV1SLMCSGEA       QEVART/CGRILADILA       DUCFOVE       KTGVAQYAP       VF       TROYL       GSBKR       SV/DECCFRPGTLD/LLSYG         B7       MXTSVMEML/VV1SLMCSGEA       QEVART/CGRILADILA       DUCFOVE       KTGGAQYAP       VF       TROYL       GSBKR       SV/DECCFRPGTLD/LLSYG         B8       XTSVIFVL1VLNLMMSGEA       QEVART/CGRILADILA       DUCFOVE       KTGGAQYAP       VF       TROYL       GSBKR       SV/DECCFRPGTLD/LLSYG         B10       XTTILFLVVISLMYSGEA       QEVART/CGRILADILA       DUCFOVE       KTGGAQYAP       VF       TROYL       GSBKR       SV/DECCFRPGTLD/LLSYG         B11       B11       B11       B11       B11       MXTTIMFML/VVISLT/SSEE       QEVART/CGRILANILA       DUCFOVE       KTGGAQYAP       VF       TROYL       GSBKR       SV/DECCFRPGTLD/LLSYGG         Subfamily       B5       MXTTIMFL/VVISLT/SSEE       QEVART/CGRILANILA       DUCFOVE       KTGGAQYAP       VF       TROYL       GSBKR       GV/DECCFRPGTLD/LLSYGG         Subfamily       B5       MXTAVMFIL/VVISLT/SSEE       QEVART/CGRILANILA       TYOFGVE       KTGGAQ	Subfamily	B2	MMITSVMFMLVFVISLMCSSEA	0EVARTYCGRHLADTLA	D CFGVE	KRSGAQYAP	F TROYL	GSRGKR	GVVDECCFRPCTLDVLLSYCG
B6       MXTSVMFMLVV1SLMCSSEA       OEVARTYCGETDADILA       DXGFGVE       KRGVAQYAP       IF       TTROYL       GS35KF       GV/DECGFRPGTLD/ULSYGG         B7       MXTSVMLMLVV1SLICSGEA       OEVARTYCGETLADILA       DXGFGVE       KRGAQYAP       VF       TTROYL       GS35KF       GV/DECGFRPGTLD/ULSYGG         B8       MXTSVIFVLIVLNLMMSGEA       OEVARTYCGETLADILA       DXGFGVE       KRGAQYAP       VF       TTROYL       GS35KF       GV/DECGFRPGTLD/ULSYGG         Subfamily       B3       MXTTIMFML/VVISLTYSSEE       OEVARTYCGETLADILA       DXGFGVE       KRGAQYAP       VF       TTROYL       GS35KF       SV/DECGFRPGTLD/ULSYGG         Subfamily       B3       MXTTIMFML/VVISLTYSSEE       OEVARTYCGATLANILA       DXGFGVE       KRGAQYAP       VF       TTROYL       GS36KF       SV/DECGFRPGTLD/ULSYGG         Subfamily       B1       B1       B1       MXTAVMFIL/VVISLTYSSEE       OEVARTYCGATLANILA       DXGFGVE       KRGAQYAP       VF       TTROYL       GS36KF       SV/DECGFRPGTLD/ULSYGG         Subfamily       B1       B1       B1       MXTAVMFIL/VVISLTYSSEE       OEVARTYCGATLANILA       YVGFGVE       KRGAQYAP       VF       TTROYL       GS36KF       SV/DECGFRPGTLD/ULSYGG         B11       B11       MXTAVMFIL/VVISLTYSSEE <td< td=""><td>BI</td><td>B4</td><td>MMATSVMFMLVIVISLMCSGEA</td><td>QEVARTY<u>CGRHLA</u>DTLA</td><td>D CFGVE</td><td>KRGGAQYAP</td><td>F TROYL</td><td>GSRGKR</td><td>GV<b>VDECC</b>FRP<mark>C</mark>TLD<b>VLLSYC</b>G</td></td<>	BI	B4	MMATSVMFMLVIVISLMCSGEA	QEVARTY <u>CGRHLA</u> DTLA	D CFGVE	KRGGAQYAP	F TROYL	GSRGKR	GV <b>VDECC</b> FRP <mark>C</mark> TLD <b>VLLSYC</b> G
B7       MXRTSVMLMLVV1SLICSGEA       OEVARTICORTINATULA DICFOVE       KRSGAQYAP       IF       THOHL       GN30KF       GV/DECGFRPGTLD/TLSTG         B8       MXRTSVIFULIVLULINUMSGEA       OEVARTICORTINATURA       DECFGVE       KRSGAQYAP       IF       THOHL       GN30KF       GV/DECGFRPGTLD/TLSTG         B10       MXRTTIMFML/VV1SLTYSSEE       OEVARTICORTINATURA       DECFGVE       KRSGAQYAP       IF       THOYL       GS36KF       SV/DECGFRPGTLD/TLSTG         B11       B1       MXRTTIMFML/VV1SLTYSSEE       OEVARTICORTINATURA       DECFGVE       KRSGAQYAP       IF       THOYL       GS36KF       SV/DECGFRPGTLD/TLSTG         Subfamily       B3       MXRTTIMFML/VV1SLTYSSEE       OEVARTICORTINATURA       DECFGVE       KRSGAQYAP       IF       THOYL       GS36KF       SV/DECGFRPGTLD/TLSTG         Subfamily       B4       B1       MXRTTMFL/VV1SLTYSSEE       OEVARTICORTINATURA       DECFGVE       KRSGAQYAP       IF       THOYL       GS36KF       SV/DECCFRPGTLD/TLSTGS         B11       B9       MXRTATMFL/VV1SLTYSSEE       OEVARTICORTINATURA       VCFGVE       KRGAQYAP       IF       THOYL       SSRK       GF2V/DECCFRPGTLD/TLSTGS         B11       MXRTATMFL/VV1SLTYSSEE       OEVARTICORTINATURA       VCFGVE       KRGAQYAP       IF		B6	MMSTSVMFMLVVVISLMCSSEA	OEVART YCGRDLADTLA	DCFGVE	KRGVAQYAP	F TROYL	GSRGKR	GVVDECCFRPCTLDVLLSYCG
B8       MXTSVIFVLIVULNUMWSGEA       OEVARTICGSTEADULA       DEGGVV       KRGGAQYAP       VF       MXAXL       GS36KF       SVDECGFRPGTLD/ULSYG         B10       MXTTIHFML/VVISLTYSSEE       OEVARTICGSTEADULA       DEGGVV       KRGGAQYAP       VF       TROYL       GS36KF       SVDECGFRPGTLD/ULSYG         B11       B1       B1       B1       MXTTIHFML/VVISLTYSSEE       OEVARTICGATEANELA       DEGGVE       KRSGAQYAP       VF       TROYL       GS36KF       SVDECGFRPGTLD/ULSYG         Subfamily       B5       MXTTIHFML/VVISLTYSSEE       OEVARTICGATEANILA       DEGGVE       KRSGAQYAP       VF       TROYL       GS36KF       SVDECGFRPGTLD/ULSYG         Subfamily       B5       MXTATMFIL/VVISLTYSSEE       OEVARTICGATEANILA       PUGFGVE       KRGGAQYAP       VF       TROYL       GS36KF       SVDECGFRPGTLD/ULSYG       SVGEC         B111       B11       MXTATMFIL/VVISLTYSSEE       OEVARTICGTEANILA       YQFGVE       KRGGAQYAP       VF       OETYL       SRK       GP       GP       SVDECGFRPGTLD/ULSYGC       SVGEC         B11       MXTATMFIL/VVISLYSSEE       OEVARTICGTEANILA       YQFGVE       KRGGAQYAP       VF       OETYL       SRK       GP       GP       GP       GP       GP       GP       GP		B7	MMATSVMLMLVVVISLICSGEA	QEVARTY <u>CGRHLADTLA</u>	DCFGVE	KRSGAQYAP	F TROHL	GNRGKR	GVVDECCFRPCTLDVLLSYCG
Leito Marti Li FLVVI SLMYSGEA OEVAHTI CORTH ADULA DEGROVE KISGAQYAP VF UTROYL GS30KF SVDEOCFRPGTLD/ULS3CO Subfamily FB3 MMKTT IMFML/VVI SLTYSSEE OEVARTI CORTH ADULA DEGROVE KISGAQYAP VF UTROYL GS30KF SVDEOCFRPGTLD/ULS3CG B11 B12 MMKTT IMFML/VVI SLTYSSEE OEVARTI CORTH ADULA DEGROVE KISGAQYAP VF UTROYL GS30KF SVDEOCFRPGTLD/ULS3CG Subfamily FB5 MMKTAVMFIL/VVI SLTYSSEE OEVARTI CORTH ADULA DEGROVE KISGAQYAP VF UTROYL GS30KF SVDEOCFRPGTLD/ULS3CG B11 B11 MKTAVMFIL/VVI SLTYSSEE OEVARTI CORTH ADULA DEGROVE KISGAQYAP VF UTROYL GS30KF SVDEOCFRPGTLD/ULS3CG B9 MKTAVMFIL/VVI SLTYSSEE OEVARTI CORTH ADILA VOGOVE KISGAQYAP V OETVL BSRK GP2VDEOCFRPGXLE/UKS7CGV B11 MKTAVMFIL/VVI SLTYSSEE OEVARTI CORTH ADILA VOGOVE KIGGAQYAP V OETVL BSRK GP2VDEOCFRPGXLE/UKS7CGV B11 MKTAVMFIL/VVI SLTYSSEE OEVARTI CORTH ADILA VOGOVE KIGGAQYAP V OETVL BSRK GP2VDEOCFRPGXLE/UKS7CGV C1 MKAVMFIL/VVI SLTYSSEE OEVARTI CORTH ADILA VOGOVE KIGGAQYAP V OETVL BSRK GP2VDEOCFRPGXLE/UKS7CGV C2 MKLVI LLVVVSAML/LGGA OTASOF KODF KRIMS STEWSDMO KISGSQYAG K G IVMLPP FSSS30KF GIVDEOCYRPGTT DVIKLYOKON LCG C3 MKLVI LLVVVSAML/LGGA OTASOF KODF KRIMS I CMPDMP KISGSQYAG G IPWLPP FSSS30KF GIVDEOCYRPGTT DVIKLYOKON LCG C4 MKLVI LLVVVSAML/LGGA OTASOF KODF KRIMS I CMPDMP KISGSQYAG G IPWLPP FSSS30KF GIVDEOCYRPGTT DVIKLYOKON LCG C5 MKLVI LLVVVSAML/LGGA OTASOF KODF KRIMS I CMPDMP KISGSQYAG G IPULPS LSEE30KF GIVDEOCYPGTT DVIKLYOKON LCG C6 MKLVI LLVVVSAML/LGGA OTASOF KODF KRIMS I CMPDMP KISGSQYAG G IPULPS LSEE30KF GI ADECOLOPGT NDVIL SYO C6 MKLVI LLVVVSAML/LGGA OTASOF KODF KRIMS I CMPDMP KISGSQYAG G IPULPS LSEE30KF GI ADECOLOPGT NDVIL SYO Family E E1 NNPPVFLVLLLTGFLCI AA OEANVAHH KORTH ANTOK DICMRAGFE KISVAHYAG I G IPLLPS LSEE30KF GI ADECOLOPGT NDVIL SYO Family F F1 KLVVI LVI SVSILVSS OE OEANVAHH KORTH ANTOK DICMRAGFE KISVAHYAG G IPULP ADAXKIF GVVDEOCIOPGT LDVILATYO Family F F1 KLVVI LVI SVSILVSS OE OEANVAHH KORTH ANTOK DICMRAGFE KISVAHYAG G IPPLP ADAXKIF GUIDECOLOPGT NDVIL SYO		B8	MATSVIFVLIVLNLMWSGEA	QEVARTIY <u>CG</u> S <u>HLA</u> DTLA	DCFGVV	KRGGAQYAP	ĨF ∐OKAYL	GSRGKR	CVVDECCFRPCTLDVLASYCG
Subfamily       FB3       M&TTIMPMLVVISLTYSSEE       OEVARTYCGATLANILA       DEGGVE       KSGAQYAP       VF       TROYL       GS10KE       SVDECOFPPGTLDVLSTGG         BII       BII       MXTAVMFILVVVISLTYSSEE       OEVARTYCGATLANILA       DEGGVE       KSGAQYAP       VF       TROYL       GS10KE       SVDECOFPPGTLDVLSTGG         Subfamily       FB       MXTAVMFILVVVISLTYSSEE       OEVARTYCGATLANILA       PUGFGVE       KSGAQYAP       VF       TROYL       GS10KE       SVDECOFPPGTLDVLSTGG         BIII       B9       MXTAVMFILVVVISLTYSSEE       OEVARTYCGATLANILA       PUGFGVE       KGGAQYAP       VF       UTROYL       SSK       GPVDECOFPPGTLDVLSTGG         BIII       B9       MXTAVMFILVVVISLTYSSEE       OEVARTYCGATLANILA       PUGFGVE       KGGAQYAP       VF       OETYL       SSK       GPVDECOFPPGTLDVLSTGGV         Family C       C1       TRUMULVVISAMLVLGGA       OTASOF COFF ARTING STONDMO       TGSOSQYAG       VG       JPMLPP       FSSS10KE       GIVDECOFPGTLDVLSTGOV         C3       TRUVILLVVSAMLVLGGA       OTASOF COFF ARTING STONDMO       TGSOSQYAG       VG       JPMLPP       FSSS10KE       GIVDECOFPGTLDVLSTGOV         C4       MXUVILLVVSAMLVLGGA       OTASOF       OTASOF       GEVARDGFE       KSVAHYAG       VG		LB10	TILIFLVVISLMYSGEA	QEVAHT <mark>YCGHILA</mark> DTLA	DEGVE	KRSGAQYAP	F TROYL	GSRGKR	SV <b>VDECC</b> FRP <mark>C</mark> TLD <u>VL</u> LSYCD
BII       LB12       MIXTAVMFILVVISLTYSSEE       OEVARTI/CGATLANILA       DIGFOVE       KSGAQYAP       IF       MIXTAVMFILVVISLTYSSEE       OEVARTI/CGATLANILA       DIGFOVE       KSGAQYAP       IF       MIXTAVMFILVVISLTYSSEE       OEVARTI/CGATLANILA       DIGFOVE       KSGAQYAP       IF       MIXTAVMFILVVISLTYSSEE       OEVARTI/CGATLANILA       VOFGVE       KGGAQYAP       IF       MIXTAVMFILVUSLA       CF       KGGAQYAP       IF       MIXTAVMFILVUSLA       CF       KGGAQYAP       IF       MIXTAVMFILVUSLA       CF       KGGAQYAP       IF       MIXTAVMFILVUSLA       CF       KGGAQYAP       IF       MIXTAVMFILVUSLA       CF       KGGAQYAP       IF       MIXTAVMFILVUSLA       CF       KGGAQYAP       IF       MIXTAVMFILVUSLANILVG	Subfamily	LB3	MMATTIMEMLVVVISLTYSSEE	OEVART <u>YCG</u> A <u>HLANILA</u>	D CFGVE	KESGAQYAP	F TROYL	GSECKE	<u>CVVDECC</u> FRPCTLD <u>XLLSYC</u> G
Subfamily       E5       Militat/WirFiL/WV1SLTYSSEE       OEVARTICGETEANILAR YQFGVE       KrGGA0YAP       MOETYL       ISRK       GP2VDECOFFPS/LEVENSYGV         B111       B111       Militat/WirFiL/WV1SLTYSSEE       OEVARTICGETEANILAR YQFGVE       KrGGA0YAP       MOETYL       ISRK       GP2VDECOFFPS/LEVENSYGV         B111       Militat/WirFiL/WV1SLTYSSEE       OEVARTICGETEANILAR YQFGVE       KrGGA0YAP       MOETYL       ISRK       GP2VDECOFFPS/LEVENSYGV         Family C       C1       Militat/WirFiL/WV1SLTYSSEE       OEVARTICGETEANILAR YQFGVE       KrGGA0YAP       MOETYL       ISRK       GP2VDECOFFPS/LEVENSYGV         C2       Militat/WirFiL/WV1SLTYSSEE       OEVARTICGETEANILAR YQFGVE       KrGGA0YAP       MOETYL       ISRK       GP2VDECOFFPS/LEVENSYGV         C3       Militat/WirFiL/WV1SAML/VGGA       OTASOF/COFF       KrSGS0YAG       G       PWLPP       FSSSTERE       GI MDECOFPS/LEVENSYCON         C4       Militat/V1L/WVSAML/LGGA       OTASOF/COFF       KrSVAHYAG       G       PWLPP       FSSSTERE       GI ADECOLOPGTID/UL/VICKLYCDK01T1         C5       Militat/V1L/WVSAML/LGGA       OTASOF/L       G       GR2       G       PWLPP       FSSSTERE       GI ADECOLOPGTID/UL/VICKLYCDK01T1       G         C6       Militat/V1L/WVSAML/LGGA       OTASOF/L       G </td <td>BII</td> <td>LB12</td> <td>MMATTIMFMLVVVISLTYSSEE</td> <td>QEVARTY<u>CGAHLANTLA</u></td> <td>DCFGVE</td> <td>KESGAQYAP</td> <td>F TROYL</td> <td>GS <u>CKR</u></td> <td><u>CVVDECC</u>FQPCTLD<u>V</u>LS<u>YC</u>G</td>	BII	LB12	MMATTIMFMLVVVISLTYSSEE	QEVARTY <u>CGAHLANTLA</u>	DCFGVE	KESGAQYAP	F TROYL	GS <u>CKR</u>	<u>CVVDECC</u> FQPCTLD <u>V</u> LS <u>YC</u> G
BIII       B9       Militatvine ILVVVI SLTYSSEE       OEVARTICOGHLANILA YQFGVE       KIGGAQYAP       MOETYL       ISRK       GPV/DECCFRP9/LEVILSUSGSV         B111       MILITAVME ILVVVI SLTYSSEE       OEVARTICOGHLANILA YQFGVE       KIGGAQYAP       MOETYL       ISRK       GPV/DECCFRP9/LEVILSUSGSV         Family C       C1       MILITAVME ILVVSAMLVLGGA       OTASOF/CODFLARMS STONSDMQ       KISGSQYAG       MOETYL       ISRK       GPV/DECCFRP9/LEVILSUSGSV         C2       MILVILLVVSAMLVLGGA       OTASOF/CODFLARMS STONSDMQ       KISGSQYAG       MG       PMULPP       FSSSTEKE       GIVECOYP9/GTD/LKLY0KG/HA/MSYDN         C3       MILVILLVVSAMLVLGGA       OTASOF/CODFLARMS STONSDMQ       KISGSQYAG       MG       PMULPP       FSSSTEKE       GIVECOYP9/GTD/LKLY0KG/HI/MSYDN         C4       MILVILLVVSAMLVLGGA       OTASOF/CODFLARMS       GEVANDAGE       GSSQYAG       G       PPULPP       FSSSTEKE       GIVECOYP9/GTD/LKLY0KG/HI/MSYDN         C6       MILVILLVVSAMLVLGGA       OTASOF/       GG       GEVANLVLGGA       OTASOF/       GG         Family D       D1       MILCFFLSWSVCAIVSA       SECOHL/GGYPAYKM/ DICWRAGFE       KSVAHYAG       G       GPLLPS       LSEE_EXKE       GIADECCLOPG/TND/L SYC         Family E       E1       NRPVFLVLLLIGGLCIAA       GEANVAHH/CGRHUNIUA <td>Subfamily</td> <td>FB5</td> <td>MMATAVMFILVVVISLTYSSEE</td> <td>OEVART<u>YCGRHLA</u>N I LA</td> <td>YVOFGVE</td> <td>KEGGAQYAP</td> <td>0ETYL</td> <td>*SRK</td> <td>GPSVVDZCCFRPCKLEVIKSYCGV</td>	Subfamily	FB5	MMATAVMFILVVVISLTYSSEE	OEVART <u>YCGRHLA</u> N I LA	YVOFGVE	KEGGAQYAP	0ETYL	*SRK	GPSVVDZCCFRPCKLEVIKSYCGV
EB11       MXXITAVWFILVVVISLIVSSE       OEVARIVCCCTUANILATVOFOVE       KCGSAQYAP       MOETYL       ISRK       GP2V0/BCCFHP2KLE/MXSIFFFFCD         Family C       C1       MXXITAVWFILVVVSAMLVLGGA       OTASOF/CCDFLATIMS STONSDMO       Krscsoyag       MG       PWLPP       FSSSTERF       CIV/DECCHP2KLE/MXSIFFFFCD         C2       MXXITAVWFILVVVSAMLVLGGA       OTASOF/CCDFLATIMS STONSDMO       Krscsoyag       MG       PWLPP       FSSSTERF       CIV/DECCHP2KTD/LKL/WSYCDN         C3       MXXITAVWFILVVVSAMLVLGGA       OTASOF/CCDFLATIMS I/COMPDMP       Krscsoyag       MG       PWLPP       FSSSTERF       CIV/DECCHP2KTD/LKL/YDK0ITI         C4       MXXITAVWFILVVSAMLVLGGA       OTASOF/CCDFLATIMS       CFW       Krscsoyag       MG       PWLPP       FSSSTERF       CIV/DECCHP2KTD/LKL/YDK0ITI         C5       MXXIVILVVSAMLVLGGA       OTASOF/CCDFL       CFW       Krscsoyag       G       PWLPP       FSSSTERF       CIV/DECCHP2KLYD/LKLYD/CK0ITI         C6       MXIVILVVSAMLVLGGA       OTASOF/CCDFL       CFW       Krscsoyag       G       PLLPS       LSEE       LSEE       Krscsoyag       G       PLLPS       LSEE       CFW	BIII	B9	MMATAVMFILVVVISLTYSSEE	QEVARTYCCHLLAN I LA	YVOFGVE	GGAQYAP	MOETYL	*SRK	GPCVVIII CCOFRPOKLEMIKSYCGV
Family C       C1       KLVMLLVVVSAMLVLGGA       OTASOF_VCOPF_ANIMS_SECVISDMO       KLSSSVAG       VG       PMLPP       FSSSMAR       STUDECCYPPOTTDVLASTODN         C2       KLVILLVVVSAMLVLGGA       OTASOF_VCOPF_ANIMS_ICOPF_ANIMS_ICOPPOTTDVLASTODN       KLSSSVAG       VG       PMLPP       FSSSMAR       STUDECCYPPOTTDVLASTODN         C3       KLVILLVVVSAMLVLGGA       OTASOF_VCOPF_ANIMS_ICOPF_ANIMS_ICOPPOTTDVLASTODN       KLSSSVAG       VG       PMLPP       FSSSMAR       STUDECCYPPOTTDVLASTODN         C4       KLVILLVVVSAMLVLGGA       OTASOF_V       C5       KLVILLVVVSAMLVLGGA       OTASOF_V       C6       KLVILLVVVSAMLVLGGA       OTASOF_V         C5       KLVILLVVVSAMLVLGGA       OTASOF_V       C6       KLVILLVVVSAMLVLGGA       OTASOF_V         C6       KLVILLVVVSAMLVLGGA       OTASOF_V       KSVAHYAG       VG       VG       VG         Family D       D1       KLVILLVVSAMLVLGGA       OTASOF_V       KSVAHYAG       VG       VG       VG       VG         Family D       D1       KLVILLVVSAMLVLGGA       OEANVAHYAG       VG       VG       KSVAHYAG       VG       VG <t< td=""><td></td><td>LB11</td><td>MMATAVMFILVVVISLTYSSEE</td><td>QEVART CGELLANI</td><td>YVG-GVE</td><td>GGAQYAP</td><td>QETYL</td><td>SRK</td><td>GPCVVIII CCCFRPCKLEVIIKSFFFFCD</td></t<>		LB11	MMATAVMFILVVVISLTYSSEE	QEVART CGELLANI	YVG-GVE	GGAQYAP	QETYL	SRK	GPCVVIII CCCFRPCKLEVIIKSFFFFCD
C2 WALVILLVVVSAMLVLGGA OTASOF C3 WALVILLVVVSAMLVLGGA OTPSOF C4 WALVILLVVVSAMLVLGGA OTASOF C5 WALVILLVVVSAMLVLGGA OTASOF C6 WALVILLVVVSAMLVLGGA OTASOF C6 WALVILLVVVSAMLVLGGA OTASOF Family D D1 WALGFFLSWVSVCAIVSA SEGHIVCGRYLAYKM DIGWRAGFE KESVAHYAG VG VPLPS LSEERKE GIADEGOLOPOTIDVILSYC Family E E1 WAPPYELVLLITGFLCIAA GEANVAHIVCGENEANITA DIGWRAGFE KESVAHYAG VG PHLPS LSEERKE GIADEGOLOPOTIDVILSYC Family F F1 WALVVILVISVSILVSA GELGGSRRVCGENEANITA DIGWRAGFE KESVAHYAG VG MPELLP ADARKE GIIDEGOLOPOTIDVILSYC	Family C		MALVMLLVVVSAMLVLGGA	QTASQFY <b>CC</b> DF <b>F</b> RIMS	SICWSDMQ	KESGSQYAG	G PWLPP	FSSS CKE	CIVELCCYRPOTIDE MSYCON
C3 WALVILLVVVSAMLVLGGA UTPSDFY C4 WALVILLVVVSAMLVLGGA UTASOFY C5 WALVILLVVVSAMLVLGGA UTASOFY C6 WALVILLVVVSAMLVLGGA UTASOFY Family D D1 WALGFFLSWVSVCATVSA SEEGHTYCOTYLVKMA DIOWRAGFE KRSVAHYAG VG VPLLPS LSEERSKE CTADECOLOPOTIDVLLSYC Family E E1 WNPVFLVLLLTGFLCTAA QEANVAHHYCOTHLANTUA DIOWRAGFE KRSVAHYAG VG VPLLPS LSEERSKE CTADECOLOPOTIDVLLSYC Family F F1 WALVVIVLLVTSVLLVSSLUSSA QELGGSRR/COTHLANTUA DIOWRDTSVE KRSESSLAS VSSRG/PWLPT PNFNKRATKKE CVVDECCTOPOTIDVLLSYC Family F F1 WALVVIVLLVTSVLLVSSLUSSA QELGGSRR/COTHLANTUA DIOWRDTSVE KRSESSLAS VSSRG/PWLPT PNFNKRATKKE CVVDECCTOPOTIDVLLSYC		C2	MALVILLVVVSAMLVLGGA	QTASOF CCDF RIMS	T CWPDMP	KESGSQYAG	G PWLPP	FSSS_C <u>KR</u>	GIVIECOYRPOIIDVIEKLYCOKOIII
C4     MILVILLYVVSAMLVLGGA     0TASOFY       C5     MILVILLVVVSAMLVLGGA     0TASOFY       L6     MILVILLVVVSAMLVLGGA     0TASOFY       Family D     D1     MILLGFLSWVSVALVSA     SEGHI VCCTVLVVKM       Family E     E1     MRPVFLVLLLTGFLCIAA     QEANVAHH       Family F     F1     MILVILVVSSILVSA     OELGGSRR/CGCHLANIUA       Family F     F1     MILVILVISSILVSA     QELGGSRR/CGCHLANIUA		C3	MALVILLVVVSAMLVLGGA						
CS WALVILLVVVSAMLVLGGA UTASUFY C6 WALVILLVVVSAMLVLGGA OTASOFY Family D D1 WALLGFLSWVSVCATVSA SEECHIVCCTVLAYKWA DICWRAGFE KRSVAHYAG VG UPLLPS LSEE <u>BSKE</u> GLADECCLOPGTNDVLLSYC Family E E1 MNRPVFLVLLLTGFLCIAA QEANVAHHVCCFHUNTUA DICWRAGFE KRSVAHYAG VG WELLP ADASKKE GVVDECGLOPGTLDVLATVC Family F F1 WALVIVLVISVSLLVSA QELGSSRRVCCFHUADIWA VICWGIDEMSAE KRNDMVYE DSG MPELLP ADASKKE GUIDECCLAGTRDVILLSYC		C4	MALVILLVVVSAMLVLGGA	UTASUF1					
ECG WALLYNLLYWSAMLUNGGA UTASUFW Family D D1 MALLEFELSWYSVCATVSA SECHTWOORFYLAYKW DIGWRAGFE KESVAHYAG IG IPLLPS LSEEISKE GLADEGOLOPOTIDWLLSYC Family E E1 MNPYVFLVLLITGFLCIAA QEANVAHWIGGEHLANILA DIGWDTSVE KESESLAS YSSROïPWLPT PNFNKRATKKE GVIDEGOLOPOTIDWLATYC Family F F1 MALYVIVLLYISVSILVSA QELGSSRPYCGEHLAQUW YGWSIDEMSAE KENSDWYF DSG MPELLP ADARKE GIJEGOLOACTROVILLSYC		105	MALVILLVVVSAMLVLGGA						
Family D DI MALLGH-LSWYSVATVSA SEEGHILGERT AND DEWHAGFE KISVAHYAG IG UPLLPS LSEEMSE GTADEGGLOPGINUAL STO Family E E1 INRPVFLVLLLTGFLCIAA OEANVAHHICGERLANILL DIGWDTSVE KISESSLAS KISSRGÜPWLPT PNFNKRAIKKE GVDEGGLOPGILDXLATIO Family F F1 MALVVIVLLVISVSILVSA OELGESRR GERLANILL DIGWDTSVE KINSDMVYE DSG MPELLP ADAKKE GTI EGOLOAGTROXILSYO	<b>F</b> 1 <b>B</b>	LCO					30 10 LDC		
Family E EI WHPVELVELLIGELCIAA GEANVARHIGGEREANDE DEGNOSVE KESSLAS KESKGWEPI PNENKRAIKKE GVØLEGCIOPOTELADATIO Family F F1 KELVVIVLLVISVILVSA OELGESRRYGGEREANDE KINSDAVYE DSG MPELLP ADASKKE EIIJEGELGAGTROVILSVO	Family D			SEEGHT CG YLAYKMA		KESVAHYAG	C PLLPS		
	Family E					MISESSLAS	MOON NOTION		
	Family F	C1				MINSUMVIE	ON MPELLP		
rainity o o <u>matter protocological and a constantino o constantino de constantino de Constantino de constantino de</u>	Family 6	01					OUN UPLLAIS		

**Fig. 2.** Amino acid sequences of *Bombyx* preprobombyxins deduced from nucleotide sequences of their genes and human preproinsulin. Solid boxes represent identical residues among all the insulin family prepropeptides listed and hatched boxes those among preprobombyxins. Amino acid residues are numbered from the N terminus of each domain of insulin. Amino acid sequences are from Bell *et al.*, 1980; Iwami *et al.*, 1989; Kondo *et al.*, 1996; Tsuzuki *et al.*, 1997; Yoshida *et al.*, 1997, 1998.

peptide and those of the signal peptide within each family show a remarkably high level of conservation between families. High similarities are found among bombyxin members in cysteine and hydrophobic residues responsible for the hydrophobic core formation of bombyxin and insulin (Jhoti et al., 1987). The surface patch formed by the central part of the B chain is of critical importance for recognition by the bombyxin receptor (Nagata et al., 1995). Intriguingly, the amino acid sequences of the surface patch are conserved or conservatively substituted among bombyxin members. Also conserved is glycine at position 8 of the B chain. In insulin, the glycine residue maintains the conformation through its contribution to the mainchain turn. The role of the N terminus of the bombyxin II A chain was investigated using bombyxin analogues with modifications (Büllesbach, 1999). Modification of the A chain dramatically reduced the affinity to the bombyxin receptor, suggesting that bombyxin exhibits a mammalian insulin-like structural sensitivity for the A chain modifications. The paired basic amino acid residues are also conserved among preprobombyxins, which suggests that the mature bombyxins are generated through excision of the C peptide, exactly the same as insulin is (Chan et al., 1981).

Interestingly, the signal peptide sequences of the subfamily BII members are more similar to those of the subfamily BIII members than to those of the subfamily BI members, whereas, in contrast, the C peptide and A chain sequences of the subfamily BII members are more similar to those of the subfamily BI members than to those of the subfamily BIII members (Fig. 2). This tendency is also pronounced in the nucleotide sequences of the three subfamily members. A subfamily BII prototype gene therefore must be generated by a crossing-over in the B chain region between prototype genes belonging to the subfamilies BI and BIII (see also Fig. 4).

### PHYSIOLOGICAL FUNCTION OF BOMBYXIN

The Bombyx brain contains two types of molecules that stimulate the synthesis and release of ecdysone, i. e. bombyxin and PTTH (Ishizaki et al., 1983). Bombyx bombyxin can induce adult development when injected into brain-removed dormant pupae of Samia. The amount of bombyxin that induces adult development is estimated to be roughly 0.1-0.4 ng (Nagasawa et al., 1984). When Bombyx bombyxin was injected into brain-removed Bombyx pupae, even at amounts than 1,000 times, it did not elicit adult development, whereas only 0.1 ng of the Bombyx PTTH was effective in the brainremoved Bombyx pupal assay. In addition, the dose required for attaining maximum activation by bombyxin was as high as 1,600 times of that of PTTH when assayed using in vitro culture of the Bombyx pupal prothoracic glands (Kiriishi et al., 1992). Therefore, it was concluded that Bombyx bombyxin is not a "pure" Bombyx PTTH active in Bombyx itself (Kawakami et al., 1990). The insulin-related structure of bombyxin both in its primary and tertiary forms, however, strongly suggests that bombyxin plays important roles in sugar metabolisms and growth of insects.

Bombyx haemolymph contains a large amount of trehalose at a concentration of about 10 mM (Oda et al., 1997). Bombyxin II was shown to have the hypotrehalosemic activity that reduces the concentration of trehalose, when injected to the neck-ligated larvae (Mizoguchi, 1994; Satake et al., 1997). Although bombyxin II shows the activity in a dose-dependent manner, the dose required to elicit a half maximal effect is 10 ng, which is higher than the amount in the feeding Bombyx larvae. It is not known whether bombyxin molecules other than bombyxin II also have a hypotrehalosemic activity. The concentration of trehalose in the haemolymph is controlled by the relative rates of trehalose synthesis in the fat body and trehalose uptake by the tissues. Bombyxin II was demonstrated to cause elevated trehalase activity in the Bombyx midgut and muscles, which suggests that bombyxin induces hypotrehalosemia by promoting the hydrolysis of trehalose to glucose and thereby facilitating its transport into tissue cells (Mizoguchi, 1994; Satake et al., 1997). Bombyxin II was also shown to reduce the glycogen content in the fat body by elevating the percentage of the active form of glycogen phosphorylase (Satake et al., 1997). In the adult moths, bombyxin II, however, does not affect the concentration of haemolymph trehalose nor the activity of trehalase in the muscles, which indicates the occurrence of a change in the action of bombyxin during metamorphosis (Satake et al., 1997).

Developmental changes in the titer of bombyxin in *Bombyx* haemolymph were investigated using a monoclonal antibody against bombyxin II by radioimmunoassay (Saegusa *et al.*, 1992) and time-resolved fluoroimmunoassay (Satake *et al.*, 1999). It was shown that after eclosion the haemolymph titers of bombyxin were low in both males and females, increased steeply in males to a very high level and that this titer was maintained for several hours, whereas the titer increment in females was small and transient. The difference in the change of bombyxin titer between males and females suggests that bombyxin is responsible for the regulation of physiological changes underlying sexually different behavior of the adult moth (Satake *et al.*, 1999).

In the nematode Caenorhabditis elegans, the daf-2 gene, which regulates longevity and diapause, was identified as encoding an insulin receptor-like molecule (Kimura et al., 1997). This indicates that the ligand of DAF-2 may be an insulin-like peptide and act in metabolic and diapause control in the nematode. It is recalled that bombyxin was discovered through its induction of adult development when injected into dormant pupae whose development had been arrested by the removal of the brain. Further, Bombyx bombyxin was reported to bind specific receptors on the ovarian cells (Fullbright et al., 1997) and to induce meiosis in the ovarian cells (Orikasa et al., 1993) and morphological changes in a cell line derived from the ovarian tissue (Tanaka et al., 1995). The control of reproductive maturation of the germ line was also shown to require DAF-2 activity in the nematode (Hsin and Kenyon, 1999). In addition, the haemolymph titer of bombyxin attains a peak value during pupal-adult development in which tissue remodeling and productive maturation are in maximum progress (Saegusa *et al.*, 1992). It is therefore highly probable that bombyxin plays critical roles in metamorphosis and reproduction through its ability to metabolize energy and growth and/or developmental controls. Further study is necessary for the elucidation of the real physiological function(s) of bombyxin.

## STRUCTURE, ORGANIZATION, AND EVOLUTION OF BOMBYXIN GENE

All insulin genes except the murine insulin-I gene have two introns, one in the 5'-untranslated region and the other in the C peptide region. The murine insulin-I gene has only one intron in the 5'-untranslated region and is inferred to be a functional processed gene due to its unique structure in respect to the processed gene (Soares et al., 1985). The typical hallmarks of processed genes are (1) the lack of introns, (2) the presence of a poly(A) tract 3' to the transcriptional terminator, and (3) the presence of direct repeats bounding the transcribed region (Vanin, 1985). All bombyxin genes lack introns both in the 5'-untranslated region and in the C peptide region (Iwami, 1990, 1995; Iwami et al., 1989, 1990; Kondo et al., 1996; Tsuzuki et al., 1997; Yoshida et al., 1997, 1998). In addition, not a few bombyxin genes have boundary short repeats and the remnants of a poly(A) tract immediately preceding the downstream repeat, though the repeats and remnants are not complete in some cases (Iwami et al, 1989; Iwami, 1990; Tsuzuki et al., 1997). The structural features of bombyxin genes thus indicate that bombyxin genes are functional processed genes.

The bombyxin genes in the *Bombyx* genome show a unique spatial organization (Kondo *et al.*, 1996; Yoshida *et al.*, 1997, 1998). Thirty-one out of the 32 bombyxin genes are clustered in the genome as shown in Fig. 3. Segment A in Fig. 3 contains 25 bombyxin genes in a 50 kb range and segment B 6 genes in a 15 kb range. The topological relationship of segment A to segment B is unknown. The arrangement of the bombyxin genes in the clusters is classified into three categories: (1) gene pairs—two genes that belong to different families are opposed to form a gene pair with opposite transcriptional orientations. Most of such gene pairs are composed of the family A and B members. Nine sets of bombyxin genepairs can be found in the *Bombyx* genome. The distance

between the two genes of the pairs ranges from 0.5 to 2.3 kb. (2) Gene triplets–all triplets are composed of the family B, C and A genes, in this order from 5' to 3', and the three genes are apposed. The transcriptional direction of the family C and A genes is opposite to that of the family B gene. All the family C genes in the triplets C3, C4, C5 and C6 are pseudogenes. Four triplet sets are found in segment A. (3) Single genes– the bombyxin D1 gene is present singly, not forming either a pair or a triplet. The bombyxin E1 gene, which is not included in either segment A or B, is also present singly (Tsuzuki *et al.*, 1997). However, the possibility cannot be excluded that the two genes form pairs or triplets with yet undetected bombyxin genes.

Nucleotide sequence comparison suggests that the gene pairs form the basic unit of bombyxin genes in the genome and that the gene triplets are generated by an unequal crossing-over between the two gene pairs (Kondo et al., 1996). By nucleotide sequence analysis of the genes and their spacers together with restriction site analysis of the surrounding regions, the present form of the tandemly arranged bombyxin gene pairs was shown to be generated by gene duplication of a putative original pair (Kawakami et al., 1989; Kondo et al., 1996). Further, the present form of the four tandemly arranged bombyxin gene triplets should be generated by two duplications. An ancestral bombyxin gene triplet duplicated to form two closely situated gene triplets, as triplets B11/C5/A9 and B12/C6/A10, and the resulting double-triplet sets duplicated again to form the present triplet-sets on segment A in Fig. 3 (Kondo et al., 1996). It has been further demonstrated that a crossing-over between a gene triplet and a gene pair occurred which resulted in the diversification of the gene triplets, as shown in step (8) of Fig. 4 (Yoshida and Iwami, unpublished results). Thus, the very high degree of diversification in structure and genomic organization of bombyxin genes results from equal and unequal crossings-over as well as gene and geneset duplications. A model for generation of the diversity is as follows (Fig. 4) (Kondo et al., 1996): (1) An ancestor bombyxin gene diverged into genes ancestral to at least 7 families, A to G, by gene duplication followed by mutations. (2) The family B gene diverged further into two genes ancestral to subfamilies BI and BIII by point mutations. (3) Two gene pairs, BIII/C and BI/A, were generated and the other family genes, not il-



**Fig. 3.** Schematic representation of the two segments, A and B, of the *Bombyx* genomic DNA showing the organization of 31 bombyxin genes. Segment A carries 25 bombyxin genes in a 50 kbp range and segment B 6 bombyxin genes in a 15 kbp range. Boxes on the maps represent the bombyxin genes. Arrows over the genes indicate their transcriptional directions. The bombyxin E1 gene is not included in either segment A or B. Combined from Kondo *et al.*, 1996; Yoshida *et al.*, 1997, 1998.



Fig. 4. A schema for the molecular evolutionary history of the bombyxin genes in the *Bombyx* genome. Boxes represent the bombyxin genes, and arrows over the genes indicate the transcriptional directions. Modified from Kondo *et al.*, 1996.

lustrated in the figure, were duplicated and diverged. (4) The gene pair BI/A was duplicated to form a tandem array of two gene pairs. (5) Subsequently, an unequal crossing-over between gene pairs BIII/C and BI/A generated a gene triplet BIII/C/A. (6) Then, a crossing-over between gene pair BI/A and gene triplet BIII/C/A generated a tandem array of the two triplets, one of which contains a new subfamily BII. (7) Gene duplication of the triplet sets BIII/C/A and BII/C/A generated four gene triplets. (8) Lastly, an unequal crossing-over between gene triplet BII/C/A and gene pair B/A generated the present form of gene triplet B3/C/A1, as seen in the *Bombyx* genome. Point mutations have occurred throughout these process to diversify the structure of bombyxin genes (see Discussion in Kondo *et al.*, 1996).

### **EXPRESSION OF THE BOMBYXIN GENE**

The bombyxin gene is expressed predominantly in the brain and at low levels in a number of other tissues (Iwami *et* 

al., 1996b), in contrast to the insulin gene which is expressed in the gastroenteric organs and is almost silent in the brain. The brain is the only tissue that expresses the bombyxin gene throughout Bombyx development from the embryonic to adult stages when analyzed by northern hybridization (Adachi et al., 1989; Iwami, 1990). Reverse transcription-PCR analysis, however, demonstrated the presence of bombyxin family A and B mRNAs in all larval tissues examined although the levels of expression were very low (Iwami et al., 1996b). The bombyxin mRNA expression was detected in ganglia, epidermis, testis, ovary, fat body, silk gland, Malpighian tubule, midgut, and hindgut of the fifth instar larvae, which indicates the ubiquitous expression of bombyxin mRNA in the larvae. In the Bombyx brain, genes of all 7 families are expressed in four pairs of the medial neurosecretory cells as shown in Fig. 5 (Iwami, 1990, 1995; Tsuzuki et al., 1997; Yoshida et al., 1997, 1998). The concentration of bombyxin mRNA in the bombyxin-producing cells (BPCs) is remarkably high; for example, the concentration of the family A mRNA is about



**Fig. 5.** a). Schematic representation of the bombyxin-producing cells, BPCs, of the *Bombyx* brain. b) Whole-mount *in situ* hybridization of a *Bombyx* fifth-instar larval brain. The bombyxin family-A mRNAs were detected only in four pairs of BPCs (shown as dark blue). Bar, 100 μm. c) Section *in situ* hybridization of a *Bombyx* fifth-instar ovary. The bombyxin B1 DNA fragment was used as a probe. The bombyxin mRNA-producing cells are shown as dark blue. Bar, 100 μm. d) Fluorescence microphotograph of a Bombyx fifth-instar brain electroporated with the bombyxin/GFP reporter, pC4/B3::EGFP. The GFP fluorescence is visible as bright green in five cells, two in the left hemisphere and three in the right. Modified from Moto *et al.*, 1999.

2.8 ×10<sup>9</sup> molecules/µg of total RNA (Adachi et al., 1989). The amounts of the family A, B, and C mRNAs decrease gradually during larval to pupal development (Iwami, 1990). Immunohistochemical analysis using anti-bombyxin I monoclonal antibody also demonstrated the same spatial specificity (Mizoguchi et al., 1987). The immunohistochemistry revealed that the axons for the BPCs innervate the corpora allata and that their terminals are preferentially localized on the surface region of the corpora allata. Thus, bombyxin is presumed to be liberated into haemolymph from the corpora allata. Developmental fluctuation of the bombyxin content in the Bombyx brain was also investigated by immunohistochemistry using the monoclonal antibody (Mizoguchi et al., 1990). Interestingly, the four bombyxin cells in a brain hemisphere behaved immunohistochemically as two groups, each of which comprised two cells. On many occasions, two cells stained heavily while the other two stained lightly. At the mRNA level, there exists no such differential expression pattern. Therefore, the differential immunostaining might represent a pulsatile secretion of bombyxin (Mizoguchi et al., 1990) as demonstrated in PTTH (Gilbert et al., 1981).

The spatial specificity is also shown in the *Bombyx* larval ovary where two or three cell layers on the outer surface of each ovariole are found as shown in Fig. 5c (Iwami *et al.*, 1996b). The cells of the inner surface of an ovariole and the cell layers beneath the ovarial capsule also express the bombyxin mRNA. It is noteworthy that the presence of bombyxin molecules has been demonstrated in the developing embryos shortly after oviposition but before appearance of neurosecretory cells in the brain (Fugo *et al.*, 1987). It is thus probable that bombyxin in the ovariole is transferred to the embryo. The presence of a large number of bombyxin gene copies in the *Bombyx* genome might meet the demand for a large quantity of bombyxins for growth, development, and reproduction at a specific stage such as pupal to adult development.

In order to analyze the BPC-specific expression of the bombyxin gene, a gene-transfer technique was developed that enables introduction of a reporter gene into the brain by electroporation (Moto et al., 1999). By combination of this technique with the green fluorescent protein (GFP) reporter assay under UV-microscopy, we developed an assay system of the cell-specific expression of bombyxin gene. The reporter gene, which consists of a bombyxin gene spacer between B3 and C4 and the GFP coding region, was introduced into Bombyx brains by electroporation (Fig. 5d). The observation of fluorescence exclusively in the BPCs indicates that the reporter is under the control of the bombyxin gene promoter in a BPCspecific manner and that the spacer between B3 and C4 contains such a promoter. The expression pattern of the reporter has also been observed when the Bombyx pupal brains were used for the recipient tissue (Moto et al., 1999). The exact promoter site is now under investigation.

A recent study showed that a limited number of bombyxin genes are expressed in the brain just after pupation (Ino *et al.*, unpublished results). Thirteen out of the 28 genes belonging to the family A, B and C are expressed while the other 15

genes are silent or almost silent. The 13 bombyxin genes expressed show a gene-arrangement-dependent manner that can be classified into two categories: (1) gene-pair-specific expression of the family A and C genes—the family A and C genes expressed are located exclusively in the gene pairs but not in the gene triplets. The family A and C genes belonging to the gene triplets are silent or almost silent. Thus, the family A and C genes belonging to the genes show a gene-pair-specific expression. (2) gene-triplet-specific expression of the family B genes—the family B genes belonging to the gene triplets are expressed. In the gene pairs, some family B genes are expressed. The mechanism of these unique expression patterns of bombyxin genes remains to be elucidated.

### **CONCLUDING REMARKS**

The study of bombyxin is an odyssey from classic experimental morphology to modern molecular biology (see also Ishizaki, 1986). The study of bombyxin began in the 1950s when H. Ishizaki, the author's supervisor at Nagoya University, set up his research with M. Ichikawa at Kyoto University. Ishizaki tried to purify the so-called brain hormone, because it controls insect postembryonic development. The hormone assay was fully dependent on the methods of classical experimental morphology and the purification techniques were primitive because of the lack of modern molecular techniques such as peptide analysis. The purification was therefore very difficult work. After a 30-year struggle, Ishizaki purified bombyxin and PTTH from Bombyx heads in collaboration with biochemists of the University of Tokyo, S. Tamura, A. Suzuki, H. Nagasawa, and H. Kataoka. Shortly after the hormone purification, the author and his colleagues succeeded in cloning the genes encoding bombyxin and PTTH.

In parallel with bombyxin study in Bombyx, the author and his colleagues isolated 10 bombyxin genes in other Lepidopteran species. In Samia, six bombyxin genes have been cloned and proven to be expressed in the medial neurosecretory cells in the Samia brain (Kimura-Kawakami et al., 1992; Yagi et al., 1995; Inoue et al., unpublished results). They have a domain organization and nucleotide sequence very similar to the Bombyx bombyxin gene. The 6 genes isolated have been classified into two families, A and B, by nucleotide sequence similarity. They form a cluster in the genome, and two genes belonging to different families are localized close to each other with opposite transcriptional orientations (Kimura-Kawakami et al., 1992), resembling the Bombyx bombyxin gene. From the hornworm Agrius convoluvuli, three cDNA clones that encode Agrius bombyxins have been isolated and classified into two families, A and B (Iwami et al., 1996a). Southern hybridization analysis demonstrated the presence of multiple gene copies of the Agrius bombyxin gene in the genome. The Agrius bombyxin genes also have the insulinlike domain structure and an expression site that is localized to four pairs of the medial neurosecretory cells in the brain as the Bombyx bombyxin genes are (Iwami et al., 1996a). An antibody against Bombyx bombyxin detected the immunoreactive molecules in the oligochaete *Eisenia foetida* (Sauber *et al.*, 1990), the locust *Locusta migratoria* (Zachary *et al.*, 1988), the wax moth *Galleria mellonella* (Žitňan *et al.*, 1990), the hornworm Manduca sexta (Dai *et al.*, 1994), and the fruit fly Drosophila melanogaster (Žitňan *et al.*, 1993). Those studies demonstrate that bombyxin and/or bombyxin-related peptides exist ubiquitously in insects. The study of bombyxin has thus expanded the insulin world.

### ACKNOWLEDGMENTS

The author is grateful to Drs. S. Sakurai and H. Ishizaki for their continuous encouragement. He also thanks all the members of his laboratory, especially Dr. K. Moto, Dr. S. Tsuzuki, Dr. Y. Oda, M. Ino, S. E. Abdel Salam, M. Miyamoto, and I. Yoshida. Work done at the author's laboratory was supported by grants from the Ministry of Education, Science and Culture of Japan, Takeda Science Foundation, Ciba-Geigy Foundation (Japan) for the Promotion of Science, and the Kato Memorial Bioscience Foundation.

#### REFERENCES

- Adachi T, Takiya S, Suzuki Y, Iwami M, Kawakami A, Takahashi SY, Ishizaki H, Nagasawa H, Suzuki, A (1989) cDNA structure and expression of bombyxin, an insulin-like brain secretory peptide of the silkmoth *Bombyx mori.* J Biol Chem 264: 7681–7685
- Bell GI, Pictet R, Rutter W, Cordell B, Tischer E, Goodman HM (1980) Sequence of the human insulin gene. Nature 284: 26–32
- Bollenbacher W, Granger N (1985) Endocrinology of the prothoracicotropic hormone. In "Comprehensive insect physiology, biochemistry, and pharmacology Vol. 7" Ed by GA Kerkut, LI Gilbert, Pergamon Press, Oxford, pp 109–151
- Büllesbach EE (1999) Bombyxin exhibits an insulin-like response to modification in the N-terminal region of the A chain. J Peptide Res 54: 12–17
- Chan SJ, Kwok SC, Steiner DF (1981) The biosynthesis of insulin: some genetic and evolutionary aspects. Diabetes Care 4: 4–10
- Dai J-D, Mizoguchi A, Gilbert LI (1994) Immunoreactivity of neurosecretory granules in the brain-retrocerebral complexes of *Manduca sexta* to heterologous antibodies against *Bombyx* prothoracicotropic hormone and bombyxin. Invert Reprod Dev 26: 187–196
- Fugo H, Chen JH, Nakajima M, Nagasawa H, Suzuki A (1987) Neurohormones in developing embryos of the silkworm, *Bombyx mori:* the presence and characteristics of prothoracicotropic hormone-S. J Insect Physiol 33: 243–248
- Fukuda S (1944) The hormonal mechanism of larval molting and metamorphosis in the silkworm. J Fac Sci Tokyo Imperial Univ IV 6: 447–532
- Fullbright G, Lacy ER, Büllesbach EE (1997) The prothoracicotropic hormone bombyxin has specific receptors on insect ovarian cells. Eur J Biochem 245: 774–780
- Gilbert LI, Bollenbacher WE, Agui N, Granger NA, Sedlak BJ, Gibbs D, Buys CM (1981) The prothoracicotropes: source of the prothoracicotropic hormone. Am Zool 21: 641–653
- Hsin H, Kenyon C (1999) Signals from the reproductive system regulate the lifespan of *C. elegans.* Nature 399: 362–366
- Ichikawa M, Ishizaki H (1963) Protein nature of the brain hormone of insects. Nature 198: 308–309
- Ishizaki H (1986) A PTTH pilgrimage from *Luehdorfia* to *Bombyx*. Zool Sci 3: 921–929
- Ishizaki H, Mizoguchi A, Fujishita M, Suzuki A, Moriya I, O'oka H, Kataoka H, Isogai A, Nagasawa H, Tamura S, Suzuki, A (1983) Species specificity of the insect prothoracicotropic hormone (PTTH): the presence of *Bombyx*- and *Samia*-specific PTTHs in

the brain of Bombyx mori. Dev Growth Differ 25: 593-600

- Iwami M (1990) The genes encoding bombyxin, a brain secretory peptide of *Bombyx mori*: structure and expression. In "Molting and Metamorphosis" Ed by E Ohnishi, H Ishizaki, Japan Sci. Soc. Press, Tokyo/Springer-Verlag, Berlin, pp 49–66
- Iwami M (1995) Multi-family genes encoding bombyxin and bombyxinrelated peptides. In "Molecular Mechanisms of Insect Metamorphosis and Diapause" Ed by A Suzuki, H Kataoka, S Matsumoto, Industrial Publishing & Consulting, Tokyo, pp 65–74
- Iwami M, Kawakami A, Ishizaki H, Takahashi SY, Adachi T, Suzuki Y, Nagasawa H, Suzuki A (1989) Cloning of a gene encoding bombyxin, an insulin-like brain secretory peptide of the silkmoth *Bombyx mori* with prothoracicotropic activity. Dev Growth Differ 31: 31–37
- Iwami M, Adachi T, Kondo H, Kawakami A, Suzuki Y, Nagasawa H, Suzuki A, Ishizaki H (1990) A novel family C of the genes that encode bombyxin, an insulin-related brain secretory peptide of the silkmoth *Bombyx mori*: isolation and characterization of gene C1. Insect Biochem 20: 295–303
- Iwami M, Furuya I, Kataoka H (1996a) Bombyxin-related peptides: cDNA structure and expression in the brain of the hornworm *Agrius convoluvuli*. Insect Biochem Molec Biol 26: 25–32
- Iwami M, Tanaka A, Hano N, Sakurai S (1996b) Bombyxin gene expression in tissues other than brain detected by reverse transcription-polymerase chain reaction (RT-PCR) and in situ hybridization. Experientia 52: 882–887
- Jhoti H, McLeod AN, Blundell TL, Ishizaki H, Nagasawa H, Suzuki A (1987) Prothoracicotropic hormone has an insulin-like tertiary structure. FEBS Lett 219: 419–425
- Kawakami A, Iwami M, Nagasawa H, Suzuki A, Ishizaki H (1989) Structure and organization of four clustered genes that encode bombyxin, an insulin-related brain secretory peptide of the silkmoth *Bombyx mori*. Proc Natl Acad Sci USA 86: 6843–6847
- Kawakami A, Kataoka H, Oka T, Mizoguchi A, Kimura-Kawakami M, Adachi T, Iwami M, Nagasawa H, Suzuki A, Ishizaki H (1990) Molecular cloning the *Bombyx mori* prothoracicotropic hormone. Science 247: 1333–1335
- Kimura KD, Tissenbaum HA, Liu Y, Ruvkun G (1997) daf-2, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. Science 277: 942–946
- Kimura-Kawakami M, Iwami M, Kawakami A, Nagasawa H, Suzuki A, Ishizaki H (1992) Structure and expression of bombyxin-related peptide genes of the moth *Samia cynthia ricini*. Gen Comp Endocr 86: 257–268
- Kiriishi S, Nagasawa H, Kataoka H, Suzuki A, Sakurai S (1992) Comparison of the *in vivo* and *in vitro* effects of bombyxin and prothoracicotropic hormone on prothoracic glands of the silkworm, *Bombyx mori.* Zool Sci 9: 149–155
- Kondo H, Ino M, Suzuki A, Ishizaki H, Iwami M (1996) Multiple gene copies for bombyxin, an insulin-related peptide of the silkmoth *Bombyx mori*: structural signs for gene rearrangement and duplication responsible for generation of multiple molecular forms of bombyxin. J Molec Biol 259: 926–937
- Kopeć S (1922) Studies on the necessity of the brain for the inception of insect metamorphosis. Biol Bull 42: 323–342
- Kramer KJ (1985) Vertebrate hormones in insects. In "Comprehensive insect physiology, biochemistry, and pharmacology Vol. 7" Ed by GA Kerkut, LI Gilbert, Pergamon Press, Oxford, pp 511–537
- Maruyama K, Hietter H, Nagasawa H, Isogai A, Tamura S, Suzuki A, Ishizaki H (1988) Isolation and primary structure of bombyxin-IV, a novel molecular species of bombyxin from the silkworm, *Bombyx mori*. Agric Biol Chem 52: 3035–3041
- Maruyama K, Nagasawa H, Isogai A, Tamura S, Ishizaki H, Suzuki A (1990) Synthesis of bombyxin-IV, an insulin-like heterodimeric peptide from the silkworm, *Bombyx mori*. Peptide 11: 169–171
- Mizoguchi A (1994) Distribution and function of bombyxin. In "Perspectives in Comparative Endocrinology" Ed by KG Davey, RE

Peter, SS Tobe, National Research Council of Canada, Ottawa, pp 215–219

- Mizoguchi A, Ishizaki H, Nagasawa H, Kataoka H, Isogai A, Tamura S, Suzuki A, Fujino M, Kitada C (1987) A monoclonal antibody against a synthetic fragment of bombyxin (4K-prothoracicotropic hormone) from the silkmoth, *Bombyx mori*: characterization and immunohistochemistry. Molec Cell Endocr 51: 227–235
- Mizoguchi A, Hatta M, Sato S, Nagasawa H, Suzuki A, Ishizaki H (1990) Developmental change of bombyxin content in the brain of the silkmoth *Bombyx mori.* J Insect Physiol 36: 655–664
- Moto K, Abdel Salam SE, Sakurai S, Iwami M (1999) Gene transfer into insect brain and cell-specific expression of bombyxin gene. Dev Genes Evol 209: 447–450
- Nagasawa H, Kataoka H, Isogai A, Tamura S, Suzuki A, Ishizaki H, Mizoguchi A, Fujiwara Y, Suzuki A (1984) Amino-terminal amino acid sequence of the silkworm prothoracicotropic hormone: homology with insulin. Science 226: 1344–1345
- Nagasawa H, Kataoka H, Isogai A, Tamura S, Suzuki A, Mizoguchi A, Fujiwara Y, Suzuki A, Takahashi SY, Ishizaki H (1986) Amino Acid sequence of a prothoracicotropic hormone of the silkworm *Bombyx mori.* Proc Natl Acad Sci USA 83: 5840–5843
- Nagasawa H, Maruyama K, Sato B, Hietter H, Kataoka H, Isogai A, Tamura S, Ishizaki H, Semba T, Suzuki A (1988) Structure and synthesis of bombyxin from the silkworm, *Bombyx mori*. In "Peptide Chemistry 1987" Ed by T Shiba, S Sakakibara, Protein Research Foundation, Osaka, pp 123–126
- Nagata K, Hatanaka H, Kohda D, Kataoka H, Nagasawa H, Isogai A, Ishizaki H, Suzuki A, Inagaki F (1995) Identification of the receptor-recognition surface of bombyxin-II, an insulin-like peptide of the silkmoth *Bombyx mori*: critical importance of the B-chain central part. J Molec Biol 253: 759–770
- Oda Y, Iwami M, Osanai M, Sakurai S (1997) Dynamics of haemolymph sorbitol-6-phosphate and its control by ecdysteroid in the larvae of the silkworm, *Bombyx mori.* Insect Biochem Molec Biol 27: 461–468
- Okada TS (1994) Experimental embryology in Japan, 1930–1960. Int J Dev Biol 38: 135–154
- Orikasa C, Yamauchi H, Nagasawa H, Suzuki A, Nagata M (1993) Induction of oocyte nurse cell differentiation in the ovary by the brain during the initial stage of oogenesis in the silkworm, *Bombyx mori* (Lepidoptera: Bombycidae). Appl Entomol Zool 28: 303– 311
- Saegusa H, Mizoguchi A, Kitahora H, Nagasawa H, Suzuki A, Ishizaki H (1992) Changes in the titer of bombyxin-immunoreactive material in hemolymph during the postembryonic development of the silkmoth *Bombyx mori*. Dev Growth Differ 34: 595–605
- Satake S, Masumura M, Ishizaki H, Nagata K, Kataoka H, Suzuki A, Mizoguchi A (1997) Bombyxin, an insulin-related peptide of insects, reduces the major storage carbohydrates in the silkworm *Bombyx mori*. Comp Biochem Physiol 118B: 349–357
- Satake S, Nagata K, Kataoka H, Mizoguchi A (1999) Bombyxin secretion in the adult silkmoth *Bombyx mori*: sex-specificity and its correlation with metabolism. J Insect Physiol 45: 939–945
- Sauber F, Reuland M, Rises E, Holder F, Charket M (1990) Presence of bombyxin-like molecules in the brain of *Eisenia fetida* (Oligochaeta, Lumbricidae). Invert Reprod Dev 17: 123–126

- Soares MB, Schon E, Henderson A, Karathanasis SK, Cate R, Zeitlin S, Chirgwin J, Efstratiadis A (1985) RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. Molec Cell Biol 5: 2090–2103
- Steiner DF, Chan SJ, Welsh JM, Kwok SCM (1985) Structure and evolution of the insulin gene. Ann Rev Genet 19: 463–484
- Suzuki A, Nagasawa H, Kataoka H, Hori Y, Isogai A, Tamura S, Guo F, Zhong X-C, Ishizaki H, Fujishita M, Mizoguchi A (1982) Isolation and characterization of prothoracicotropic hormone from silkworm *Bombyx mori*. Agric Biol Chem 46: 1107–1109
- Tanaka M, Kataoka H, Nagata K, Nagasawa H, Suzuki A (1995) Morphological changes of BM-N4 cells induced by bombyxin, an insulin-related peptide of *Bombyx mori*. Regul Peptides 57: 311– 318
- Tsuzuki S, Masuta T, Furuno M, Sakurai S, Iwami M (1997) Structure and expression of bombyxin E1 gene, a novel family gene that encodes bombyxin-IV, an insect insulin-related neurosecretory peptide. Comp Biochem Physiol 117B: 409–416
- Vanin EF (1985) Processed pseudogenes: characteristics and evolution. Ann Rev Genet 19: 253–272
- Wigglesworth VB (1940) The determination of characters at metamorphosis in *Rhodnius prolixus* (Hemiptera). J Exp Biol 17: 201– 222
- Williams CM (1947) Physiology of insect diapause. II. Interaction between the pupal brain and prothoracic glands in metamorphosis of the giant silkworm, *Platysamia cecropia*. Biol Bull 93: 89–98
- Yagi Y, Ishibashi J, Nagata K, Kataoka H, Suzuki A, Mizoguchi A, Ishizaki H (1995) The brain neurosecretory cells of the moth *Samia cynthia ricini:* immunohistochemical localization and developmental changes of the *Samia* homologues of the *Bombyx* prothoracicotropic hormone and bombyxin. Dev Growth Differ 37: 505–516
- Yoshida I, Tsuzuki S, Abdel Salam SE, Ino M, Korayem AM, Sakurai S, Iwami M (1997) Bombyxin F1 gene: structure and expression of a new bombyxin family gene that forms a pair with bombyxin B10 gene. Zool Sci 14: 615–622
- Yoshida I, Moto K, Sakurai S, Iwami M (1998) A novel member of the bombyxin gene family: structure and expression of bombyxin G1 gene, an insulin-related peptide gene of the silkmoth *Bombyx mori.* Dev Genes Evol 208: 407–410
- Zachary D, Goltzene F, Holder FC, Berchtold JP, Nagasawa H, Suzuki A, Mizoguchi A, Ishizaki H, Hoffmann J (1988) Presence of bombyxin (4K-PTTH)-like molecules in neurosecretory granules of brain-corpora cardiaca complexes of *Locusta migratoria*: developmental aspects. Int J Invert Reprod Dev 14: 1–10
- Žitňan D, Schenal F, Mizoguchi A, Ishizaki H, Nagasawa H, Suzuki A (1990) Developmental changes in the bombyxin- and insulin-like immunoreactive neurosecretory system in the wax moth, *Galleria mellonella*. Dev Growth Diff 32: 637–645
- Žitňan D, Schenal F, Bryant P (1993) Neurons producing specific neuropeptides in the central nervous system of normal and pupariation-delayed *Drosophila*. Dev Biol 156: 117–135

(Received May 24, 2000 / Invited Review)