Chapter – 5

Paralytic peptide binding protein (PP-BP) gene expression during egg diapause and its multi-gene organization
5.1. Introduction

During insect development the central nervous system regulates growth, metamorphosis, diapause and reproduction by secreting peptides. Insect hemolymph contains various proteinaceous components viz., transport proteins, storage proteins, peptide hormones which are essential for growth and development, and defensive peptides while induced by bacterial invasion or injury. Injection of the hemolymph obtained from larvae of other Manduca species into that of Manduca sexta, the latter become paralyzed. This factor from the hemolymph was isolated from several kinds of lepidopteran hemolymph and characterized as 23 amino acid peptides which were named as Paralytic peptide (Ha et al., 1999).

This ENF peptide family shares similar amino acid sequence and conserved at C-terminal region. It is named after the N-terminal-conserved three amino acids (Glu-Asn-Phe) (Strand et al., 2000). ENF family proteins are synthesized as inactive precursors that are activated by a serine protease cleavage. Mature ENF family proteins, including Bombyx mori paralytic peptide (BmPP), have been reported to show multiple effects, such as induction of morphological changes of plasmatocytes, inhibition of larval growth, promotion of cell growth and local muscle contraction (Ha et al., 1999; Sasagawa et al., 2001; Miura et al., 2002; Nakahara et al., 2003).

In Bombyx mori the embryonic diapause is normally induced by the diapause hormone (Bommo-DH), which is produced in the pupae from the suboesophageal ganglion. Injection of Antheraea yamamai-paralytic peptide (Antya-ParP) into non-diapausing pupae of B. mori, however, converted the pupae into diapauses egg producers. Even extirpation of brain along with SG
had no influence on the induction of diapausing egg production by Antya-ParP, which demonstrated that Antya-ParP does not act via endogenous
*B. mori* diapause hormone (Bombo-DH) but mimics its action at the level of
target tissues. No homology exists between the primary sequences of the
Bombo-DH and ENF peptide families, but this does not exclude that one can
mimic the action of the other (An *et al.*, 2007).

In *B. mori*, 30K proteins are synthesized by the fat body and then
secreted into hemolymph during the last instar larval stage, where they get
accumulated to a high concentration and then are gradually absorbed into the
oocyte (Chen and Yamashita, 1990). About 35% of the total yolk proteins
consists of 30K proteins (Zhu *et al.*, 1986). These proteins have a variety of
biological functions, providing a source of nutrients for embryogenesis (Zhong
*et al.*, 2005), binding to glucan to defend fungal infection (Ujita *et al.* 2002;
2005) and inhibiting apoptosis (Kim *et al.*, 2003; Park *et al.*, 2003).

There are several 30KPs as well as their homologs reported in
silkworm *B. mori*. Ten 30KP genes have been identified in the silkworm (Sun
*et al.*, 2007), and more genes show sequence homologous to 30KPs like
*B. mori* larvae serum protein (LSP) and LSP-T protein that are homologs of
30KPs (Fujiwara and Yamashita, 1992; Miyagawa *et al.*, 2004).
Microvitellogenin of *Manduca sexta* and the C terminus (170th–430th amino
acids) of growth- blocking peptide binding protein (GBP-BP) of *Pseudaletia separata* also show some similarity to 30KPs (Matsumoto *et al.*, 2003; Wang
*et al.*, 1989). Both Growth blocking peptide (GBP) and Paralytic peptide (PP)
belong to ENF peptide family in view of the consensus N-terminal sequence
(Glu-Asn-Phe). A cytokine-like factor, *Bombyx mori* paralytic peptide (BmPP),
which was purified from the silkworm hemolymph, belongs to the ENF peptide family. Like other Paralytic peptides, the Paralytic peptide of B. mori (BmPP) is also multifunctional ENF-peptide (Ha et al., 1999). The C-terminal amino acid sequence (190th–430th amino acids) of B. mori paralytic peptide binding proteins is homologous to 30KPs (Hu et al., 2006). Growth-blocking peptide binding protein (GBP-BP) was reported to function for silencing the GBP action (Matsumoto et al., 2003).

In a recent study by Zhang et al., (2012), 73 genes of Lipoprotein_11 family were retrieved from 12 lepidopteran species and the structural analysis showed that these genes could be divided into three distinct subfamilies according to their differential N-terminal domain. A novel subfamily of Lipoprotein_11 was reported for the first time in the study and named serine/threonine-rich 30KP according to its exclusive domain.

The abundant level of 30KPs expression suggests that they may serve as storage proteins (Fujiwara and Yamashita, 1992; Izumi et al., 1981; Wang et al., 1989). Not much is known about the molecular mechanisms controlling the terminal process in the cellular immune response. The expression of PP and PP-BP has not been studied till date in diapause induced eggs of multivoltine silkworm B. mori.

In this study we report a high expression of the paralytic peptide and paralytic peptide binding protein in the diapause induced eggs of multivoltine silkworm B. mori which may suggest that these proteins might also play a role in regulation of diapause induction.
5.2. Materials and methods

5.2.1. Insect rearing

The multivoltine strain MW13 (Indian origin) was selected for the study. The larvae were reared as per the standard rearing method of Krishnaswami (1978) up to last instar to obtain non-diapausing eggs. The late stage (4th & 5th instars) larvae were then reared under low temperature (18ºC) and photoperiod (6L: 18D) up to cocooning stage and the moths were made to lay diapausing eggs at normal room temperature (25 ºC) (Saravanakumar et al., 2008).

5.2.2. RNA isolation

After oviposition, the diapause and non-diapause egg samples were collected from 0 to 48h at every 6h time interval, while, other tissues were collected from the day 3 of 5th instar larvae. Total RNA was extracted from the diapause and non-diapause eggs using TRlpol reagent (Invitrogen, USA), denatured in formaldehyde, formamide and electrophoresed in 2.0% agarose gels.

5.2.3. cDNA preparation

The first strand cDNA was synthesized utilizing RNA (2μg) treated with 0.5μl of DNase buffer and 0.5μl of DNase (Invitrogen, USA) for 15 minutes. Then, the reaction was terminated by heating at 75ºC for 10 minutes, to the above DNase treated sample, 1μl 10 mM dNTP, 1μl oligo (dT)18 (0.01mM) (Eurofin India Pvt Ltd, Bangalore) was added followed by incubation at 65ºC for 5 min. Finally 1X reverse transcriptase buffer (4μl), 5mM DTT (1μl) and 1μl of M-MLV Superscript III reverse transcriptase (Invitrogen, USA) was added to
obtain a final volume of 20μl. The reaction was terminated by heating at 75°C for 10 min according to the manufacturer’s protocol.

5.2.4. Microarray experiment and data analysis

A genome wide oligonucleotide microarray containing 24,924 probes were used to investigate the gene expression profiles of diapause induced and non-diapause eggs of multivoltine silkworm *B. mori* at 18 and 30hrs after oviposition. The complete sets of raw and normalized data from this study have been deposited in the NCBI Gene Expression Omnibus (GEO) repository (accession number GSE35622).

5.2.5. Reverse transcription polymerase chain reaction (RT PCR) analysis for tissue specific gene expression

PCR amplification was performed in a 25μl reaction mixture containing 2.0μl of 10X reaction buffer (100mM Tris-HCL, pH 8.3, 500mM KCl), 0.2 mM dNTPs, 1.5 mM MgCl₂, 10 picomoles of forward and reverse primers, 0.3 U of Taq DNA polymerase (MBI Fermentas) with 1μl first strand cDNA as template. β actin (FP 5’cactgaggctccctgaac 3’ and RP 5’ ggagtgcgtatccctcgtag 3’) (Eurofins, Bangalore) was used as an internal standard. The PCR amplification was carried out under the following conditions: 94°C for 3min followed by 27 cycles of 94°C for 30s, 54°C for 30s, 72°C for 2 min and a final extension of 7 min at 72°C.

5.2.6. Real time PCR analysis

One μl of first strand cDNA synthesized from diapause and non-diapause eggs from 6 to 48h after oviposition was used as template for qPCR analysis in a 25μl reaction mixture containing SYBR green mastermix (ABI, CA, USA) and the specific primers designed for Realtime PCR (qPCR). The
reactions were conducted on a STRATAGENE Mx 3005P realtime PCR system. The experiment was performed in triplicate and results were standardized to the expression level of the constitutive β actin gene. A Non-template control (NTC) sample was also run to detect contamination if any.

5.2.7. Identification of paralogous gene sequences in *B. mori*

The *B. mori* cDNA sequences were BLAST searched with Silkworm Genome Database (http://silkowrm.genomics.org.cn) to identify paralogous multigene family. Using the silkworm database, the functional annotation of genes, paralogous gene sequences, gene products and chromosome mapping were determined. Further, the tools provided in the database were utilized to perform specific genomic BLAST search as well as Map view (a visualization tool that provides a graphical view of selected genes). The organization of paralogous 30KP multigene family on individual scaffold was also analyzed using BLAST search with Gene ID. Phylogenetic analyses were performed with the multiple sequence alignment using ClustalW through MEGA 4 (Kumar et al., 2004). The Bootstrap consensus NJ tree for 30KP paralogous gene sequences was constructed with the Bootstrap values. Signal peptide cleavage sites were predicted using SignalIP algorithm (www.cbs.dtu.dk/services/SignalIP), based on the Neural Network and Hidden Markov Model.

5.3. Results

Survey of the microarray gene expression data for comparison of diapause induced and non-diapause eggs of polyvoltine silkworm *B. mori* after oviposition showed a few genes specifically up-regulated in the diapause
induced eggs, especially paralytic peptide (PP) and paralytic peptide binding protein (PP-BP) were significantly up-regulated at 12 and 18 hours (Fig. 5.1).

5.3.1. PP-BP gene expression in different tissues

RT-PCR analysis was performed to determine the tissue specific expression of the *B. mori* PP-BP gene. Different tissues viz., fat body, brain, mid gut of 5\(^{th}\) instar 3\(^{rd}\) day larvae as well as, diapause induced eggs were analyzed for tissue distribution of PP-BP transcripts. The results indicated that PP-BP was highly expressed in fat body followed by brain and eggs, while no expression was observed in mid gut tissue (Fig. 5.2). *B. mori* \(\beta\)-actin gene was used as an internal control. Further the results validated through realtime PCR analysis (Fig. 5.3) revealed that the PP-BP gene expression was higher in diapause induced eggs, compared to the non-diapause eggs.

5.3.2. Differential expression of *B. mori* PP-BP and PP in diapause induced eggs

The expression level of PP-BP gradually increased from 0 to 24h, where, the expression was at maximum, followed by a very steep decrease from 30 to 48h in diapause induced eggs. The gene expression in non-diapause eggs was very low from 6 to 48h time intervals compared to diapause eggs. The expression level of paralytic peptide was almost similar to PP-BP in diapause eggs while in non-diapause eggs they were slightly higher compared to PP-BP. (Fig. 5.4, Fig. 5.5).
Fig. 5.1. Microarray expression profile of paralytic peptide and paralytic peptide binding protein in diapause induced eggs of silkworm *B. mori* at different time intervals.

![Graph showing expression profiles of paralytic peptide binding protein at different time intervals.]

Fig. 5.2. Expression profiles of paralytic peptide binding protein in different tissues. *B. mori* β-actin was used as internal control.

![Graph showing expression profiles of paralytic peptide binding protein in different tissues.]

Fig. 5.3. Relative gene expression patterns of paralytic peptide binding protein in different tissues.

![Graph showing relative gene expression of paralytic peptide binding protein in different tissues.]
Fig. 5.4. Relative gene expression patterns of paralytic peptide binding protein gene upregulated during diapause.

Fig. 5.5. Relative gene expression patterns of paralytic peptide gene upregulated during diapause.
5.3.3. Genomic organization of *B. mori* PP-BP paralogous genes

The cDNA sequence of *B. mori* PABP was subjected to blast analysis and several paralogous contigs were identified. A total of 46 paralogous sequences were identified which were distributed on chromosomes 7, 20, 22 and 24 (Fig. 5.6. a, b, c, d). Based on the structural characteristics of the above identified 46 sequences they have been classified into ENF-BP, typical 30KP, and serine/threonine-rich 30KP. ClustalW analyses were performed with the multiple sequence alignment using ClustalW through MEGA 4 which revealed that these paralogous sequences formed three major clusters. Of the 46 sequences identified, 10 fall under the category of ENF-BP cluster, 12 fall under S/T-rich 30KP cluster and 24 of them fall under typical 30KPs clusters (Fig. 5.7).

**Fig. 5.6.a.** Schematic representation of the ENF-BP gene cluster in *B. mori*.

The two paralogous sequences are located in nscaf 2986 between 1470Kb to 1490Kb on chromosome 7.
**Fig. 5.6.b.** Schematic representation of the 30KP and S/T rich 30KP gene clusters in *Bombyx mori*. The two major clusters are located in nscaf 2795. Cluster one flanked from 1350Kb to 1700Kb and cluster two from 2100Kb to 2400Kb on chromosome 20 with intergenic region of 400Kb.
**Fig. 5.6.c.** Schematic representation of the ENF-BP gene clusters in *B. mori.* The two clusters are located in nscaf 2891 and nscaf 2962. Cluster one flanked from 1350Kb to 1700Kb and cluster two from 2100Kb to 2400Kb on chromosome 24 with intergenic region of 996Kb.

![Schematic representation of the ENF-BP gene clusters in B. mori.](image)

**Fig. 5.6.d.** Schematic representation of the ENF-BP gene cluster in *B. mori.* The three paralogous sequences are located in nscaf 3005 between 14000Kb to 22000Kb on chromosome 22.

![Schematic representation of the ENF-BP gene cluster in B. mori.](image)
**Fig. 5.7.** Neighbor-joining tree of paralogous genes of paralytic peptide binding protein, percentage of bootstrap values (based on the 1000 replication) for the main branching nodes are shown on the tree. The paralogous gene sequences retrieved from silkworm genome database are indicated by Gene ID number.
5.4. Discussion

Since the finding of the first member of the ENF peptide family, eleven lepidopteran species including *B. mori* have been found to have such peptides (Ha *et al.*, 1999). Antya-ParP when injected in to the non-diapause egg producing pupae, the pupae are converted into diapause egg producers even lacking sub esophageal ganglion. Even though Bommo-DH and ENF peptides share no homology in the primary sequences, the later can mimic the action of the former. ENF peptides are regarded as insect cytokines and have multiple biological activities, including mitogenic, paralytic, hemocyte-spreading, and growth blocking activities (Aizawa *et al.*, 2002; Kamimura *et al.*, 2001). However, limited information is available on diapause eggs.

In the present study, the gene expression levels of paralytic peptide (PP) and PP-BP were compared in diapause and non-diapause eggs after oviposition. The expression of PP was higher at 18h similarly the expression of both PP and PP-BP genes was higher at 12 and 18h along with few other genes in microarray experiment. These genes were taken up for further study, which were validated through q-PCR.

Most of the reports emphasize the role of ENF peptides in cellular immunity and cell proliferation. An *et al.*, (2007) proved that *Antheraea yamamai* paralytic peptide induces egg diapause even in pupae where subesophageal ganglion had been extirpated and also causes a rapid and rigid larval paralysis in *B. mori*. It was also inferred that AnyParP interacts with other pathways to excite diapause rather than those involved in paralysis. It was also speculated that AnyParp and homologs act on the same channel in the cascade of diapause egg induction as the DH.
The structural and phylogenetic analyses of paralogous sequences revealed that these sequences can be categorized into three categories ENF-BP, typical 30k proteins and S/T rich 30K proteins. The 30k proteins are involved in the diapause mechanism through regulating the action of paralytic peptide.

In a study by Zhang et al., (2012), the ENF-BP genes were also expressed in the gonads (ovary and testis), fat body, head, integument, silk gland, etc which was different from the hemocyte specific expression pattern of *P. separate* ENF-BP. The Bm ENF peptides, also show extensive tissue expression profiles, such as in gonads, fatbody, integument, central nervous system, etc. (Aizawa et al., 2002; Kamimura et al., 2001). Similarly in the present study it is found that the PP-BP is highly expressed in fat body followed by eggs and brain while, no expression is observed in midgut. As ENF peptides are reported to play a role in regulation of the cell proliferation as growth factors, it can also be assumed that ENF-BPs may be involved in development by regulating the expression or activities of ENF peptides.

**5.4.1. Multigene organization of *B. mori* PP-BP gene**

Genes that have originated by gene duplication retained a certain degree of similarity forming multigene families. The multigene family members are often arranged in a compact cluster due to chromosomal rearrangements subsequent to gene duplications. The members of a multigene family can be functional or nonfunctional which are known as pseudo genes. Multigene families containing paralogous gene sequence that evolve in different ways, and assumption of the multigene family evolution is essential for estimation of its phylogeny. It is assumed that members of a multigene family may evolve at
a faster rate and such members are designated as fast evolving genes. This 
phenomenon takes place when one gene member of a multigene family takes 
on a novel function, which is essential for survival and thus encounters 
significantly different selective pressure from other multigene family members. 
Another usual assumption of molecular tree of multigene family is that, each 
branch of the tree evolves independently from other branches. These families 
often show co-incidental evolution, either indirectly through biased mutational 
and selective force or directly by mechanism such as gene conversion (Roach 
et al., 2005).

A total of forty six *B. mori* paralytic peptide binding protein homologs 
were retrieved from the silkworm genome database (http://silkowrm.genomics.org.cn). These forty six paralogous gene 
sequences, belonging to PP-BP gene family diverged into three different 
groups, viz., ENF-BP, S/T rich, Typical 30KP. Ten sequences belonged to 
ENF–BP, twelve to S/T rich 30KP proteins and Twenty four to typical 30KPs. 
The PP-BP has conserved domain similar to ENF-BP and hence, ENF-BP is 
also considered as paralytic peptide binding protein. The twenty four 30KP 
paralogous genes are organized in a single locus of chromosome 20; 
however, two paralogous genes of 30KPs formed a separate cluster from the 
remaining 30KP gene family. The 30KPs can be divided into 2 groups one as 
typical 30KP and S/T rich 30KP. The S/T rich rich 30KP have a 
serine/threonine rich domain between signal peptides and N-terminal domain, 
while the N terminal of typical 30KPs is composed of the signal peptides and 
an un-conserved region. Twenty four 30KPs as well as Twelve S/T rich 30KPs 
formed a major cluster located in nscaf 2795 of chromosome 20. This locus
was grouped into two clusters with intergenic region of 400K. The cluster 1 consists exclusively of 30KP; however, the cluster 2 consists both 30KP and S/T rich 30KP. Though a major cluster is formed, the 30KP and S/T rich 30KP grouped independently in this major cluster. The paralogous gene 04451 was located approximately 150kb from the remaining cluster, indicating, that this paralogous gene has diverged from the single ancestral gene prior to multigene family formation, to further support this hypothesis the gene 4451 formed a separate cluster in the phylogenetic tree. Comparison of the molecular distance of the paralogous genes within the species indicates the amount of coincidental evolution (Walsh, 1995). The presence of S/T rich subgroup in the 30KP group indicates that both the groups have evolved recently from the ancestral gene. The original gene might have originated from the common ancestors by gene duplication and later the individuals within the multigene family were formed during the process of evolution. Unlike the typical 30KP and ST rich 30KPs the multigene cluster of ENF-BP was spread over different chromosomes viz., chromosome number 7, 22 and 24. The long intergenic regions between two paralogous gene sequences as well as presence of these genes in different chromosomes indicate that the multi gene family evolved much ahead of typical 30KP and S/T rich 30KPs. Similar phenomenon was observed for cecropin multigene family in *B. mori*. Cecropin B locus on genomic DNA consists of 6 paralogous gene groups as well as Cecropin D and E (Ponnudhevan *et al.*, 2010). However, the maximum intergenic region of 23Kb was observed between Cecropin D and remaining Cecropin B family indicates that the two gene families were distantly evolved. The ENF-BP located on chromosome 24 had 4 paralogous gene sequences
in a single locus of opposite orientation. These 4 genes of ENF-BP located on
the same position of nscaf 3005 on chromosome 22 indicates that these 3
genes were closely evolved. Similarly, 2 genes present in nscaf 2986 of
chromosome 27 also formed a subcluster from the remaining group. It was
considered as 30KPs only based on the conserved sequence present in the c
terminal region of the protein.

In the current study it is observed that paralytic peptide is specifically
up-regulated at 18h time point suggesting its possible role in induction of
diapause along with other factors that contribute to the same. Excessive
amount of PPs cause serious damage to insect self organization. Paralytic
peptide binding proteins are also up-regulated during the diapause period,
which suggests that PP-BP act as active regulators of the paralytic peptide
(Matsumoto et al., 2003), as even the over expression of the paralytic peptide
might cause damage to the cells.

Paralytic peptide and paralytic peptide binding protein interaction has
to be further exploited to find out whether it is an alternative mechanism of
diapause induction and also its possibility to design novel growth regulators in
other insects.