Proteomic analysis of pre-diapause, diapause and post-diapause larvae of the wheat blossom midge, *Sitodiplosis mosellana* (Diptera: Cecidomyiidae)

WEI-NING CHENG¹, XIU-LIAN LI¹, FENG YU², YI-PING LI¹, JIAN-JUN LI¹ and JUN-XIANG WU¹,*

¹The College of Plant Protection, Key Laboratory of Plant Protection Resources and Pest Management, Ministry of Education, Northwest A & F University, Yangling, 712100 Shaanxi, China; e-mail: juxw@yahoo.com.cn
²The College of Animal Science and Technology, Northwest A & F University, Yangling, 712100 Shaanxi, China

Key words. *Sitodiplosis mosellana*, diapause, proteome, 2-dimensional electrophoresis, mass spectrometry

Abstract. To determine the relationship between protein expression and insect diapause, a proteomic approach was used to investigate the proteins extracted from larvae of the wheat blossom midge *Sitodiplosis mosellana* Gehin at different developmental stages, including pre-diapause, over-summering diapause, over-wintering diapause and post-diapause. Using 2-DE gels stained with coomassie brilliant blue, about 300 protein spots were detected in the extracts of pre-diapause larvae and 275 for those in each of the other stages. There were 91, 92 and 95 protein spots that showed more than a 2-fold change in abundance in the over-summering diapause, over-wintering diapause and post-diapause stages compared with pre-diapause. Eight protein spots, which showed the greatest difference in the larvae at different stages of diapause, were analyzed using Matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS). Seven of them were successfully identified from their peptide mass fingerprints using the NCBInr database. They were proopiomelanocortin, NADH dehydrogenase subunit 1 and F10F2.5, which were up-regulated or unique to pre-diapause larvae, IKK interacting protein isoform 2 up-regulated in diapause and post-diapause larvae, GA10647-PA unique to pre-diapause larvae, purple CG16784-PB isoform B and B0228.6 up-regulated in over-summering and over-wintering diapause larvae. The potential functions of these proteins during wheat blossom midge diapause are discussed.

INTRODUCTION

Diapause is a genetically controlled period of developmental arrest that enables insects to survive adverse environmental conditions and synchronize their life cycles with the availability of food (Salama & Miller, 1992; Li et al., 2008). Because of its central role in insect pest survival and reproduction, an understanding of the mechanism of diapause is important for agricultural pest management.

The molecular mechanisms of diapause have been studied extensively recently, and several proteins associated with diapause were found in some insect species. Salama & Miller (1992) reported a 490 kDa glycolipoprotein in the hemolymph of last instar larva of the pink bollworm *Pectinophora gossypiella* during diapause. Other proteins, which have distinct expression patterns in relation to insect diapause, include a 41 kDa protein in the larval fat body of the southwestern corn borer *Diatraea grandiosella* (Brown & Chippendale, 1978), four proteins with molecular weights 30 kDa, 35 kDa, 55 kDa and 60 kDa in the gut of pharate first instar larvae of the gypsy moth *Lymantria dispar* (Lee & Denlinger, 1996), a 77 kDa protein (AgSP-1) in the haemolymph of the boll weevil *Anthonomus grandis* (Lewis et al., 2002) and a 23 kDa heat shock protein in the brain of diapausing pupae of the flesh fly *Sarcophaga crassipalpis* (Li et al., 2007).

The wheat blossom midge, *Sitodiplosis mosellana* (Gehin), is a serious intermittent pest damaging wheat in the northern hemi-sphere (Chen & Ni, 1998; Ding & Guo, 1992). Generally it has one generation per year. The mature larvae fall to the ground from middle to late May from wheat ears and form round cocoons in which the larvae over-summer and over-winter in obligatory diapause. Although diapause in the wheat blossom midge has been intensely investigated, most studies have focused on the ecology (Basedow, 1977; Hinks & Doane, 1988) and the molecular control mechanism is unknown.

Proteomics is a large-scale study of gene expression at the protein level, which ultimately provides a direct measurement of protein expression levels and insight into the activity of relevant proteins. Recently, this approach was successfully used in studies of different aspects of entomology, such as induced immunity (Levy et al., 2004), resistance mechanisms (Jiang et al., 2004), molecular pathology (Kaeslin et al., 2005), developmental process (Li et al., 2006; Jia et al., 2007) and diapause (Joplin et al., 1990; Li et al., 2007). In this report, a proteomic approach was used to investigate the proteome of wheat blossom midge larvae in different stages of diapause in order to determine the relationship between protein expression and insect diapause. The results should facilitate further research into diapause regulation in the wheat blossom midge, which may lead to new ways of effectively controlling this insect pest.
**MATERIAL AND METHODS**

**Insect collection**
Wheat ears containing late instar larvae of the wheat blossom midge were harvested and put on soil in a field insectary in mid-May, 2006. Pre-diapause larvae were collected when the wheat was ripe in the experimental fields in mid-May, 2006. Over-summering and over-wintering larvae were collected in late July and late December, 2006, respectively. Post-diapause larvae were collected in mid-March, 2007. All larvae collected were stored at −80°C until used.

**Main reagents**
Low molecular weight markers, SDS, acrylamide, bis-acrylamide, TEMED, Tris and PlusOne 2D Quant protein assay kit were purchased from Amersham Biosciences (Uppsala, Sweden). Trifluoroacetic acid, acetonitrile, iodacetamide and alpha-cyano-4-hydroxycinnamic acid were purchased from Sigma (St. Louis, MO). CHAPS, urea, dithiothreitol, immobilized pH gradient strip 13cm pH3-10, immobilized pH gradient buffer, ammonium persulphate and agaro agar were purchased from Sino-American Biotechnology (Luoyang, China). Mercaptoethanol, glycine, bromophenolblue, Coo massacre Brilliant Blue R-250, methanol, HCl, acetic acid and glycerol were obtained from Beijing Dingguo Biotechnology (Beijing, China).

**Preparation of protein sample**
One hundred larvae, previously stored at −80°C, were rinsed with double-distilled H2O. Excess water was removed with the aid of filter paper. The larvae were then homogenized in 1.2 ml ice cold 0.02 M phosphate-buffered saline (PBS, pH 7.0) using a pre-chilled glass homogenizer. Insect debris was removed by centrifuging at 8,000×g for 15 min at 4°C, and the supernatant was then transferred into a 1.5 ml centrifuge tube. After freezing overnight at −80°C, the supernatant was lyophilized in a FlexDry™ Lyophilizer (FTS Systems, Stone Ridge, NY). The lyophilized sample was rehydrated in a buffer containing 8 M urea, 2% CHAPS, 0.5% IPG buffer and 1% DTT. Insoluble materials were removed by centrifugation. Protein concentration was determined using the PlusOne 2D Quant protein assay Kit following the procedure recommended by the manufacturer. At least three replicates were performed on each developmental stage.

**2D gel electrophoresis**
First dimension gel electrophoresis was performed using an IPGphor isoelectric focusing (IEF) system (Amersham Biosciences, Uppsala, Sweden) by applying 200 µg of protein extract in 250 µl buffer to a 13 cm, pH3-10 linear IPG strip with rehydration at 30 V for 12 h followed by isoelectric focusing at 500 V for 1 h, 1000 V for 1 h and 8000 V for 2 h. After IEF separation, the strips were equilibrated for 15 min in an equilibration buffer (50 mM Tris-HCl pH 8.8, 6 M urea, 30% glycerol, 0.002% bromophenol blue, 2% SDS and 1% DTT). The strips were then treated for another 15 min with the same buffer, except 1% DTT was replaced with 2.5% iodoacetamide. The equilibrated strips were subjected to a second-dimension electrophoresis after transferring onto a 15% SDS-polyacrylamide gel. Electrophoresis was run on the Hoefer SE 600 system (Amersham Biosciences, Uppsala, Sweden) at 10 mA per gel for 20 min, followed by 20 mA until the bromophenol blue front reached the bottom of the gel. After separation, the gel was stained using colloidal Coo massacre Brilliant Blue R-250.

**Image analysis**
Protein images on each gel were recorded using an Image-Scanner (Amersham Biosciences, Uppsala, Sweden) and analyzed using PDQuest software release 7.4 (BioRad, Hercules, CA), including background subtraction, spot detection, spot matching and spot intensity normalization analysis (total density of gel image), following the protocols provided by the manufacturer.

**In-gel digestion and mass spectrometry**
Protein spots were excised from gels and subjected to in-gel digestion using trypsin following published procedures (Hellman et al., 1995; Courchesne & Patterson, 1999) with slight modifications. Briefly, the gel pieces were destained with 50 mM ammonium bicarbonate, 50% and 100% acetonitrile and then treated with 10 mM dithiothreitol (in 25 mM ammonium bicarbonate) and 55 mM iodoacetamide (in 25 mM ammonium bicarbonate) to reduce and alkylate the cysteines. After that, the gel pieces were washed with 25 mM ammonium bicarbonate in 50% acetonitrile, dehydrated with 100% acetonitrile and dried in a vacuum centrifuge. The gel pieces were then rehydrated in 25 mM ammonium bicarbonate containing trypsin and digestion carried out overnight at 37°C. The enzymatic reactions were stopped by the addition of 2% trifluoroacetic acid. After centrifugation, the peptide mixture was transferred into a clean 0.5 ml centrifuge tube and then concentrated in a SpeedVac (Savant) to 10 µl. 0.5 µl of the concentrated mixture was mixed again with 0.5 µl of saturated α-cyano-4-hydroxycinnamic acid matrix solution, applied to a stainless steel target MALDI plate and air-dried before analysis in a mass spectrometer.

Mass spectrometry was performed on a MALDI-TOF spectrometer (Bruker Autoflex) equipped with a nitrogen laser (337 nm). Peptide mass fingerprint spectra were obtained in the reflectron positive mode by summing 200 laser shots with delayed extraction (90 ns), an accelerated ion source voltage 1 of 19.1 kV and accelerated ion source voltage 2 of 16.4 kV, calibrated using matrix and trypsin autolysis peaks (at m/z 842.51 and 2211.11) as internal standards.

**Protein identification**
Proteins were identified using peptide mass fingerprinting (PMF) performed by Mascot on-line search engine (http://www.matrixscience.com) and reference to the National Center for Biotechnology non-redundant (NCBI-nr) protein database. The search parameters used were enzyme of trypsin; monoisotopic masses; fixed modifications of carbamidomethyl (C); variable modifications of Glu→pyro-Glu (N-term Q), oxidation (M); peptide charge state of 1+; peptide mass tolerance of ± 100 ppm and one max missed cleavages. Only significant hits, defined by mascot probability analysis (P < 0.05), were accepted.

**RESULTS**
2-DE analyses of extracts of larvae in different stages of diapause
Analysis of the three replicates of each sample produced consistent results. Representative images of 2-DE gels of the protein extracts from larvae in different stages of diapause are shown in Fig. 1. The analysis done by PDQuest software indicated that about 300 protein spots were repeatedly detected in the extracts of pre-diapause larvae and 275 in each of those of other diapause stages in the molecular weight range from approximately 7.0–94.0 kDa and isoelectric points range from about 3.0–10.0, and that the majority of the proteins had...
molecular weights ranging from 8.0–50.0 kDa and isoelectric points from about 4.0–9.0.

**Differential expression of proteins among larvae in different stages of diapause**

A comparison of the location and volume of each spot revealed that the majority of proteins were expressed at similar levels in all the larvae. However, there was a distinct portion of proteins that were expressed differentially with at least a two-fold difference among larvae in different stages of diapause. Specifically, 91 proteins differed in abundance in the extracts of larvae in over-summering diapause compared to those in pre-

<table>
<thead>
<tr>
<th>Protein no.</th>
<th>Diapause stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-diapause</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>–</td>
</tr>
</tbody>
</table>

**Note:** – missing protein, + unique protein, ≈ change in protein abundance was less than twice that in pre-diapause larvae. The data in columns 3–5 of this table were the rates of protein abundance in other stages to that in pre-diapausine stage.
diapause. Of these 91 proteins, 26 were unique or upregulated during over-summering diapause, while 65 were either missing or downregulated. During over-wintering diapause, 92 proteins exhibited different expression levels compared to pre-diapause. Of these 92 proteins, 39 were unique or upregulated whereas 53 were either missing or downregulated in over-wintering larvae compared to pre-diapause. Similarly, the expression levels of 95 proteins differed at post-diapause compared to pre-diapause. Of these 95 proteins, 22 were unique or upregulated, whereas 73 were either missing or downregulated in post-diapause larvae compared to pre-diapause larvae.

Overall, the differences in levels of protein expression were smaller among larvae in over-summering diapause, over-wintering diapause and post-diapause than between the levels in them and the level in pre-diapause larvae. The main differences between the level in pre-diapause larvae and other larvae were the low abundance of unique proteins in pre-diapause larvae. Molecular weights of these proteins were approximately 28.0–40.0 kDa, with isoelectric points of 6.6–9.3 and 30.0–46.0 kDa, with isoelectric points of 4.2–6.4.

Among the differentially expressed proteins, eight were selected for MALDI-TOF-MS analysis because of their consistent differences among pre-diapause, diapause and post-diapause larvae (Table 1). Protein 1 and 3 were significantly downregulated during both over-summering and over-wintering diapause as well as post-diapause. Protein 2 was expressed during pre-diapause, but absent during diapause and post-diapause. Protein 4 and 7 were dramatically upregulated during diapause and post-diapause. Protein 5 and 6 were upregulated during over-summering and over-wintering, but not significantly so in post-diapause larvae. Protein 8 was unique to larvae in over-wintering diapause. These eight protein spots are indicated in Fig. 1A, B, C and D, and six of them at a great magnification in Fig. 2.

**MALDI-TOF-MS analysis and protein identification**

The eight spots selected were excised from the gels and subjected to in-gel digestion with trypsin and MALDI-TOF-MS analysis. Fig. 3 shows the separation profile of protein 5. All other proteins yielded satisfactory peptide mass fingerprints (PMF).

The PMF data for individual proteins were used to search the NCBI nr database using the Mascot search...
engine. Seven of the eight proteins were successfully identified with $P < 0.05$. They were proopimelanocortin for protein 1, F10F2.5 for protein 2, NADH dehydrogenase subunit 1 for protein 3, IKK interacting protein isoform 2 for protein 4, purple CG16784-PB isoform B for protein 5, B0228.6 2 for protein 6 and GA10647-PA for protein 8, respectively. Protein 7 was not identified and is possibly a novel protein. The NCBI accession numbers, sequence coverage of peptides, Mowse scores, and theoretical molecular weights and pIs for these proteins are shown in Table 2.

**DISCUSSION**

The complete genome sequencing of many organisms and new advances in proteomics make it possible to analyze global differences in protein among organisms or of the same organism at different developmental stages or experiencing different environmental conditions. Insect diapause, a biological process that allows insects to survive adverse conditions, involves extensive physiological changes (Joanisse & Storey, 1994; Li et al., 2002), and potentially proteins are either upregulated or downregulated during the different stages of diapause. Indeed, preliminary analysis using SDS-PAGE has detected qualitative and quantitative differences during diapause in several insects (Osir et al., 1989; Ichimor et al., 1990; Mao & Cao, 2001). However, very few proteins that may play a role in insect diapause have been identified. In this study, protein expression profiles of the wheat blossom midge larvae during pre-diapause, over-summering diapause, over-wintering diapause and post-diapause were compared. About 300 protein spots were detected in extracts of pre-diapause larvae and 275 protein spots in those in other stages. Of these protein spots, 91, 92 and 95 were more than twice as abundant in larvae in over-summering diapause, over-wintering diapause and post-diapause compared to pre-diapause. The general trend was that the majority of these differentially expressed proteins, for example, proteins 53 and 65, were downregulated or absent during diapause. Only a couple of

**TABLE 2. Proteins identified by PMF.**

<table>
<thead>
<tr>
<th>Protein no.</th>
<th>NCBI GI</th>
<th>Homologous protein name</th>
<th>Species</th>
<th>Sequence coverage (%)</th>
<th>Mowse score</th>
<th>Theoretical Mr (kDa)/pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>213469</td>
<td>Proopi melanocortin</td>
<td><em>Oncorhynchus keta</em></td>
<td>65</td>
<td>87</td>
<td>11.70/9.93</td>
</tr>
<tr>
<td>2</td>
<td>71984420</td>
<td>F10F2.5</td>
<td><em>Caenorhabditis elegans</em></td>
<td>28</td>
<td>60</td>
<td>61.22/5.78</td>
</tr>
<tr>
<td>3</td>
<td>3021562</td>
<td>NADH dehydrogenase subunit 1</td>
<td><em>Pinus sylvestris</em></td>
<td>100</td>
<td>85</td>
<td>4.89/9.21</td>
</tr>
<tr>
<td>4</td>
<td>42491362</td>
<td>IKK interacting protein isoform 2</td>
<td><em>Homo sapiens</em></td>
<td>28</td>
<td>84</td>
<td>39.40/9.21</td>
</tr>
<tr>
<td>5</td>
<td>19550074</td>
<td>purple CG16784-PB, isoform B</td>
<td><em>Drosophila melanogaster</em></td>
<td>53</td>
<td>88</td>
<td>19.50/6.71</td>
</tr>
<tr>
<td>6</td>
<td>25150027</td>
<td>B0228.6</td>
<td><em>Caenorhabditis elegans</em></td>
<td>50</td>
<td>74</td>
<td>17.50/6.15</td>
</tr>
<tr>
<td>7</td>
<td>—</td>
<td>Unknown</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>125808983</td>
<td>GA10647-PA</td>
<td><em>Drosophila pseudoobscura</em></td>
<td>25</td>
<td>61</td>
<td>69.68/8.38</td>
</tr>
</tbody>
</table>

Fig. 3. Peptide mass fingerprint of protein 5.
proteins, i.e., proteins 26 and 39, were upregulated or unique to over-summering and over-wintering diapause larvae, respectively. This result is consistent with the fact that the metabolism of insects is low during diapause, resulting in a lower synthesis of proteins. Whereas some chemicals that are necessary for diapause survival, such as glycerol, trehalose and antifreeze proteins, increase during diapause. Thus the levels of the enzymes necessary for the synthesis of these compounds increase (Joannis & Storey, 1994; Li et al., 2000). Identification of these differentially expressed proteins and elucidation of the regulatory mechanism will advance our understanding of their role in insect diapause.

Through MALDI-TOF-MS analysis, we have for the first time identified seven proteins that were either up- or downregulated during the different diapause stages. Among the seven proteins, four were dramatically upregulated or uniquely expressed during diapause (Tables 1, 2). These upregulated or unique proteins include IKK interacting protein isoform 2 (IKIP2), GA10647-PA, purple CG16784-PB isoform B and B0228.6. The IKK interacting protein can promote apoptosis and inhibit cell division (Hofer-Warbinek et al., 2004). Since insect diapause is a dormant stage during which it is likely cell division is inhibited, the strong upregulation of IKIP2 during diapause may suggest a role of this protein in inhibiting cell division during diapause. Therefore, we speculate that this protein may play a regulatory role in diapause activation of wheat blossom midge larvae. GA10647-PA functions like phosphoenolpyruvate carboxykinase, which is a critical enzyme in gluconeogenesis (Richards et al., 2005). This protein is necessary for the production of phosphoenolpyruvate and glyceraldehyde-3-phosphate, which are precursors of the cryoprotectant glycerol. The fact that this protein was uniquely expressed during over-wintering suggests that it may be involved in cold tolerance of wheat blossom midge by reducing the supercooling point. CG16784-PB isoform B is highly homologous with 6-pyruvyl tetrahydropterin synthase, which is involved in the biosynthesis of pteridine and tetrahydrobipterin. The function of B0228.6 is unclear. Both CG16784-PB and B0228.6 were highly upregulated during both over-summering and over-wintering, and their roles in diapause remain to be resolved.

In contrast to the four proteins that were upregulated during diapause, three proteins, including F10F2.5, proopiomelanocortin and NADH dehydrogenase subunit 1, were strongly downregulated during diapause. F10F2.5 is a protein involved in increasing growth rate. This protein is not expressed during diapause. This observation is consistent with the fact that insect development is suppressed transiently during diapause. Proopiomelanocortin is the precursor of a number of biologically active peptides in the pituitary of animals, and its expression is regulated by several signal regulatory molecules in the cytoplasm and nucleus (Seo et al., 2008). Our results showed that this protein was markedly less abundant during and after diapause, suggesting that this protein may be connected with diapause related signal transduction. NADH dehydrogenase is a key metabolic enzyme in the mitochondrion respiratory chain. The downregulation of this enzyme may reflect the fact that metabolism is less active during diapause. Diapause is characterized by a major shut-down of metabolism and it is not surprising that the proteins associated with metabolic activities were less abundant during diapause. The downregulation of metabolic enzymes was also recorded in the flesh fly Sarcophaga crassipalpis (Li et al., 2007). Several proteins involved in primary metabolism, including a fatty acid binding protein and an endonuclease, were downregulated in diapauing pupae. Proteins that are less abundant in diapausing larvae could potentially be just as important as proteins that are unique to diapause or are more abundant during diapause. At the transcript level, for example, down-regulation of the cell cycle regulator, Proliferating Cell Nuclear Antigen, appears to be a key feature in shutting down the cell cycle during diapause in the flesh fly (Hayward et al., 2005).

In summary, we used a proteomic approach to analyze global differences in protein in the larvae of the wheat blossom midge in different stages of diapause. Numerous proteins that were differentially expressed in the different diapause stages were identified. With the aid of MALDI-TOF-MS analysis, a further seven proteins that were either up or downregulated during diapause were identified. These results should facilitate further analysis with the objective of revealing the molecular mechanism that regulates the diapause process in the wheat blossom midge.

ACKNOWLEDGMENTS. This research was supported by grants from The Ministry of Education of China (105165), the National Science Foundation of China (30370936) and the Research Program of Science and Technology in the Eleventh Five-year Plan of China (2006BAD08A05). We deeply thank J.-Z. Liu of Hebei Normal University, China for his technical assistance; M.-S. Chen in the Plant Science and Entomology Research Unit, USDA-ARS and Department of Entomology, Kansas State University, USA and H. Chen, H.-X. Zhao from Northwest A & F University, China for their suggestions and comments.

REFERENCES


