Cloning and characterization of protease inhibitor-like cDNAs from the Hessian fly *Mayetiola destructor* (SAY)

A. A. Maddur*¶, X. Liu*, Y. C. Zhu**, J. P. Fellers†, B. Oppert‡, Y. Park,* J. Bai§, G. E. Wilde* and M.-S. Chen*†

*Department of Entomology, 123 Waters Hall, Kansas State University, Manhattan, USA; **USDA-ARS-JWDSRC, PO Box 346/141 Exp Stn Road, Stoneville, USA; †USDA-ARS-GMPCR, Plant Science and Entomology Research Unit, 4008 Throckmorton Hall, Kansas State University, Manhattan, USA; ‡USDA-ARS-GMPCR, Biological Research Unit, 1515 College Avenue, Manhattan, USA; §Department of Plant Pathology, Kansas State University, Manhattan, USA ¶Current address: Center for Molecular Biology of Oral Diseases, College of Dentistry, University of Illinois at Chicago, 801 South Paulina Street, Chicago, USA

Abstract

Analysis of transcriptomes from the salivary glands and midgut of Hessian fly larvae *Mayetiola destructor* (say) identified a set of diverse cDNAs that encode proteins with a relatively high percentage (over 10%) of cysteinyl residues. Structural comparison of these putative proteins with known sequences in GenBank revealed that the positions of the cysteinyl residues in the identified proteins were highly conserved within a family of proteinase inhibitors despite very little overall sequence similarity. Phylogenetic analysis sorted this set of cDNAs into five different groups. To determine if these cDNAs indeed encode proteinase inhibitors, recombinant proteins were generated with two cDNAs from two different groups. Biochemical analysis of the recombinant proteins against commercial and insect gut proteinases demonstrated that the recombinant proteins are strong proteinase inhibitors with different specificities. Northern blot and real-time PCR analysis revealed that the different genes were expressed at different developmental stages and in different tissues. The overall results indicated that *M. destructor* contains a complex of genes that code for proteinase inhibitors which may regulate proteinase activities in different regulatory pathways. The GenBank accession numbers for the cDNAs in this paper were DQ232690 to DQ232718.

Keywords: protease inhibitors, wheat, Hessian fly, *Mayetiola destructor*.

Introduction

Proteolysis, the breakdown of proteins into smaller peptides or amino acids by proteinases, is involved in various biological processes such as food digestion, gene replication, cell-cycle progression, neuronal outgrowth, wound healing, and immunity (Sternlicht & Werb, 2001). Proteinases include all enzymes that catalyse hydrolysis of peptides, and include both exopeptidases (amino- and carboxypeptidases) and endopeptidases which cleave internal peptide bonds, also referred to as ‘proteinases’ (Neurath et al., 1989). Because of its vital role in critical cellular processes, proteolysis is strictly regulated. In mammals, deficiency in proteolytic regulation can lead to serious pathological conditions such as cancer, arthritis, neurodegenerative, and cardiovascular diseases (Sternlicht & Werb, 2001). In insects, deficiency in proteolytic regulation may cause abnormalities in physiological processes such as morphogenesis, diapause, and immunological response (Suzuki & Natori, 1985; Yamamoto et al., 1999; Meier et al., 2000; Huang et al., 2001a; Goto & Denlinger, 2002; Liao et al., 2002; Liu et al., 2003).

One of the mechanisms for regulation of proteolysis is through the inhibition of proteinases by proteinaceous inhibitors (Dickinson et al., 1995; Armstrong et al., 1996; Turk et al., 2002). A classical example of proteolytic regulation through proteinase inhibitors is the inhibition of serine proteinases that are involved in blood clotting in vertebrate animals. In this cascade, activated enzymes are regulated by corresponding inhibitors that are also circulating in the blood. In humans, deficiency in proteinase inhibitors circulating in the blood can lead to various clinical disorders. For example, deficiency in α-1 antitrypsin, a serine proteinase inhibitor of neutrophil elastase, is the cause of emphysema (Huber & Carrell, 1989). In insects, it is well known that proteinase inhibitors play a regulatory role in melanization, one of the defensive mechanisms adopted by insects against
various pathogens (Jiang & Kanost, 1997). Melanization involves the activation of proteinases, leading to the activation of phenoloxidase, which catalyses formation of quinones. The quinones and reactive oxygen species produced during melanization are toxic to invasive microorganisms as well as insect cells (Boigregrain et al., 1992). Under normal conditions, to avoid unwanted production of these reactive molecules, insects have evolved various serine proteinase inhibitors to control the activity of proteolytic enzymes (Sugumaran et al., 1985; Aspan et al., 1990; Boigregrain et al., 1992; Jiang & Kanost, 1997).

The Hessian fly, *Mayetiola destructor* (say), is one of the most destructive pests of wheat crops (Hatchett et al., 1987). Initial analysis of the *M. destructor* gut indicates that there is a complex proteinase network present (Shukle et al., 1985; Zhu et al., 2005). Understanding the regulation of proteolysis by proteinase inhibitors in *M. destructor* will not only contribute to the basic understanding of the proteolysis process, but may also have practical application in identifying new control products for this insect pest. Identification and characterization of proteinases and proteinase inhibitors is the first step to understand the proteolytic system and its regulation. To date, proteinase inhibitors have not been identified in *M. destructor*.

To identify genes that encode potential proteinase inhibitors, we followed an Expressed-Sequence-Tag (EST) approach by randomly sequencing cDNA clones derived from different tissues. A total of 8000 ESTs were produced from cDNA libraries made from either whole body or dissected tissues of *M. destructor* larvae. Sequence analysis of the cDNAs revealed that 8 of these clones encode proteins that are likely proteinase inhibitors because these putative proteins share 40% or higher sequence identity with pancreatic trypsin inhibitor II (Kraunsoe et al., 1996). In addition, there are 56 cDNAs that encode proteins with little sequence similarity (24% or less amino acid identity) with known proteinase inhibitors. However, these putative proteins contain a high content of cysteinyl residues in conserved positions. Since the presence of numerous cysteinyl residues is one of the characteristics of some proteinase inhibitors, we hypothesized that these putative proteins are also proteinase inhibitors. Here we report the molecular characterization of the cDNAs and proteinase inhibitory activity of recombinant proteins generated with representative cDNAs.

**Results**

**Characterization of cDNA clones**

An analysis of random sequences from cDNAs, made with the transcriptomes isolated from the salivary glands and midgut of *M. destructor* larvae, revealed a set of diverse sequences that encode proteins containing a number of cysteinyl residues. These cDNA clones were selected for further sequencing and characterization. From that sequencing 51 (of 56 clones) cDNAs were full length. Twenty-nine of the full length cDNAs encoded unique proteins. The remainders of the clones were either redundant or contained minor nucleotide differences in the non-coding regions. Phylogenetic analysis of the 29 unique proteins encoded by the cDNAs sorted the proteins into three lineages and five groups (Fig. 1). Lineage I and III each represented an independent group. Lineage II was further separated into three sublineages, each of which represented distinct groups. Sequence analysis revealed that members of the same group shared greater than 85% sequence identity with E-values < 10^-50 (Fig. 2) while members from different groups share less than 51% identity (data not shown). All putative

![Figure 1. Phylogenetic analysis of the putative *Mayetiola destructor* proteinase inhibitors. Bootstrap values are indicated at the branching sites. Phylogenetic tree shows three lineages. One of the lineages (Lineage II) contains three sublineages. The three sublineages together with lineage I and lineage III were designated five different groups, Group I to V.](https://example.com/figure1.png)
proteins contained a secretion signal peptide at the N-terminal according to PSORT and/or SignalP analysis. There were 10 members in group I, which encoded eight unique proteins (Fig. 2A). cDNAs L7G8 and 29E7 encode identical proteins, but with nucleotide differences in the non-coding regions. All members in group I contained the same number of amino acid residues. The differences between the various members of this group were single amino acid substitutions. There were nine members in group II (Fig. 2B). Group II was more diversified than other groups. In addition to amino acid substitutions, small insertions/deletions (indels) and different endings also existed among group II members. Group III contained seven members that encoded six unique proteins (Fig. 2C).
were many single amino acid substitutions among members from this group. In addition, there was a three amino acid indel in one of the members. Group IV contained two members (Fig. 2D). Again, these two proteins were similar proteins with several amino acid differences. Only one member was identified in group V (Fig. 2E).

Conservation of cysteinyl residues

A common feature in the protein primary structure of members from different groups was that all members contained a relatively high content of cysteinyl residues (7–10 cysteines out of 60–92 residues) in the mature proteins (Figs 2 and 3). Sequence alignment with representative proteins from each group and two known proteinase inhibitors using ClustalW (http://www.ebi.ac.uk/clustalw/) demonstrated that the positions for most of the cysteinyl residues are conserved (Fig. 3). In addition to the six completely conserved cysteinyl residues, two other cysteines (at position 14 and 47) are also conserved in six of the seven proteins. Despite the highly conserved cysteines, the remainder of the sequences has little sequence similarity. A pair-wise comparison using BLASTP revealed overall sequence identity of 40% (29% excluding conserved cysteines) between G14A4 (group II) and Lg2F7 (group IV), 36% overall (26% excluding conserved cysteines) between G14A4 (group II) and Lg2A3 (group III), and 53% overall (37% excluding conserved cysteines) between Lg2F7 (group IV) and Lg2A3 (group III). Except for these three groups of lineage II, there was no sequence similarity detected between other groups using BLASTP. The highly conserved cysteines suggested a similar tertiary structure among these proteins. Since Ascaris trypsin inhibitor (ATI) and Bombina skin trypsin inhibitor (BSTI) are two serine proteinase inhibitors in the Ascaris family (Mignogna et al., 1996), the newly identified proteins may also possess proteinase inhibitory activity.

Inhibitory activity of recombinant proteins towards serine proteinases

To test the hypothesis that the newly identified proteins are also proteinase inhibitors, we selected two cDNAs, G11A6 from group I and G14A4 from group II, to generate recombinant proteins. Recombinant proteins were expressed at levels that an extra protein band could be seen on SDS/PAGE for each construct in comparison with controls (Fig. 4A).
The recombinant proteins were purified to homogeneity, and the fusion tag was removed from the recombinant proteins using enterokinase (Fig. 4B).

To determine if the recombinant proteins are functional protease inhibitors, inhibition assays were carried out by testing the recombinant proteins G11A6 and G14A4 towards a number of commercial proteases as well as towards protease activities present in the midgut of insect pests. Recombinant protein G11A6 inhibited trypsin activity by as much as 62% at a molar ratio of 1.54–1 (Fig. 5A), with IC50 0.52 μM and Ki 3.5 nM. Interestingly, G11A6 also exhibited ~30% inhibition to papain (a cysteine protease). Even though the inhibition of G11A6 is weak inhibitor to papain, the percentage of inhibition was consistent from assay to assay under our experimental conditions. No inhibition was detected in the assays against chymotrypsin. In addition to these commercial enzymes, G11A6 also inhibited protease activity in gut extracts derived from M. destructor, red flour beetle (Tribolium castaneum), and yellow meal worm (Tenebrio molitor) by 35%, 82%, and 60%, respectively, under the defined conditions of the assay.

In comparison, recombinant protein G14A4 inhibited trypsin activity by as much as 68% and chymotrypsin activity by 54% at a molar ratio of 3.2–1 (Fig. 5B). The IC50 values for trypsin and chymotrypsin were 0.9 and 2.32 μM, and Ki values were 8.5 and 15.2 nM, respectively. G14A4 had no significant impact on the activity of papain. Similar to G11A6, G14A4 also inhibited protease activity in the gut extracts from M. destructor, T. castaneum, and T. molitor by 62%, 82%, and 86%, respectively.

Developmental regulation of gene expression

Northern blot analysis was conducted to study the differential gene expression profile in different developmental stages.
of the insect with probes derived from cDNAs representing each of the five groups of *M. destructor* inhibitors. Since cDNA sequences from the same group were very similar, the probe were designed to hybridize with transcripts of members in each group.

The expression patterns of mRNA recognized by group specific probes varied widely among the different inhibitor groups (Fig. 6). Group I was abundantly expressed in 4- to 12-days old larvae, and was weakly expressed in pupa. The highest RNA level was detected in 6-day old larvae when the larva transited from 1st to 2nd instar. Group II was detected in 0-day old larvae, but the RNA level was significantly elevated in 2- to 6-day old larvae. No RNA was detected in 12-day larvae, pupae, and adults. The expression pattern of group III and group IV was similar, with a low level in newly hatched larvae, a slight increase in 2-day old larvae, and maximum level in 4- to 12-day old larvae. Maximum expression was detected by group III probe in 4- and 6-day old larvae, but group IV expression was maximum in 4-day old larvae. Group V had a quite different expression pattern. It was only expressed in late instar larvae and pupae.

**Tissue specific expression**

To quantitatively determine transcript levels in different tissues of *M. destructor* larvae, real-time PCR was conducted with primers specific to each of the five groups. Transcripts of group I genes were detected in all tissues, with the highest level in malpighian tubules, followed by the midgut and salivary glands (Fig. 7A). Only low levels were detected in fat bodies and carcass. The expression profiles for gene groups II, III, and IV was very similar. These three gene groups were predominantly expressed in the malpighian tubules, were expressed at a relatively much lower level in the midgut and salivary glands, and were barely detectible in fat bodies and carcass (Fig. 7B, 7C,D). Similar to the results from the Northern blot analysis, in the expression pattern of group V transcripts were very different. This gene group was predominantly expressed in salivary glands, and only low levels of expression were detected in other tissues (Fig. 7E).
Immunohistochemistry

Since all of the putative inhibitors have a sectional signal peptide, the production sites of the proteins may be different from the actual physiological function sites for these proteins. To determine where the proteins are located in the larva, indirect immunostaining of different tissues with polyclonal antibodies against a synthetic peptide was carried out. Due to the limitation of resources, only one antibody was generated. The antibody was produced against a 20 amino acid peptide synthesized from a sequence in the G14A4 protein (group II). The antibody was affinity purified and demonstrated to recognize the recombinant protein (data not shown).

G14A4 antibodies recognized proteins that were abundantly present in the malpighian tubules, particularly at the distal tip and in the middle segment of the tubules (Fig. 8A, 8B,C). Very low levels of the protein were detected in the salivary glands (Fig. 8E, 8F,G). The protein was detected in the cells located at the bottom of the basal region and also in the tubular cells. A low level of the immunoreactivity was also detected in gut cells, where the protein was located throughout the cytoplasm as small speck of granules. No protein was detected in the nucleus.
Discussion

In this report, we characterized five groups of cDNAs from the salivary glands and midgut of M. destructor larvae that encoded proteins with homology to proteinase inhibitors. Members from the same groups encoded similar proteins with minor differences in amino acids and small indels. These group members may represent multiple alleles of the same gene or different genes of recent duplicates. In contrast to highly conserved sequences within the same group, members from different groups encoded highly diversified proteins. Among the five groups, the sequence from group V was so different that no sequence similarity was detected to any of the other groups (data not shown). Sequences from groups II, III, and IV were less diversified. The sequence identity between groups II and III, II and IV, and III and IV, were 36%, 41%, and 53% in the mature protein region, and 76%, 95%, and 80% in the signal peptide region, respectively. The lower degree of conservation in the mature protein region compared to the signal peptide region indicates that the coding region for the mature protein has been under higher selection pressure for mutations, and these changes may result in different proteins for different functions or specificities. The rapid diversification among these genes was further supported by the fact that no apparent sequence similarities of these cDNAs and their corresponding proteins could be found against databases of related species such as Drosophila except those cysteinyl residues (data not shown). Diversification in gene duplicates was also observed in proteinase inhibitors in other organisms (Harrison et al., 2002). It is believed that divergent evolution of proteinase inhibitors is the response of the insect to counteract the rapidly evolving targeted proteinases, especially for those involved in insect defense against microorganisms or detrimental changes in host plants (Boigregrain et al., 1992; Zhang & Maizels, 2001). From that point of view, the rapid diversification of the protease inhibitors described here could indicate that some of them may be involved in defense response to microorganism invasion or in insect–host interaction in addition to their potential role in developmental regulation. Even though these inhibitor genes have undergone rapid divergent evolution, their inhibitory activity of encoded proteins is maintained and may be due to conservation of the cysteinyl residues at their respective positions (Harrison et al., 2002; Zhang & Maizels, 2001).

So far, 15 families of serine proteinase inhibitors have been identified (Laskowski & Kato, 1980; Reeck et al., 1997). Each family is characterized by the number of cysteinyl residues and disulphide bonds. For example, the pancreatic trypsin inhibitor Kunitz family contains 4–8 disulphide bonds. The Bowman-Birk inhibitors contain seven disulphide bonds. The M. destructor proteinase inhibitors are most closely related to those of the Ascaris family (Gronenborn et al., 1990). Members in the Ascaris family typically have 10 cysteines that form five disulphide bonds (Bernard & Peanasky, 1993; Gronenborn et al., 1990; Mignogna et al., 1996). The M. destructor proteinase inhibitors contain various numbers of cysteinyl residues. In general, group I and group IV contain 10 cysteines in the predicted mature proteins (Fig. 2 A and 2D), while group II and group V contain nine cysteines (Fig. 2B,E) and group III contains eight cysteines (Fig. 2C). There are two exceptions; one in group I and another in group III, in which one of the cysteines is substituted with an arginine (Fig. 2 A and 2C). Despite the variation in the numbers, the positions of the cysteinyl residues are highly conserved. Cysteines at positions 29, 34, 59, 61, 76, and 85 are conserved in all the aligned sequences (Fig. 3). In addition, cysteines at position 14 and 48 are also conserved among all sequences except one in each case. According to the conserved positions of the cysteines, we conclude that the M. destructor inhibitors belong to the Ascaris family. The significance of the variation in the numbers of cysteines among different groups as related to function remains to be determined.

Northern blot and real-time-PCR analyses demonstrated that M. destructor inhibitor genes were expressed at different developmental stages and in different tissues. Group I genes were transitorily expressed in 4- to 12-day larvae, and its transcripts were detected in multiple tissues, including the gut, salivary glands, and Malpighian tubules. This transitory expression could indicate that this gene group is involved in developmental regulation. Under our experimental conditions, M. destructor larvae transform from 1st into 2nd instar on the 4th to 6th day, and from 2nd into 3rd instar on the 10th to 11th day. It is possible that group I proteins play a regulatory role in these transitions during larval morphogenesis. Further correlation with more developmental time points is necessary to test this postulate. In contrast to the other groups, group V genes were transitorily expressed in the late larval stage and in pupae, and its transcripts were detected nearly exclusively in salivary glands. In M. destructor, the 3rd instar larva is a non-feeding stage that is also called prepupa. Pre-pupation and pupation take place on the 9th to 10th day, and on the 13th to 15th day, respectively, under our experimental conditions. Thus, proteins from this group may be involved in preprepation and pupation, where massive remodeling of tissue takes place and presumably the regulation of proteinases is critical. Proteinase inhibitors are involved in the regulation of insect development. For example, in mosquitoes inhibitors of chymotrypsin are known to regulate the development of the insect (Horler & Briegel, 1997). In Drosophila, the inhibitor gene kil-1 is expressed during preupal stage (Kress et al., 2004). The KIL-1 protein is synthesized in the preupal salivary glands and released into the molting fluid during pupation, which coincides with loss of proteinase activity in pupae.

The other three groups (II, III, and IV), which belong to the same sublineage (Fig. 1), had similar expression patterns. They were all abundantly expressed in the larval
stage and their transcripts were almost exclusively present in Malpighian tubules. The Malpighian tube specific expression was further confirmed by indirect immunostaining with an antibody specific to a member (G14A4) of group II. The *M. destructor* Malpighian tubules are freely floating in the haemolymph. Considering the fact that all of the inhibitors have a secretion signal peptide, the synthesized inhibitors may be secreted at the same site and transported through the haemolymph to other tissues to perform their biological functions. Western blot analysis with the G14A4 antibody demonstrated that this protein was abundantly present in the haemolymph (data not shown). Recently, the expression of aminopeptidase genes and the presence of proteinase activities were documented in the Malpighian tubules of *Trichoplusia ni* (Wang et al., 2005). Insect Malpighian tubules are known to be involved in urine secretion and excretion of organic metabolites (Sözen et al., 1997). Recently, the studies of Malpighian tubules of *Drosophila* found that many immune-related genes are also expressed in Malpighian tubules (Wang et al., 2004). Considering the fact that these genes were under selection pressure for rapid diversification, they may also have regulatory roles in immunological process controlled by protease cascades. Therefore, studies of the biological impact of the *M. destructor* proteinases and proteinase inhibitors expressed in Malpighian tubules may reveal important features of insect defense, development control, and maintenance of homeostasis.

**Experimental procedures**

**Insects**

Three *M. destructor* populations, a Kansas population (biotype GP), biotype L, and a Syrian population, were used in this study. The Kansas population was collected from Ellis County (Gagane and Hatchett, 1989) and maintained on susceptible wheat in a greenhouse. Biotype L was provided by S. Cambrown at the USDA *Mayetiola destructor* laboratory in West Lafayette, Indiana. The Syrian population was maintained at the International Center for Agricultural Research in Dry Area in Aleppo, Syria. Dissected tissues were shipped in dry ice to the USA for processing.

cDNA libraries, sequencing, and sequence analysis

cDNA libraries from dissected tissues were constructed and sequenced as described previously (Chen et al., 2004; Zhu et al., 2005). The clones derived from biotype GP, biotype L, and the Syrian population were named starting with G, L, and S, respectively. Sequence analysis was performed using programs at http://www.ncbi.nlm.nih.gov. Alignments were generated with ClustalW (http://www.ebi.ac.uk/clustalw/).

**Phylogenetic analysis**

Full length putative protein sequences were used to construct the phylogenetic tree with softwares GENEBEE and PHYLIP (v4.0) available at (http://www.genebee.msu.su/services/phtree_reduced.html).

**RNA extraction and Northern analyses**

Total RNA was extracted from the different stages of biotype GP larvae as described previously (Chen et al., 2004). 100 mg of tissues were homogenized in 1 ml of TRI reagent™ (Molecular Research Center, Inc., Cincinnati, OH, USA). Total RNA was then extracted following the protocol provided by the manufacturer. 5 g of total RNA was separated on 1.2% agarose gel containing formaldehyde in MOPS buffer and blotted on to the GeneScreen Plus™ membrane after the separation. The RNA was cross linked on to the membrane by baking the membrane at 80 °C for 2 h. The membranes were then hybridized separately to individual cDNA probes generated using a random labelling kit from Stratagene (La Jolla, CA, USA). Hybridization was carried out overnight at 42 °C in a plastic bag containing a 15 ml hybridization solution (10% dextran sulphate/1% SDS/1 M NaCl, pH 8.0). After hybridization, the membranes were washed twice with 2xSSC at room temperature for 30 min, twice with 2xSSC plus 1% SDS at 65 °C for 30 min, and twice with 0.1xSSC plus 1% SDS at room temperature for 30 min. The membranes were then exposed to Kodak SR-5 X-ray film overnight.

**Real-Time PCR**

Salivary glands, midgut, Malpighian tubules, fat bodies, and carcass were dissected from 1st biotype GP instar larvae. Total RNA was then isolated from these tissues as well as from whole larvae using TRI reagent™ following the protocol provided by the manufacturer. The total RNA was reverse-transcribed into cDNA using super-script reverse transcriptase (BD Biosciences, San Jose, CA, USA) following the manufacturer’s procedure. After removing the RNA in the reaction by DNase-free RNase from the same company, the cDNAs were used as template for real-time PCR analysis.

Real time PCR reactions were performed with iQ SYBR Green Supermix on an iCycler real time detection system (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer’s instructions. Each reaction was carried out with 2 μl of a 1/40 (v/v) dilution of the first cDNA strand. 0.5 μM of each primer in a total volume of 10 μl. The cycling conditions were the following: 95 °C for 5 min followed by 45 cycles of denaturation at 95 °C for 20 s, annealing at 62 °C (63 °C for G14A4, 64.5 °C for Lg2F7) for 45 s. At the end of the cycles, PCR amplification specificity was verified by obtaining the dissociation curve, from which samples were cooled to 55 °C after denaturing, and then treated for 10 s with 0.5 °C increment for each cycle, for a total of 80 cycles until reaching 95 °C. The PCR products were analysed on 1.5% agarose gels, and subsequently purified and sequenced. Primers were designed using the Beacon Designer (v2.0) software from Biosoft International (Palo Alto, CA, USA) and the primer sequences were listed in Table 1. PCR fragments corresponding to each gene were cloned into pCR® 2.1-TOPO plasmid (Invitrogen, Carlsbad, CA, USA), sequenced, and serial dilutions of these plasmids were used to generate a calibration standard curve, where Ct (Cycle threshold) values are plotted to serve as standard concentrations. The transcript concentration for each sample was calculated based on these standard curves. A negative control reaction without template was always included for each primer combination. Template concentrations of different tissue samples were normalized by the ribosomal protein S30 (RPS30) gene. For each sample analysed, results represent the mean of values obtained from at least two independent PCR reactions, and from at least three independent biological experiments.
**Table 1. Primers for real-time PCR analysis. The primers were designed using the Beacon Designer (v2.0) software from Biosoft International (Palo Alto, CA, USA)**

<table>
<thead>
<tr>
<th>Clone Name</th>
<th>Primer Type</th>
<th>Sequence</th>
<th>Product Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>G11A6</td>
<td>F</td>
<td>5′-GCTGCCTTGACAGTATTGAGG-3′</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5′-TCGTCAGCGCCACGATTTG-3′</td>
<td></td>
</tr>
<tr>
<td>G14A4</td>
<td>F</td>
<td>5′-TACGAGGTAGCTGCGGAGCTC-3′</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5′-CAGGGCCGACTCCAATTTACTG-3′</td>
<td></td>
</tr>
<tr>
<td>Lg2A3</td>
<td>F</td>
<td>5′-TCTGCTGCTTCTTGCTACATC-3′</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5′-TTGCACAAGTGCTTGACGCG-3′</td>
<td></td>
</tr>
<tr>
<td>Lg2F7</td>
<td>F</td>
<td>5′-GCACTTGGTGCTGACGCTGAACG-3′</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5′-ACGGAACAGAGCGATTCC-3′</td>
<td></td>
</tr>
<tr>
<td>G11C4</td>
<td>F</td>
<td>5′-CACCAGAAAGCTGCGACGAC-3′</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5′-GATTGACGTTGACGTTGACACG-3′</td>
<td></td>
</tr>
<tr>
<td>RPS30</td>
<td>F</td>
<td>5′-ACCAAATGGCCACCTTGAGGAG-3′</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5′-TCAAAGCTGTCAATCTGTTGAG-3′</td>
<td></td>
</tr>
</tbody>
</table>

**Generation of recombinant proteins**

PCR® T7/NT-TOPO vector (Invitrogen) was used to make expression constructs. Two representative cDNAs, G11A6 (group I) and G14A4 (group II), were selected to generate recombinant proteins. The inserts that encode mature proteins as predicted by PSORT (http://psort.bii.a-star.edu.sg/) were produced through PCR using primer pairs that cover the specific regions (5‘→3‘). Inserts were then ligated into the pCR® T7/NT-TOPO vector (Invitrogen) to make expression constructs. Two representative cDNAs, G11A6 (group I) and G14A4 (group II), were selected to generate recombinant proteins.

**Inhibitory activity assay**

Commercial trypsin, chymotrypsin and papain were purchased from Worthington Biochemical Corporation (Lakeview, NJ, USA). Bovine pancreatic trypsin and chymotrypsin were dissolved in 0.001 N HCl at 1 mg/ml and 10 mg/ml, respectively. Papain was dissolved in a buffer containing 50 mM Cysteine-HCl, pH 6.2, 10 mM EDTA, and 60 mM mercaptoethanol. Insect gut extracts were obtained from M. destructor larvae as described by Zhu et al. (2004) and from other insects similar to the procedure by Oppert et al. (1997). Briefly, midguts were collected in deionized water, homogenized, and pooled (20 guts per 20 ml) into eppendorf tubes on ice. The extract was centrifuged at 15,000 g for 5 min, and the supernatant was aliquoted to 10 ml and frozen at −20°C.

Inhibitory activity assays of the recombinant proteins were conducted by measuring the hydrolysis of casein conjugated to a fluorescent probe, BODIPY-TR-X (Molecular Probes, Eugene, OR, USA) according to a method previously described (Oppert et al., 1997). Inhibitor proteins were preincubated with trypsin, chymotrypsin, papain, or with gut extracts from M. destructor, T. molitor, and T. castaneum (0.5 gut equivalents) for 10 min at 37°C, respectively, before the addition of substrate. At time 0, 10 ng of substrate was added to the reaction. The plates were incubated at 37°C for 4 h and inhibitory activity was measured using fluorescence microplate reader (Fluoroskan Ascent FL, Labsystems, Thermo Electron Corp., Milford, MA, USA), with an excitation coefficient of 584 nm and emission of 620 nm.

The percentage of inhibition was calculated by dividing the treatment value with the corresponding control. IC50, the amount of the inhibitor concentration resulting in 50% inhibition, was calculated using linear regression of data points. The equilibrium dissociation constants (Ki) for trypsin, chymotrypsin and papain were determined using the concentrations of active enzyme molecules in the solution, which were determined by titration following the methods described by Chase & Shaw (1969).

**Indirect immunostaining**

For indirect immunostaining, a polyclonal antibody against a peptide (KDYYDHSSNECKLPATFAK) derived from G14A4 was produced commercially by Bethyl Laboratories Inc. (Montgomery, TX, USA). The antibodies were affinity purified by passing the hyperimmune serum from the rabbits against immunosorbents. The specificity of the antibody against the G14A4 recombinant protein and proteins from different tissues of M. destructor larvae was determined through Western blots before immunostaining.

Tissues for the immunostaining including salivary glands, midgut, and malphigian tubules were collected by dissecting first instar Mayetiola destructor larvae in PBS. The tissues were immediately fixed in 4% paraformaldehyde overnight at 4°C. The tissues were washed 6 times, 15 min each with PBS containing 1% Triton-X-100 (PBST). Then the tissues were incubated overnight at 4°C in 1:1000 primary antibody diluted in PBST. After incubation, tissues were washed 8 times, 15 min each with PBST. Tissues were then blocked with 5% goat serum overnight at 4°C. After washing (PBST for 6 times, 15 min each), the tissues were incubated with 1:500 secondary antibody (goat anti-rabbit conjugated with Cy3) in PBST overnight at 4°C and the samples were covered with aluminum foil to avoid light. The samples were then washed eight times in PBST. Tissues were mounted on the glass slide using the mounting medium and the tissues were excited at 550 nm under a confocal microscope (Leica Microsystems, Bannockburn, IL, USA).
Acknowledgements

This is contribution no. 06–172-J from the Kansas Agricultural Experiment Station. Mayetiola destructor voucher specimens (No. 150) are located in the KSU Museum of Entomological and Prairie Arthropod Research, Kansas State University, Manhattan, Kansas. The authors thank Drs Srini Kambhampati and Sukhwinder Singh for reviewing an earlier version of the manuscript. The authors also thank Drs Philine Wangemann and Mariam Burton for assistance in confocal microscopy. This work was supported by a grant from the National Research Initiative Competitive Grant Program, the U.S. Department of Agriculture (USDA 04-35607-14861). The real-time PCR work was carried out at the KSU Gene Expression Facility supported by NSF-MRI grant DBI-0421427.

References


