



# **Chymotrypsin-like peptidases in insects**

## **Dissertation**

zur Erlangung des akademischen Grades

**Doctor rerum naturalium**

(Dr. rer. nat.)

Fachbereich Biologie/Chemie der Universität Osnabrück

vorgelegt von

**Gunnar Bröhan**

Osnabrück, Mai 2010



## Table of contents

<b>1. Introduction</b>	<b>1</b>
<b>1.1. Serine endopeptidases</b>	<b>1</b>
<b>1.2. The structure of S1A chymotrypsin-like peptidases</b>	<b>2</b>
<b>1.3. Catalytic mechanism of chymotrypsin-like peptidases</b>	<b>6</b>
<b>1.4. Insect chymotrypsin-like peptidases</b>	<b>9</b>
1.4.1. Chymotrypsin-like peptidases in insect immunity	9
1.4.2. Role of chymotrypsin-like peptidases in digestion	14
1.4.3. Involvement of chymotrypsin-like peptidases in molt	16
<b>1.5. Objective of the work</b>	<b>18</b>
<b>2. Material and Methods</b>	<b>20</b>
<b>2.1. Material</b>	<b>20</b>
2.1.1. Culture Media	20
2.1.2. Insects	20
<b>2.2. Molecular biological methods</b>	<b>20</b>
2.2.1. Tissue preparations for total RNA isolation	20
2.2.2. Total RNA isolation	21
2.2.3. Reverse transcription	21
2.2.4. Quantification of nucleic acids	21
2.2.5. Chemical competent <i>Escherichia coli</i>	21
2.2.6. Ligation and transformation in <i>E. coli</i>	21
2.2.7. Preparation of plasmid DNA	22
2.2.8. Restriction enzyme digestion of DNA	22
2.2.9. DNA gel-electrophoresis and DNA isolation	22
2.2.10. Polymerase-chain-reaction based methods	23
2.2.10.1. RACE-PCR	23
2.2.10.2. Quantitative Realtime PCR	23
2.2.10.3. Megaprimer PCR	24
2.2.11. Cloning of insect CTLPs	25
2.2.12. Syntheses of Digoxigenin-labeled DNA and RNA probes	26
2.2.13. RNA gel-electrophoresis	26
2.2.14. Northern blotting	27
2.2.15. Genomic DNA isolation and Southern blotting	27
2.2.16. <i>In situ</i> hybridization	27
2.2.17. Generation of double stranded RNA	28
2.2.18. Double stranded RNA injection and RNA interference	28
<b>2.3. Biochemical methods</b>	<b>30</b>
2.3.1. Tissue preparations	30
2.3.2. Determination of protein concentration	30
2.3.3. SDS-Page and Coomassie blue staining	30
2.3.4. Immunological methods	31
2.3.5. Detection of glycoproteins	31

2.3.6. Heterologous expression of recombinant proteins	31
2.3.7. Affinity purification of proteins	32
2.3.8. Generation of antibodies	33
2.3.9. Peptidase activity assays	34
<b>2.4. Other methods</b>	<b>34</b>
<b>3. Results</b>	<b>35</b>
<b>3.1. Interaction between MsCTLP1 and MsCHS2</b>	<b>35</b>
3.1.1. Introduction	35
3.1.2. Isolation and sequencing of the cDNA encoding MsCTLP1	36
3.1.3. Localization of MsCTLP1 and MsCHS2	38
3.1.4. Immunological detection of MsCTLP1	39
3.1.5. Co-immunoprecipitation reveals that MsCTLP1 interacts with MsCHS2	40
<b>3.2. MsCTLP1 is a glycoprotein</b>	<b>41</b>
<b>3.3. Midgut MsCTLPs are encoded by a multigene family</b>	<b>42</b>
<b>3.4. Estimation of gene copy numbers by Southern blotting</b>	<b>45</b>
<b>3.5. MsCTLPs are expressed in the midgut and the Malpighian tubules</b>	<b>46</b>
<b>3.6. <i>MsCTLP</i> mRNA expression is not affected by starvation but down-regulated during molt</b>	<b>47</b>
<b>3.7. MsCTLP1 secretion is suspended during molt and starvation and stimulated upon feeding</b>	<b>49</b>
<b>3.8. Identification of CTLP-encoding genes in the <i>Tribolium</i> genome</b>	<b>52</b>
<b>3.9. Isolation and sequencing of <i>TcCTLP</i> cDNAs</b>	<b>52</b>
<b>3.10. Expression profiles of <i>TcCTLPs</i> genes</b>	<b>56</b>
<b>3.11. <i>TcCTLP</i> gene expression in carcass and midgut</b>	<b>57</b>
<b>3.12. Purification and recombinant expression of insect CTLPs</b>	<b>59</b>
3.12.1. Purification of MsCTLP1 from the midgut	59
3.12.2. Recombinant expression of MsCTLP1	60
3.12.3. Exchange of the hydrophobic region of MsCTLP1	62
3.12.4. Expression of recombinant TcCTLPs and antibody generation	64
<b>3.13. RNAi of <i>Tribolium</i> CTLPs</b>	<b>67</b>
3.13.1. Effects of dsRNA induced knockdown of <i>TcCTLP</i> transcripts	67
3.13.2. Knockdown of <i>TcCTLP-5C</i> transcripts affects larval-pupal and pupal-adult molt	69
3.13.3. Knockdown of <i>TcCTLP-6C</i> transcripts affects pupal-adult molting	72

---

<b>4. Discussion</b>	<b>73</b>
<b>4.1. The structure function link</b>	<b>73</b>
4.1.1. Structure of <i>Manduca</i> chymotrypsin-like peptidases	73
4.1.2. Maintenance of catalysis at alkaline conditions	75
4.1.3. Structure of <i>Tribolium</i> chymotrypsin-like peptidases	76
<b>4.2. Involvement in chitin synthesis</b>	<b>80</b>
<b>4.3. Digestion and beyond</b>	<b>81</b>
<b>4.4. Role in molt</b>	<b>84</b>
<b>5. Summary</b>	<b>89</b>
<b>6. References</b>	<b>91</b>
<b>7. Appendix</b>	<b>103</b>
7.1 Abbreviations	103
7.2. Bacterial strains and insect cell lines	105
7.3. Antibodies	105
7.4. Oligonucleotides	106
7.5. Plasmids	108
7.6. Kits	109
7.7. Curriculum vitae	110
7.8. Publications	111
7.9. Erklärung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistung	112
<b>8. Acknowledgements</b>	<b>113</b>



## 1. Introduction

### 1.1. Serine endopeptidases

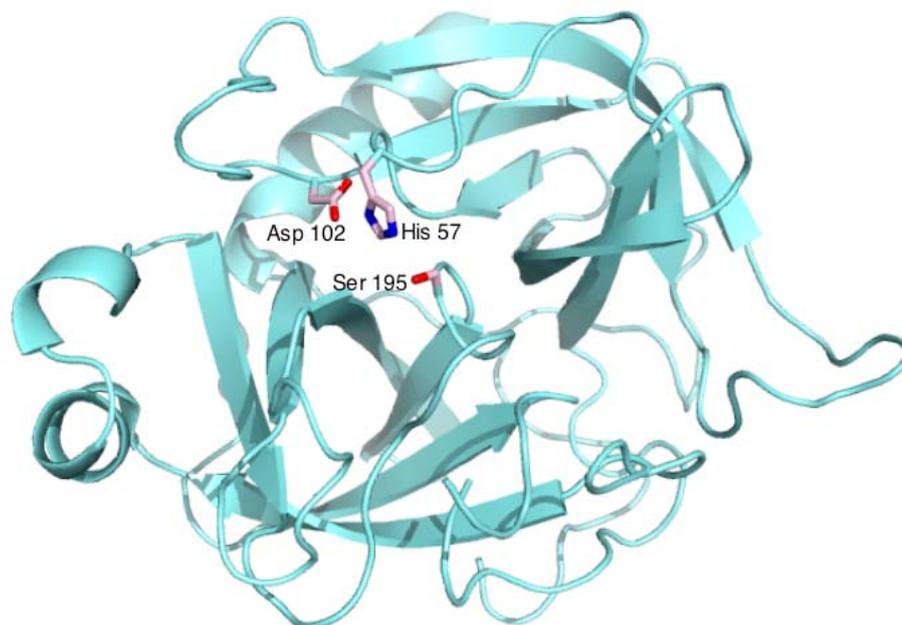
At present, over 140000 peptidases have been recorded in the 8.5 release of the MEROPS database (<http://merops.sanger.ac.uk>). This classification system introduced by Rawlings and Barrett (2010) divides these peptidases into 52 clans based on catalytic mechanisms and into 208 families based phylogeny criteria. Over one third of all peptidases are serine peptidases, categorized into 13 clans and 46 families. According to the MEROPS nomenclature, each clan is identified by two letters, the first refers to the catalytic type of the families included and the latter numbers the clan. Consequently, clans denoted with S- include serine peptidases whereas P- clans contain families with more than one catalytic type (Rawlings *et al.*, 2010). For instance, the PA clan contains serine as well as cysteine peptidases.

The name serine peptidase originates from the nucleophilic Ser residue at the active site, which is part of a catalytic triad of Asp, His and Ser residues. This triad, also referred to as the charge relay system, can be found in the distinct peptidase clans PA, SB, SC and SK (Hedstrom, 2002). In the serine peptidase clans SH and SN catalysis is carried out by different catalytic triads such as His-Ser-His and Ser-His-Glu, respectively. Dyad mechanisms, where the catalytic Ser is paired with Lys or His, can be found in the clans SE, SF and SM. Even serine peptidases with a single catalytic serine at the amino-terminus have been described for the subclan PB(S) (Hedstrom, 2002).

Though serine peptidases can be found throughout all taxa of living organisms as well as in viruses, yet the distribution of each clan differs from species to species. For example, clan PA serine peptidases are widespread throughout invertebrate and vertebrate genomes, whereas clan SB and SC serine peptidases are predominant in archaea, prokaryotes, plants and fungi (Di Cera, 2009; Hedstrom, 2002; Page and Di Cera, 2008). The majority of serine peptidases belongs to the S1 family of clan PA, which includes the two phylogenetically distinct but structurally related S1A and S1B subfamilies (Rawlings *et al.*, 2010). While S1B peptidases function in intracellular protein turnover, S1A peptidases act mostly in the extracellular space where they serve different functions (Di Cera, 2009). The subfamily S1A, which is typified by chymotrypsin (EC 3.4.21), is by far the largest serine peptidase family, by both the number of proteins that have been sequenced and the number of diverse peptidase activities (Rawlings *et al.*, 2010).

In mammals, functional diversity in the chymotrypsin subfamily is reflected by its involvement not only in digestion but also in apoptosis, signal transduction and reproduction. Cascades of sequential activation of S1A peptidases are further involved in immune responses, development, matrix remodeling, differentiation, and blood coagulation (Hedstrom, 2002; Page and Di Cera, 2008). Distinct physiological roles require S1A peptidases that differ greatly in specificity, varying from digestive peptidases that non-specifically cleave proteins to those who precisely detect a particular amino acid motif or even a single protein (Hedstrom, 2002).

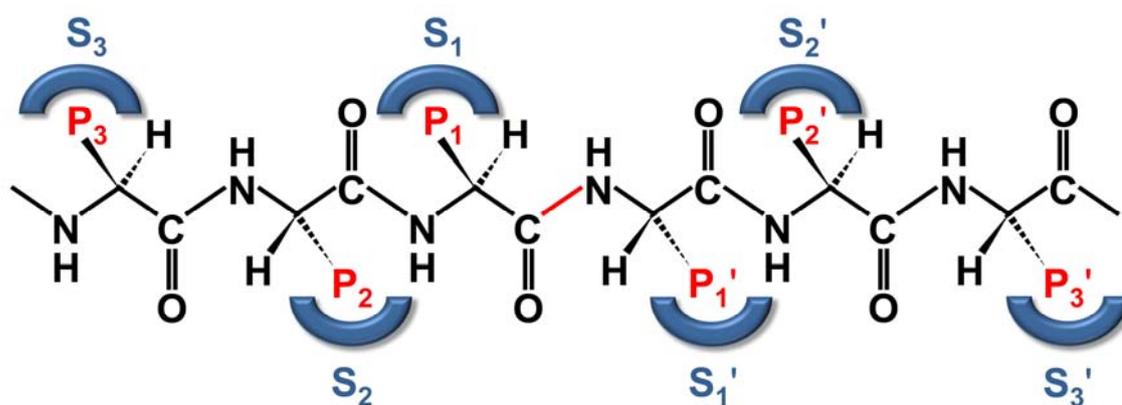
## 1.2. The structure of S1A chymotrypsin-like peptidases



**Fig.1.** The X-ray crystal structure of bovine chymotrypsin (Protein data bank code 4CHA), which is representative for the clan PA, family S1, subfamily A of chymotrypsin-like peptidases (Tsukada and Blow, 1985). The residues of the catalytic triad (sticks; His 57, Asp 102 and Ser 195) are hosted at the interface of two similar  $\beta$ -barrels.

The structures of S1A peptidases (Fig. 1) have been determined for various mammalian serine peptidases and are characterized by the prototypic chymotrypsin fold of two six-stranded  $\beta$ -barrels (Kraut, 1977; Page and Di Cera, 2008; Tsukada and Blow, 1985). Most notably, the catalytic residues are distributed evenly across the entire sequence. While His57 and Asp102 (numbering according to Greer (1990) used throughout) of the catalytic triad are located at the amino-terminal  $\beta$ -barrel, the nucleophilic Ser195 and the oxyanion hole (Gly193 and Ser195) are generated by the carboxy-terminal  $\beta$ -barrel. The residues determining topology of the amino-terminal and carboxy-terminal  $\beta$ -strands run in opposing directions with respect to their N to C termini so that the two  $\beta$ -barrels come together asymmetrically with the residues of the catalytic triad in the cleft between them. This cleft is surrounded by eight surface loops, as the end of each  $\beta$ -barrel donates three loops to the active site's entrance in addition to two loops which emerge from the back of the molecule. These loops are arranged on the two  $\beta$ -barrels in a symmetrical orientation whose axis rotates around the catalytic triad residues His57 and Ser195 (Di Cera, 2009; Page and Di Cera, 2008). The active site cleft is spanned by the residues of the catalytic triad, with His57 and Asp102 on side and Ser195 in the other side (Fig. 1). The catalytic triad is part of an extensive network of hydrogen bonds (Hedstrom, 2002; Rawlings and Barrett, 1994). Usually His57 forms hydrogen bonds with Asp102 and Ser195 although the latter bond is lost when His 57 is protonated. An additional hydrogen bond between His57 and Ser214 might prevent localization of a positive charge on His57 (Hedstrom, 2002). Ser214 is further involved in a hydrogen bond with Asp102 in almost all CTLPs and was thus considered as the

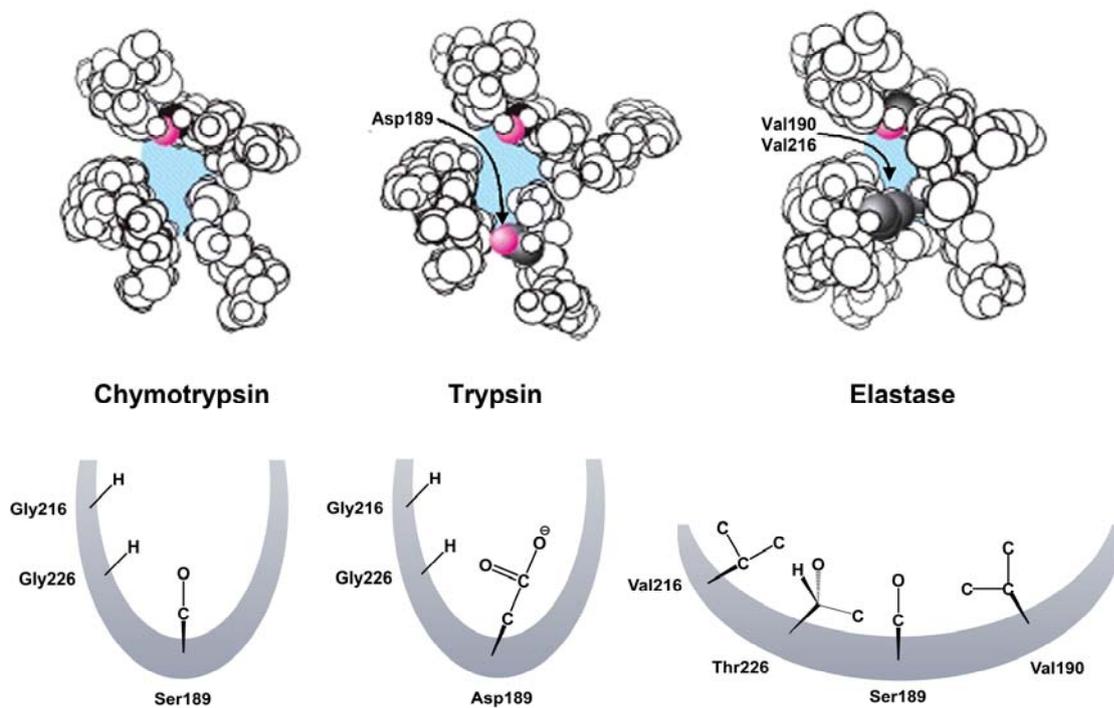
fourth member of the catalytic triad. However, studies showed that Ser214 is not essential for catalytic function (Epstein and Abeles, 1992; Hedstrom, 2002; McGrath *et al.*, 1992). Additional hydrogen bonds between Asp102 and Ala56 as well as His57 are thought to align Asp102 and His57 in the right orientation. The catalytic triad is structurally linked to the oxyanion hole, which also plays an important role in catalysis and is formed by the main chain NH-groups of Gly193 and Ser195. This pocket creates a positive charge, which activates the carbonyl of the scissile bond and helps in stabilizing the negatively charged oxyanion of the tetrahedral intermediate (Henderson, 1970).



**Fig. 2.** Specificity nomenclature for peptidase-substrate interactions. The potential sites of interaction of the substrate with the enzyme are designated P (shown in red), and corresponding binding sites on the enzyme are designated S (shown in blue). The scissile bond (also shown in red) is the reference point. Modified according to Berg *et al.* (2007).

Polypeptide binding is mediated by the backbone of the residues 214-216, which form an anti-parallel  $\beta$ -sheet with the main chain of the P1-P3 residues of the peptide substrate. In common to all CTLPs, binding occurs by the formation of hydrogen bonds between the carbonyl oxygen of Ser214 and the NH of the P1 residue, the NH of Trp215 and the carbonyl of P3 and the carbonyl of Gly216 and the NH of P3 (Hedstrom, 2002). The backbone conformation of Gly216 differs between chymotrypsin, trypsin and elastase (Perona *et al.*, 1995) so that the hydrogen bond to the P3 residue might vary in strength among these peptidases. Moreover, Ser214 also forms a hydrogen bond to His57, allowing communication between the substrate and the catalytic triad. Selectivity of substrate binding is mainly mediated by the S1 specificity pocket (Czapinska and Otlewski, 1999), which lies adjacent to the nucleophilic Ser195. The pocket is named in accordance to the nomenclature of Schechter and Berger (1967), which denotes the residues on the acyl and leaving group side of the scissile bond with Pn-P1, P1'-Pn', respectively, whereas Sn-S1, S1'-Sn' denotes the corresponding binding sites of the peptidase (Fig. 2). Three  $\beta$ -sheet regions (residues 189-192, 214-216, 225-228) form the walls of the specificity pocket. Since one of the walls also forms the polypeptide binding side, which is connected by the His57-Ser214 hydrogen bond to the catalytic triad, substrate binding and recognition as well as catalysis seem to be intrinsically tied to each other. The walls of the S1 pocket are connected by two surface loops (loop1: residues 185-188; loop2: residues 221-224) and stabilized by disulfide

bonds between Cys191 and Cys220 (Hedstrom, 2002; Perona and Craik, 1997; Perona *et al.*, 1995). The key determinants of substrate specificity are usually the residues at positions 189, 216, and 226 (Fig. 3; Czapinska and Otlewski, 1999; Perona and Craik, 1995). In the case of chymotrypsin a deep hydrophobic pocket is created by Gly216, Gly226 at the pocket's wall and Ser189 at the pocket's bottom, which leads to a specificity for large aromatic amino acids like phenylalanine, tryptophan or tyrosine. An analogously shaped pocket in which Asp189 introduces a negative charge at its base, accounts for trypsin's specificity for positively charged amino acids like arginine and lysine. Elastase prefers substrates with small aliphatic residues at P1 as its pocket exhibits only a shallow hydrophobic depression due to the presence of Val190, Val216 and Thr226 (Hedstrom, 2002; Perona and Craik, 1997; Perona *et al.*, 1995).



**Fig. 3.** Model of the S1 specificity pockets of chymotrypsin, trypsin and elastase. Certain residues play key roles in determining the specificity of these enzymes. The side chains of these residues, as well as those of the active-site serine residues, are shown in color. Modified according to Berg *et al.* (2007).

Next to the shape of the S1-pocket, substrate binding affinity is further determined by the surface loops 1 and 2 which do not contact the substrate (Botos *et al.*, 2000; Hedstrom *et al.*, 1992; Perona and Craik, 1997). These loops presumably do have an indirect effect on peptide binding, as they help to orientate Gly216 in the right conformation so that its carbonyl group points toward the P3 amino group in an optimal orientation for the formation of a hydrogen bond, which is essential for substrate binding. Furthermore the loops are also thought to play a role in adjusting the dynamic properties of the active site cleft (Hedstrom, 2002). While loop1 is thus far known to determine only the S1 site specificity, loop3 (residues 169-175) is thought act in concert with loop2 in modulating the S1- and sub-site preferences (Perona and Craik, 1997). In chymotrypsin, interaction between loop2 and loop3 is presumably mediated

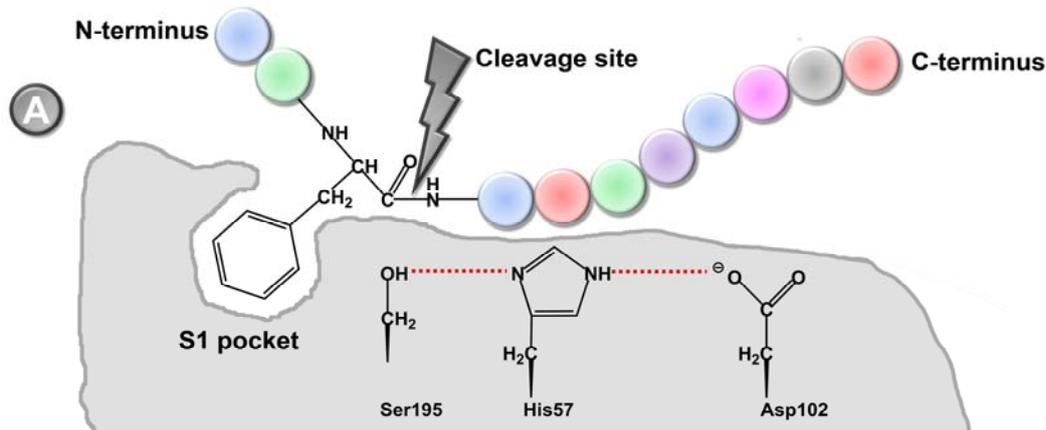
by Trp172 (loop3), as this residue stabilizes loop2 and hence influences the conformation of the S1 pocket and eventually substrate specificity (Czapinska and Otlewski, 1999). Experiments converting trypsin into chymotrypsin showed that Trp172 acts synergistically with the S1 site, loop1 and loop2, as a single Tyr172Trp mutation did not improve chymotrypsin activity of the hybrid protein (Hedstrom *et al.*, 1994).

However, the picture that peptidase specificity is determined by only a few amino acids is not sufficient to explain this phenomenon completely. Experiments to convert trypsin into chymotrypsin by replacing the S1-site residues and the two surface loops with their chymotrypsin counterparts as well as a Tyr172Trp substitution could only reconstitute 15% of chymotrypsin activity (Hedstrom *et al.*, 1994). Furthermore, the same strategy failed in the opposite case converting chymotrypsin into trypsin (Venekei *et al.*, 1996) or converting trypsin into elastase (Hung and Hedstrom, 1998). This implies that the structure-function link is much more complex and involves interactions with distal residues as well as the Sn-S2 and S1'-Sn' sites. To complicate matters further, the S1 pocket residues of chymotrypsin-like peptidases are generally less conserved in contrast to trypsin-like peptidases (Perona and Craik, 1995) and vary particularly among lepidopteran gut chymotrypsins at the subsites S2, S3 and S10 (Sato *et al.*, 2008). Additionally, the S2-S3 sites are thought play only a minor role in the substrate discrimination of chymotrypsin (Hedstrom, 2002; Schellenberger *et al.*, 1991). By contrast these sites can be also the major determinant of specificity as for example shown for elastase where P3/S3 interaction is predominant (Bode *et al.*, 1989; Stein *et al.*, 1987) and enteropeptidase, which precisely recognizes the P5-P1 sequence Asp-Asp-Asp-Lys (Lu *et al.*, 1999).

Most S1A peptidases are secreted into the extracellular space as inactive, zymogenic precursor proteins, which become activated by proteolytic cleavage at conserved sites (Bode *et al.*, 1978; Neurath and Dixon, 1957; Page and Di Cera, 2008). The zymogens of chymotrypsin-like peptidases exhibit low activity, as four regions, including the S1-site and the oxyanion hole, are deformed in the precursor peptidases. Proteolytic processing releases the amino-terminal Ile16, which successively forms a salt bridge with Asp194 in the GDSGGP motif. The resulting conformational change orders the deformed residues, which creates the active peptidase by forming the S1-site and the oxyanion hole. At high alkaline pH the enzyme returns in an inactive zymogen-like state, as the amino-terminal Ile16 is deprotonated, leading to disruption of the salt bridge with Asp194 (Hedstrom, 2002). In this manner, Asp194 acts as a gate for the S1 pocket, triggered by the ionization state of Ile16 (Bender and Kézdy, 1964). However, zymogen activation reflects a powerful regulatory mechanism that is frequently embedded in a cascade of successive proteolytic steps that together initiate physiological processes in due time at the right place.

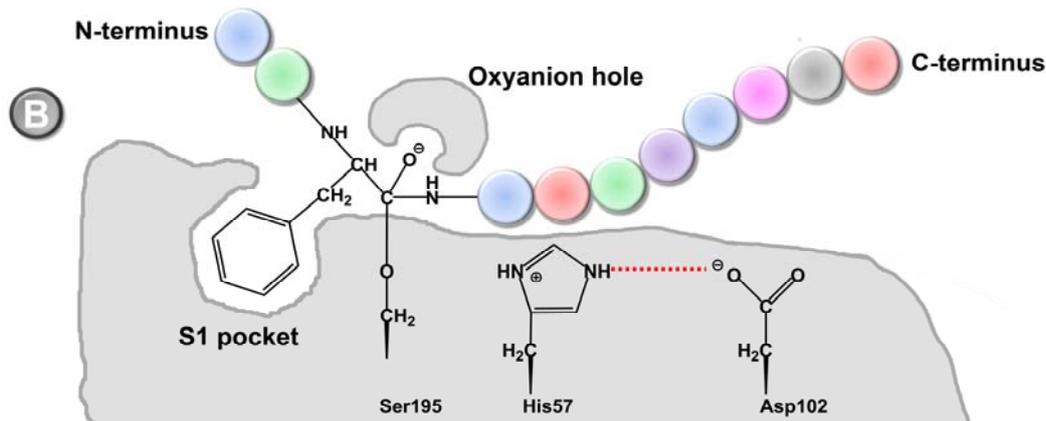
### 1.3. Catalytic mechanism of chymotrypsin-like peptidases

The catalysis of peptide hydrolysis usually involves the activation of the carbonyl residue of the substrate by a general acid, the activation of water with a general base and protonation of the amine leaving group (Hedstrom, 2002).



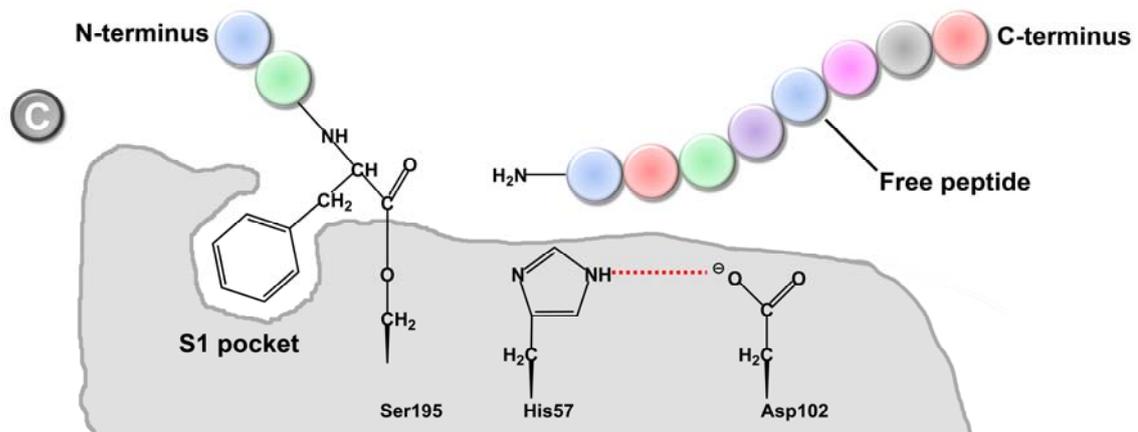
**Fig. 4A.** Chymotrypsin reaction mechanism: Polypeptide binding. Hydrogen bonds are indicated by dotted red lines. Modified according to Cooper and Hausman (2009).

After substrate binding (Fig. 4A) catalysis is initiated by the charge relay system of Ser195, His57 and Asp102. His57, acting as the general base, facilitates the abstraction of the proton from Ser195 and generates a potent nucleophile (Blow, 1997). The resulting protonated His57 is stabilized by a hydrogen bond to Asp102 (Fig. 4B).



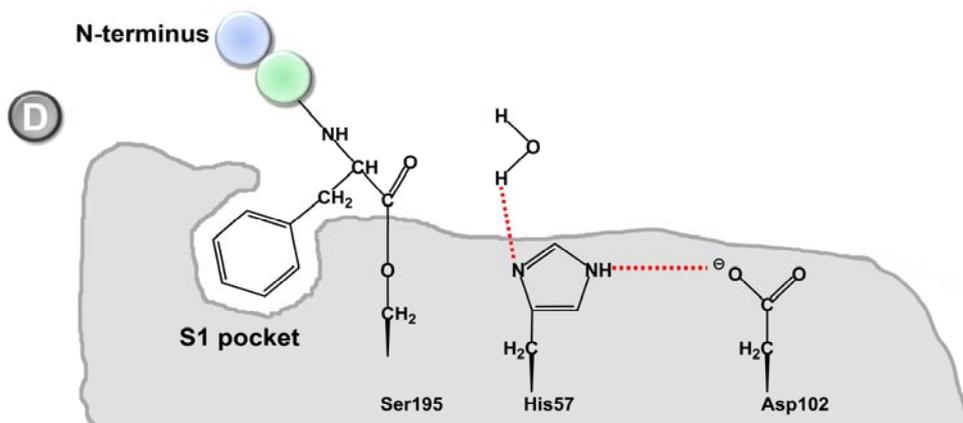
**Fig. 4B.** Chymotrypsin reaction mechanism: Creation of a nucleophile and formation of a tetrahedral intermediate. Hydrogen bonds are indicated by dotted red lines. Modified according to Cooper and Hausman (2009).

The hydroxyl oxygen of Ser195 attacks the carbonyl of the peptide substrate, leading to the formation of a tetrahedral intermediate. The oxyanion of the tetrahedral intermediate is stabilized by the positive charge created by the backbone NH-groups of Gly193 and Ser195 of the oxyanion hole (Fig. 4B).



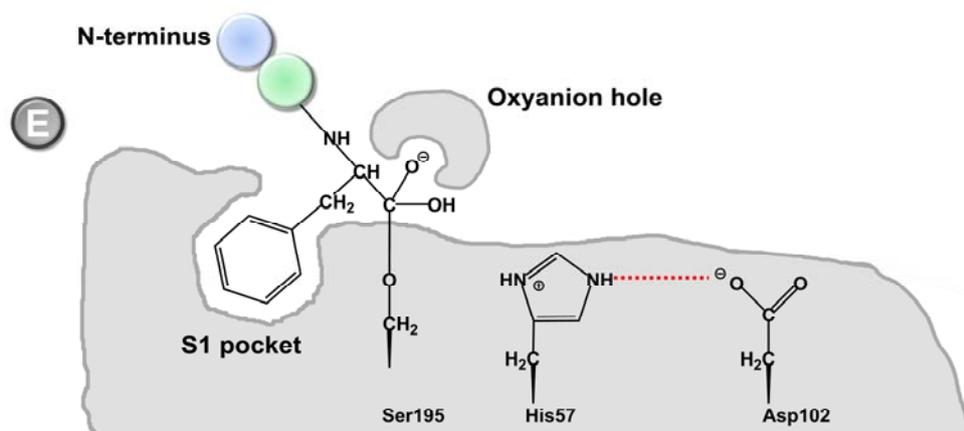
**Fig. 4C.** Chymotrypsin reaction mechanism: Formation of an acyl-enzyme intermediate and cleavage of the peptide bond. Hydrogen bonds are indicated by dotted red lines. Modified according to Cooper and Hausman (2009).

The breakdown of the tetrahedral intermediate leads to the cleavage of the peptide bond and to the formation of an acyl-enzyme intermediate (Hedstrom, 2002; Kraut, 1977; Page and Di Cera, 2008). His57, now acting as the general acid, stabilizes the leaving amine group by transferring a proton, which originates from the serine hydroxyl-group (Fig. 4C; Polgar, 2005).



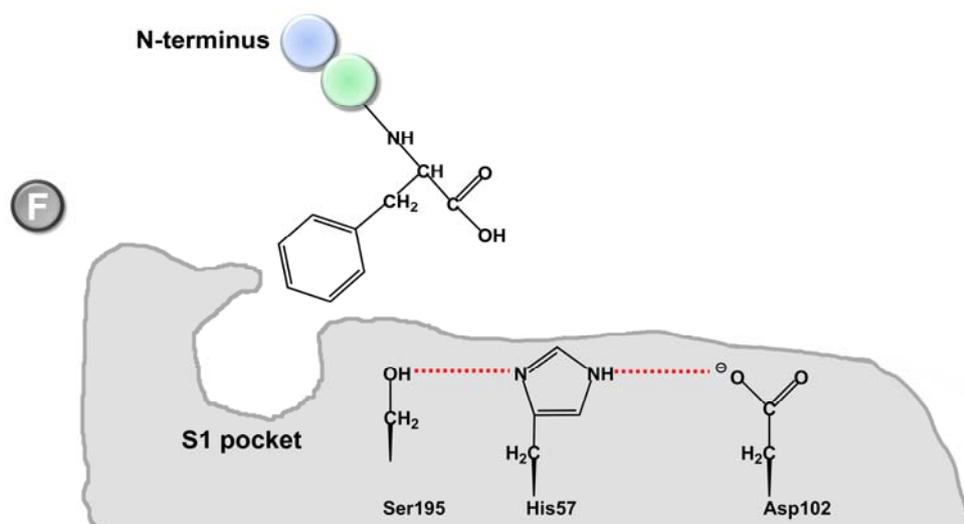
**Fig. 4D.** Chymotrypsin reaction mechanism: formation of a hydrogen bond with a water molecule. Hydrogen bonds are indicated by dotted red lines. Modified according to Cooper and Hausman (2009).

The deacylation half of the reaction begins by the formation of another nucleophile by the enzyme using His57 to withdraw a proton from a water molecule to form a hydroxide ion (Fig. 4D).



**Fig. 4E.** Chymotrypsin reaction mechanism: formation of a second tetrahedral intermediate. Hydrogen bonds are indicated by dotted red lines. Modified according to Cooper and Hausman (2009).

This nucleophilic hydroxide attacks the acyl carbon, leading to the formation of a second tetrahedral intermediate, which is again stabilized by the oxyanion hole (Fig. 4E).



**Fig. 4F.** Chymotrypsin reaction mechanism: release of the amino-terminal peptide fragment. Hydrogen bonds are indicated by dotted red lines. Modified according to Cooper and Hausman (2009).

The second tetrahedral intermediate collapses when His57 donates a proton to Ser195, displacing the acyl group and regenerating the hydroxyl group of Ser195. This releases the amino-terminal peptide fragment and concludes the reaction (Fig. 4F; Hedstrom, 2002; Kraut, 1977; Page and Di Cera, 2008).

## 1.4. Insect chymotrypsin-like peptidases

Although S1 peptidases are one of the best characterized enzyme families, their precise functions in insects are still poorly understood. Recent genome annotations have allowed the first comprehensive studies of this peptidase family. The *Drosophila* genome contains more than 200 genes encoding serine peptidases and homologous enzymes, of which 37 contain regulatory CLIP domains at the amino-terminus (Ross *et al.*, 2003). The vast majority of these genes encode S1 peptidases with a predicted trypsin-like specificity. In contrast, only 29 of these genes were predicted to encode S1A peptidases with a predicted chymotrypsin-like specificity. Also in other available insect genomes the total number of genes encoding CTLPs is rather low compared to the number of genes that encode TLPs. In *Apis mellifera*, 57 genes encoding serine peptidases and homologous enzymes were reported, of which only six had a predicted chymotrypsin-like specificity (Zou *et al.*, 2006). Similar results were obtained by screening the *Anopheles gambiae* genome for serine peptidases. From more than 300 genes encoding S1 peptidases, less than a tenth encode CTLPs. Thus, it appears that genes encoding serine peptidases with a chymotrypsin-like specificity constitute a smaller subset within the large gene family encoding insect S1A peptidases.

The determination of the atomic structure of a chymotrypsin from the fire ant *Solenopsis invicta* demonstrated that the overall S1A fold applies also to insect serine peptidases (Botos *et al.*, 2000). However, structural differences to vertebrate chymotrypsin-like peptidases can be found with respect to the activation of the enzymes. A structural property of vertebrate chymotrypsins is that they possess conserved cysteine residues that allow the formation of intramolecular disulfide bridges joining the chymotrypsin fragments resulting from additional proteolytic processing (Birktoft and Blow, 1972). In contrast to the vertebrate enzymes insect chymotrypsins contain only six conserved cysteine residues and appear not to be further processed after cleavage of the activation peptide (Davis *et al.*, 1985).

Serine peptidases of insects serve numerous physiological functions that are essential for growth and development. In addition to their well-established function in digestion, they function in molting and metamorphosis, and act as highly specific regulators of innate immune responses, including melanization and the expression of antimicrobial peptides (Kanost and Gorman, 2008; Law *et al.*, 1977; Neurath, 1984; Srinivasan *et al.*, 2006). The following subchapters will give an outline on the different functions CTLPs accomplish in insects.

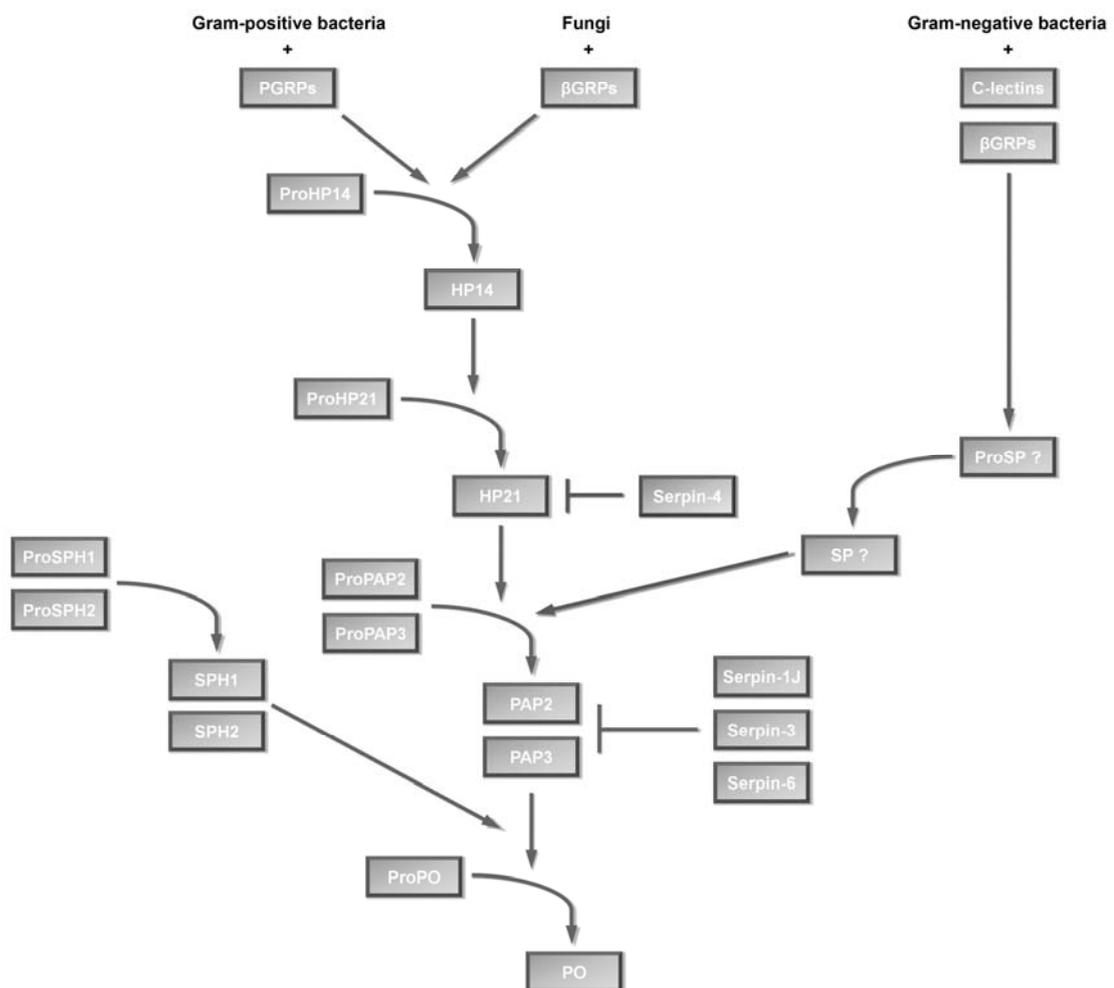
### 1.4.1. Chymotrypsin-like peptidases in insect immunity

As part of the innate immune response, phenoloxidases (PO), which are zymogenic enzymes in the insect's hemolymph, become activated upon wounding and/or infection (Kanost and Gorman, 2008). These enzymes catalyze the production of quinones, which can nonspecifically cross-link neighboring molecules to form melanin at wound sites or all over the surface of invading microorganisms. Additionally cytotoxic reactive nitrogen and oxygen intermediates produced during melanin

synthesis appear to help kill invading pathogens (Christensen *et al.*, 2005). Phenoloxidases are expressed as an inactive precursor, the prophenoloxidases (proPO), predominantly expressed by hemocytes (Cerenius and Söderhäll, 2004) and are presumably released into plasma by cell lysis. ProPOs are activated at a specific site near their amino-terminus by so called proPO-activating factors (PPAFs) or enzymes (PPAEs), or alternatively proPO-activating proteins (PAPs) (Jiang *et al.*, 1998; Jiang *et al.*, 2003a; Jiang *et al.*, 2003b; Lee *et al.*, 1998; Satoh *et al.*, 1999; Wang *et al.*, 2001). These enzymes contain a carboxyl-terminal peptidase domain with trypsin specificity and one or two amino-terminal clip domains. Clip domains possibly mediate interaction between members of peptidase cascade pathways (Jiang and Kanost, 2000). The proPO-activating clip domain serine peptidases are themselves synthesized as zymogens and are activated by other proteases as a part of a serine peptidase cascade. Many of the proPO-activating serine peptidases require certain cofactors, which turned out to be proteins with a clip domain and a serine peptidase domain in which the active site serine was replaced by a glycine (Jiang *et al.*, 1998; Kwon *et al.*, 2000; Yu *et al.*, 2003). These so called serine peptidase homologues (SPHs) lack proteolytic activity and need to be activated by so far uncharacterized hemolymph serine peptidases possibly including CTLPs (Kim *et al.*, 2002; Lee *et al.*, 2002; Yu *et al.*, 2003). SPHs from *M. sexta* have been described to bind to lectins, which recognize bacterial lipopolysaccharides, thus functioning as a mediator to recruit proPOs and proPO-activating serine peptidases to the site of infection (Yu *et al.*, 2003). This assumption is supported by the finding of Piao *et al.* (2005), who describe a cleft in the clip domain of a SPH of *Holotrichia diomphalia* (PPAFII) possibly responsible for the binding to the phenoloxidase. Furthermore PPAFII tends form oligomers, which consist of two stacked hexameric rings (Piao *et al.*, 2005). These large oligomers, also observed in *M. sexta* (Wang and Jiang, 2004), possibly bring components of the peptidase cascades in close proximity to each other on the surface of invading microorganisms. Initiation of the proPO activation cascade is thought to occur with the help of pattern recognition proteins (C-type lectins; peptidoglycan recognition proteins, PGRP; glucan recognitions proteins), which interact with self-activating zymogenic peptidases (Kanost and Gorman, 2008; scheme see Fig. 5).

Although the innate immune response leading to the activation of proPOs seems to be primarily controlled by clip domain serine peptidases whose peptidase domain is of the trypsin type, involvement of CTLPs in these regulatory cascades cannot be fully excluded. In this context direct activation of proPO with chymotrypsin (Onishi *et al.*, 1970) or partially purified peptidases from insect tissues (Asada and Sezaki, 1999; Aso *et al.*, 1985; Hall *et al.*, 1995; Pau and Kelly, 1975), possibly also including CTLPs, has been shown. Furthermore systemic melanization of *M. sexta* can be elicited by injecting small amounts of metalloproteases (micrograms) or more than 1 mg of chymotrypsin (Kanost and Gorman, 2008). The reason that higher quantity of chymotrypsin than metalloproteases is needed to trigger proPO activation might be explained by the high concentration of serine peptidase inhibitors including members of the serpin family in the insect's hemolymph (Polanowski and Wilusz, 1996). These serine peptidase inhibitors are essential for the tight regulation of melanization to the wound sites or the

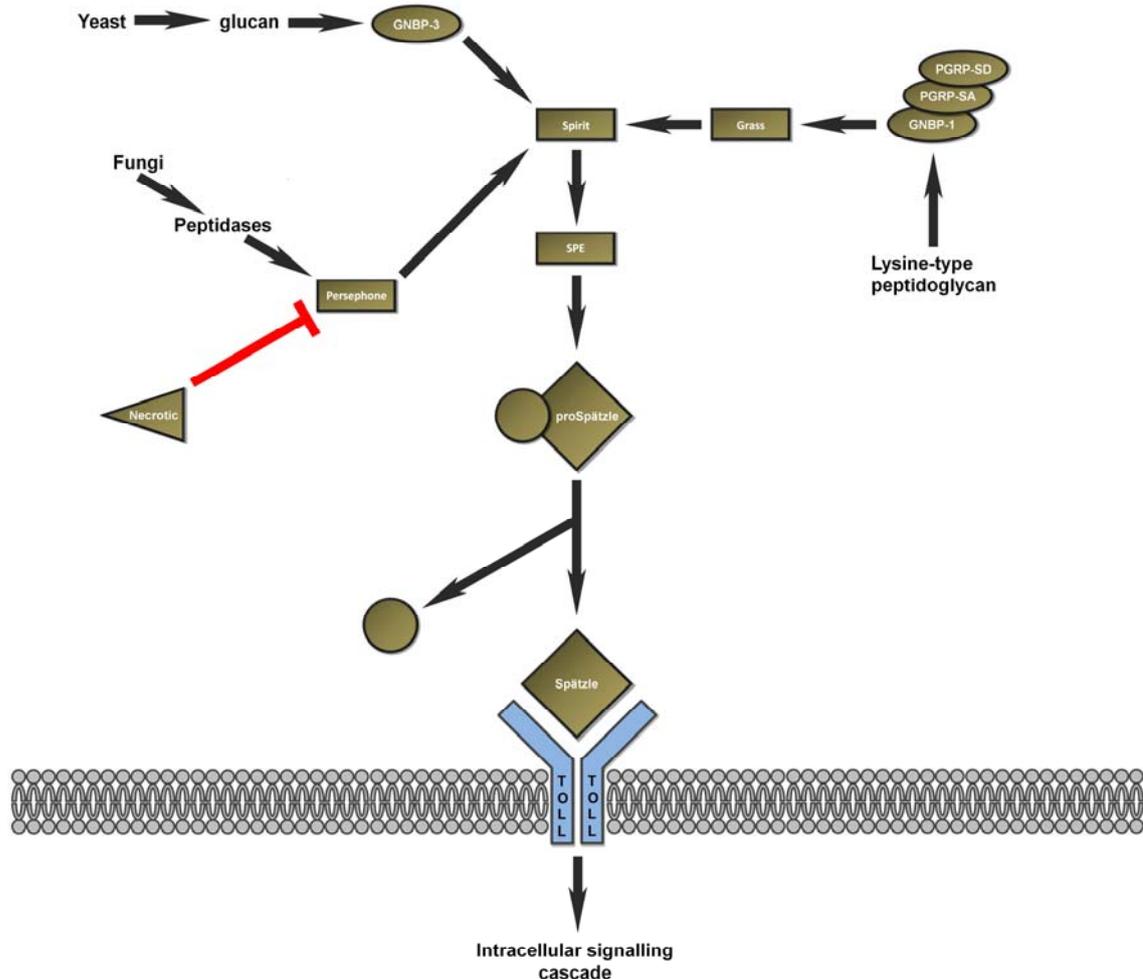
surface of invading microorganisms (see Fig. 5), since products of phenoloxidases activity are potentially toxic to the host. However, serine peptidase inhibitors may provide additional hints of the involvement of CTLPs in peptidase cascades in insect immunity. In this regard a proPO-activating enzyme of *Bombyx mori* has been described to be inhibited by an endogenous chymotrypsin inhibitor 13 (CI-13; Aso *et al.*, 1994). Chymotrypsin inhibitor b1 (CI-b1) a protein interacting with lipopolysaccharides of Gram negative bacteria, is thought to play a regulatory role in the melanin formation (He *et al.*, 2004). Furthermore, infection of *B. mori* with nucleopolyhedrovirus (BmNPV) increased the activity of the CI-b1 promoter (Zhao *et al.*, 2007), providing additional evidence to the involvement of CTLPs in the innate immunity of insects.



**Fig. 5.** A model for a proPO activating pathway in *M. sexta*. Pattern recognition proteins (PGRP,  $\beta$ GRP, C-type lectins) bind to polysaccharides on the surface of microorganisms triggering a peptidase cascade by the activation of initiator peptidases (HP14). HP14 activates HP21 which cleaves zymogenic proPAPs. Together with the SPHs, which are activated by yet unknown peptidases, PAPs lead to the activation of proPOs. POs catalyze catechols to corresponding quinones resulting in the formation of melanin. Peptidases in the cascade are regulated by serine peptidase inhibitors (serpins). Abbreviations:  $\beta$ GRP,  $\beta$ -1,3-glucan recognition protein; HP hemolymph peptidase; PAP prophenoloxidases activating peptidase; PGRP, peptidoglycan recognition protein; PO, phenoloxidase; SPH, serine peptidase homolog; SP ?, unknown serine peptidase. Modified according to Kanost and Gorman (2008).

The expression of antimicrobial peptides and the initiation of innate immune responses has been extensively studied in *Drosophila* and is regulated by two additional recognition and signaling cascades. Firstly the Toll pathway, which moreover establishes the embryonic dorsoventral axis (Morisato and Anderson, 1995) and secondly the immune deficiency (Imd) pathway. Whereas the Imd pathway mainly responds to Gram-negative bacteria, the Toll pathway is activated upon infections with Gram-positive bacteria and fungi (Aggarwal and Silverman, 2008). Stimulation of the Toll pathway leads to activation of two NF- $\kappa$ B-like factors, Dorsal and Dif (Dorsal-related immunity factor), regulating a large set of genes. This includes antimicrobial peptide genes as well as components of the melanization and clotting cascades expressed by the fat body and in hemocytes (De Gregorio *et al.*, 2002). Activation of the toll pathway occurs after the cleavage of the cytokine Spätzle by serine peptidase cascades.

The proteolytic cascade activating Spätzle during dorsoventral patterning has been well characterized and involves the serine peptidase Gastrulation defective as well as the trypsin type clip domain serine peptidases Snake and Easter. The latter directly cleaves the zymogenic Spätzle into its active form (Dissing *et al.*, 2001; LeMosy *et al.*, 2001). This peptidase cascade is negatively regulated by the serpin SPN27A which acts at the level of Easter (Ligoxygakis *et al.*, 2003).



**Fig. 6.** Extracellular components of the Toll pathway involved in insect immunity. Arrows indicate the Toll signaling cascade, including the three microbial recognition systems. The red arrow highlights the negative regulator Necrotic and its likely target. Serine peptidases involved in the signaling cascade are displayed in rectangular boxes. Modified according to Aggarwal and Silverman (2008).

The cascades activating Spätzle during the immune response are much more complex with branches also involving CTLPs (Fig. 6). Recognition of Gram-positive bacteria is mediated through the detection of lysine-containing peptidoglycan by various pattern recognition proteins (Bischoff *et al.*, 2004; Gobert *et al.*, 2003; Michel *et al.*, 2001; Pili-Floury *et al.*, 2004). Fungal  $\beta$ -glucans are thought to be recognized by the receptor GNBP-3 (Gottar *et al.*, 2006), whereas another pathway involves a serine peptidase known as Persephone which probably senses proteolytic activities elicited by both fungal and Gram-positive bacterial infections (El Chamy *et al.*, 2008). Pathogen recognition leads to Spätzle cleavage by the activation of serine peptidase cascades that converge on the two CTLPs Spirit and the Spätzle-processing enzyme (SPE), which both contain amino-terminal clip domains (Kambris *et al.*, 2006). Grass, another clip domain CTLP has been proposed to act upstream of Spirit upon gram-positive bacteria infection (Kambris *et al.*, 2006) and recognition of  $\beta$ -glucans (El Chamy *et al.*, 2008). Involvement of CTLPs upon fungal infections has also been shown for the larvae of *Tenebrio molitor*. The *Tenebrio* modular serine peptidase (MSP) contains a carboxy-

terminal chymotrypsin-like catalytic serine peptidase domain and triggers the Spätzle-processing enzyme-activating enzyme (SAE). SAE in turn activates SPE, leading to the cleavage of Spätzle (Roh *et al.*, 2009).

Together these findings indicate that CTLPs indeed play a role in the innate immunity of insects, though their precise involvement in cascades leading to melanization or the expression of antimicrobial peptides needs to be further characterized.

#### 1.4.2. Role of chymotrypsin-like peptidases in digestion

Digestive enzymes of phytophagous insects have been the subject of many studies due to their importance as potential targets for insect control. Digestion in insects is primarily carried out by serine peptidases belonging to the PA clan, cysteine peptidases of the C- clan as well as by metallopeptidases of the M- clan (Rawlings *et al.*, 2010; Srinivasan *et al.*, 2006). This chapter will be about the largest serine peptidase family S1A, with focus on midgut chymotrypsin-like peptidases of lepidopteran and coleopteran species, represented in this work by *Manduca sexta* and *Tribolium castaneum*.

Herbivorous lepidopteran larvae are rapidly growing organisms (Reynolds *et al.*, 1985), which take up large amounts of plant material in order to maintain amino acid homeostasis as plants are comparatively low in protein. Proteins are basically digested by serine peptidases which contribute to about 95% of the total digestive activity of the lepidopteran midgut. For lepidopteran several midgut serine peptidases have been described in closer detail. For example, the midgut of the polyphagous lepidopteran larvae of *Helicoverpa amigera* contains about twenty different isoforms of active serine peptidases (Srinivasan *et al.*, 2006). This multitude of digestive enzymes has been suggested to be advantageous for a polyphagous insect in order to adapt to the protein composition of different host-plants and to the exposure to peptidase inhibitors (Brito *et al.*, 2001; Jongasma *et al.*, 1995; Lopes *et al.*, 2004; Paulillo *et al.*, 2000; Srinivasan *et al.*, 2006). In this context, polyphagous lepidopteran larvae are known to quantitatively and qualitatively alter the composition of their gut peptidases when digestion becomes inefficient and in response to changes in the protein content of their diet (Brito *et al.*, 2001; Patankar *et al.*, 2001; Paulillo *et al.*, 2000). These modifications include a specific or general up-regulation of gut peptidases in order to maintain optimal protein utilization (Broadway, 1997; Gatehouse *et al.*, 1997). This further involves the synthesis of peptidases that are insensitive to certain peptidase inhibitors (Bown *et al.*, 1997; Jongasma *et al.*, 1995; Mazumdar-Leighton and Broadway, 2001a; Mazumdar-Leighton and Broadway, 2001b; Volpicella *et al.*, 2003), or of those peptidases which are able to digest specific peptidase inhibitors (Ishimoto and Chrispeels, 1996; Telang *et al.*, 2005).

For monophagous insects a lower variety of midgut peptidases has been suggested since feeding on a single host plant supplies the animals with a sufficient amount of their preferential diet. Thus animals like *M. sexta*, which is monophagous at its later larval stages, might manage well on a basic set of peptidases without any stress on its

digestive system arising due to host plant peptidase inhibitors. While this means a reduced flexibility in the choice of the host plant, the advantage is that peptidase synthesis becomes less costly in comparison to polyphagous larvae, possibly giving monophagous larvae a benefit in fitness (Srinivasan *et al.*, 2006). For *M. sexta* so far five cDNAs encoding one CTLP (MsCT, GenBank|AAA58743.1; Peterson *et al.*, 1995), three closely related trypsins (MsTRP-A/B/C, GenBank|AAA29339.1/AAA29340.1/AAA29341.1; Peterson *et al.*, 1994) and one trypsin-like peptidase (MsTLP1, emb|CAM84320.1; Broehan *et al.*, 2008) have been described. The peptidases described by the Peterson group are expressed in the anterior and median midgut in the case of chymotrypsin (Peterson *et al.*, 1995) and in the median and posterior midgut in the case of trypsin (Peterson *et al.*, 1994).

A common property of all lepidopteran midgut serine peptidases is their resistance to the extreme alkaline milieu of pH 10–12 present in the lepidopteran midgut (Dow, 1984). The role of high luminal pH has been discussed to be due to a limited protection against tannins, whose complexes are easier dissociated at strong alkaline conditions. A high midgut pH might further be beneficial in the protection against pathogens. (Dow, 1992). As an adaptation to the alkaline midgut conditions, most lepidopteran CTLPs exhibit pH optima greater than nine. How the enzyme's functionality is maintained at such a high alkaline pH still remains elusive (see 4.1.2). Possibly, the larger number of arginine residues (pKa = 12.5) found in lepidopteran chymotrypsins, stabilizes the enzyme at an alkaline pH, as only this amino acid remains positively charged at conditions prevailing in the lepidopteran midgut (Peterson *et al.*, 1995).

Whereas serine peptidases account for the major digestive activity in *M. sexta* (80% chymotrypsin, 10% trypsin and 1% elastase) and other lepidopteran species (Srinivasan *et al.*, 2006), *Tenebrionidae* seem to follow heterogeneous strategies. Similar to lepidopterans, serine peptidases are the predominate digestive enzymes of 9<sup>th</sup> instar larvae of the black flour beetle *Cyaneus angustus* (Oppert *et al.*, 2006). In the larvae of *T. molitor*, however, cysteine peptidases account for 64% of the total digestive activity in the anterior midgut (serine peptidases about 36%), whereas major proteolysis in the posterior midgut was due to serine peptidases (serine peptidases 80%; cysteine peptidases 20%; Vinokurov *et al.*, 2006). In *T. castaneum* approximately 80% of total digestive activity was found in the anterior midgut and 20% in the posterior midgut. Activity was shown to be basically due to cysteine peptidases but nevertheless serine peptidases contribute to about 10% digestive activity in the anterior midgut and 31% in the posterior midgut (Vinokurov *et al.*, 2009).

The midgut of *T. castaneum* exhibits slightly acidic conditions in the anterior, rising to neutral in the median and further increasing to weakly alkaline in the posterior third of the midgut. Chymotrypsin activity was determined to be maximal at pH 8.0 – 10.5, approving the alkaline pH-optima of *Tribolium* chymotrypsins. For this reason these enzymes probably participate mainly in the final stages of digestion as their activity was largely decreased by the acidic pH of the anterior midgut. However, chymotrypsin fractions were discovered in the anterior as well as in the posterior midgut of *T. castaneum* (Vinokurov *et al.*, 2009). Localization of several serine peptidases in the anterior midgut was also observed for *T. molitor* (Vinokurov *et al.*, 2006), probably

due to the secretion of the enzymes in this tissue and successive transport to the posterior midgut with the food bolus, where their activity considerably increased. In the more alkaline conditions of the posterior midgut the serine peptidases are thought to act on partially digested products initially cleaved by cysteine peptidases in the anterior midgut (Vinokurov *et al.*, 2009).

### 1.4.3. Involvement of chymotrypsin-like peptidases in molt

In der Regel ist der Leib der Larven noch wenig entwickelt; sie fressen jedoch meistens sehr viel und wachsen sehr schnell. Da die Haut nicht mitwächst, so wird sie für den Leib nach und nach zu enge; das Insekt fühlt sich in diesem Zustande unbehaglich, stellt das Fressen auf kurze Zeit ein, sucht sich dann gewöhnlich ein passendes und bequemes Plätzchen, wo es die Haut abstreift und dann wieder im Fressen das nachzuholen sucht, was vorher darin versäumt worden ist. Solche Häutungen kommen mehrere Male, in der Regel aber vier Mal, im Larvenzustande vor.

Nach Herold's trefflichen Beobachtungen ist bei der jungen Larve gar keine Spur von der neuen Haut, sondern diese entsteht erst am Ende des ersten Abschnittes im Raupenleben, einige Tage vor der Abstreifung der alten. Man bemerkt zu dieser Zeit, wie sich die Schleim- und Muskellage der Haut ringsum von der Epidermis (Oberhaut) ablöst, und erst sich alsdann auf der Oberfläche mit einer neuen Epidermis bekleidet. Die Entwicklung dieser neuen Oberhaut dauert zwei bis drei Tage, während welcher die Raupe zu kränkeln scheint und wenig, ja fast gar keine Nahrung zu sich nimmt. Endlich spaltet sich die alte Haut der Länge nach auf dem Rücken, und die Raupe befreit sich von ihrer nunmehr abgetrennten Haut durch schlängelnde, rasche Bewegung, erst das Kopfende, dann das Leibende daraus hervorziehend.

**Fig. 7.** Description of insect molt in an old German textbook by Dr. M. Bach (1870).

The article of Dr. M. Bach (Fig. 7) from the 19<sup>th</sup> century nicely describes the behavioral and physiological changes that accompany insects molt without knowing the biochemical processes these effects underlie. Today we know that molting is initiated by the steroid hormone 20-hydroxy-ecdysone (Truman, 2005; Zitnan *et al.*, 1999) and coordinated by peptide hormones, which regulate the physiological responses and the sequential behaviors in ecdysis (Ewer, 2005; Truman, 2005; Zitnan *et al.*, 2007). During this process, the insects detach their old cuticle from the epidermis, creating an exuvial space. Molting fluid enzymes are secreted in the exuvial space between the old and the new cuticle to degrade the inner layers of the old cuticle as part of the molting process

before ecdysis. Following apolysis, the epidermal cells synthesize a new cuticle and the remains of the old cuticle are shed (Law *et al.*, 1977).

The molting fluid secreted into the exuvial space by the epidermis exhibits both chitinolytic and proteolytic activities (Samuels and Reynolds, 1993). Degradation of the old cuticle is probably a sequential process in which the peptidases initially cleave the proteinous proportion of the endocuticle to make chitin accessible for the digestion by chitinases (Fukamizo and Kramer, 1985). However, molting fluid peptidases may also activate pro-chitinases already present in the endocuticle (Law *et al.*, 1977).

Several studies provided evidence for the existence of proteolytic enzymes in insect molting fluids. The molting fluid of the fruit fly *Ceratitis capitata* exhibits two cysteine peptidases (MCAP-I/-II), which can be detected at the transition from pre-pupa to pupa where they are thought to play a role in metamorphosis (Rabossi *et al.*, 2008). A cysteine peptidase (cathepsin-L) was also detected in the larval cuticle and the epidermis layer of *H. armigera*. Cathepsin-L gene expression was correlated closely with ecdysone and was highest during larval molt. This enzyme probably plays an important role in molt, as injection of cysteine peptidase inhibitors delayed molting from 5<sup>th</sup> to 6<sup>th</sup> instar (Liu *et al.*, 2006). Apolysis in *H. armigera* is further performed by a metallo-peptidase (carboxypeptidase-A) present in the integument from 5<sup>th</sup> instar larvae. Expression of carboxypeptidase-A was highest when molting from 5<sup>th</sup> to 6<sup>th</sup> instar and its expression could be up-regulated by 20-hydroxy-ecdysone. Furthermore the peptidase could be detected in the molting fluid indicating that carboxypeptidase-A participates in larval molting and metamorphosis (Sui *et al.*, 2009a). Expression of three metallo-peptidases (carboxypeptidase-A, BmADAMTS-1; BmADAMTS-like) was found to be induced during pupal ecdysis in the wing discs of *Bombyx mori* (Ote *et al.*, 2005a; Ote *et al.*, 2005b). Carboxypeptidase-A was expressed during molt in epithelial tissues and secreted as a zymogen which was processed in the molting fluid (Ote *et al.*, 2005a). The BmADAMTS-1 gene expression was detected in the midgut and hemocytes of the wandering stage and was highly induced during the pupal ecdysis in the hemolymph. Expression of the BmADAMTS-like gene was found in the epithelial tissues of the wandering stage and reached its maximum slightly later than the BmADAMTS-1 gene (Ote *et al.*, 2005b). Whereas carboxypeptidase-A probably degrades the proteins from the old cuticle, the two ADAMTS-type metallo-peptidases of *B. mori* are thought play a role in degenerating and remodeling tissues during the molting periods (Ote *et al.*, 2005a; Ote *et al.*, 2005b).

Next to cysteine and metallo-peptidases, also serine peptidases were shown to play an important role as molting fluid enzymes. For example, the molting fluid of pharate adult *Antheraea polyphemus* hydrolyzes certain trypsin specific substrates (Katzenellenbogen and Kafatos, 1970). From the molting fluid of this silkworm, Katzenellenbogen and Kafatos (1971b) partially purified two serine peptidases with trypsin specificity. The enzymes were shown to be secreted into the exuvial space as inactive precursors (Katzenellenbogen and Kafatos, 1971a), probably activated by other molting fluid enzymes (Reynolds and Samuels, 1996). RNAi of a trypsin-like serine peptidase in *Locusta migratoria* led to reductions in peptidase and cuticle degrading activity. As animals injected with dsRNA for this TLP showed severe defects when

molting from 4<sup>th</sup> to 5<sup>th</sup> instar and in adult molt, this peptidase probably plays an important role in ecdysis of *L. migratoria* (Wei *et al.*, 2007). Another trypsin-like serine peptidase (Ha-TLP2) was detected in the integument of *H. amigera* (Liu *et al.*, 2009). The peptidase is most similar to the trypsin-like serine peptidase Ha-TLP which is expressed in the midgut of *H. amigera* (Sui *et al.*, 2009b). Ha-TLP2 could be down-regulated by 20-hydroxy-ecdysone and up-regulated by the juvenile hormone analog methoprene. The finding that Ha-TLP2 is constitutively expressed with activity in the integument during larval feeding, molting, and metamorphosis suggests an involvement in the remodeling of the integument (Liu *et al.*, 2009). Degradation of cuticular proteins in *M. sexta* involves a complex mixture of exo- and endopeptidases. By the usage of specific substrates and inhibitors, Brookhart and Kramer (1990) showed the presence of an aminopeptidase, at least three carboxypeptidases as well as trypsin- and chymotrypsin-like enzymes in the molting fluid of *M. sexta*. However, Samuels *et al.* (1993b) demonstrated that the major cuticle-degrading enzyme found in the molting fluid of pharate adult *M. sexta* is a trypsin like peptidase (MFP-1). MFP-1 was shown to degrade *Manduca* cuticle protein *in vitro*. MFP-2, a metallo-peptidase additionally detected in the molting fluid of *M. sexta* was not able to degrade pupal cuticle. However, MFP-2 acting together with MFP-1 was found to synergistically enhance degradation (Samuels *et al.*, 1993a).

The above mentioned findings demonstrate that serine peptidases indeed play an important role in insect molt. However, so far no CTLP has been identified to be involved in molting, although chymotrypsin-like activities have been reported in molting fluids from lepidopteran species (Brookhart and Kramer, 1990). The occurrence of trypsin and chymotrypsin inhibitors in the molting fluid of *M. sexta* (Reynolds and Samuels, 1996) further emphasizes the possibility of CTLP involvement in molting. These inhibitors probably regulate molting processes as described for the hemolymph peptidases involved in the activation of prophenoloxidasases (see 1.4.1). In this context Samuels and Richards (2000) detected an inhibitor of MFP-1 in the molting fluid of all preecdysial stages except in the stage of highest MFP-1 activity.

## 1.5. Objective of the work

This work will deal with chymotrypsin-like peptidases of insects with focus on those CTLPs, which have diversified in function. To begin with, this question shall be elaborated with the help of *M. sexta*. Easy dissectible midguts with high tissue yields, makes *Manduca* the ideal research object to investigate digestive peptidases. The circumstance, however, that its genome is not fully accessible, adversely affects a holistic approach in the characterization of the *Manduca* midgut CTLPs. For this reason further characterization of CTLPs shall be conducted in *T. castaneum*, which genome is fully sequenced and well aligned.

Previous results obtained by Lars Maue demonstrated, that activation of the midgut specific chitin synthase 2 of *M. sexta* (MsCHS2) is trypsin mediated but additionally requires a soluble factor from midgut extracts (Zimoch *et al.*, 2005). This soluble factor was identified to be a chymotrypsin-like peptidase (MsCTLP1) which

showed an interaction with the carboxy-terminal region of MsCHS2 in a Yeast-two-hybrid screening performed by Anton Wessels and Beyhan Ertas (Broehan *et al.*, 2007).

Biochemical proof that MsCTLP1 directly interacts with MsCHS2 shall be provided by recombinant expression or purification of the peptidase from the midgut. Furthermore co-immunoprecipitation experiments shall confirm the data obtained by Yeast-two-hybrid screening. Due to the specific role of MsCTLP1, this work for the first time describes a midgut specific serine peptidase which has diversified in its function next to digestion. For this reason additional screenings shall identify further peptidases which also might exhibit functional diversity. To get first hints on the functions of the newly discovered peptidases, their expression shall be studied in different tissues at various physiological states.

To gain more insights into the functions of CTLPs, the gene family encoding these peptidases in the red flour beetle, *T. castaneum*, a well-established genetic and genomic insect model, shall be characterized. Following the identification of genes that encode for CTLPs in the *Tribolium* genome, the expression patterns of the corresponding enzymes shall be characterized with respect to differences in tissue distribution and differences in development. Finally, systemic RNAi experiments of the identified peptidases shall provide first evidence of the role of CTLPs in *Tribolium*.

## 2. Material and Methods

### 2.1. Material

#### 2.1.1. Culture Media

- LB-media: 1.0% (w/v) Bactotryptone, 1.0% (w/v) NaCl and 0.5% (w/v) yeast extract
- LB-Agar: LB-media + 1.5% (w/v) Agar
- ΨB-media: 2.0 % (w/v) Peptone, 0.5% (w/v) MgSO<sub>4</sub>, 0.5% (w/v) yeast extract, pH7.6 with KOH
- Insect-XPRESS Protein-free Insect Cell Medium with L-glutamine (Lonza)

#### 2.1.2. Insects

Larvae of *M. sexta* Linné 1763 (Lepidoptera, Sphingidae) were reared under long-day conditions (16 h of light) at 27°C using a synthetic diet modified as described previously (Bell and Joachim, 1974). *T. castaneum* used in this work were from the GA-1 strain (Haliscak and Beeman, 1983). Insects were reared at 30°C under standard conditions (Beeman and Stuart, 1990).

## 2.2. Molecular biological methods

### 2.2.1. Tissue preparations for total RNA isolation

For the isolation of total RNA, *Manduca* 5<sup>th</sup> instar larvae that were either feeding, starved for 36 h, re-fed 8 h after 36 h of starvation or molting from the 4<sup>th</sup> to the 5<sup>th</sup> instar (molting stage F according to Baldwin and Hakim, 1991) were dissected on cooling plates. The midguts were isolated, opened and the gut content was removed by peeling it off together with the peritrophic matrix. The anterior, median and posterior parts of the midguts were separated. Malpighian tubules were obtained by carefully detaching them from the midgut without injuring the tissues. Abdominal tracheae were excised between the midgut and the epidermis. The fat body was prepared from subepidermal regions. Total RNA from *Tribolium* was isolated from about 40 embryos and 3-5 whole animals of young larval, last instar larval, pharate pupal, pupal and adult stages, respectively. For tissue specific expression analysis in *Tribolium*, total RNA was prepared from midgut and carcass tissues of young and last instar larvae as well as prepupae. The carcass tissue was obtained by removing the head, the two last abdominal segments and the complete gut. To avoid contaminations, midguts were carefully removed from the insects and the remaining carcass was washed repeatedly in a phosphate buffered saline (PBS; 0.1 M potassium phosphate, pH 7.0). To enable clear differentiation between carcass and midgut expression, the legs of 10 young larvae were dissected additionally, serving as a non-gut control.

### 2.2.2. Total RNA isolation

Total RNA was isolated by using either TRIZOL Reagent (Invitrogen) or the RNA-easy Mini Kit (Qiagen) following the manufacturer's recommendations. TRIZOL Reagent is a mono-phasic solution of phenol and guanidine isothiocyanate. The method is an improvement to the single-step RNA isolation method developed by Chomczynski and Sacchi (1987).

### 2.2.3. Reverse transcription

First strand cDNA synthesis was accomplished by using either the ThermoScript RT-PCR System (Invitrogen) or the SuperScript III First-Strand Synthesis System (Invitrogen) following the manufacturer's recommendations. Transcription was performed by using total RNA as a template and oligo(dT) primers. For complete hydrolyzation of the RNA component of DNA/RNA hybrids, samples were incubated for 20 min with 2U *E. coli* RNase H. The cDNA was stored at -20°C or directly used for PCR.

### 2.2.4. Quantification of nucleic acids

For the quantification of DNA and RNA 1 µl of solutions containing nucleic acids were submitted to the NanoPhotometer (Implen) using the LabelGuard technology. The concentration was calculated according to Sambrook *et al.* (1989).

### 2.2.5. Chemical competent *Escherichia coli*

Generation of chemical competent *E. coli* was performed according to Hanahan (1983). For this purpose 5 ml ΨB-media were inoculated with a single colony of *E. coli* and incubated overnight at 37°C and shaking at 220 rpm. On the next day this preculture was transferred into 100 ml ΨB-media and incubated at 37°C and shaking at 220 rpm until the extinction reached  $A_{600} \approx 0.5$ . Following chilling on ice for 15 min, bacteria were centrifuged at 3500 x g and 4°C for 10 min. The pellet was resuspended in 100 ml Tfb1 (30 mM KAc, 100 mM RbCl, 10 mM CaCl<sub>2</sub>, 50 mM MnCl<sub>2</sub>, 15% (v/v) Glycerin, pH 5.8). The bacterial suspension was once more centrifuged at 3500 x g and 4°C for 10 min and the pellet resuspended in 10 ml Tfb2 (10 mM MOPS, 75 mM CaCl<sub>2</sub>, 10 mM RbCl, 15% (v/v) Glycerin, pH 6.5) and chilled on ice for 15 min. Bacteria were aliquotted, frozen in liquid nitrogen and stored at -80°C.

### 2.2.6. Ligation and transformation in *E. coli*

Ligation of cohesive DNA ends was performed at a total volume of 10 µl using 50-200 ng of vector DNA as and a threefold molar excess of insert DNA. The reaction was carried out overnight at 16°C using 3 weiss units T4-DNA-Ligase (New England Biolabs) following the manufacturer's recommendations. To transform competent *E. coli* cells, 200 µl of the suspension were quickly thawed and mixed with the ligation

reaction followed by 10 min incubation on ice. After a heat shock at 42°C for 90 sec, the reaction was chilled again on ice for 2 min. Then 800 µl LB-media were added and the bacteria were incubated for 1h at 37°C with shaking. The cells were pelleted and plated on LB-agar plates containing the appropriate antibiotics for selection. If the vector as well as the bacterial strain allowed  $\alpha$ -complementation of the *lacZ* gene, 10 µl 1 M IPTG and 2% (w/v) X-Gal were additionally plated onto the LB-agar.

### 2.2.7. Preparation of plasmid DNA

Isolation of plasmid DNA was accomplished by using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's recommendations. Therefore bacteria were lysed under alkaline conditions. Following neutralization and adjustment to high salt binding conditions, the lysate was cleared by centrifugation. The cleared lysate was submitted to QIAprep columns where plasmid DNA selectively binds to a silica membrane under high salt conditions, whereas RNA, cellular proteins, and metabolites are not retained on the membrane. After column washing, plasmid DNA was eluted under low salt conditions. Plasmids were used immediately or stored at -20°C.

### 2.2.8. Restriction enzyme digestion of DNA

For the enzymatic restriction of DNA 3-5 U per µg DNA of the individual restriction enzyme were used. Digestion was performed for 2 h using appropriate buffers and temperature. To test the efficiency of the reaction and to purify the DNA from the enzymes, DNA fragments were electrophoretically separated and the respective band was excised from the gel and purified by using the QIAquick Gel Extraction Kit (Qiagen).

### 2.2.9. DNA gel-electrophoresis and DNA isolation

Separation of DNA-fragments by size was performed in 0.7 – 2.0% agarose (Serva) gels. Gels were prepared by cooking the respective amount of agarose in TAE buffer (40 mM TRIS, 10 mM NaAc, 1 mM EDTA, pH 8). Following electrophoresis at a field intensity of 5-7 V/cm the gel was stained in homidium-bromide (2 µg/ml). After washing in deionized water, nucleic acid fragments were visualized under UV-Light exposure at 312 nm with the help of the VersaDoc MP Imaging System (Bio-Rad). For further characterization of DNA fragments, bands were excised under UV-Light exposure and purified with the help of the QIAquick Gel Extraction Kit (Qiagen) following the manufacturer's recommendations.

## 2.2.10. Polymerase-chain-reaction based methods

### 2.2.10.1. RACE-PCR

When only partial cDNA sequences were available, the FirstChoice RLM-RACE Kit (Ambion) was used to amplify full length cDNAs. This kit is an improvement of the classic RACE technique (Maruyama and Sugano, 1994; Schaefer, 1995) amplifying cDNAs only from full-length, capped mRNA.

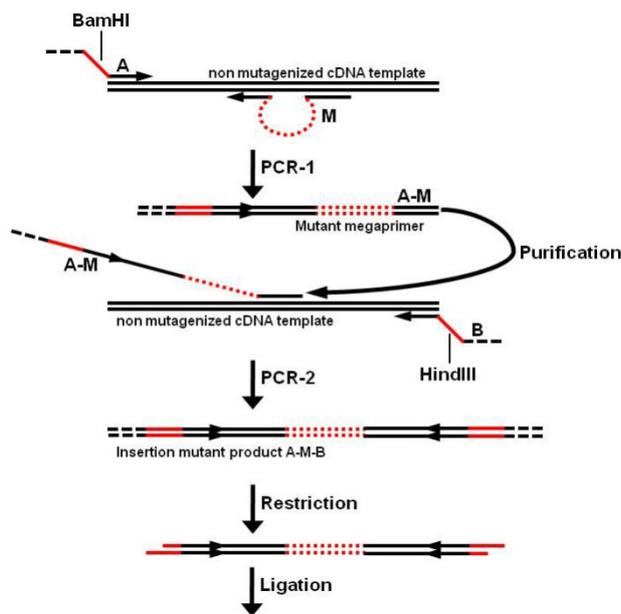
In the first step of the 5'RACE the oligo-nucleotide mixture is treated with calf intestine alkaline phosphatase (CIP) removing free 5'phosphates from fragmented RNA, contaminating DNA and ribosomal RNA, whereas intact mRNA is protected by the cap structure. In a second step this structure is removed from full length mRNA by the tobacco acid pyrophosphatase leaving a single 5'phosphate. In the final step this restricts the ligation of a 45 bp 5'RACE adapter by T4 RNA ligase only to intact mRNA since all other oligo-nucleotides are dephosphorylated and therefore lack the 5'phosphate necessary for the ligation. The RNA was then used in a RT-PCR followed by a nested PCR using 5'RACE-adapter specific primes and gene specific primers. To amplify and clone unknown 3' regions, a 3'RACE was performed. For the 3'RACE, RNA is reverse transcribed using the 3'RACE adapter. This adapter contains a specific sequence as well as a poly-T stretch which will bind to the poly-a-tail of the mRNA. The obtained cDNA was used in a PCR using 3'RACE-adapter specific primes and gene specific primers.

### 2.2.10.2. Quantitative Realtime PCR

Quantitative Realtime PCR is a technique to simultaneously amplify and quantify a target DNA. The usage of a fluorescent dye, which intercalates into double stranded DNA allows quantification during the exponential phase of the PCR by the measurement of fluorescent signals during a single PCR cycle (Realtime). Fluorescent signals are proportional to the amount of the accumulating PCR products. For the quantitative Realtime-PCR total RNA was isolated from the midgut epithelium or the Malpighian tubules of *Manduca* 5<sup>th</sup> instar larvae. Larvae were either feeding, starved for 36 h, re-fed 8 h after 36 h of starvation or molting from the 4<sup>th</sup> to the 5<sup>th</sup> instar (molting stage F according to Baldwin and Hakim, 1991). First strand cDNA synthesis was performed using the ThermoScript RT-PCR System (Invitrogen). About ~200 bp fragments were amplified using the primer pairs MsCHS2-F/-R, MsCT-F/-R, MsCTLP1-F/-R, MsCTLP2-F/-R, MsCTLP3-F/-R and MsCTLP4-F/-R (Table 7). Fluorescent-based Realtime PCR was carried out with the iCycler iQ Realtime PCR Detection System and iQ SYBR Green Supermix following the manufacturer's recommendations (Bio-Rad). The conditions for the PCR reactions were the following: initial denaturation at 95°C for 10 min followed by 42 cycles of denaturation at 95°C for 30 s, annealing at 59°C for 45 s and elongation at 72°C for 45 s. The specificity of each reaction was tested by melting curve analyses. Expression levels are given as relative CT-values normalized to 10<sup>6</sup> copies of the ribosomal protein S3 from *M. sexta*

(*MsRPS3*; GenBank|U12708.1; *MsRPS3-F/-R*; Table 7). Background fluorescence was subtracted using the CT-values obtained for the fat body, which showed no CTLP expression in Northern blots or RT-PCR.

### 2.2.10.3. Megaprimer PCR



**Fig. 9.** The megaprimer PCR method. Restriction sites in primer A and B are marked in red lines. The inserted region in the mutant primer M and in subsequent PCR products is indicated by the dotted red line. Modified after Burke and Barik (2003).

megaprimer in a PCR using a pGEM-T vector containing the whole coding region of *MsCTLP3* as a template and the *MsCTLP3* specific primer MsCT3-R (Table 7). After gel electrophoresis the PCR product was excised, purified with the QIAquick Gel Extraction Kit (Qiagen) and cloned into the pGEM-T vector. Following sequencing the vector was used as template in PCR with the primers MsCTLP1-BH-F and MsCTLP3/1-R (Table 7). The latter primer was again containing the last 15 bps of the exchanged *MsCTLP3* region as well as the base pair positions 641-655 of *MsCTLP1*. The resulting PCR fragment was applied as the second megaprimer in a PCR using the primer MsCTLP-BH-R (Table 7) and the pGEM-T vector containing the full coding region of *MsCTLP1* as a template. Following gel extraction, the PCR product was digested with BamHI and HindIII and ligated into the pFBD vector. After amplification in DH5 $\alpha$  and plasmid preparation (QIAprep Spin Miniprep Kit; Qiagen), the vector was send in for sequencing. Sequencing confirmed that the desired *MsCTLP3* region has been exchanged against those base pairs coding for the hydrophobic region in *MsCTLP1*. To exclude any unwanted mutations, megaprimer PCR was carried with proofreading Phusion High-Fidelity DNA Polymerase (New England Biolabs).

For the exchange of a hydrophobic region of *MsCTLP1* against a more hydrophilic region of *MsCTLP3*, a modified protocol of the megaprimer PCR was performed (Burke and Barik, 2003). The method is a simple application in mutagenesis and gene fusion (general principle displayed in Fig. 9). Starting with a pGEM-T vector carrying the full coding region of *MsCTLP1*, a PCR was carried out using the primers MsCTLP1-BH-F and MsCTLP1/3-R (Table 7). The latter primer was binding at base pair positions 527-541 of *MsCTLP1* and additionally contained the initial 15 bps of the *MsCTLP3* region that should be exchanged. The resulting PCR fragment was used as the forward

### 2.2.11. Cloning of insect CTLPs

A partial sequence for MsCTLP1 (embl|CAL92020.1) was identified by a yeast two hybrid screen of a midgut cDNA library of *Manduca* 5<sup>th</sup> instar larvae using the C7 domain of MsCHS2 as bait (Broehan *et al.*, 2007). The isolated cDNAs were 409 base pairs (bps) in length and encoded 109 amino acids of the carboxy-terminal half of MsCTLP1 and the complete 3' untranslated region but lacked the amino-terminal end and the 5' untranslated region. The missing 5' region of MsCTLP1 was completed by 5'RACE and the sequence specific primers MsCTLP1-5'-In and MsCTLP1-5'-Out. Nested PCR yielded a single product of about 600 bps, which was cloned and sequenced. The RACE product included the 5' untranslated region and 59 nucleotides were overlapping with the 5' end of the truncated *MsCTLP1* cDNA. For the isolation of the cDNA for *MsCTLP2* (embl|CAM84317.1) a *Manduca* midgut cDNA library (Gräf *et al.*, 1994) was screened with a digoxigenin-labeled cDNA probe generated by PCR using the primer pair MsCTLP2-DIG-F/-R (Table 7). After plaque hybridization, cloning and in vivo excision a cDNA fragment of 842 bps was obtained encoding the carboxy-terminal half of the coding region and the complete 3' untranslated region of *MsCTLP2*. Completion of the 5' region was performed by 5' RACE using the sequence specific primers MsCTLP2-5'-In and MsCTLP2-5'-Out (Table 7). The 3' end of the resulting PCR fragment, which was 942 bps in size perfectly overlapped with the 5' end of the previous nucleotide sequence. A central cDNA fragment encoding *MsCTLP3* (embl|CAM84318.1) was amplified from the midgut total RNA using the degenerate primers MsCTLP3-Deg-F/-R (Table 7). As this cDNA was truncated at the 5' and 3' region, 5'RACE was performed with sequence specific primers MsCTLP3-5'-In and MsCTLP3-5'-Out and 3'RACE was performed with sequence specific primers MsCTLP3-3'-In and MsCTLP3-3'-Out (Table 7). The resulting fragments perfectly overlapped with the central cDNA fragment giving the complete cDNA of 943 bps encoding *MsCTLP3*. The cDNA sequence of *MsCTLP4* (embl|CAM84319.1) was obtained by screening the midgut cDNA library with a *MsCTLP3* specific digoxigenin-labeled cDNA probe obtained by PCR with the primer pair MsCTLP3-DIG-F/-R (Table 7). After plaque hybridization, cloning and in vivo excision a full-length cDNA of 983 bps was obtained encoding *MsCTLP4*. The nucleotide sequences for *MsCTLP1-4* were confirmed by sequencing the amplification products for the complete coding sequence, which were obtained by RT-PCR with the corresponding 5' and 3' primers and mRNA from the larval midgut. The cDNAs for the seven *TcCTLPs*, *TcCTLP-5A* (embl|CBC01177.1), *TcCTLP-5B* (embl|CBC01166.1), *TcCTLP-5C<sub>1</sub>* (embl|CBC01175.1), *TcCTLP-5C<sub>2</sub>* (embl|CBC01172.1), *TcCTLP-6A* (embl|CBC01171.1), *TcCTLP-6C* (embl|CBC01169.1), *TcCTLP-6D* (embl|CBC01179.1) and *TcCTLP-6E* (embl|CBC01181.1) identified in the *Tribolium* genome were amplified by RT-PCR using gene-specific primers (Table 7) and total RNA from the larval midgut. The PCR products were ligated into the pGEM-T vector (Promega) and amplified in  $\alpha$ -select competent cells (Bioline). Plasmid DNA was isolated using the Plasmid Mini Kit (Qiagen) and the inserts were sequenced (Sequencing Core Facility, Dept. of Plant Pathology, Kansas State University). Sequences matched those published in Beetlebase (<http://beetlebase.org/>) and NCBI's GenBank (<http://blast.ncbi.nlm.nih.gov>) except for a

few nucleotide substitutions. A near-full-length *TcCTLP-5C* cDNA could not be amplified in the first attempt. However, a central 263 bp cDNA fragment of *TcCTLP-5C* was amplified using the PCR primers TcCTLP-5C-F and TcCTLP-5C-R (Table 7). Completion of the missing 5'- and 3'- sequences of *TcCTLP-5C* was accomplished with the FirstChoice RLM-RACE Kit and the gene-specific primers TcCTLP-5C-5'-Out and TcCTLP-5C-5'-In for 5'RACE and TcCTLP-5C-3'-Out and TcCTLP-5C-3'-In for 3'RACE (Table 7).

### 2.2.12. Syntheses of Digoxigenin-labeled DNA and RNA probes

Synthesis of Digoxigenin-labeled (DIG-labeled) DNA probes was performed with the DIG DNA Labeling Mix (Roche). The labeled DNA probes were generated by PCR using vectors carrying the designated region for the probe as a template and the specific primers MsCT-P-F/-R, MsCTLP1-SP-F/-R, MsCTLP1-NP-F/-R, MsCTLP2-SP-F/-R, MsCTLP2-NP-F/-R, MsCTLP3-SP-F/-R, MsCTLP3-NP-F/-R and MsCTLP4-P-F/-R (Table 7). The dNTP mixture used contained 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 0.65 mM dTTP as well as 0.35 mM DIG-11-dUTP, allowing the incorporation of DIG-labeled dUTP into the probe. For the purification of the probes and to check whether DIG-labeled dUTP was incorporated, the samples were separated on 2% SeaPlaque low melting point agarose (Lonza) gels. Since the incorporation of labeled dUTP increases the molecular mass, DIG-labeled DNA probes are retained in gel electrophoresis. Probes were excised from the low melting point agarose gel and used either directly for hybridization or were frozen at -20°C.

Synthesis of Digoxigenin-labeled RNA probes was performed with the DIG RNA Labeling Kit (Sp6/T7) (Roche) according to the manufacturer's recommendations. As a template for *in vitro* transcription pGEM-T vectors carrying MsCTLP1 (nucleotide positions 1–178) or MsCHS2 cDNA (nucleotide positions 2094–2303) were linearized and purified by gel electrophoresis (see 2.2.6). Linearization was performed by cleaving either upstream or downstream of the cDNA-insert. *In vitro* transcription of linearized plasmids was performed with Sp6 or T7 RNA polymerases to synthesize the sense or antisense probes, respectively. After 2 h incubation at 37°C, the reaction was stopped by the addition of 2 µl 0.2 M EDTA. RNA probes were precipitated for 30 min incubation at -80°C by the addition of 2.4 µl 4 M LiCl<sub>2</sub> and 75 µl 99.8 % ethanol. Following centrifugation for 15 min at 12000 x g and 4°C the RNA pellet was washed with 50 µl 70% ice-cold ethanol. After another centrifugation step for 5 min at 12000 x g the pellet was resuspended in 50 µl DEPC-H<sub>2</sub>O at 37°C. Purified DIG-labeled RNA probes were either directly used for hybridization or stored at -80°C.

### 2.2.13. RNA gel-electrophoresis

Separation of RNA-fragments was performed in 1.0% agarose (Serva) gels containing 2.0% formaldehyde. For the preparation of the gels 1 g agarose was boiled in 84 ml DEPC-H<sub>2</sub>O. After cooling to about 60°C, 6 ml of 37% formaldehyde and 10 ml 10 x RNA-gel running buffer (80 mM NaAc, 196 mM MOPS, 10 mM EDTA, pH 7.0)

were added and the solution was poured into a casting tray treated with RNase AWAY (ROTH). Gel electrophoresis was carried out in a gel running buffer (70 ml 10 x RNA-gel running buffer filled up to 700 ml with DEPC-H<sub>2</sub>O + 40 ml of 37% formaldehyde) at 5 V/cm for 15 min. RNA samples were prepared by the addition of 2.5 µl RNA-gel running buffer, 3.5 µl of 37% formaldehyde, 10.0 µl formamide and filled up with DEPC-H<sub>2</sub>O to 23.0 µl. Before loading, the RNA was denatured by heating for 15 min at 68°C and chilling on ice for 2 min. Following the addition of 2 µl RNA sample buffer (50% (v/v) Glycerin, 1 mM Na<sub>2</sub>EDTA (pH 8.0), 0.25% (w/v) bromophenol blue), samples were electrophoresed at 4 V/cm. RNA was stained in Radiant Red (Bio-Rad) following the manufacturer's recommendations. RNA-fragments were visualized at 302 nm with the help of the VersaDoc MP Imaging System (Bio-Rad).

#### 2.2.14. Northern blotting

For Northern hybridization 5 µg of total RNA from different tissues of *M. sexta* were loaded onto a 1% agarose/formaldehyde gel. Sample preparation and gel electrophoresis were performed as described above. Northern transfer, hybridization and chemiluminescent detection were performed according to Merzendorfer *et al.* (2000).

#### 2.2.15. Genomic DNA isolation and Southern blotting

For Southern hybridization genomic DNA was isolated from total midguts of *M. sexta* 5th instar larvae as described previously (Gräf *et al.*, 1994). Each 5 mg of genomic DNA was digested overnight with 5–10 U/mg of four different restriction enzymes (New England Biolabs). The restriction enzymes did not cut in the regions selected for probe hybridization. The resulting DNA fragments were electrophoresed in a 1% agarose gel and blotted onto a Hybond N nylon membrane (GE Healthcare) according to Sambrook *et al.* (1989). Moderate or high stringency hybridization was performed at 52°C in the absence of formamide or at 58°C in the presence of 50% formamide using the DIG-labeled probe at a concentration of 100 ng/ml. Chemiluminescent detection with CSPD (Roche) was carried out as described previously (Gräf *et al.*, 1994; Merzendorfer *et al.*, 1999).

#### 2.2.16. *In situ* hybridization

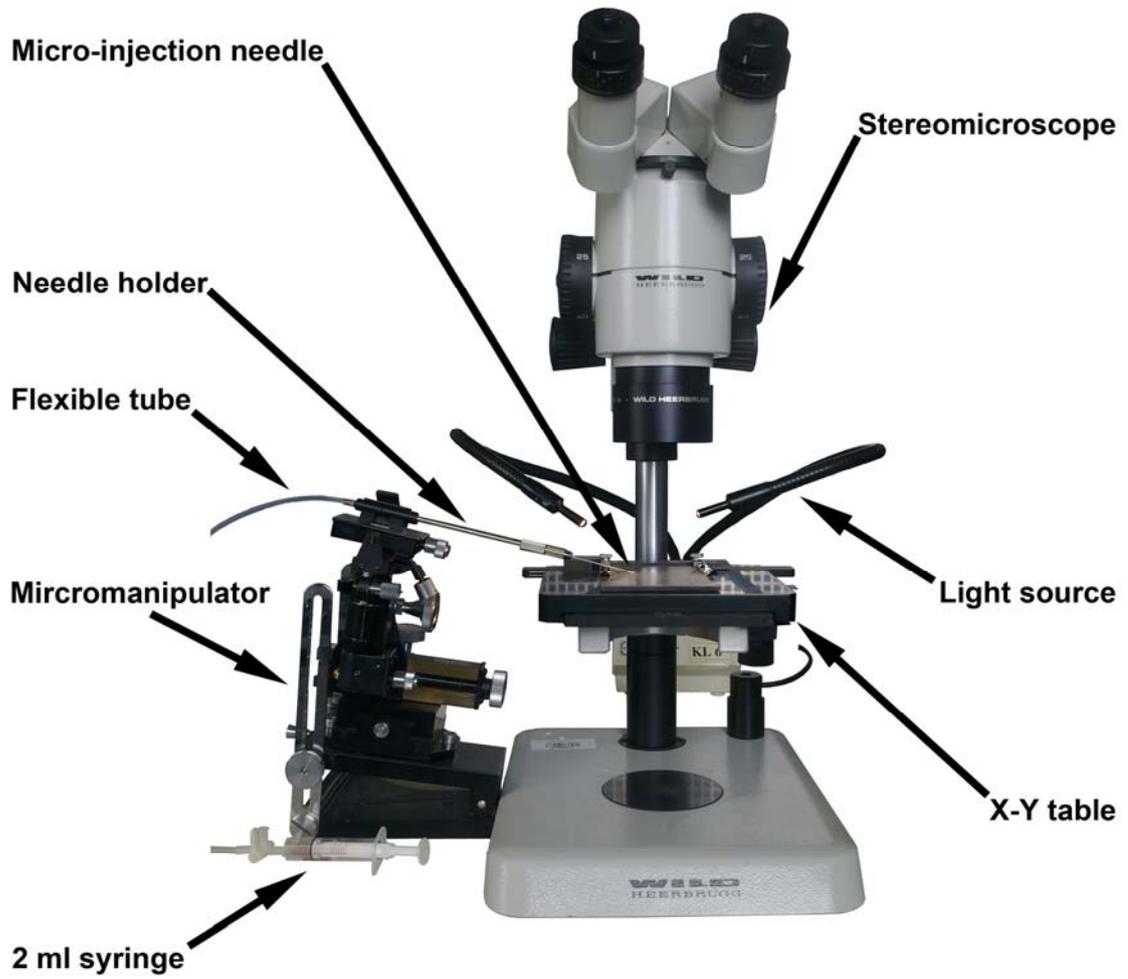
Synthesis of RNA probes was performed as described above (2.2.11). *In situ* hybridization was performed according to previous protocols (Zimoch and Merzendorfer, 2002), with the exception that the detection of RNA probes was carried out with Anti-Digoxigenin-Gold antibodies (Roche) and silver enhancement (IntenSETM silver enhancement kit, GE Amersham).

### 2.2.17. Generation of double stranded RNA

For the synthesis of dsRNA used in knockdown experiments in *T. castaneum*, full coding regions of the respective *TcCTLPS* were cloned into the pGEM-T vector. The pGEM-T-*TcCTLPS* vectors were used as templates in a PCR with primers containing T7 promoter sequences at the 5'-end and a sequence specific for the corresponding *TcCTLPS* cDNA (Table 7). After agarose gel electrophoresis, PCR products were excised and purified with the QIAquick Gel Extraction Kit (Qiagen), subsequently used as templates for dsRNA synthesis performed with the AmpliScribe T7-Flash transcription kit (Epicentre) following the manufacturer's recommendations. After *in vitro* transcription and hybridization of the complementary RNA strands, 30  $\mu$ l of DEPC-H<sub>2</sub>O was added to the reaction leading to a total volume of 50  $\mu$ l. Following the addition of phenol-chloroform-isoamyl alcohol in a 1:1 proportion the sample was vortexed and centrifuged for 5 min at 13000 x g and 4°C. The RNA was precipitated from the aqueous layer by the addition of 10  $\mu$ l 2 M ammonium acetate and 50  $\mu$ l isopropanol as well as 1 h incubation at -20°C. Following the centrifugation for 12 min at 13000 x g and 4°C, the pellet was washed with 1 ml 75% ice-cold ethanol and centrifuged again for 3 min at 13000 x g and 4°C. The pellet was resuspended in 30  $\mu$ l DEPC-H<sub>2</sub>O. Purified dsRNA was either used directly for injection into *Tribolium* or was stored at -80°C.

### 2.2.18. Double stranded RNA injection and RNA interference

For RNAi experiments at least 30 penultimate instar larvae, prepupae, pupae (two days old) were anesthetized by exposure to ether for 3 minutes. Injection occurred with the help of a stereo-microscope (Wild Heerbrugg; experimental setup see Fig. 10) and a micromanipulator holding the needle holder with the micro-injection needle in a fixed angle. The needle-holder was connected by a flexible tube with a 2 ml syringe. In case of larval or prepupal injections, the anesthetized insects were arranged laterally on double sided Scotch-tape. By carefully moving the X-Y table of the stereomicroscope against the immobile injection needle, the latter was inserted between the second and third abdominal segment of the insects. For pupal injections, insects were positioned with the ventral side facing the Scotch-tape and injection occurred between the first and second abdominal segments. Adults (shortly after eclosion) were anesthetized by chilling on ice and injection into the abdominal segments was performed after carefully lifting the elytra and the hindwings with forceps. In every case about 200 nl of the indicated *Tribolium CTLPS* dsRNA (1 $\mu$ g/ $\mu$ l; in 0.1 mM potassium phosphate buffer containing 5 mM KCl, pH 7; Tomoyasu and Denell, 2004) was injected. For visual evaluation of the dsRNA amount injected, filter sterilized food coloring was added to the injection buffer. After injection, the insects were kept under standard conditions for visual monitoring of phenotypes and further analysis. Phenotypes were monitored on a daily basis with the help of a stereomicroscope (Leica). Buffer containing dsRNA for the gene encoding *tryptophan oxygenase* from *T. castaneum* (*TcVer*; GenBank|NP\_001034499.1; Arakane *et al.*, 2009) was injected as a control.



**Fig. 10.** Experimental setup used for the injection of dsRNA into *Tribolium*.

## 2.3. Biochemical methods

### 2.3.1. Tissue preparations

Different kind of tissues from *M. sexta* were prepared as described above (2.2.1). To obtain crude extracts, the tissues were washed twice in TBS-Low (20 mM TRIS-HCl, 20 mM NaCl, pH 7.6) and homogenized in 1 ml of TRIS buffer (20 mM TRIS-HCl, 20 mM NaCl, 20 mM Na<sub>2</sub>EDTA, 5 mM Pefabloc SC, pH 7.6). Larval gut content was obtained by separating it from the peritrophic matrix and centrifugation for 30 min at 20000 x g and 4°C. The supernatant was directly used for further experiments. Faeces was resuspended in F-buffer (20 mM TRIS-HCl, 10 mM NaCl and 10 mM Pefabloc SC, pH 8.0) to a final concentration of 0.2 mg/ml, vortexed for 5 min and subsequently mixed with SDS-PAGE sample buffer (2.3.3). For the preparation of *Tribolium* midgut tissues, midguts from 20 larvae were dissected in dissection buffer (0.1 M potassium phosphate, pH 7) and homogenized. Exuviae of 40 larvae were homogenized in dissection buffer and subsequently boiled for five minutes in SDS-PAGE sample buffer.

### 2.3.2. Determination of protein concentration

Protein concentrations were determined by the Amido Black method (Wieczorek *et al.*, 1990). The method is relatively independent of the use of detergents in the solution and is based on the reaction of acidic Amido Black with alkaline amino groups. Following the addition of the Amido Black solution (26 mg Amido Black/100 ml in 10% (v/v) acetic acid, 90% (v/v) methanol) to the protein samples as well as to a BSA calibration series (2, 4, 6 µg) samples were incubated for 5 min at room temperature and then centrifuged at 16000 x g for 4min. After washing twice in a washing solution (10% (v/v) acetic acid, 90% (v/v) methanol), the pellet was solubilized in 100 mM NaOH and its extinction was measured photometrically at 615 nm. Protein concentration was calculated with the help of the BSA calibration series.

### 2.3.3. SDS-Page and Coomassie blue staining

Discontinuous electrophoretical separation of proteins according to their molecular radius was performed with the help of SDS polyacrylamide gels as previously reported by Laemmli (1970). Separating gels contained 17% acrylamide at 0.5% degree of cross linking in 330 mM TRIS-HCl (pH 8.7) and 0.1% (w/v) SDS. Following the polymerization, the separating gel was covered with the loading gel (5.4% acrylamide, 2.3% degree of cross linking, 125 mM TRIS-HCl, pH 6.8, 0.2% (w/v) SDS). Samples were prepared by cooking for 1 min in SDS-PAGE sample buffer (2% SDS (w/v), 5% sucrose (w/v), 0.125 M TRIS-HCl, pH 6.8, 0.005% bromophenol blue (w/v), 2% 2-mercaptoethanol (v/v)). Electrophoresis was performed in a TRIS-glycine gel running buffer (100 mM TRIS, 100 mM glycine, 0.1% (w/v) SDS) until the bromophenol blue front left the gel. Proteins were either transferred to nitrocellulose membranes or stained directly with Coomassie blue (Bollag and Edelstein, 1991).

### 2.3.4. Immunological methods

Semi-dry electroblotting of the polyacrylamide gels onto nitrocellulose membranes (0.2 mm, SERVA; Towbin *et al.*, 1979) was carried out with a threepartic buffer system according to previous protocols (Kyhse-Andersen, 1984), modified by the addition of 20% methanol. Blot membranes were stained with 0.02% (v/v) Ponceau S (SERVA). Immunoblots were performed as described previously (Zimoch and Merzendorfer, 2002). The immunoreactions were carried out with polyclonal or monoclonal antibodies and alkaline phosphatase conjugated secondary antibodies listed in Table 5 and 6. Cryosectioning of tissues and immunocytochemistry were carried out as described previously (Zimoch and Merzendorfer, 2002) using anti-*Ms*CHS2 and *Ms*CTLPI antibodies. Immunoprecipitation was performed with the protein G Immunoprecipitation kit (Sigma) following the instructions of the manufacturer.

### 2.3.5. Detection of glycoproteins

Detection of glycoproteins was performed with the DIG Glycan Differentiation Kit (Roche) following the manufacturer's recommendations. The kit contains several digoxigenin-labeled lectins (Table 1) binding discriminately to specific carbohydrate structures of proteins transferred onto blot membranes. Detection of the lectins was performed using anti-digoxigenin alkaline-phosphatase conjugates. The digoxigenin-labeled lectins enable differentiation between complex and high-mannose chains, differentiation between  $\alpha(2-3)$  and  $\alpha(2-6)$  linkage of terminal sialic acids, and identification of the core disaccharide Gal $\beta(1-3)$ GalNAc of O-glycans.

**Table 1.** Lectins used for the detection of glycoproteins.

DIG labeled lectin	Function
GNA ( <i>Galanthus nivalis</i> agglutinin)	Recognizes terminal mannose, (1–3), (1–6) or (1–2) linked to mannose
SNA ( <i>Sambucus nigra</i> agglutinin)	Recognizes sialic acid linked (2–6) to galactose; correspondingly linked sialic acids in O-glycan structures are also recognized
MAA ( <i>Maackia amurensis</i> agglutinin)	Recognizes sialic acid linked (2–3) to galactose; identifies (2–3)-linked sialic acids in O-glycans
PNA (Peanut agglutinin)	Recognizes the core disaccharide galactose- $\beta(1-3)$ -N-acetylgalactosamine
DSA ( <i>Datura stramonium</i> agglutinin)	Recognizes Gal- $\beta(1-4)$ GlcNAc in complex and hybrid N-glycans, in O-glycans and GlcNAc in O-glycans

### 2.3.6. Heterologous expression of recombinant proteins

Expression of recombinant proteins was either performed in prokaryotic protein expression systems or with the help of baculoviruses in eukaryotic cells. For the bacterial expression of proteins, the corresponding cDNA was ligated into the pET29b vector (Novagen). Protein expression was performed in *E. coli* Rosetta (DE3) pLysS cells as described previously (Merzendorfer *et al.*, 2000). For baculoviral expression of recombinant proteins in eukaryotic cells, RT-PCR was performed, using primers

containing 5' *Bam*HI (TcCTLP-5C<sub>2</sub>-BH-F; TcCTLP-6C-BH-F; MsCTLP1-BH-F; MsCTLP3-BP-F), 5' *Spe*I (MsCTLP1-Z-Bip-F; MsCTLP1-A-Bip-F), 3' *Pst*I (MsCTLP3-BP-R) and 3' *Hind*III (TcCTLP-5C<sub>2</sub>-BH-R; TcCTLP-6C-BH-R; MsCTLP1-BH-R) restriction sites (Table 7). These primers were amplifying the target region plus a sequence coding for a carboxy-terminal octa-His-tag. PCR products were separated by gel electrophoresis, excised and purified with the QIAquick Gel Extraction Kit (Qiagen). After cleavage with the respective restriction enzymes, the products were ligated into the pFast-Bac-Dual (pFBD) vector, which additionally contained a sequence coding for an enhanced green fluorescence protein (EGFP) under the control of a p10 promoter (a kind gift of Dr. Monique van Oers, Wageningen University). To enable secretion of peptidases lacking the endogenous signal peptide for the secretory pathway, an alternative signal peptide was assembled into the pFBD vector before the insertion of the target region. For this reason the BiP region of the pMT/BiP/V5-His vector was cut out with the restriction enzymes *Bam*HI and *Sac*I and ligated into the pFBD vector, which had been opened with the respective restriction enzymes previously. The vectors were amplified in DH5 $\alpha$  cells, isolated with the Plasmid Mini Kit (Qiagen) and used for transformation of DH10Bac cells. After recombination, high molecular weight bacmid DNA was isolated and used for the infection of *Sf*21 insect cells. *Sf*21 cells were either used to directly express the recombinant protein or for the amplification of recombinant viruses. In the latter case the *Sf*21 cell cultures supernatants were used to infect Hi-5 cells at 70-80% confluence, which were incubated at 27°C. After inspection of EGFP expression with a fluorescence microscope (Olympus IX70), both, Hi5 as well as *Sf*21 cells were harvested five days after infection.

### 2.3.7. Affinity purification of proteins

Bacterial as well as insect cell culture supernatants were obtained by a 600 x g centrifugation for 10 min at 4°C. Cleared supernatants were submitted to a Ni-NTA agarose column (Qiagen). After washing the column in washing buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM imidazole, pH 8.0) elution of bound proteins was accomplished by a step gradient of imidazole in washing buffer from 50 to 500 mM (50 mM steps). Purification of recombinant proteins from the pellet fraction of *Sf*21-cells was accomplished by homogenizing the latter in washing buffer with the addition of 0.1 % N-lauroyl sarcosinate (NLS). Following a centrifugation step at 20000 x g for 30 min the pellet was solubilized in washing buffer with the addition of 6 M guanidinium hydrochloride (GuHCl), incubated for 30 min at room temperature and vortexing every now and then. After another centrifugation step at 20000 x g for 30 min, the supernatant was loaded to the Ni-NTA column. Elution was carried out as described above with the addition of 6 M GuHCl to the elution buffer. Subsequently the recombinant peptidases were dialyzed against 20 mM N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) pH 11.0, 140 mM KCl.

In case Ni-NTA purification was not sufficient, the recombinant proteins were further purified on a Mono Q anion exchange column (GE Healthcare) equilibrated with

running buffer (20 mM TRIS-HCl, 50 mM NaCl, pH 8.1). Bound protein was eluted with a linear NaCl gradient (0.05-1 M NaCl, 20 mM per min) in running buffer.

Purification of native MsCTLP1 from midgut tissues was performed with the Affi-Gel Hz Immunoaffinity Kit (Bio-Rad). The method allows an orientated IgG coupling to an agarose support matrix for affinity purification. Following the purification of the anti-MsCTLP1 antibody (2.3.8) the hydroxyl groups of the carbohydrates of the F<sub>c</sub> region of the antibody were oxidized by periodate. This leads to the formation of aldehyde groups for specific coupling to the Affi-Gel Hz gel, forming stable, covalent hydrazone bonds. This coupling through the carbohydrate allows the correct orientation of the antibody and maintains its functionality. After antibody coupling the column was equilibrated with 10 ml equilibration buffer (20 mM MOPS, 0.5 M NaCl, pH 6.5) and checked for the bleeding of antibodies. Two midguts were homogenized in sample buffer (100 mM MOPS, 125 mM NaCl, pH 6.5, 20 mM EDTA, 0.5 µg/µl Soybean Trypsin inhibitor (Sigma)) and centrifuged at 20000 x g for 10 min at 4°C. The supernatant was diluted 1:10 in sample buffer and submitted to the Affi-matrix loaded with the anti-MsCTLP1 antibody. Following a wash step at high salt conditions (20 mM MOPS, 0.5 M NaCl, pH 6.5) and a wash step at low salt conditions (20 mM MOPS, 25 mM NaCl, pH 6.5) the protein was eluted with 10 ml elution buffer (0.1 M glycine-HCl, pH 3.5). 500 µl fractions were collected into 100 µl of neutralization buffer (1 M TRIS-HCl, 125 mM NaCl, 5 mM Pefabloc SC, pH 8.0).

### 2.3.8. Generation of antibodies

Polyclonal antibodies to the recombinant catalytic domain of the *Manduca* chitin synthase were generated previously by Zimoch and Merzendorfer (2002). The monoclonal antibodies to the V-ATPase subunit A (also named 221-9) were generated previously by Klein *et al.* (1991). To generate anti-MsCTLP1 antibodies, the peptide IVGGTQAPSGSHPH (amino acid positions 41–54) was synthesized, coupled to Keyhole limpet haemocyanin and used for the immunization of rabbits (Charles River). Since the MsCTLP1 peptide used for antibody generation shares some similarities with homologue regions from MsCT and MsCTLP2–4, the possibility of cross-reactions with other *Manduca* CTLPs cannot be excluded. For the generation of TcCTLP-5C<sub>2</sub> and TcCTLP-6C antibodies, proteins were recombinantly expressed in Hi-5 cells (2.3.6). Following the purification (2.3.7) and SDS-PAGE to check their integrity and correct size (Fig. 49, lane 1), the peptidases were lyophilized. Recombinant TcCTLP-5C<sub>2</sub> was used to immunize mice (Charles River) and recombinant TcCTLP-6C was used to immunize rabbits (Pineda, Germany). If necessary, antibodies were purified using a HiTrap Protein A or G HP Column (GE Healthcare) following the manufacturer's recommendations.

### 2.3.9. Peptidase activity assays

Chymotrypsin and trypsin activity assays were performed using the substrates N-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide (AAPF) or N $_{\alpha}$ -benzoyl-D,L-arginine 4-nitroanilide hydrochloride (BAPNA) following the protocols of Peterson *et al.* (1995). For enzymatic activity assays of midgut contents of *M. sexta* 5<sup>th</sup> instar larvae, gut contents of two larvae were prepared as described above (see 2.3.1) and centrifuged for 30 min at 20000 x g and 4°C. Cleared supernatants were diluted in activity buffer (50 mM TRIS-HCl, 150 mM NaCl, pH 10.5) either 1:1000 in case of chymotrypsin or 1:100 in case of trypsin activity assays. Activity assays of purified midgut proteins and recombinant peptidases were carried out in activity buffer (pH 8.5 or pH 10.5) using 0.1 µg/µl of either the respective peptidase or bovine chymotrypsin (Sigma-Aldrich) for control reactions. For the proteolytic activation of recombinant precursor peptidases the reaction was supplied with 0.01 µg recombinant trypsin (proteomics grade; Roche). Following the mixture with 5 µl 100 mM AAPF resuspended in Dimethylformamide (DMF) or 5 µl 10 mM BAPNA resuspended in Ethylene glycol methyl ether (EGME) the reactions were carried out in a total volume of 150 µl. The samples were incubated at RT for 60 minutes while simultaneously reading the absorbance values at 410 nm. On the basis of the absorbance line slope and the protein concentration of the samples, the activity was calculated based on the respective extinction coefficients.

Alternatively to the photometric assays, the digestive activity of a sample was determined by its ability to digest BSA in a time dependent manner. Samples were prepared by mixing gut contents in a 1:10 ratio with TBS (100 mM TRIS-HCl, 125 mM NaCl, pH 10.0) or by mixing anterior midgut tissue homogenates in a 1:10 ratio with MOPS buffered saline (MBS; 100 mM MOPS, 125 mM NaCl, pH 6.5) with the addition of 0.5 % (w/v) soybean trypsin inhibitor. The samples were used to digest 5% (w/v) (BSA) in a total reaction volume of 100 µl. Samples that were heat inactivated for 15 min served as a control. Reactions were carried out applying a time series from 0 to 15 min at 4°C and the samples were subsequently separated by SDS-Page. To test the activity of the recombinantly expressed MsCTLPI, the peptidase was purified (see 2.3.7) and used at a concentration of 75 ng/µl to digest 5% (w/v) BSA. The reaction was carried out in total volume of 100 µl in a CAPS buffered saline (20 mM CAPS, 140 mM KCl, pH 11.0) and incubated for 2 h at 30°C. As a positive control, bovine chymotrypsin was used at a concentration of 75 ng/µl to digest 5% (w/v) BSA in TBS (20 mM TRIS-HCl, 140 mM KCl, pH 8.0) and incubated for 2 h at 30°C.

## 2.4. Other methods

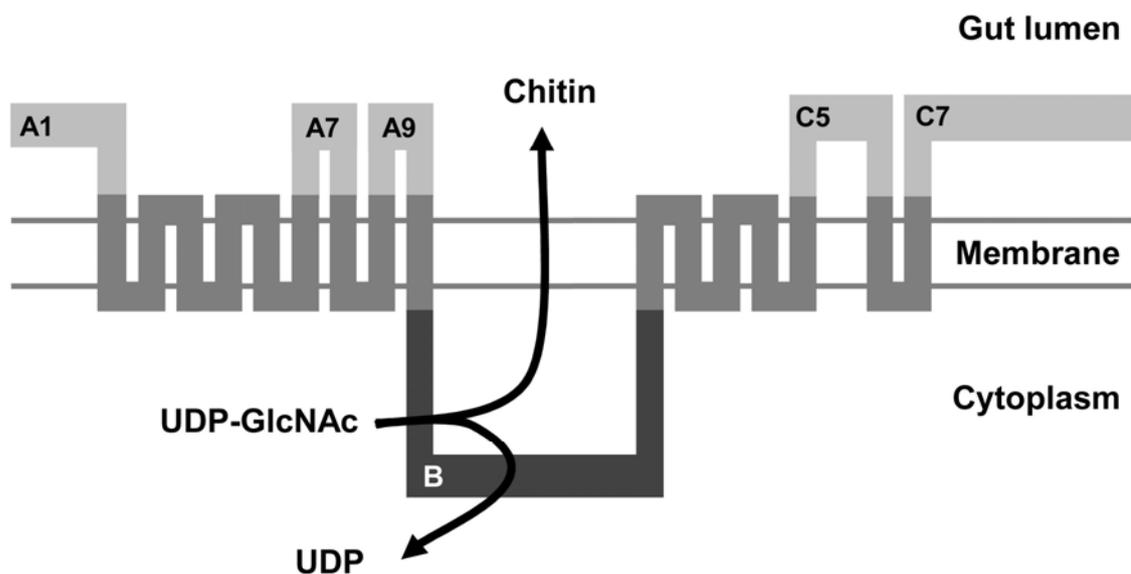
Sequencing was performed by SEQLAB, Göttingen or the Sequencing Core Facility, Dept. of Plant Pathology, Kansas State University. Alignment of amino acid sequences was performed with ClustalW (Higgins, 1994), and tree calculation was done with MEGA 4.1 software (Tamura *et al.*, 2007).

### 3. Results

#### 3.1. Interaction between MsCTLP1 and MsCHS2

##### 3.1.1. Introduction

Chitin synthases of several fungal and insect species have been discussed to be synthesized as zymogens, which are activated by trypsin or other serine peptidases (Merzendorfer, 2006). To find potential interaction partners of the midgut specific chitin synthase 2 of *M. sexta* (MsCHS2; GenBank|AAX20091.1), a yeast two hybrid screening of the midgut transcriptome was performed using the soluble domains B, C5 and C7 of MsCHS2 as a bait (Fig. 11; Broehan *et al.*, 2007). Screening for putative interaction partners of the catalytic B domain and the extracellular C5 domain did not reveal positive colonies whereas an interaction was detected for the C7 domain. The plasmids from positive yeast colonies were isolated and sequence analysis revealed the carboxy-terminal half of a previously unidentified serine peptidase. As BLAST and ClustalW (NCBI/EBI) analysis revealed highest similarities to insect chymotrypsins (Fig. 13), the putative protein was termed *M. sexta* chymotrypsin-like peptidase 1 (MsCTLP1). Two-hybrid tests using MsCTLP1 as the bait and MsCHS2-C7 as the pray as well as the converse experiment yielded positive colonies on selective media suggesting that the carboxy-terminal, extracellular C7 domain of MsCHS2 interacts with the carboxy-terminal half of MsCTLP1 in yeast.



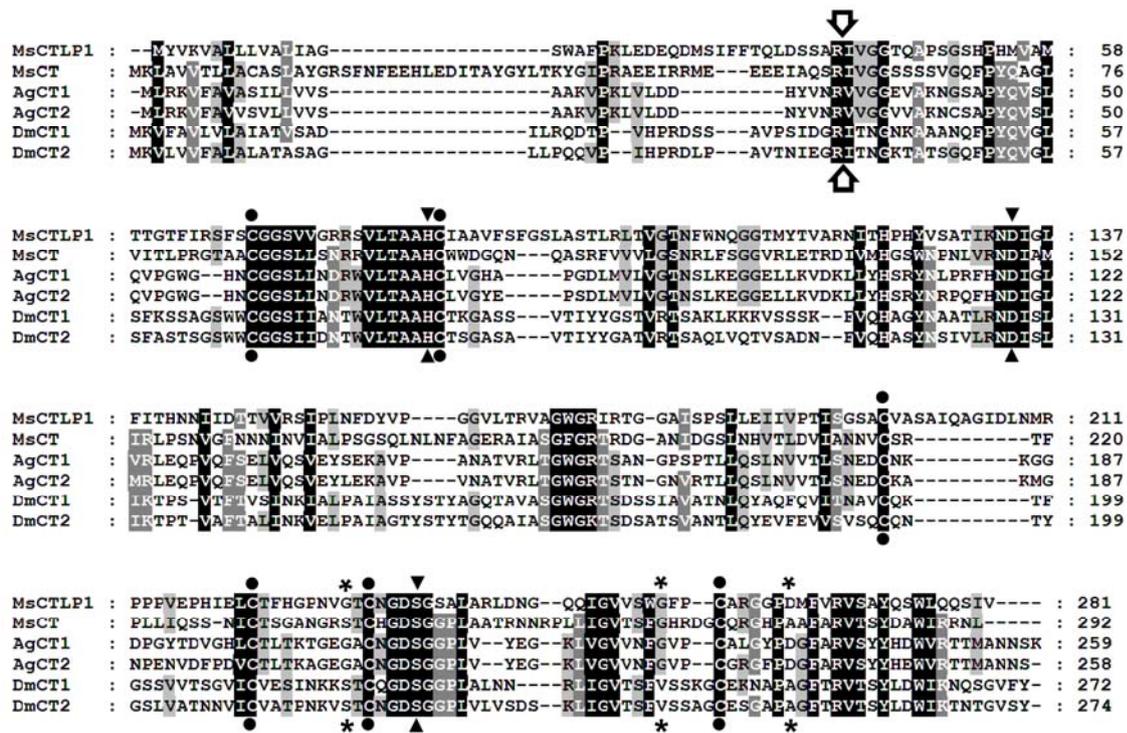
**Fig. 11.** Domain architecture of the midgut chitin synthase MsCHS2. The two horizontal lines represent the apical brush border membrane of midgut columnar cells. Horizontal bars represent intra- or extracellular domains and vertical bars transmembrane helices. (A) Aminoterminal transmembrane region; (B) cytoplasmic catalytic domain; (C) carboxy-terminal transmembrane region; the numbers refer to single domains within each region. The catalytic B domain as well as the carboxy-terminal domains C5 and C7 were used as baits in the yeast two-hybrid system. UDP, uridine 5'-diphosphate; GlcNAc, *N*-acetyl-D-glucosamine.

### 3.1.2. Isolation and sequencing of the cDNA encoding MsCTLP1

The plasmid isolated from positive yeast colonies contained a cDNAs of 409 base pairs (bps) in length, which encoded 109 amino acids of the carboxy-terminal half of MsCTLP1 and the complete 3' untranslated region, but lacked the amino-terminal end and the 5' untranslated region. Completion of the cDNA sequence was performed by 5' RACE. The *MsCTLP1* cDNA comprises 939 nucleotides encoding an open reading frame of 281 amino acids, yielding a soluble protein with a deduced molecular mass of 29.8 kDa (Fig. 12). BLAST and ClustalW analysis revealed highest similarities to unassigned insect S1A peptidases (serine peptidases of the chymotrypsin family) of the MEROPS peptidase database (Fig. 13). The deduced amino acid sequence of *MsCTLP1* showed the highest similarity to two *Anopheles gambiae* chymotrypsins (AgCT1 and AgCT2; emb|CAA79325.1 and emb|CAA79326.1), exhibiting 33% identical amino acids. The similarity to a previously reported *Manduca* chymotrypsin (MsCT, emb|AA58743.1; Peterson *et al.*, 1995) was somewhat lower, with 27% identical amino acids, and was thus comparable to that of two *Drosophila* chymotrypsin-like peptidases (*DmCT1* and *DmCT2*; emb|AAF50717.2 and emb|AAM12248.1). MsCTLP1 meets all essential structural requirements of digestive serine peptidases, as it possesses conserved histidine, aspartate and serine residues forming the catalytic triad (Kraut, 1977; Law *et al.*, 1977). Moreover, it contains six cysteine residues typically present in invertebrate serine peptidases (Fig. 13; Yan *et al.*, 2001). The presence of the three amino acids glycine, glycine and aspartate at conserved positions of the primary specificity pocket suggests that MsCTLP1 may exhibit a chymotrypsin-like substrate specificity with a glycine at the bottom of the pocket as the primary determinant (Perona and Craik, 1997). However, the occurrence of an aspartate residue at the side of the pocket may alter the specificity of MsCTLP1 significantly. MsCTLP1 is a secretory protein, since the premature form contains a signal peptide, which is predicted to be cleaved at the carboxy-terminal end of the amino acid at position 18 (Bendtsen *et al.*, 2004). Furthermore, like all members of this peptidase family, MsCTLP1 appears to be a zymogen proteolytically activated by tryptic cleavage between amino acid positions 40 and 41 (Figs. 12 and 13; Kraut, 1977; Law *et al.*, 1977; Lehane and Billingsley, 1996; Peterson *et al.*, 1995).

1	5'	TACTTCGGGCTACACAATGTACGTGAAAGTAGCACTTCTGTTGGTAGCCCT	51
		<i>M Y V K V A L L L V A L</i>	12
52		CATTGCTGGGAGCTGGGCCTTCCCAAAGCTCGAAGATGAGCAGGACATGTC	102
		<i>I A G S W A F P K L E D E Q D M S</i>	29
103		CATCTTCTTCACGCAGCTCGATTTCGAGCGCGGTATCGTGGGTGGTACCCA	153
		<i>I F F T Q L D S S A R I V G G T Q</i>	46
154		GGCCCCCAGCGGAAGTCACCCTCACATGGTGGCGATGACCACCGGTACCTT	204
		<b>A P S G S H P H M V A M T T G T F</b>	63
205		CATCAGGAGCTTCAGCTGTGGAGGCTCAGTTGTCGGTAGACGTTCCGTTCT	255
		<b>I R S F S C G G S V V G R R S V L</b>	80
256		GACTGCGGCTCATTGCATCGCTGCTGTTTTTCAGTTTCGGTTCCCTCGCCAG	306
		<b>T A A H C I A A V F S F G S L A S</b>	97
307		TACCTCCGCTTGACGGTCGGCACCAACTTCTGGAACCAGGGAGGCACCAT	357
		<b>T L R L T V G T N F W N Q G G T M</b>	114
358		GTACACCGTCGCTCGCAACATAACCCACCCCCACTACGTCTCTGCGACCAT	408
		<b>Y T V A R N I T H P H Y V S A T I</b>	131
409		CAAGAACGACATCGGTCTGTTCATCACTCACAACAACATCATCGACACGAC	459
		<b>K N D I G L F I T H N N I I D T T</b>	148
460		TGTCGTCCGCAGCATCCCTCTTAACTTTGACTATGTGCCCGGTGGTGTCT	510
		<b>V V R S I P L N F D Y V P G G V L</b>	165
511		CACTAGAGTCGCCGGATGGGGCAGGATCAGGACCGGCGGTGCCATCTCTCC	561
		<b>T R V A G W G R I R T G G A I S P</b>	182
562		CTCTCTGCTGGAGATCATTGTGCCTACTATCAGTGGAAGCGCATGCGTAGC	612
		<b>S L L E I I V P T I S G S A C V A</b>	199
613		CAGTGCAATCCAAGCTGGCATCGATCTGAACATGAGACCACCTCCCGTCGA	663
		<b>S A I Q A G I D L N M R P P P V E</b>	216
664		GCCTCACATCGAGCTGTGCACCTTCCACGGTCCCTAACGTAGGCACTTGTA	714
		<b>P H I E L C T F H G P N V G T C N</b>	233
715		TGGTGACTCCGGCAGCGCTCTTGCCCGCTAGACAACGGCCAGCAGATCGG	765
		<b>G D S G S A L A R L D N G Q Q I G</b>	250
766		TGTGGTATCGTGGGGCTTCCCGTGCGCACGCGGCGGTCCCGACATGTTTCT	816
		<b>V V S W G F P C A R G G P D M F V</b>	267
817		CAGGGTCAGCGCCTACCAATCCTGGCTGCAGCAGAGCATCGTATAATCTCC	867
		<b>R V S A Y Q S W L Q Q S I V *</b>	281
868		GAACCGAACCTGATCTTAACTGTAAATAAAAAAAGCGATACCCAAAAAAA	918
919		AAAAAAAAAAAAAAAAAAAAA 3'	939

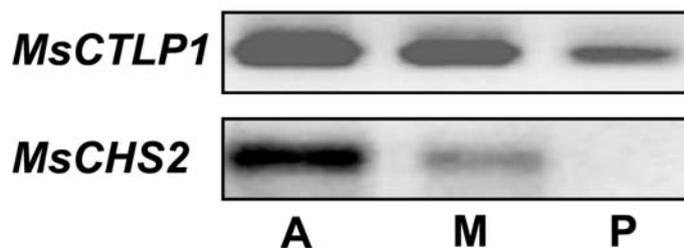
**Fig. 12.** Primary structure of MsCTPL1. cDNA sequence and deduced amino acid sequence (standard: zymogen; bold: mature peptidase) of the *M. sexta* CTPL1. Italic characters indicate the 5' and 3' untranslated regions of the cDNA and the putative signal peptide for the secretory pathway of the protein. Open arrow marks the predicted cleavage site of the signal peptidase and the closed arrow marks the putative trypsin cleavage site. Red letters indicate the putative *N*-glycosylation site for MsCTPL1.



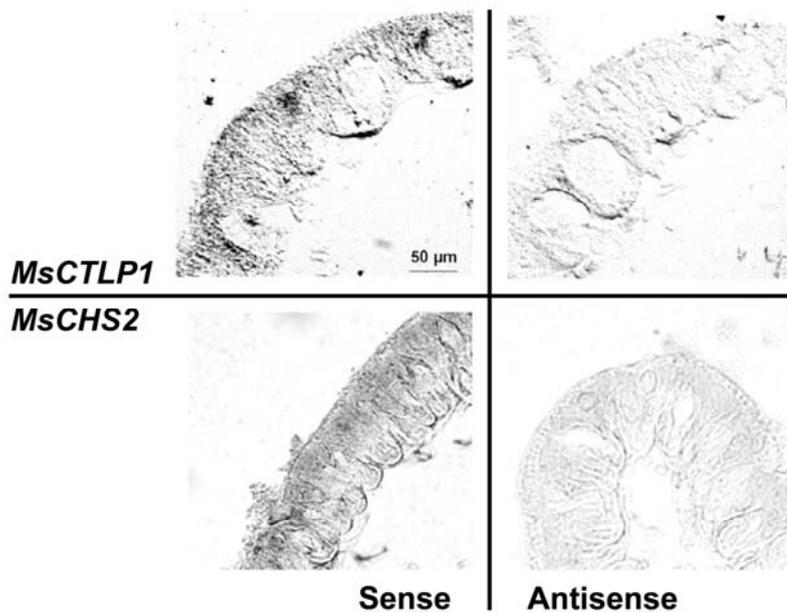
**Fig. 13.** ClustalW alignment of insect chymotrypsin-like peptidases. Amino acids that are conserved, highly conserved or identical in all sequences are highlighted in light gray, gray or black, respectively. Open arrow, conserved trypsin cleavage site; asterisks, conserved residues of the S1 specificity pocket; filled triangles, residues of the catalytic triad of serine peptidases; dots, conserved cysteines. Ms, *Manduca sexta*; Ag, *Anopheles gambiae*; Dm, *Drosophila melanogaster*; CT, chymotrypsin.

### 3.1.3. Localization of MsCTLP1 and MsCHS2

To determine the localization of MsCTLP1 and MsCHS2 expression, RT-PCR and *in situ* hybridization was performed. RT-PCR indicated that both genes are mainly expressed in the anterior and median midgut of fifth instar larvae (Fig. 14). *In situ* hybridization showed, moreover, that *MsCTLP1* and *MsCHS2* transcripts are co-localized at the apical half of midgut columnar cells (Fig. 15). As the C7 domain of MsCHS2 faces the extracellular space, MsCTLP1 has to be secreted by the columnar cells in order to reach and bind to the C7 domain.



**Fig. 14.** Expression of *MsCTLP1* and *MsCHS2* in the midgut of *M. sexta*. Total RNA from the anterior (A), median (M) and posterior (P) midgut was reverse transcribed and used as a template for PCR (25cycles) to amplify the cDNAs encoding *MsCTLP1* and *MsCHS2*.

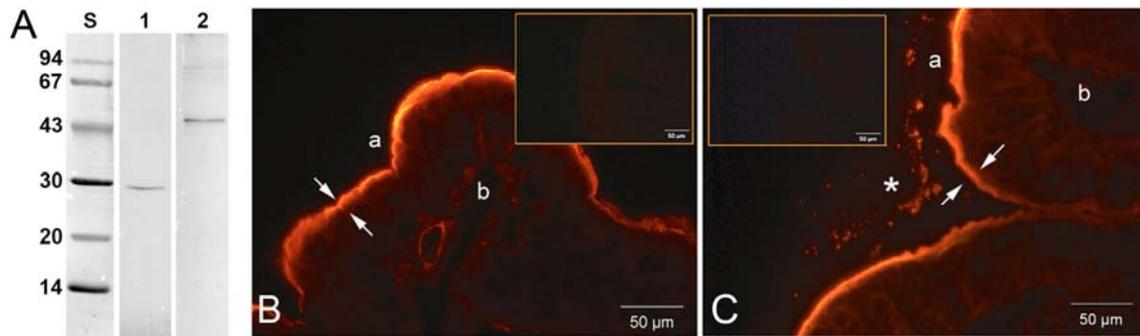


**Fig. 15.** Localization of *MsCTLP1* and *MsCHS2* mRNAs in the anterior midgut of *M. sexta*. *In situ* hybridization was performed under high stringency conditions using 20 µm cryosections of the anterior midgut and antisense RNA probes to detect the sense RNAs for *MsCTLP1* and *MsCHS2* (left). As a negative control, sense RNA probes were used to detect the corresponding antisense RNAs (right).

### 3.1.4. Immunological detection of MsCTLP1

To analyse the localization of MsCTLP1 at the protein level, monospecific antibodies were generated. The antibody was purified using a Protein A column (GE Healthcare) and used to stain proteins from midgut contents and tissue extracts transferred to nitrocellulose membranes. In the gut contents, a single protein of about 30 kDa was detected indicating that MsCTLP1 is secreted (Fig. 16 A, lane 1). However, the apparent molecular mass was in this case higher than expected masses for the secreted (27.9 kDa) or the activated peptidase (25.3 kDa). Since only a single band could be detected in Western blot, MsCTLP1 is probably proteolytically processed (Fig. 16 A, lane 1), as is well established for zymogenic chymotrypsins (Kraut, 1977). In tissue extracts of the anterior midgut the antibodies detected a single protein band (Fig. 16 A, lane 2), which represents the zymogenic form of MsCTLP1, as it exhibits a significantly higher molecular mass than the mature peptidase. In this case, the apparent molecular mass was also higher than the theoretical molecular mass of 29.8 kDa.

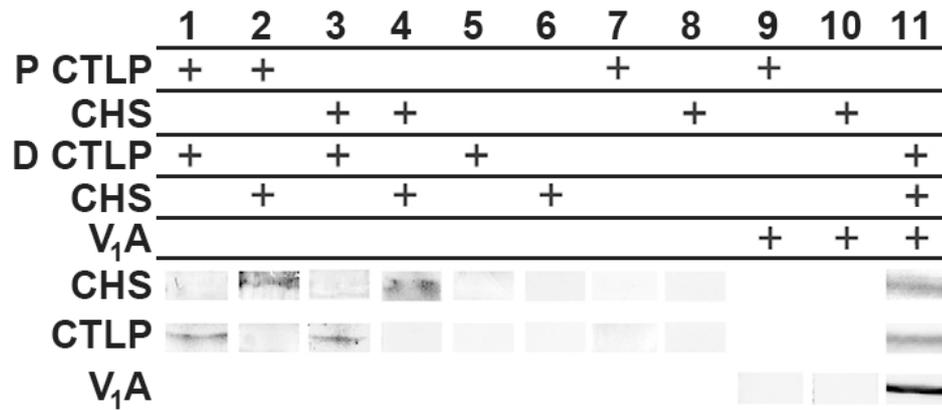
When analyzing the distribution of MsCTLP1 and MsCHS2 in cryosections of the anterior midgut from fifth instar larvae using the anti-MsCTLP1 and anti-MsCHS antibodies (Zimoch and Merzendorfer, 2002), both proteins were detected at the brush border microvilli formed by columnar cells (Figs. 16 B and C, arrows). In contrast to MsCHS2, which is a membrane-integral protein, MsCTLP1 is also detected within the gut contents, supporting the idea that MsCTLP1 is secreted into the midgut lumen (Fig. 16 C, asterisk).



**Fig. 16.** (A) Immunoblots demonstrating that MsCTLP1 is secreted and proteolytically processed. Proteins from the gut contents (1) and the anterior midgut (2) were separated by SDS-PAGE, blotted onto nitrocellulose and stained with monospecific antibodies to MsCTLP1. (B+C) Immunolocalization of MsCHS2 and MsCTLP1 in the midgut of *M. sexta*. Midgut cryosections of 10  $\mu\text{m}$  thickness from fifth instar larvae were stained with either anti-MsCHS antibodies (B) directed to the recombinant catalytic domain of chitin synthase (Zimoch and Merzendorfer, 2002) or anti-MsCTLP1 antibodies (C). Arrows mark the apical brush border of the midgut epithelium, the asterisk marks immunoreactive material in the gut lumen. The insets show corresponding negative controls performed in the absence of primary antibodies. a, apical; b, basal.

### 3.1.5. Co-immunoprecipitation reveals that MsCTLP1 interacts with MsCHS2

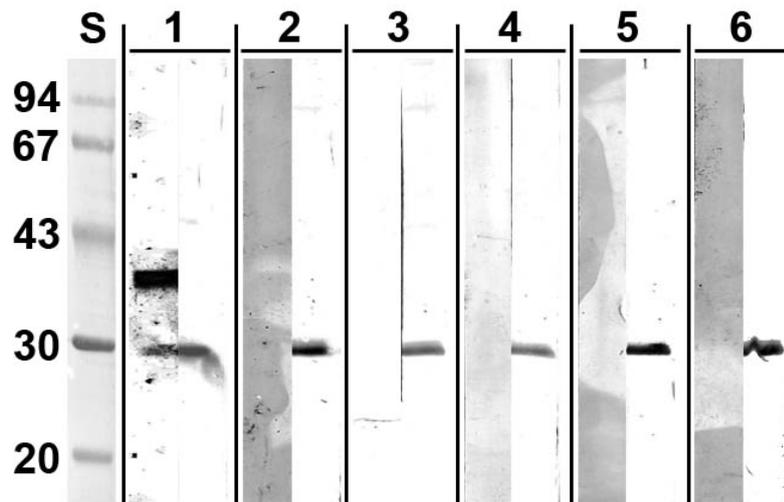
To test whether MsCTLP1 and MsCHS2 interact *in vitro*, co-immunoprecipitation assays were performed, taking advantage of the availability of monospecific antibodies to both proteins (Fig. 17). Therefore anterior midgut tissues were homogenized, incubated with the precipitating antibodies and subsequently bound to the protein-G-agarose. After several wash steps, bound proteins were eluted by cooking and analyzed by SDS-Page. After transfer to nitrocellulose the membranes were stained with the detecting antibodies. In contrast to negative control reactions in the absence of precipitating antibodies (Fig. 17, lanes 5 and 6), co-immunoprecipitation of MsCHS2 was observed when the anti-MsCTLP1 antibodies were added to midgut cell lysates (Fig. 17, lanes 1 and 2). When the anti-MsCHS antibodies were used for immunoprecipitation, co-immunoprecipitation of MsCTLP1 (Fig. 17, lanes 3 and 4) could be observed. As V-ATPases are expressed in the midgut of the tobacco hornworm in exceptionally high densities (Beyenbach and Wiczorek, 2006), the antibodies to the A subunit of the V-ATPase ( $V_1A$ ) were used as a control for non-specific precipitation by the anti-MsCHS and anti-MsCTLP1 antibodies. As shown in Fig. 17 (lanes 9 and 10), V-ATPase A subunits were not co-immunoprecipitated by either anti-MsCHS or anti-MsCTLP1 antibodies.



**Fig. 17.** Co-immunoprecipitation of MsCTLP1 and MsCHS2. For immunoprecipitation, cell lysates of the anterior midgut were incubated with the indicated precipitating antibodies (P) and then bound to protein-G-agarose. Unbound proteins were washed away, then bound proteins were eluted, separated by SDS-PAGE and analysed by immunoblotting using detecting antibodies (D) to *Manduca* CTLP1, CHS2 and V-ATPase subunit A (V<sub>1</sub>A). Lanes 1–4, co-immunoprecipitation of MsCHS2 using anti-MsCTLP1 antibodies (lanes 1,2) and of MsCTLP1 using anti-MsCHS antibodies (lanes 3,4); lanes 5–8, control reactions in the absence of precipitating (lanes 5,6) or detecting antibodies (lanes 7,8); lanes 9,10, control reactions for non-specific precipitation of V-ATPase. As a positive control, midgut cell lysates were used (lane 11).

### 3.2. MsCTLP1 is a glycoprotein

The finding that secreted MsCTLP1 shows a decreased motility in SDS-PAGE might be explained by *N*-glycosylation, as MsCTLP1 exhibits the putative *N*-glycosylation site NITH at amino acid positions 120–123 (Fig. 12). To test this hypothesis, detection of glycoproteins in midgut contents was performed by using the DIG Glycan Differentiation Kit (Fig. 18). After separation by SDS Page, proteins of the gut content were transferred to nitrocellulose. Each single lane was separated in to equal halves and stained either with the anti-MsCTLP1 antibody or with the lectins contained in the kit. Only for the *Galanthus nivalis* agglutinin (GNA), a signal at the same height as the MsCTLP1 immunosignal could be detected (Fig. 18, lane 1). This finding suggests that the proteolytically activated peptidase is most likely *N*-glycosylated since the positive reaction with GNA confirms the presence of *N*-glycosidically-linked high mannose or hybrid-type carbohydrate chains. Missing of a signal for SNA and MAA (Fig. 18, lanes 2, 3 and 6) indicates that the sugar moiety of MsCTLP1 lacks sialic acids, which are terminally linked (2–3) or (2–6) to galactose or *N*-acetylgalactosamine. Furthermore, the absence of a signal for PNA (Fig. 18, lane 4) shows that MsCTLP1 is not *O*-glycosylated since PNA detects the core disaccharide galactose- $\beta$ (1–3)-*N*-acetylgalactosamine of most *O*-glycans. Finally the sugar chain of MsCTLP1 contains no galactose- $\beta$ (1–4)-*N*-acetylglucosamine disaccharides as implied by the missing of the signal for DSA. The strong signal additionally detected by GNA at about 40 kDa possibly reflects a different dietary or midgut content protein (Fig. 18, lane 1).

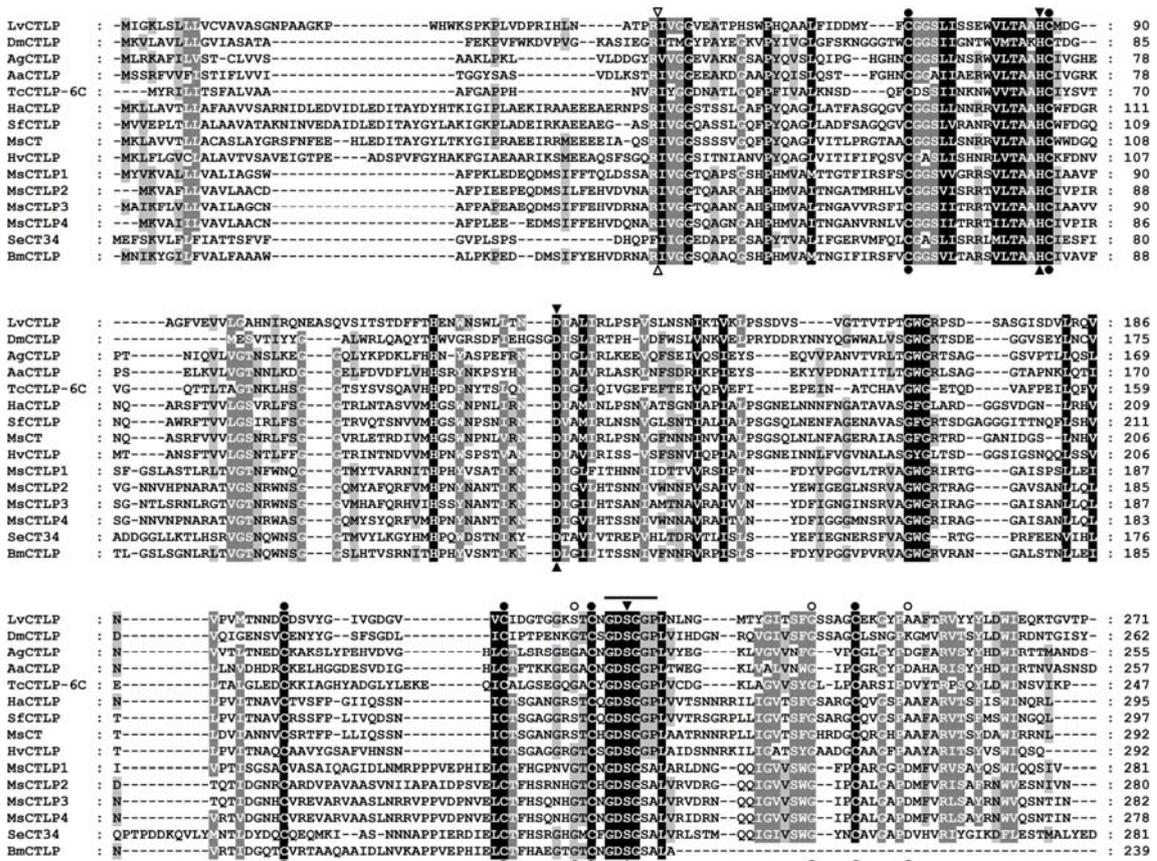


**Fig. 18.** Immunoblots to test if MsCTLP1 is glycosylated. Proteins from the gut contents (1-6) were separated by SDS-PAGE and blotted onto nitrocellulose. The lanes were separated in equal halves and stained either with different lectins (left half; 1 GNA; 2 SNA; 3 MAA; 4 PNA; 5 DSA; 6 SNA and MAA) or monospecific antibodies to MsCTLP1 (right half). Standard proteins with molecular masses indicated in kDa are also shown (S).

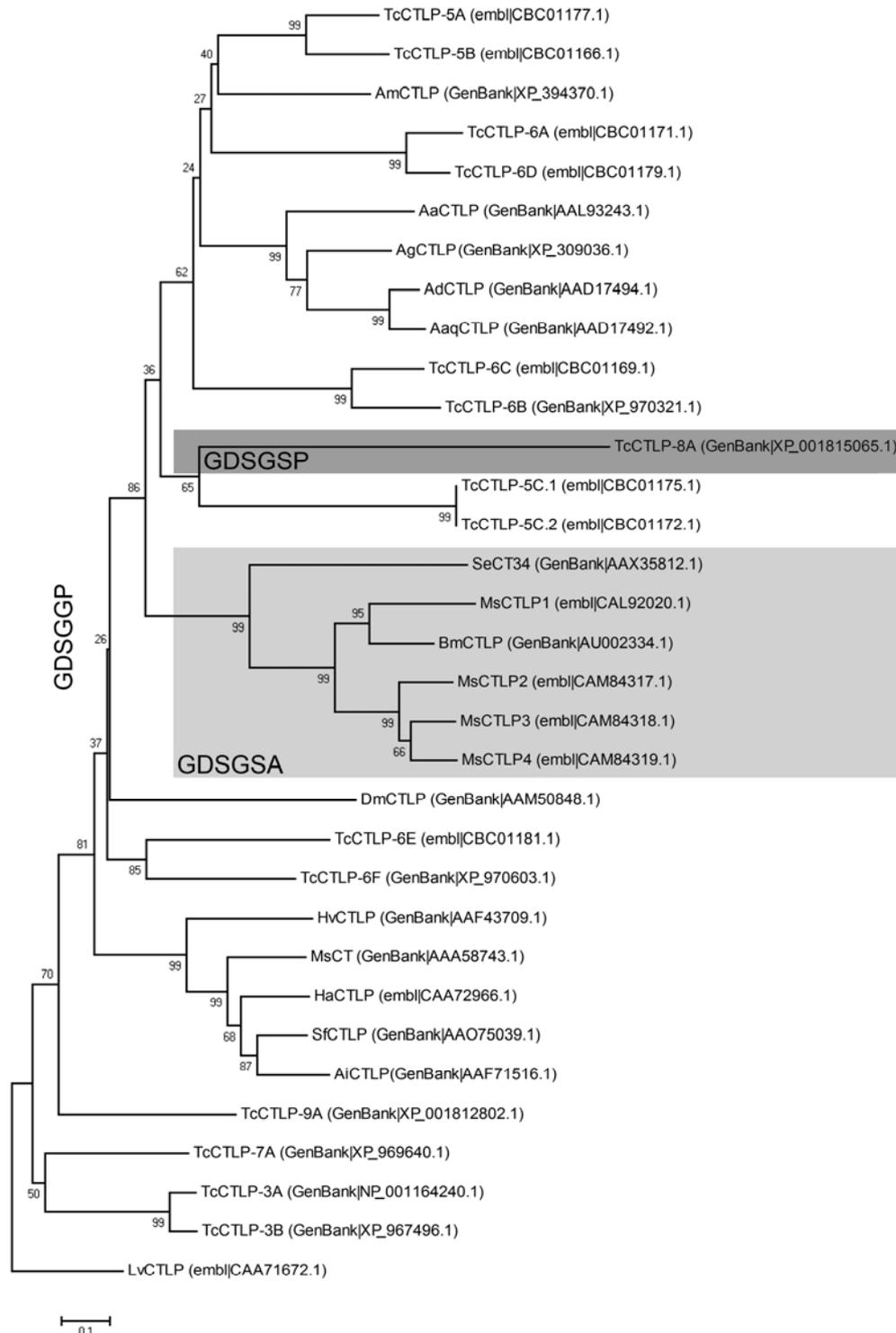
### 3.3. Midgut MsCTLPs are encoded by a multigene family

In order to identify new CTLPs in *Manduca*, a midgut cDNA library was screened and RT-PCR based strategies were performed using degenerate primers. Eventually, three different cDNAs encoding closely related *MsCTLPs* were isolated and termed *MsCTLP2-4*. The cDNAs of the three newly discovered *MsCTLPs* shared high similarity among each other (*MsCTLP2* vs. *MsCTLP3*: identity = 73%, *MsCTLP2* vs. *MsCTLP4*: identity = 78%, *MsCTLP3* vs. *MsCTLP4*: identity = 81%) and were most similar to *MsCTLP1* (*MsCTLP1* vs. *MsCTLP2-4*: identities = 53%, 58% and 55%, respectively). The only other known *Manduca* chymotrypsin (MsCT; Peterson *et al.*, 1995) was less closely related to *MsCTLP2-4* (*MsCT* vs. *MsCTLP2-4*: identities = 30%, 30% and 31%, respectively). A ClustalW alignment revealed high similarities between the amino acid sequences of MsCT and MsCTLP1-4 and chymotrypsins from different crustacean and insect species (Fig. 19). As described above for MsCTLP1 also the newly discovered MsCTLPs meet the structural requirements of digestive serine peptidases. Like MsCTLP1, MsCTLP2-4 do contain the conserved amino acids Gly189, Gly216 and Asp226, which are believed to determine substrate specificity. The first of the three amino acids is regarded to be the primary determinate located at the pocket's bottom (Fig. 19; Perona and Craik, 1997) numbering according to Greer (1990). In contrast, the specificity pocket of MsCT contains the amino acids Ser189, Gly216 and Ala226, which is frequently found in vertebrate chymotrypsins (Peterson *et al.*, 1995). Like all members of the serine peptidase family the *Manduca* CTLPs seem to be zymogens activated upon tryptic cleavage as they possess the highly conserved tryptic cleavage site RIVVG (Fig. 19; Lehane and Billingsley, 1996; Peterson *et al.*, 1995). Furthermore, all MsCTLPs seem to be secretory proteins, since the premature forms contain signal peptides for the secretory pathway, which are predicted to be

cleaved carboxy-terminal of amino acid positions 18 for MsCTLP3, and 16 for MsCTLP2 and MsCTLP4 (Bendtsen *et al.*, 2004). Construction of a phylogenetic tree for chymotrypsin-like serine peptidases from various species reveals well-defined clusters reflecting evolutionary distances (Fig. 20). Supported by a very high bootstrap value, MsCTLP1–4 fit perfectly into a novel subgroup of insect CTLPs as reported previously by Herrero *et al.* (2005). Next to sequence similarity this subgroup is defined by the GP→SA replacement in the highly conserved GDSGGP motif, which is also found in the chymotrypsins from *Spodoptera exigua* (SeCT34, emb|AAX35812.1; Herrero *et al.*, 2005) and *Bombyx mori* (GenBank|AU002334).



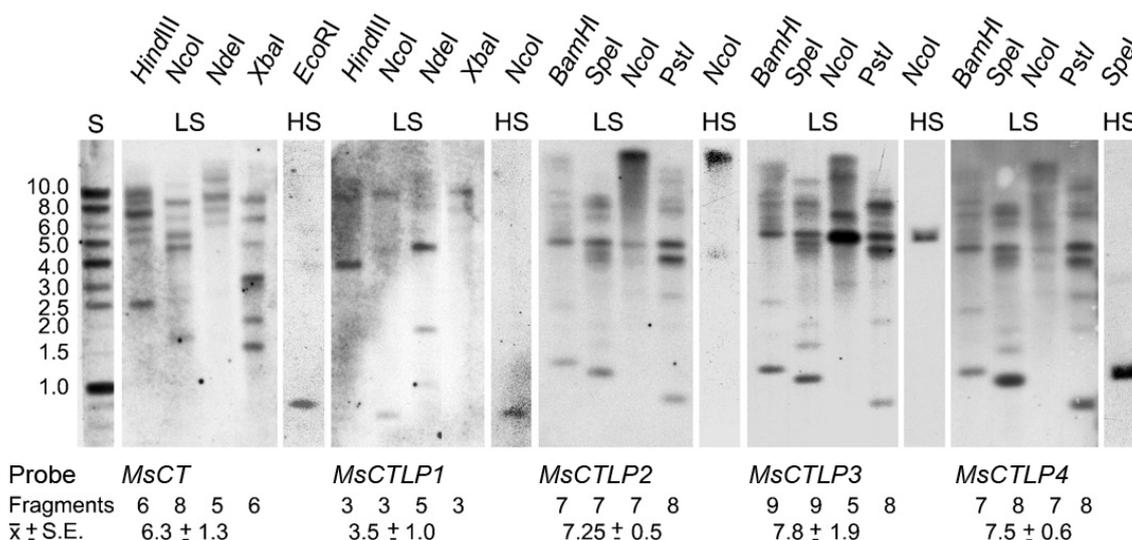
**Fig. 19.** ClustalW alignment of chymotrypsin-like peptidases from various species. Conserved, highly conserved or identical amino acids are highlighted in light-gray, gray or black, respectively. Open triangle, conserved trypsin cleavage site; open circle, conserved residues within the S1 specificity pocket; closed circle, conserved cysteines; closed triangles, conserved residues of the catalytic triad of serine peptidases; black line, highly conserved GDSGGP motif. CT, chymotrypsin; CTLP, chymotrypsin-like peptidase; Lv, *Litopenaeus vannaemey* (emb|CAA71672.1); Dm, *Drosophila melanogaster* (GenBank|AAM50848.1); Ag, *Anopheles gambiae* (GenBank|XP\_309036.1); Aa, *Aedes aegypti* (GenBank|AAL93243.1); Tc, *Tribolium castaneum* (emb|CBC01169.1); Ha, *Helicoverpa armigera* (emb|CAA72966.1); Sf, *Spodoptera frugiperda* (GenBank|AAO75039.1); Ms, *Manduca sexta* (MsCT: GenBank|AAA58743.1, MsCTLP1: emb|CAL92020.1, MsCTLP2: emb|CAM84317.1, MsCTLP3: emb|CAM84318.1, MsCTLP4: emb|CAM84319.1); Hv, *Heliothis virescens* (GenBank|AAF43709.1); SeCT34, *Spodoptera exigua* (GenBank|AAX35812.1); Bm, *Bombyx mori* (GenBank|AU002334).



**Fig. 20.** Phylogenetic tree for insect chymotrypsin-like peptidases. The neighbourhood joining tree was calculated on the basis of a ClustalW alignment (Blosom62) of CTLPs from different insect species (accession numbers are given in parentheses) and rooted with the amino acid sequence of chymotrypsin B from the crustacean *Litopenaeus vannamei*. The scale bare indicates an evolutionary distance of 0.1 amino acid substitutions per site. Bootstrap values are given in percentages at the internodes. The GDSGSA and GDSGSP subgroups of insects CTLPs are highlighted with gray shading. The remaining CTLPs exhibit the highly conserved GDSGGP motif. CT, chymotrypsin; CTLP, chymotrypsinogen like peptidase; Lv, *Litopenaeus vannamei*; Hv, *Heliothis virescens*; Ms, *Manduca sexta*; Ha, *Helicoverpa armigera*; Sf, *Spodoptera frugiperda*; Ai, *Agrotis ipsilon*; Dm, *Drosophila melanogaster*; Am, *Apis mellifera*; Tc, *Tribolium castaneum*; Aa, *Aedes aegypti*; Ag, *Anopheles gambiae*; Ad, *Anopheles darlingi*; Aaq, *Anopheles aquasalis*; Se, *Spodoptera exigua*; Bm, *Bombyx mori*.

### 3.4. Estimation of gene copy numbers by Southern blotting

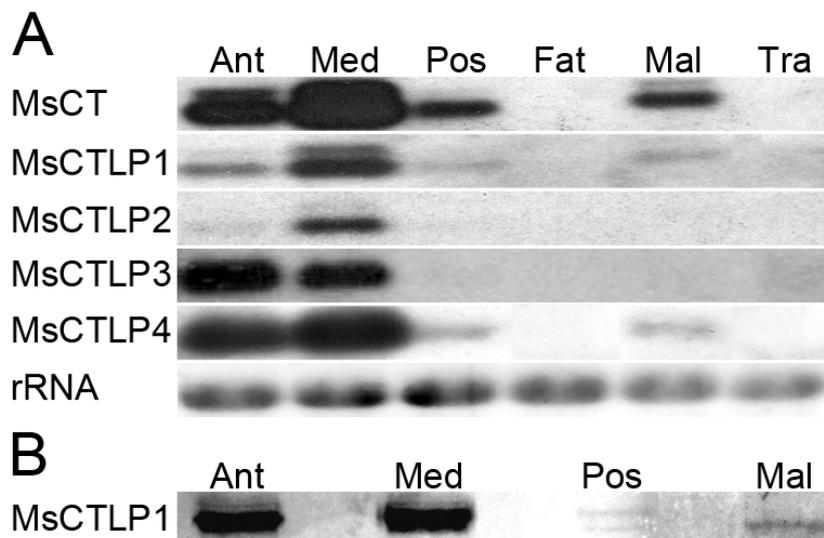
To evaluate the number of CTLP related genes in *Manduca*, Southern blot analysis was performed using DIG-labeled cDNA probes of equal length and CG content as well as genomic DNA isolated from total midguts of *M. sexta* 5<sup>th</sup> instar larvae. Genomic DNA was cut with four different restriction enzymes, separated by gel electrophoresis, blotted and finally hybridized with denatured DNA probes to the respective gene. High stringency hybridization (58°C, 50% formamide) led to the detection of single bands (Fig. 21, HS). At a lower hybridization stringency (52°C, no formamide present) averaged  $6.4 \pm 1.9$  ( $n = 20$ ) genomic DNA fragments were intensively stained with every of the probes used, suggesting that CTLPs identical or closely related to *MsCT* or *MsCTLP1-4* are encoded by up to eight genes in the *Manduca* genome (Fig. 21, LS). Although the detection of polymorphic genes cannot be excluded entirely, this seems to be rather unlikely, as the laboratory *Manduca* population is reared under inbreeding conditions since years. Weakly stained bands that are visible in some blots may represent genes encoding more distantly related serine peptidases.



**Fig. 21.** Southern blot hybridization of *Manduca* genomic DNA. In each case 5 mg of genomic DNA were digested with the indicated restriction enzyme and separated on an agarose gel. The DNA fragments were transferred onto a nylon membrane and hybridized with DIG-labeled DNA probes specific to the indicated genes encoding chymotrypsin-like peptidases from *M. sexta*. Low (LS) and high stringency (HS) hybridization was performed at temperatures of 52°C and 58°C, respectively. High stringency hybridization was performed in the presence of 50% formamide. Hybridized DNA fragments were detected with Anti DIG-AP antibodies and CSPD as a chemiluminescent substrate. The blots were exposed on a Kodak Biomax XAR X-ray film. The number of labeled bands per lane as well as the mean value  $\pm$  S.E. for each hybridization probe is given as specified. S, standard fragments with molecular masses indicated in kb.

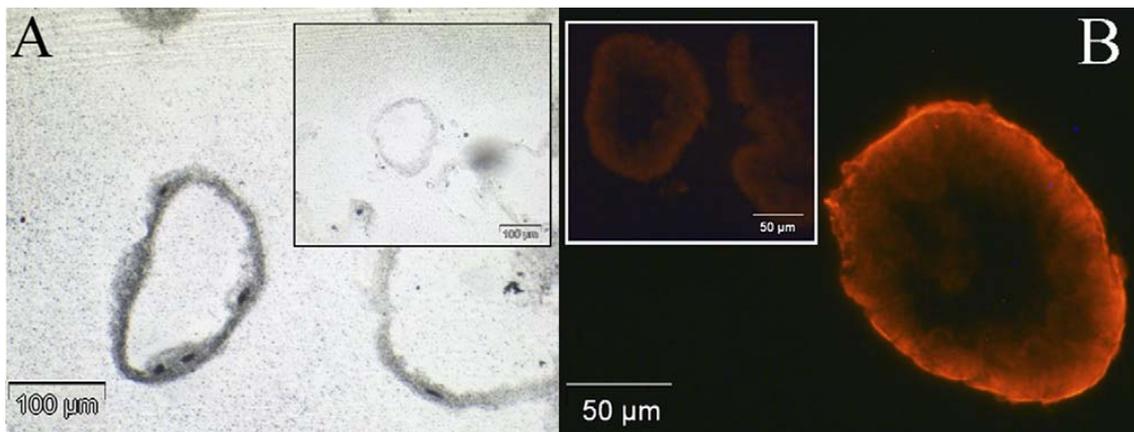
### 3.5. MsCTLPs are expressed in the midgut and the Malpighian tubules

The finding that multiple chymotrypsin-like peptidases are expressed in the midgut raises the question whether they differ in their biological functions. To get first hints on their potential physiological roles, mRNA expression was examined by Northern blots and *in situ* hybridization. Northern blot analysis revealed highest expression of *MsCT* and *MsCTLPs* in the median midgut and somewhat lower expression in the anterior midgut, except for *MsCTLP3*, which was highest in the anterior midgut and lower in the median midgut (Fig. 22 A). The weaker bands, which are additionally visible close to the strong bands in the Northern blots for *MsCT* and *MsCTLP1* in Fig. 22 A, are most likely the result from cross-hybridization with other *MsCTLP* mRNAs. *MsCT* and *MsCTLP1–4* expression was very low in the posterior midgut, if detectable at all. Surprisingly, significant expression of *MsCT*, *MsCTLP1* and *MsCTLP4* was found in Malpighian tubules. The detected bands were of the same size as those observed in the other tissues. The apparent higher molecular weight is the result of a smile-effect emerged during gel-electrophoresis. In contrast, *MsCTLP* expression could neither be detected in the tracheae nor in the fat body (Fig. 22 A), salivary glands, ganglia and epidermis (not shown).



**Fig. 22.** Northern blot and immunoblot analysis of *Manduca* CTLPs. (A) Northern blot analysis of CTLPs transcripts in different tissues of 5th instar larvae. Each 5 mg of total RNA isolated from anterior (Ant), median (Med) and posterior midgut (Pos) as well from the fatbody (Fat), Malpighian tubules (Mal) and the tracheae (Tra) were electrophoresed, blotted onto a nylon membrane and hybridized with a DIG-labeled DNA probes, which were complementary to a conserved region of the respective CTLP cDNA. As a loading control for gel electrophoresis rRNA was stained with Radiant Red. The 18S rRNA migrating at about 1.9 kb is shown. (B) Immunoblots of midgut and Malpighian tubule proteins using anti-*MsCTLP1* antibodies. Crude tissue extracts isolated from anterior (Ant), median (Med), posterior midguts (Pos) as well as from Malpighian tubules (Mal) were separated by SDS-PAGE and transferred to nitrocellulose membrane.

To confirm mRNA expression in the Malpighian tubules, *in situ* hybridization was performed with digoxigenin-labeled RNA probes specific to *MsCTLP1* (Fig. 23 A). In contrast to the sense RNA (Fig. 23 A, inset) the antisense RNA detected *MsCTLP1* mRNA in the cytoplasm of the Malpighian tubule epithelial cells (Fig. 23 A). To evaluate the expression at the level of proteins, immunoblots were performed with the anti-*MsCTLP1* antibody staining a single protein band predominantly in the anterior and median midgut (Fig. 22 B). However, the *MsCTLP1* peptide used for antibody generation shares some similarities with homologue regions from *MsCT* and *MsCTLP2–4* and cross reactions with other *Manduca* CTLPs cannot be entirely excluded. In line with the results from Northern blots and *in situ* hybridization, a *MsCTLP1* signal was detectable in Malpighian tubules, although in significantly lower amounts as in the anterior and median midgut. Finally, *MsCTLP1* expression was examined by immunocytochemistry staining cryosections of Malpighian tubules from fifth instar larvae with the anti-*MsCTLP1* antibodies. Again a *MsCTLP1* signal was detected in Malpighian tubule cells. The signal was strongest at the basal site facing the hemolymph (Fig. 23 B).

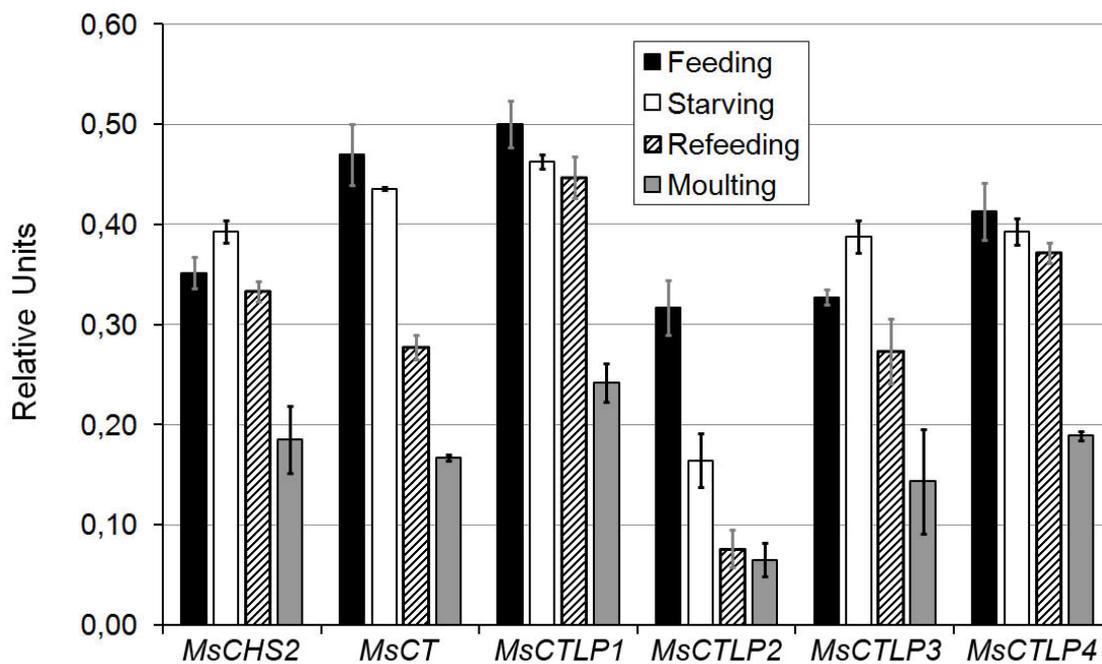


**Fig. 23.** *MsCTLP1* expression in Malpighian tubules. (A) *In situ* hybridization to detect *MsCTLP1* mRNA in the Malpighian tubules. *In situ* hybridization of 20 mm cryosections of the Malpighian tubules from fifth instar larvae was performed under high stringency conditions using a digoxigenin-labeled antisense RNA probe for *MsCTLP1* that was complementary to a conserved part of *MsCTLP1*. The inset shows the negative control performed with the corresponding digoxigenin-labeled *MsCTLP1* sense RNA probe. (B) Immunohistochemical labeling of *MsCTLPs* in the Malpighian tubules. 10 mm cryosections of Malpighian tubules from *M. sexta* fifth instar larvae were immuno-stained with the anti-*MsCTLP1* antibodies and visualized by the reaction with Cy3-conjugated anti-rabbit IgG. The inset shows the corresponding negative control performed with the corresponding pre-immune serum.

### 3.6. *MsCTLP* mRNA expression is not affected by starvation but down-regulated during molt

For some haematophagous insects such as *Anopheles gambiae* or *Aedes aegypti* it has been shown that the mRNA levels of early or late trypsins are transcriptionally up-regulated in response to a blood meal and down-regulated during periods of starvation (Barillas-Mury and Wells, 1993; Muller *et al.*, 1995). However, also more continuously feeding insects such as lepidopteran larvae undergo periods of starvation, either during

forced feeding breaks or during molt. To test whether the expression of midgut serine peptidases from *M. sexta* is regulated, Realtime PCR, Western blots and immunocytochemistry was performed comparing mRNA and protein amounts in the midgut epithelium from 5<sup>th</sup> instar larvae that were feeding, starved for 36 h, re-fed 8 h after 36 h of starvation or molting from the 4<sup>th</sup> to the 5<sup>th</sup> instar. Comparison of the CT-values obtained by Realtime PCR revealed no considerable differences in the amounts of mRNAs encoding *MsCTLP1* and *MsCTLP3–4* in the anterior midgut of feeding, starving or re-fed larvae (Fig. 24). The mRNA amounts for *MsCT* did not recover when the larvae were re-fed after starvation. Similarly, *MsCTLP2* mRNA amounts stayed relatively low upon re-feeding, but additionally dropped significantly upon starvation. In all cases, the mRNA amounts were significant lower during molting from the 4<sup>th</sup> to the 5<sup>th</sup> instar as it is also observed for chitin synthase 2 (Fig. 24). Similar results were obtained in a Realtime PCR study performed with mRNA preparations from the median midgut (data not shown). Consequently, mRNA amounts for *MsCT*, *MsCTLP1* and *MsCTLP3–4* in the larval midgut of *M. sexta* appear to be down-regulated by molting signals rather than by starving signals. By contrast, mRNA amounts for *MsCTLP2* appear to be also reduced by starving signals.

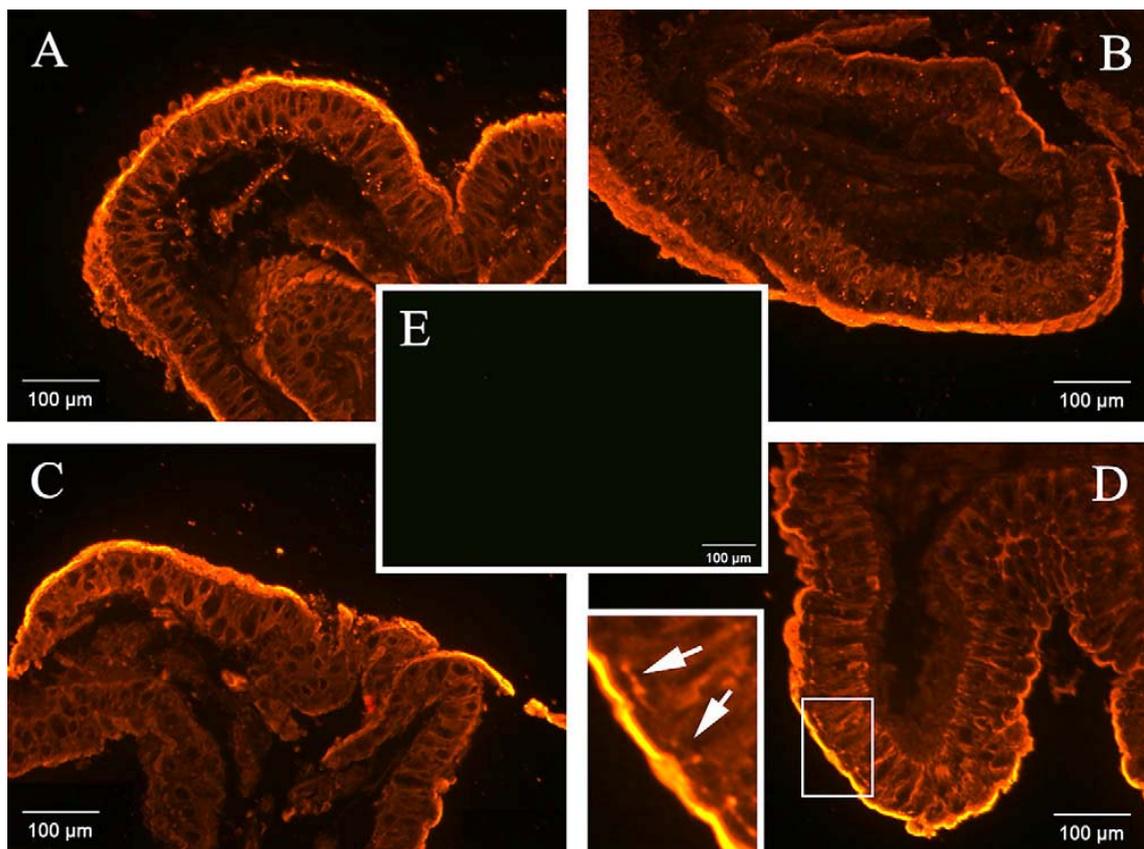


**Fig. 24.** *MsCTLP* expression in the anterior midgut tissues at different physiological states. Total RNA of anterior midguts was isolated from larvae facing different physiological conditions and the cDNA was synthesized by reverse transcription. Aliquots corresponding to 50 ng of total RNA were used as templates for quantitative Realtime PCR with 42 cycles. Black, feeding 5<sup>th</sup> instar larvae; white, 5<sup>th</sup> instar larvae after 36 h starvation; diagonal stripes, 36 h starved larvae re-fed for 8 h; gray, molting larvae (4<sup>th</sup> to 5<sup>th</sup> instar, molting stage F). Expression levels are given as relative mRNA amounts determined on the basis of the CT-values (mean values  $\pm$  S.E.,  $n = 4$ ). Normalization was performed with reactions amplifying the cDNA of the ribosomal protein S3 from *M. sexta*.

### 3.7. MsCTLP1 secretion is suspended during molt and starvation and stimulated upon feeding

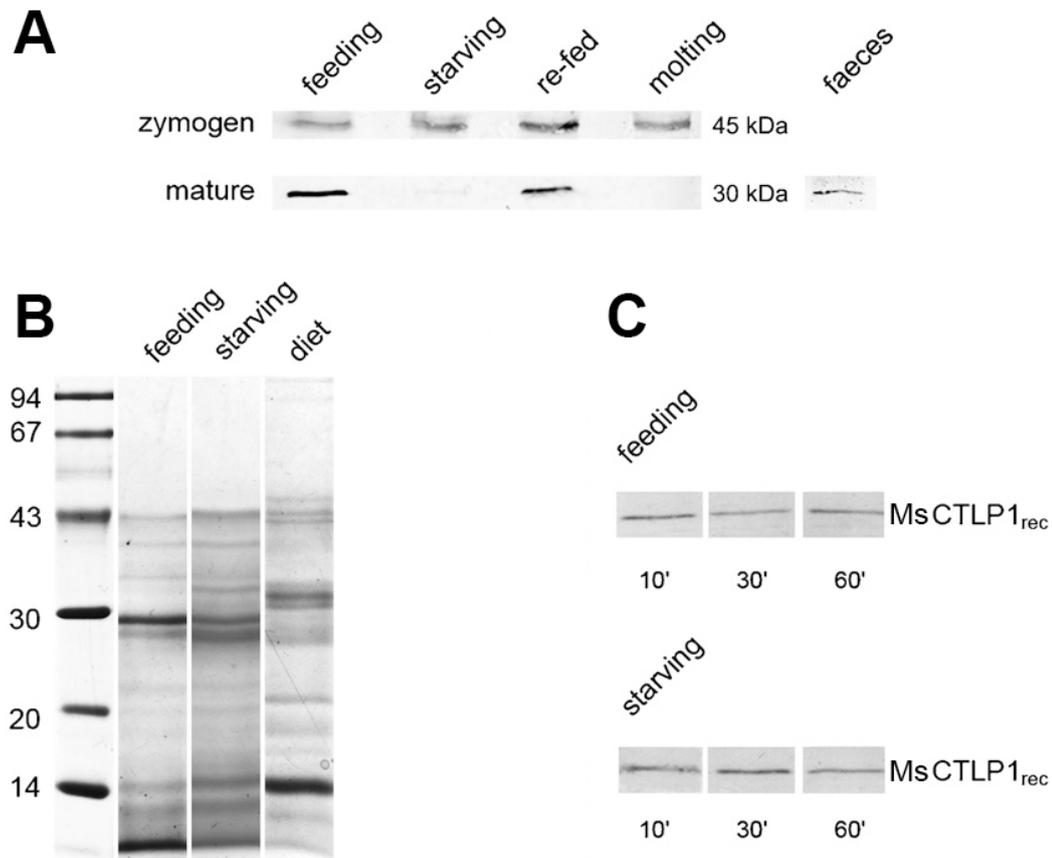
To investigate MsCTLP production at different physiological states, immunocytochemistry and immunoblots were performed using the anti-MsCTLP1 antibody.

To look for the overall localisation of MsCTLPs in the midgut epithelium, immunocytochemistry was performed using the anti-MsCTLP1 antibody to stain cryosections of anterior midguts from feeding, starving, re-fed and molting larvae. The overall distribution of MsCTLP1 immunosignals did not differ considerably for the tested physiological states. In all cases, most of the immunosignals were detected at the brush border formed by midgut columnar cells (Fig. 25). However, during starvation and molt additional immunosignals were observable in undefined compartments below the apical brush border suggesting that MsCTLP1 is not secreted during starvation and molt but may be stored in subapical vesicles, which fuse with the plasma membrane upon feeding (Fig. 25 B and D and inset).

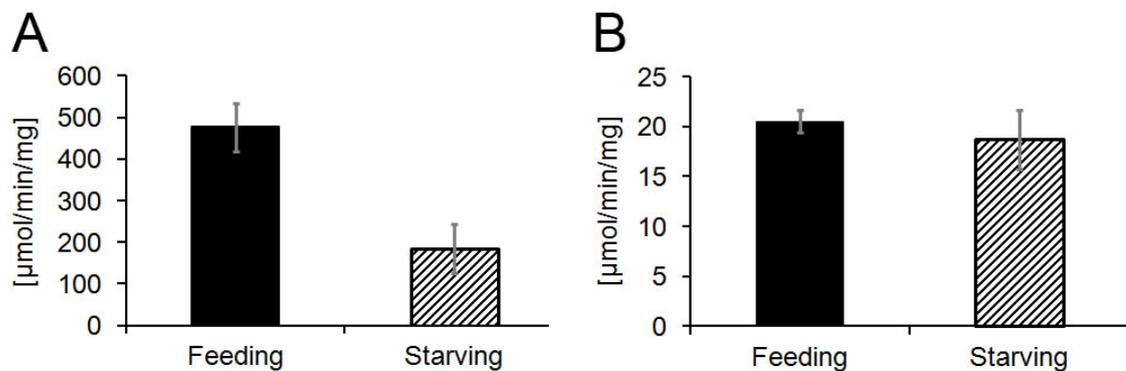


**Fig. 25.** Immunohistochemical labeling of MsCTLP1 in the anterior midgut at different physiological states. 20 mm cryosections of anterior midguts were immuno-stained with the anti-MsCTLP1 antibodies, and visualized by the reaction with Cy3-conjugated anti-rabbit IgG. (A) Feeding 5th instar larvae; (B) 5th instar larvae after 36 h starvation; (C) 36 h starved larvae re-fed for 8 h; (D) molting larvae (the inset zooms in the subapical region); (E) negative control performed with the corresponding pre-immune serum.

To address this question in closer detail, secretion and maturation of MsCTLTP1 was examined. For this purpose extracts were prepared from the midgut epithelium and the midgut content, and tested for the presence of zymogenic and mature MsCTLTP1 in immunoblots. As shown in Fig. 26 A, zymogenic MsCTLTP1 from the midgut epithelium migrates at about 45 kDa and is present in similar amounts in feeding, starving, re-fed and molting animals, which is in line with the results obtained from Realtime PCR (Fig. 24). However, MsCTLTP1 amounts did also not change significantly during molt, although MsCTLTP1 mRNA amounts decreased significantly (Fig. 24). This finding may be explained by storage of premature MsCTLTP1 in intracellular vesicles as also observed by immunocytochemistry (Fig. 25). Mature MsCTLTP1 migrates at about 30 kDa in SDS-PAGE suggesting that the activation peptide is rapidly cleaved by trypsin once secreted into the midgut lumen. In contrast to the situation in feeding and the re-fed larvae, no MsCTLTP1 signals were detectable in the gut content from starving or molting larvae (Fig. 26 A). During molt and starvation MsCTLTP1 secretion into the gut lumen appears to be suspended and this may apply also for other proteolytic enzymes. This finding appears to be plausible for molting larvae as they purge the bowels before molting. However, starving larvae still have gut content and MsCTLTPs could also be degraded due to the absence of other substrates. If starving larvae indeed suspend MsCTLTP secretion, chymotrypsin activity should be significantly reduced in starving guts. Indeed, when the chymotrypsin activity was measured in feeding and starving guts, a decrease in total chymotrypsin activity was observed in the gut content from  $477 \pm 58 \mu\text{mol}/\text{min}/\text{mg}$  ( $n = 6$ ) to  $144 \pm 59 \mu\text{mol}/\text{min}/\text{mg}$  ( $n = 6$ ; Fig. 27 A). Trypsin activity of feeding larvae was measured to be about 20-fold lower ( $20 \pm 1 \mu\text{mol}/\text{min}/\text{mg}$ ;  $n = 6$ ; Fig. 27 B) than the chymotrypsin activity. This result is to some degree in line with the findings of Srinivasan *et al.* (2006) who report chymotrypsins and trypsins to be involved in 80% and 10% of the total digestive activity of the *Manduca* midgut, respectively. However, the trypsin activity of the midgut content of starving larvae, showed no significant decrease when compared to that of feeding larvae ( $19 \pm 3 \mu\text{mol}/\text{min}/\text{mg}$ ;  $n = 6$ ; Fig. 27 B). Moreover, the activated form of MsCTLTP1, recombinantly expressed in *Sf21* cells (see 3.12.2), is not degraded by midgut extracts from feeding and starving larvae (Fig. 26 C), although BSA is still completely digested (not shown), and intact MsCTLTP1 can still be detected in the faeces of the larvae indicating its endogenous stability to proteolysis (Fig. 26 A). Finally, the gut content of starving larvae, although partially different in its protein composition from feeding larvae and completely different from diet, still contains numerous proteins, which might act as a substrate for peptidases (Fig. 26 B). Thus, missing of an immunosignal for MsCTLTP1 in the gut content of starving larvae appears not to be due to proteolytic degradation but due to the suspension of secretion.



**Fig. 26.** MsCTLPs in the midgut epithelium, the gut content and the faeces. (A) 10 mg of proteins from the anterior midgut epithelium producing the MsCTLP zymogens, or 1 mg of proteins isolated from the content of the median midgut containing mature MsCTLPs were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Immunostaining was performed with anti-MsCTLP1 antibodies. Feeding, feeding 5<sup>th</sup> instar larvae; starving, 5<sup>th</sup> instar larvae after 36 h starvation; re-fed, 36 h starved larvae re-fed for 8 h; molting, molting larvae; faeces, faeces of 5<sup>th</sup> instar feeding larvae (B). 2 mg of proteins from the gut contents of feeding and starving 5<sup>th</sup> instar larvae, and 2 mg of dietary proteins (diet) were separated by SDS-PAGE and stained with Coomassie Blue. The molecular masses of standard proteins are given in kDa (left lane). (C) Recombinant MsCTLP1 was expressed in insect *Sf21* cells and 4.5 ml of the protein solution (1 mg/ml) were incubated with 0.5 ml gut content extracts from feeding and starving larvae at 28°C for the indicated periods of time.



**Fig. 27.** Enzymatic activity assays. Photometric chymotrypsin (A) and trypsin (B) activity assays were performed using the substrates AAPF and BAPNA, respectively. The samples were incubated at RT for 60 minutes while simultaneously reading the absorbance values at 410 nm. The activity was calculated based on the absorbance line slope, the protein concentration of the samples and the respective extinction coefficients ([μmol/min/mg]; mean values ± S.E., n = 6).

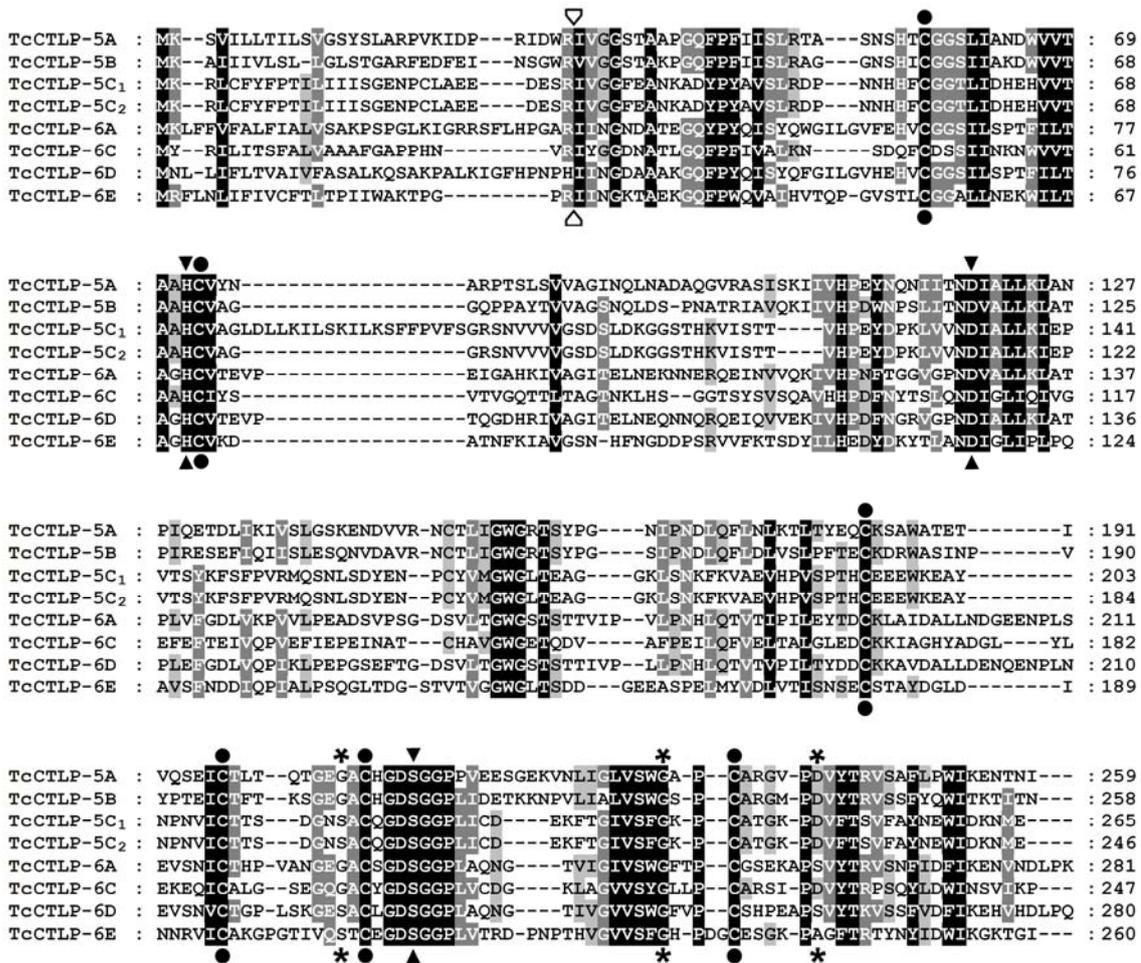
### 3.8. Identification of CTLP-encoding genes in the *Tribolium* genome

To gain more insights into the functions of CTLPs, the gene family, encoding these peptidases in the red flour beetle, *T. castaneum*, was characterized. *Tribolium* is a well established genetic insect model and has the big advantage of an almost fully assembled genome (Richards *et al.*, 2008). To identify those genes encoding chymotrypsin-like peptidases, a tblastn search of the *Tribolium* genome database at NCBI was performed using MsCTLP1 (Broehan *et al.*, 2007) as query. From the resulting list of homologous sequences, those were selected that showed a chymotrypsin-like substrate specificity pocket as defined by Perona and Craik (1995). Accordingly, proteins containing the residues Ser/Thr190, Gly216, and Gly/Ala/Ser226 in the primary substrate-binding pocket were classified as chymotrypsin-like peptidases. However, this definition was extended by including serine peptidases with the residues Gly190, Gly216, and Asp226, as also found in MsCTLP1-4. Moreover, these specificity pocket residues have been reported for *Spodoptera exigua* SeCT34 and *Anopheles gambiae* Anchym1 as well, two serine peptidases that evidently exhibit chymotrypsin-like activity (Herrero *et al.*, 2005; Vizioli *et al.*, 2001). By extending the search pattern to Gly/Ser/Thr190, Gly216, and Asp/Gly/Ala/Ser226 and using the sequences identified by the initial search as queries in individual PSI-BLAST searches, a total of 14 *TcCTLP* genes was obtained. The genes were denoted according to their linkage group by number and their physical order on the chromosome by letter. For instance, *TcCTLP-6C* is the third gene on chromosome 6 encoding a chymotrypsin-like peptidase. Comparison of the *Tribolium* and *Drosophila* genomes revealed no conservation of synteny of CTLP genes. Only small gene clusters of CTLPs can be observed in both species.

### 3.9. Isolation and sequencing of *TcCTLP* cDNAs

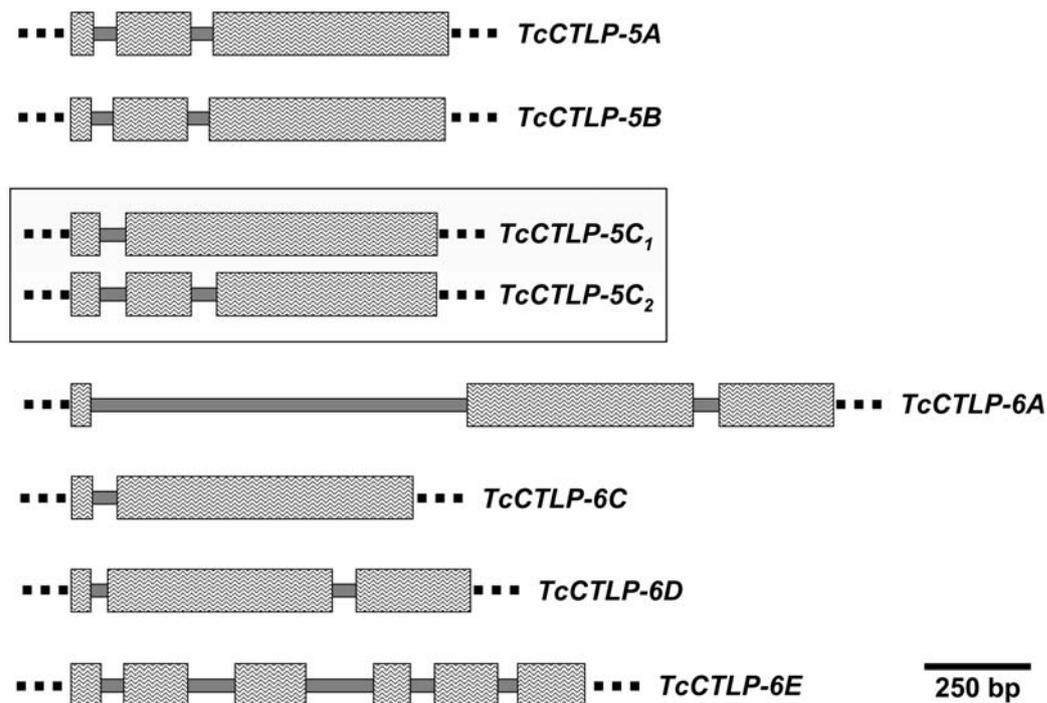
The cDNAs for seven selected *TcCTLPs*, *TcCTLP-5A*, *TcCTLP-5B*, *TcCTLP-5C<sub>1</sub>*, *TcCTLP-5C<sub>2</sub>*, *TcCTLP-6A*, *TcCTLP-6C*, *TcCTLP-6D* and *TcCTLP-6E* identified in the *Tribolium* genome were amplified by RT-PCR using gene specific primers and total RNA from the larval midgut. The PCR products were ligated into the pGEM-T vector and sequenced. Sequences matched those published in Beetlebase and NCBI's GenBank except for a few nucleotide substitutions. Some differences in the nucleotide sequences turned out to be silent nucleotide substitutions, but a few led to changes in the deduced amino acid sequences. These changes include conserved and non-conserved amino acid substitutions: *TcCTLP-5A*: S-19-R, D-110-E and L-215-P; *TcCTLP-6C*: T-8-S, G-14-A and G-51-S; *TcCTLP-6D*: G-148-K, I-175-L, S-189-T and S-210-N; *TcCTLP-6E*: A-22-T, Q-98-K, N-109-D, S-151-G, G-192-R, V-194-I, S-220-P and I-227-V. The latter substitutions might be due to the fact that the *T. castaneum* strain used for genomic sequencing was the highly inbred GA2 strain (Richards *et al.*, 2008) and not the more polymorphic GA1 strain used in this work. Furthermore, midgut serine peptidases may exhibit a high evolutionary rate as an adaptative mechanism to counteract plant serine peptidase inhibitors (Srinivasan *et al.*, 2006).

Because initially only a central fragment of the cDNA for *TcCTLP-5C* could be amplified by RT-PCR, 5'- and 3'-RACE was performed to complete the sequence. The 5'-RACE experiment yielded both long and short products, which turned out to be identical in their nucleotide sequences with the exception of a 57 bp region that was present only in the longer RACE product (Figs. 28 and 29). Also the 3'-RACE experiment performed with a forward primer that included the 57 bp region of the longer 5'-RACE product yielded longer and shorter products for *TcCTLP-5C*. The 5'- as well as the 3'-fragments perfectly overlapped with the initially obtained central fragment of *TcCTLP-5C*. RT-PCR with corresponding 5'- and 3'-primers to amplify the complete coding sequence, and subsequent nucleotide sequencing finally confirmed the existence of two *TcCTLP-5C* variants, a long two exon-spanning cDNA of 798 bp (*TcCTLP-5C<sub>1</sub>*) and a shorter three exonspanning cDNA of 741 bp (*TcCTLP-5C<sub>2</sub>*; Figs. 28 and 29). These variants are probably the result of alternative splicing of the 57 nucleotides spanning intron, which is excised via conserved 5'- and 3'-splicing sites only in *TcCTLP-5C<sub>2</sub>*.



**Fig. 28.** ClustalW alignment of chymotrypsin-like peptidases from *Tribolium castaneum*. Conserved, highly conserved or identical amino acids are highlighted in light-gray, gray or black, respectively. Open arrow head, conserved trypsin cleavage sites; asterisks, conserved S1 pocket residues; closed circles, conserved cysteines; triangles, conserved residues of the catalytic triad; TcCTLP-5A (embl|CBC01177.1), TcCTLP-5B (embl|CBC01166.1), TcCTLP-5C<sub>1</sub> (embl|CBC01175.1), TcCTLP-5C<sub>2</sub> (embl|CBC01172.1), TcCTLP-6A (embl|CBC01171.1), TcCTLP-6C (embl|CBC01169.1), TcCTLP-6D (embl|CBC01179.1), TcCTLP-6E (embl|CBC01181.1).

As described above for the MsCTLTPs, all seven TcCTLTPs meet the structural requirements of S1 family peptidases (Fig. 28) as they possess the conserved histidine, aspartate and serine residues that form the catalytic triad. Additionally, they all share the highly conserved GDSGGP motif around the catalytic serine (Table 2) as well as the six conserved cysteine residues typically present in invertebrate S1A peptidases (Yan *et al.*, 2001). Like all members of the S1A peptidase subfamily, the *Tribolium* CTLTPs are probably expressed as zymogens that are activated upon tryptic cleavage (Table 2 and Fig. 28; Lehane and Billingsley, 1996; Peterson *et al.*, 1995). All TcCTLTPs are presumably secretory proteins because the predicted proteins contain amino-terminal signal peptides for the secretory pathway predicted to be cleaved after amino acid position 18 for TcCTLTP-5A, position 19 for TcCTLTP-5B, position 23 for TcCTLTP-5C, position 16 for TcCTLTP-6A and TcCTLTP-6D, position 17 for TcCTLTP-6C and position 20 for TcCTLTP-6E (Table 2 and Fig. 28; SignalP v3.0; Emanuelsson *et al.*, 2007). The overall gene structure for the seven selected *TcCTLTP*s revealed striking differences. While *TcCTLTP-5A*, *TcCTLTP-5B*, *TcCTLTP-6A* and *TcCTLTP-6D* contain three exons and two introns, *TcCTLTP-6C* has only two exons and one intron, and *TcCTLTP-6E* has six exons and five introns. In contrast, *TcCTLTP-5C* is alternatively spliced and hence contains either two or three exons along with one or two introns, depending on which transcript is generated (Fig. 29). Furthermore the nucleotide sequence identity of the coding regions of the seven *TcCTLTP*s is rather low (about 50%; Table 3) with the exception of *TcCTLTP-6A* and *TcCTLTP-6D*, which share about 70% identical nucleotides. In line with this finding, phylogenetic analysis revealed no clustering of the TcCTLTPs as observed for the *Manduca* CTLTPs, suggesting that their divergence is more ancient (Fig. 20).



**Fig. 29.** Exon-intron organization of *TcCTLTP* genes. The exons are given by broader rectangles that enclose respective introns; intergenic regions are indicated by dotted lines. The box highlights two splice variants of *TcCTLTP-5C*.

**Table 2.** Chymotrypsin-like serine peptidases of *Tribolium*. Either GeneBank (XP\_...) or embl (CB...) accession numbers are given. If not listed, sequences are identical to the conserved motifs TAAHC, DIAL, or GDSGGP. The putative activation site is indicated by an arrow head and the numbering refers to the P4 position. The cleavage site for TcCTLP-8A could not be predicted. The lengths of the precursor and mature peptidases are indicated.

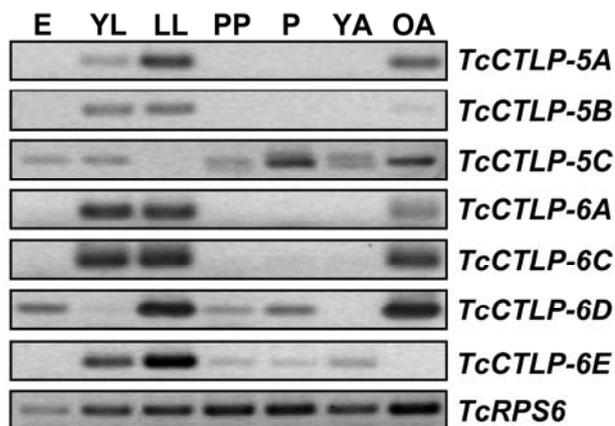
Name	Accession No.	Linkage Group	Conserved motifs			Length (aa)		Putative activation Site	Specificity pocket residues
			TAAHC	DIAL	GDSGGP	Precursor	Mature		
TcCTLP-3A	XP_967407.2	3		DICL		274	256	31 SKLH <sup>^</sup> IIGG	GGG
TcCTLP-3B	XP_967496.1	3	TAGHC	DICL		276	257	33 SKVR <sup>^</sup> IIGG	GGG
TcCTLP-5A	CBC01177.1	5				259	241	27 IDWR <sup>^</sup> IVGG	GGD
TcCTLP-5B	CBC01166.1	5		DVAL		258	239	26 SGWR <sup>^</sup> VVGG	GGD
TcCTLP-5C <sub>1</sub>	CBC01175.1	5				265	242	26 DESR <sup>^</sup> IVGG	SGD
TcCTLP-5C <sub>2</sub>	CBC01172.1	5				246	223	26 DESR <sup>^</sup> IVGG	SGD
TcCTLP-6A	CBC01171.1	6	TAGHC			281	265	32 PGAR <sup>^</sup> IING	GGG
TcCTLP-6B	XP_970321.1	6		DIGL		248	231	21 PDVS <sup>^</sup> IHGG	GAD
TcCTLP-6C	CBC01169.1	6		DIGL		247	230	21 HNVR <sup>^</sup> IYGG	GGD
TcCTLP-6D	CBC01179.1	6	TAGHC			280	264	31 PNPV <sup>^</sup> IING	SGS
TcCTLP-6E	CBC01181.1	6	TAGHC	DIGL		260	240	23 PGPR <sup>^</sup> IING	SGA
TcCTLP-6F	XP_970603.1	6				274	258	41 IGNR <sup>^</sup> IING	SGS
TcCTLP-7A	XP_969640.1	7		DLAL		277	261	38 PGVR <sup>^</sup> ITGG	SGS
TcCTLP-8A	XP_001815065.1	8		DYSL	GDSGSP	288	265	-	TGS
TcCTLP-9A	XP_001812802.1	9	TAGHC	DVAV		277	260	41 PSPK <sup>^</sup> IIGG	GGG

**Table 3.** Nucleotide sequence identities of *TcCTLPs*. The identities of the coding regions of the *CTLPs* were calculated in percent with the help of the ALIGN program at SDCS Biology Workbench (<http://workbench.sdsc.edu/>). For the purpose of comparison *MsCTLP1* has been included in this calculation.

	<i>MsCTLP1</i>	<i>TcCTLP-5A</i>	<i>TcCTLP-5B</i>	<i>TcCTLP-5C<sub>1</sub></i>	<i>TcCTLP-5C<sub>2</sub></i>	<i>TcCTLP-6A</i>	<i>TcCTLP-6C</i>	<i>TcCTLP-6D</i>	<i>TcCTLP-6E</i>
<i>MsCTLP1</i>	100								
<i>TcCTLP-5A</i>	46.1	100							
<i>TcCTLP-5B</i>	44.3	47.7	100						
<i>TcCTLP-5C<sub>1</sub></i>	49.1	48.4	46.0	100					
<i>TcCTLP-5C<sub>2</sub></i>	49.9	47.7	47.1	93.7	100				
<i>TcCTLP-6A</i>	48.2	51.3	51.1	45.7	46.2	100			
<i>TcCTLP-6C</i>	46.4	53.9	52.5	46.2	46.8	49.0	100		
<i>TcCTLP-6D</i>	48.1	50.6	51.0	46.7	47.4	70.1	50.7	100	
<i>TcCTLP-6E</i>	45.6	50.4	49.2	46.7	46.5	48.8	49.9	48.6	100

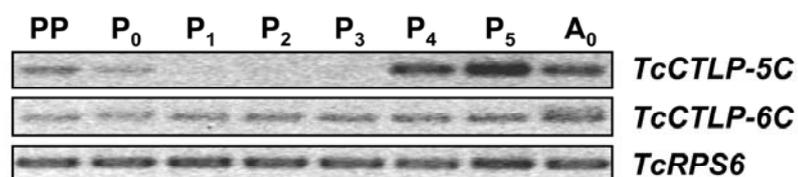
### 3.10. Expression profiles of *TcCTLPs* genes

To determine *TcCTLP* gene expression in different tissues, RT-PCR was performed using sequence-specific primer pairs (Table 7) and total RNA isolated from eggs, young and last instar larvae, prepupae, pupae, young adults and old adults (Fig. 30). Transcripts were detected in the larval stages for all *TcCTLPs* with the exception of *TcCTLP-5C* and *TcCTLP-6D*, both of which showed only very low expression in last instar and younger larvae, respectively. Furthermore, *TcCTLP-5C* and *TcCTLP-6D* were the only *TcCTLPs* expressed at the embryonic stage. *TcCTLP-5C*, *TcCTLP-6C*, *TcCTLP6D* and *TcCTLP-6E* were expressed in the pupal stage (Fig. 30). However, *TcCTLP-6C* expression in pupae was rather low as it could be amplified only by using high cycle numbers (Fig. 31). While *TcCTLP-5C*, *TcCTLP-6D* and *TcCTLP-6E* showed considerable expression in nonfeeding pupae, gene expression for *TcCTLP-5A*, *TcCTLP-5B* and *TcCTLP-6A* could be detected only in feeding stages (larval and old adult stages), indicating that these *TcCTLPs* may play important roles in digestion (Fig. 30). The finding that *TcCTLP-5C*, *TcCTLP-6C*, *TcCTLP-6D* and *TcCTLP-6E* are also expressed in pupae suggests that these enzymes may have additional functions beyond digestion of food. To achieve a higher time resolution in analyzing stage-specific expression in pupae, total RNA was isolated from prepupae, pupae of different ages and young adults immediately after the pupal-adult molt, and those RNAs were used as templates for RT-PCR with primers specific for *TcCTLP-5C* or *TcCTLP-6C* (Table 7). *TcCTLP-5C* showed highest expression shortly before the larval-pupal and pupal-adult molts. *TcCTLP-6C* showed only low levels of expression at all stages tested, with expression increasing shortly after the pupal-adult molt (Fig. 31).



**Fig. 30.** Stage specific *TcCTLPs* expression profiles as determined by RT-PCR. Total RNA was prepared from different developmental stages and cDNAs were synthesized. RT-PCR was carried out with primers specific to the indicated genes. PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide. E, eggs; YL, young (penultimate) larvae; LL, last instar larvae; PP, prepupae; P, pupae; YA, young adults; OA, old adults.

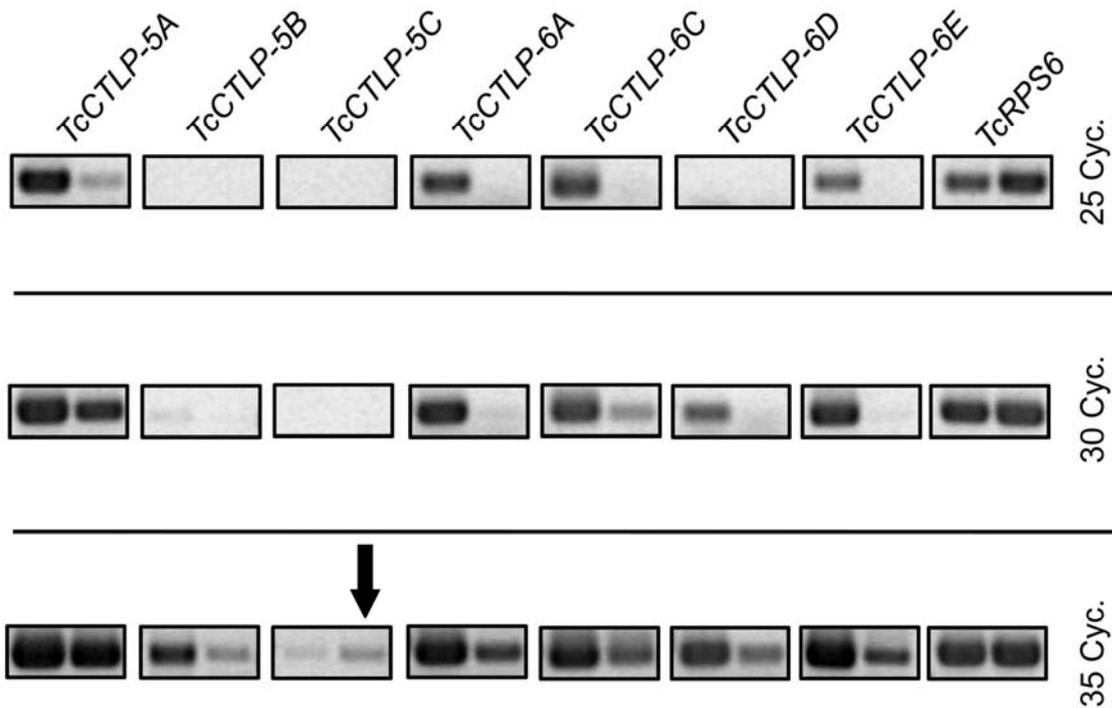
**Fig. 31.** Expression profiles of *TcCTLP-5C* and *TcCTLP-6C* during the pupal stage as determined by RT-PCR. Total RNA was prepared from different



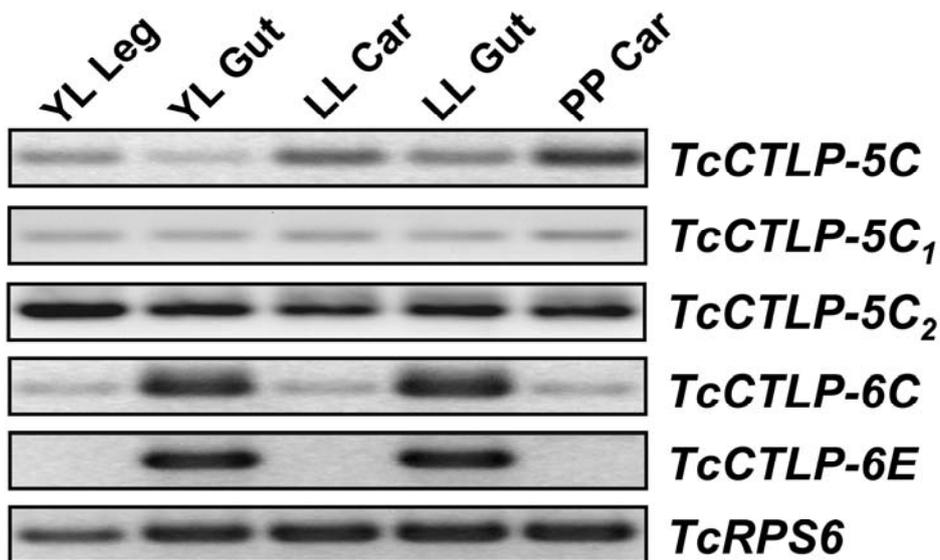
developmental stages and cDNAs were synthesized. RT-PCR was carried out with primers specific to the indicated genes. PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide. Lanes: PP, prepupa; P<sub>0</sub>, Pupae right after larval-pupal molt; P<sub>1</sub>-P<sub>5</sub>, pupae 1-5 days after larval-pupal molt; A<sub>0</sub>, young adults right after pupal-adult molt.

### 3.11. *TcCTLP* gene expression in carcass and midgut

To determine the tissue specificity of *TcCTLP* gene expression, carcass and midgut tissues were prepared from penultimate instar larvae. Following reverse transcription, the cDNA was used as a template in a PCR with different cycle numbers and *TcCTLP* gene specific primers (Table 7). RT-PCR revealed that *TcCTLP* expression in almost all cases was higher in the midgut tissue than in the carcass with the exception of *TcCTLP-5C*, which was highest in the carcass (Fig 32). Since it cannot be excluded that the carcass preparation has been contaminated with midgut tissue, midgut and carcass tissues were once more prepared from young and late larva as well as prepupae. To avoid contamination with midgut material, carcass tissues were washed repeatedly in PBS buffer. Additionally samples of legs from young larvae were prepared serving as a non-gut control to enable clear differentiation between epidermal and gut expression. Following the isolation of the total RNA, RT-PCR was performed with *TcCTLP* gene specific primers (Table 7). However, RT-PCR did reveal *TcCTLP-6E* expression exclusively in midgut tissue and not in carcass indicating that the carcass preparation was free of midgut contamination (Fig. 33). In contrast, *TcCTLP-6C* was not only highly expressed in the midgut, but it was also expressed at lower levels in the carcass. This suggested additional functions of *TcCTLP-6C* besides nutrient digestion in the midgut, an inference also made from the observation of pupal expression of this gene (Fig 31). *TcCTLP-5C* exhibited weak expression in midgut tissue but a somewhat higher level of expression in the carcass (Fig. 33, upper row). To test if either one of the two splice variants of *TcCTLP-5C* was specifically expressed in the midgut or carcass, RT-PCR was performed with primers specific for *TcCTLP-5C<sub>1</sub>* and *TcCTLP-5C<sub>2</sub>* (Table 7) using cDNAs from the midgut and carcass as templates. The results showed that both splice variants of *TcCTLP-5C* are expressed in all of the tissues analyzed (Fig. 33).



**Fig. 32.** Expression of *TcCTLPs* in midgut and carcass. Total RNA was prepared from midgut and carcass tissues of penultimate instar larvae and cDNAs were synthesized. RT-PCR was carried out at different cycle numbers with primers specific to the indicated genes. PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide (Midgut left, carcass right). *TcCTLP* expression was found mainly in the midgut in almost all cases with the exception of *TcCTLP-5C*, which was highest in the carcass (arrow).

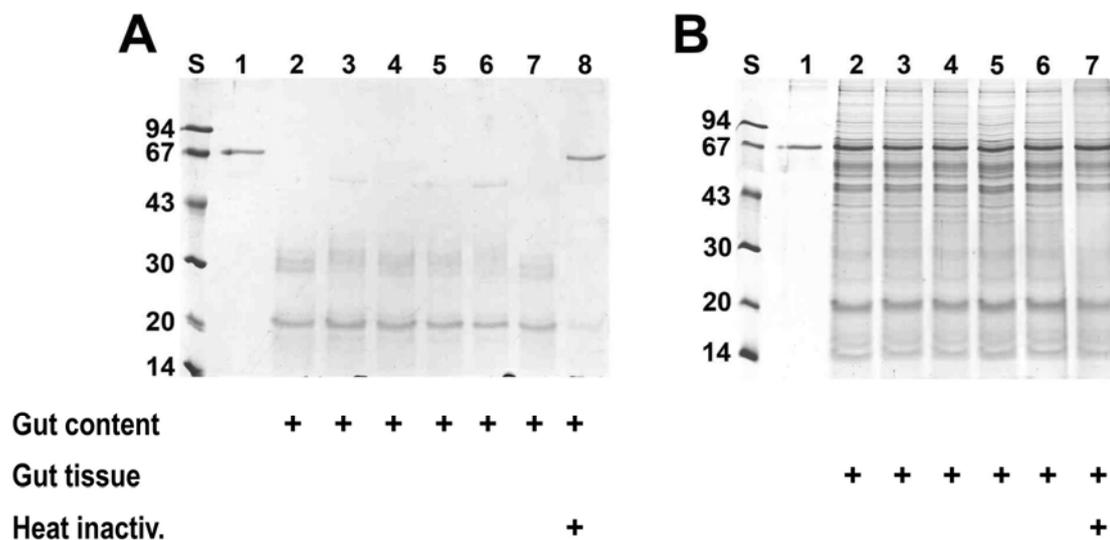


**Fig. 33.** Expression of *TcCTLP-5C*, *TcCTLP-6C* and *TcCTLP-6E* in midgut and carcass. Total RNA was prepared from midgut and carcass tissues and cDNAs were synthesized. RT-PCR was carried out with primers specific to the indicated genes and splice variants. PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide. YL Leg, legs from young larvae; YL Gut, midguts from young larvae; LL Car, carcass from late larvae; LL Gut, midguts from last instar larvae; PP Car, carcass from prepupae.

### 3.12. Purification and recombinant expression of insect CTLPs

#### 3.12.1. Purification of MsCTLP1 from the midgut

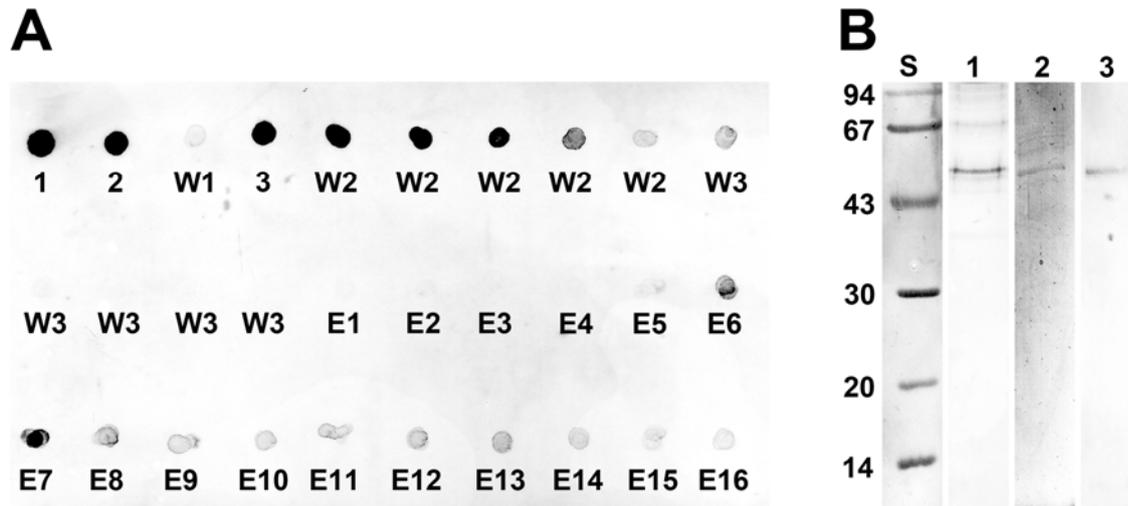
Purification of MsCTLP1 from the midgut content was performed by using the Affi-Gel Hz Immunoaffinity Kit (Bio-Rad). For this purpose the anti-MsCTLP1 antibody was purified by Protein A column (GE Healthcare), desalted and coupled to the Affi-Gel Hz matrix. To rule out the possibility that the midgut peptidases digest the anti-MsCTLP1 antibody of the column-material, pretests were made in advance to the protein purification (Fig. 34). The pretests suggested that none of the tested conditions was suitable to purify MsCTLP1 from the gut content because of its strong proteolytic activity (Fig. 34 A). For this reason midgut homogenates were used for the purification of zymogenic MsCTLP1 since no proteolytic degradation of BSA could be detected (Fig. 34 B).



**Fig. 34.** SDS-Page stained with Coomassie blue showing the digestive activity of the gut content (A) and the midgut tissue (B). Gut contents (A) or midgut tissue extracts (B) were used to digest 5% BSA (w/v). Midgut contents or tissue extracts heat inactivated for 15 min (Heat inactiv.) served as the control respectively. Experiments were either carried out in TBS (pH 10.0; A) or in a MOPS buffered saline (pH 6.5) with the addition of 0.5 % (w/v) soybean trypsin inhibitor (B). Incubation times were the following: A. 2, 30 sec; 3, 1 min; 4, 2 min; 5, 5 min; 6, 9 min; 7, 15 min; 8, 15 min; B. 2, 30 sec; 3, 1 min; 4, 3 min; 5, 8 min; 6, 15 min; 7, 15 min. Standard proteins with molecular masses indicated in kDa are also shown (S).

Midguts extracts were prepared as described above (see 2.3.1) and were purified by affinity chromatography using the anti-MsCTLP1 antibody coupled Affi-Gel Hz matrix. Following two wash steps at high and low salt conditions, bound proteins were eluted by 10 ml elution buffer and collected into 500  $\mu$ l fractions. To detect MsCTLP1 in the resulting fractions, 1  $\mu$ l of each sample was dotted onto a nitrocellulose membrane and stained with anti-MsCTLP1 antibodies (Fig. 35 A). Immuno-reactive fractions E6-E9 were pooled, concentrated by ultrafiltration and submitted to SDS-Page. After gel-electrophoresis and Coomassie blue staining a signal at the height of zymogenic MsCTLP1 could be detected (Fig. 35 B lane 2). To confirm that the band

corresponds to MsCTLTP1, an immunostaining with anti-MsCTLTP1 antibodies was performed, detecting a signal at the height of zymogenic MsCTLTP1 (Fig. 35 B lane 3). Nevertheless protein yield was too low for further characterization of MsCTLTP1. For this reason MsCTLTP1 was expressed in heterologous expression systems.



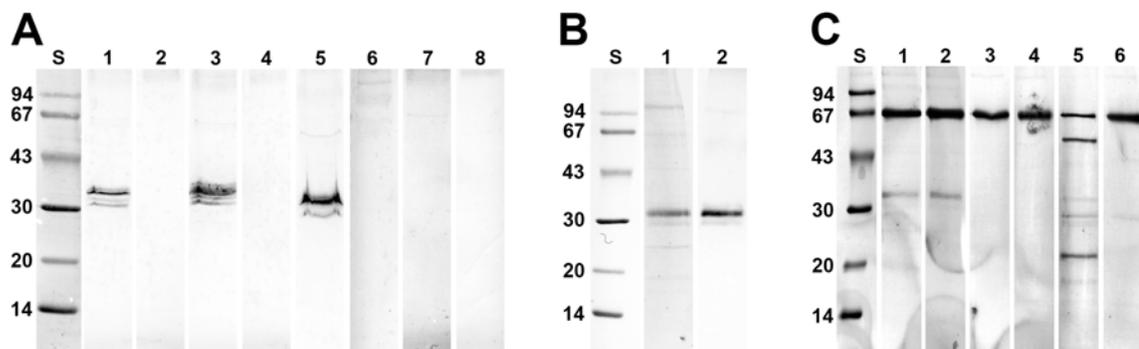
**Fig. 35.** Affi-Gel matrix purification of MsCTLTP1 from midgut tissue. (A) 1  $\mu$ l of each fractions from the purification of MsCTLTP1 was dotted onto nitrocellulose membrane and incubated with the anti-MsCTLTP1 antibodies. 1, anti-MsCTLTP1 antibody; 2, midgut homogenate; 3, column flow through; W1-3, wash steps; E1-E16, elution fractions. (B) Elution fractions 6-9 were pooled, concentrated and submitted to SDS-Page (2) as well as transferred to nitrocellulose membrane and stained with the anti-MsCTLTP1 antibodies (3). Purified proteins migrated at the same height as MsCTLTP1 from midgut tissue extracts made visible with the anti-MsCTLTP1 antibodies (1).

### 3.12.2. Recombinant expression of MsCTLTP1

Since purification of MsCTLTP1 from either the gut content or the midgut tissue proved to be difficult, it was expressed as a recombinant protein. To accomplish this, several approaches were made. In all expression systems used, MsCTLTP1 was expressed as the immature (amino acid positions 1–281), zymogenic (amino acid positions 19–281) or activated form (amino acid positions 41–281). Expression performed in *Sf9* cells using the pIB/V5-His vector system (Invitrogen), in COS cells using the pFLAG-CMV-3 and pFLAG-CMV-4 vector systems (Sigma Aldrich), in *Pichia pastoris* using the pPICZa vector system as well as expression in S2 cells using the pMT/BiP/V5-His vector system (Invitrogen) proved not to be successful. Expression of activated MsCTLTP1 in Rosetta (DE3) pLysS was achieved by using the pET29b vector system. In immunoblots, the anti-MsCTLTP1 antibodies specifically stained the recombinant MsCTLTP1. Its apparent molecular mass was about 30 kDa, which is in good agreement with the expected mass of the activated MsCTLTP1 (25.3 kDa) plus that of two tags (4.3 kDa) at the amino- and carboxy-terminal ends provided by the pET29b expression vector (data not shown). However, MsCTLTP1 expressed by Rosetta cells turned out to be insoluble and could not be solubilized.

Using a prokaryotic expression system like Rosetta for the expression of eukaryotic proteins has the disadvantage that post-translational modifications as well as

the proper protein folding might be incorrect, leading to non functional and/or insoluble proteins. By expressing MsCTLP1 in *Sf21* cells, the negative side effects of prokaryotic expression systems can be circumvented. For this reason, the full coding region of *MsCTLP1* was expressed in *Sf21* cells. To test whether secretion into the media had occurred, the cell culture was detached from the flask surface and the cell suspension was pelleted by centrifugation. Aliquots of the pellet and the supernatant fractions were separated by SDS-Page and transferred to nitrocellulose membrane. Staining with the anti-MsCTLP1 antibody revealed that the protein is expressed by the *Sf21* cells but not into the media since it was only found in the pelleted fraction (Fig. 36 A, lane 1). Zymogenic and active MsCTLP1 recombinantly expressed by *Sf21* cells were also not found in the supernatant when the endogenous signal peptide of MsCTLP1 was replaced by the BiP peptide encoded by the pMT/BiP/V5-His vector (Fig. 36 A, lane 3 and 5). To purify the active form of MsCTLP1 from the cell pellet via its His-Tag, *Sf21* cells were homogenized in the presence of 0.1% NLS and purified by Ni-NTA chromatography in the presence of 6 M GuHCl. After dialysis, the purified protein was soluble and showed an immunoreaction with the anti-MsCTLP1 antibody (Fig. 36 B). The signal migrated at about 32 kDa, hence it was slightly larger than the calculated theoretical mass for secreted, activated MsCTLP1 with 26.4 kDa (including the carboxy-terminal octa-His-Tag). To test whether recombinant MsCTLP1 is indeed active, the purified peptidase was used to digest BSA (Fig. 36 C). In parallel, the purified enzyme was applied in a photometric assay using the synthetic substrate AAPF (data not shown). None of the two methods could show any digestive activity of purified MsCTLP1.



**Fig. 36.** (A) Western Blot of *Sf21* cells expressing MsCTLP1, stained with the anti-MsCTLP1 antibody. The expression media was separated into pellet (1, 3, 5 and 7) and supernatant fractions (2, 4, 6 and 8). 1+2 immature form; 3+4, zymogenic form; 5+6, activated form; 7+8, untransfected control. (B) The pellet of *Sf21* cells expressing the activated form of MsCTLP1 (see A. lane 5) was solubilized and purified by Ni-NTA column. After dialysis the purified peptidase was separated by SDS-Page (1), transferred to nitrocellulose and stained with the anti-MsCTLP1 antibodies (2). (C) Activity assay of recombinant MsCTLP1. Purified MsCTLP1 expressed by *Sf21* cells was used to digest BSA in the absence (1) or the presence (2) of 5 mM Pefa Bloc. Buffer without (3) or with (4) 5 mM Pefa Bloc, served as a negative control. No digestive activity could be observed except for bovine chymotrypsin used to digest BSA (5). The ability of bovine chymotrypsin to digest BSA was blocked by the addition of 5 mM Pefa Bloc to the saline (6). The molecular masses of standard proteins are given in kDa (left lanes).

### 3.12.3. Exchange of the hydrophobic region of MsCTLP1

Hydropathicity plots of the MsCTLP1-4 performed by ProtScale (ExPASy Proteomics Server) using the algorithm of Kyte and Doolittle (1982), showed a high amount of hydrophobic amino acids in MsCTLP1 when compared to the other MsCTLPs (Fig. 38; red region). According to the Kyte and Doolittle scale, especially the amino acid positions 176-208 show a high average hydropathy index of 0,994 (comparable regions: MsCTLP2, amino acid positions 174-206: 0.103; MsCTLP3, amino acid positions 176-208: 0.038; MsCTLP4, amino acid positions 172-204: 0.148). This region might contribute to the inability to express MsCTLP1 as a soluble protein in Rosetta as well as *Sf21* cells. To prove this assumption the hydrophobic region of *MsCTLP1* (base pairs 542-640) was exchanged by megaprimer PCR (see 2.2.10.3) against the equivalent region of *MsCTLP3* (base pairs 551-649), which was the least hydrophobic of all MsCTLPs (Fig. 37). After construction of a pFBD expression vector containing the MsCTLP1/MsCTLP3 hybrid construct, *Sf21* cells were transformed in order to express the less hydrophobic peptidase. After detachment of the cell culture, the cell suspension was separated into pellet and supernatant fractions by centrifugation. The proteins were separated by SDS-Page and transferred to nitrocellulose membranes. Membranes were stained with the anti-MsCTLP1 antibody in the case of the hybrid protein and with the anti-His antibody in the case of MsCTLP3, which was additionally expressed in *Sf21* cells as a control. Single bands of the correct size confirmed the expression of both proteins by *Sf21* cells but the signals were detected exclusively in the pellet fractions (Fig. 39). In summary, other aspects than hydropathicity seem to be the crucial factor for successful expression of MsCTLPs in *Sf21* cells. It appears very likely that the pH of the Insect-XPRESS cell medium (pH 6.8) influences protein functionality. Since the midgut pH of *Manduca* is highly alkaline (up to pH 12; Dow, 1984), MsCTLPs expressed in *Sf21* cells might be folded incorrectly when secreted into the mild acidic environment of the cell culture medium.

```

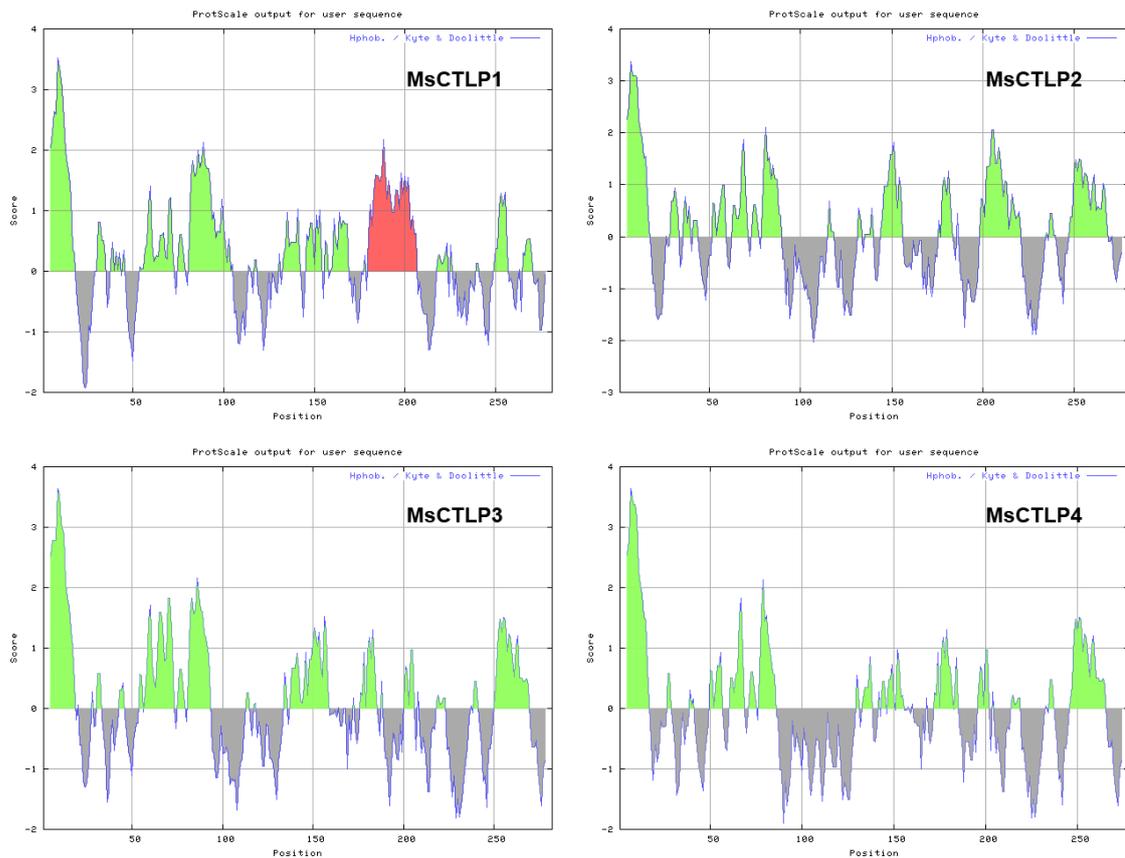
MsCTLP1 : MYVKVALLLVALLAGCNAPFAPEAEQDMSIFPEHVDNRNARIVGGTQANCAHPHMVALINCAVVRSEFLCGGSIITRRRVLTAAHCIAAVVSGNT : 94
MsCTLP3 : MYVKVALLLVALLAGCNAPFAPEAEQDMSIFPEHVDNRNARIVGGTQANCAHPHMVALINCAVVRSEFLCGGSIITRRRVLTAAHCIAAVVSGNT : 94

MsCTLP1 : IASTLRLLTVGTRNWNQGGTLYTVARNITPEEYVSNITIKNDIGLEITHNNTIIDITVVRSTPEINQDYVFGCVLDRVAGWGRIRTGGAI SPSLLEIT : 188
MsCTLP3 : IASRNLKTVGTRNWNQGGVWHAFOQHVHIESSINANTIKNDIGLEITHSANIAMTNAVRAIVVNDPIGNCINSRVAGWGRIRAGGAI SANLLQLN : 188

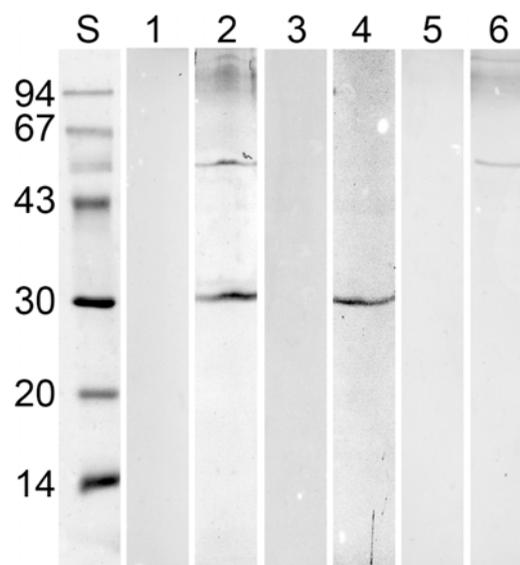
MsCTLP1 : VPTISGSACVASAIQAGIDINMRFPFVPEFHIELCTFFHGFNVGTCNGDSGSALRFDNQCQIGVVSNGEPCARGCPDMFVVRVSAYQSNLQOSIV- : 281
MsCTLP3 : TQTIDGNHCVREVARVAASINRREVPFVDENVELCTFFHSONHGTCNGDSGSALVRVDRNQOIGVVSNGEPCALGAPDMFVVRVSAYRNWVCSNTIN : 282

```

**Fig. 37.** ClustalW alignment of MsCTLP1 and MsCTLP3. The basepairs highlighted in green and red mark the regions, which were exchanged from MsCTLP3 to MsCTLP1.



**Fig. 38.** Hydropathicity plots of MsCTLPs. The plots for MsCTLP1-4 were performed with ProtScale (ExpASY Proteomics Server) using the algorithm of Kyte and Doolittle. Green areas represent hydrophobic regions whereas gray areas represent hydrophilic regions. The red area marks the hydrophobic region of MsCTLP1, which has been replaced by a less hydrophobic region of MsCTLP3.

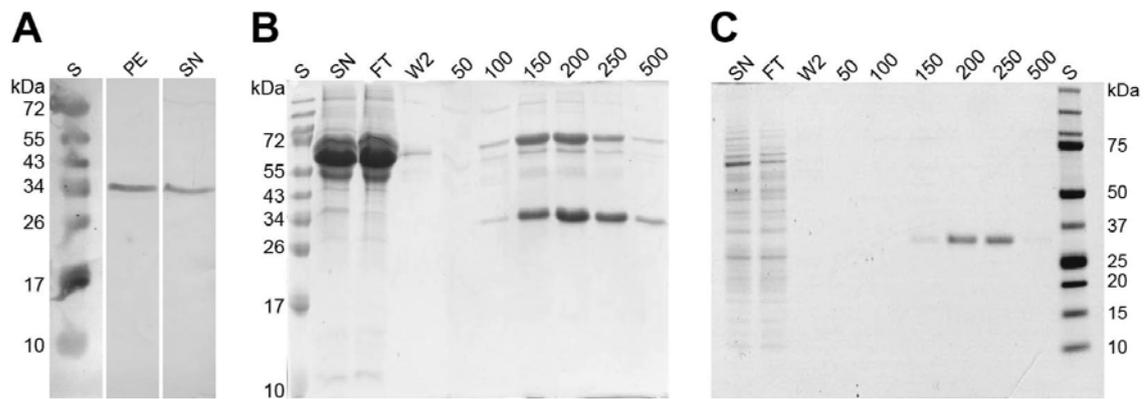


**Fig. 39.** Recombinant expression of a MsCTLP1/MsCTLP3 hybrid construct (1+2) and MsCTLP3 (3+4). Aliquots of the supernatant fractions (1+3) and the pellet fractions (2+4) of *Sf21* cells expressing the respective peptidase were separated by SDS-Page and transferred to nitrocellulose membranes. Membranes were either stained with the anti-MsCTLP1 antibody (1+2) or the anti-His antibody (3+4). Proteins of supernatant (5) and pellet fractions (6) of an untransfected control were stained with the anti-MsCTLP1 antibody. The molecular masses of standard proteins are given in kDa (left lane).

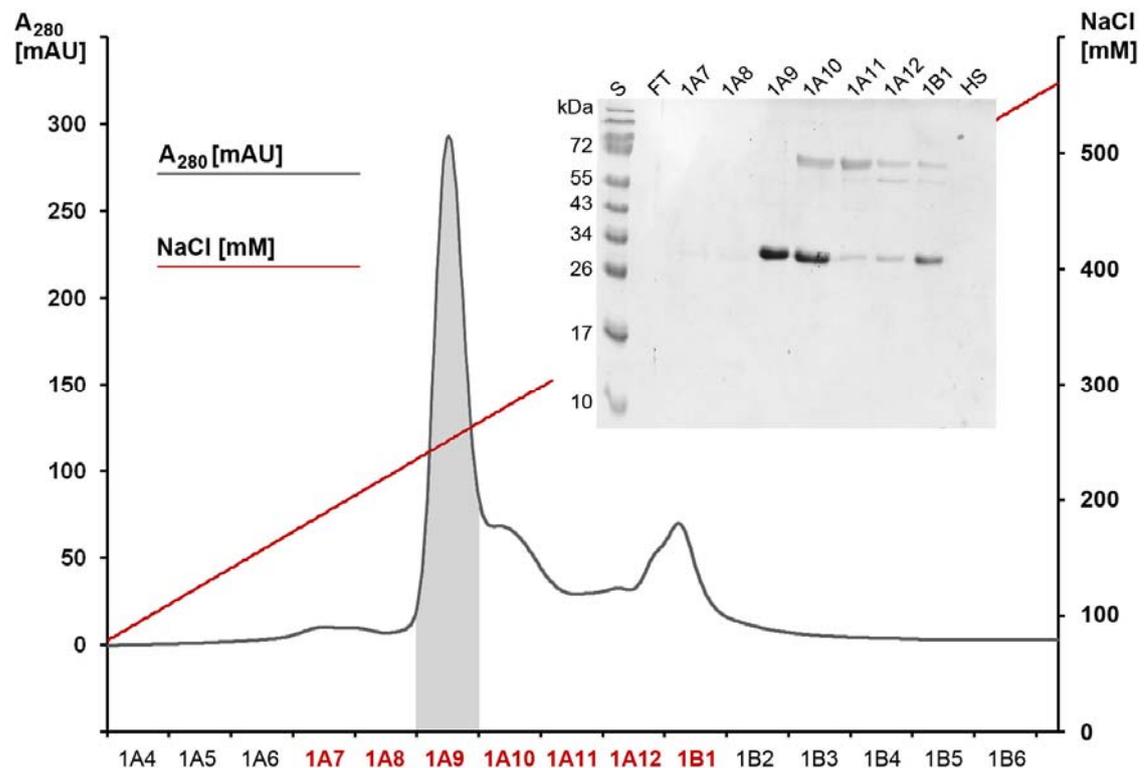
#### 3.12.4. Expression of recombinant TcCTLTPs and antibody generation

For the generation of recombinant TcCTLTP-5C<sub>2</sub> and TcCTLTP-6C proteins, pFBD-vectors were constructed containing the respective complete coding regions and used for the amplification of recombinant viruses in *Sf21* cells. Virus stocks were used to infect Hi-5 cells, which were incubated for 5 days before harvesting. The cell suspension was separated in pellet and supernatant fractions by centrifugation. After SDS-Page, proteins of both fractions were transferred to nitrocellulose membranes. Staining of the blot membranes with the anti-His antibody revealed a single immunosignal for TcCTLTP-5C<sub>2</sub> as well as TcCTLTP-6C in the pellet and in the supernatant fraction (exemplarily shown for TcCTLTP-5C<sub>2</sub>; Fig. 40 A). The immunosignals were migrating at about 30 kDa, which is in good agreement with the calculated theoretical mass of 28.0 kDa for TcCTLTP-5C<sub>2</sub> and 27.7 kDa for TcCTLTP-6C. Since the proteins were secreted into the insect cell media, the supernatants were purified by a Ni-NTA chromatography. Recombinant TcCTLTP-6C eluted at an imidazole concentration of about 0.2 M (Fig. 40 C). Ni-NTA purification was not sufficient for TcCTLTP-5C as the expression media contained 2.5% fetal calf serum (FCS), which apparently also bound to the column (upper band; Fig. 40 B). For this reason, recombinant TcCTLTP-5C<sub>2</sub> was additionally purified by anion exchange chromatography and was eluted at about 0.25 M NaCl (Fig. 41). Aliquots of the purified proteins were lyophilized and used as antigens for the generation of polyclonal antibodies. The specificity of each antibody was tested in immunoblots using supernatants of Hi5-cell cultures expressing TcCTLTP-5C<sub>2</sub> and TcCTLTP-6C. Both antibodies showed a strong preference for their own antigens, but the anti-TcCTLTP-6C antibody additionally detected the TcCTLTP-5C<sub>2</sub> antigen to some degree (Fig. 42).

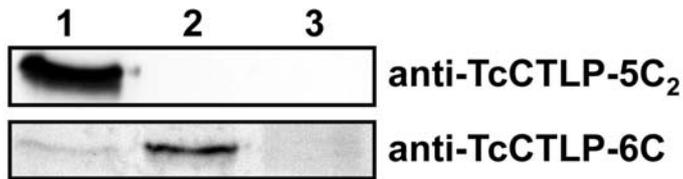
To test whether the purified peptidases exhibit chymotrypsin substrate specificity, recombinant TcCTLTP-5C<sub>2</sub> and TcCTLTP-6C were applied in a photometric assay using the synthetic substrate AAPF. Since the peptidases had been expressed as immature precursors, both of the TcCTLTPs had to be activated by trypsin. Recombinant TcCTLTP-6C showed no chymotrypsin activity and could also not be activated by the addition of trypsin to the reaction (data not shown). This indicates that either the wrong substrate was chosen in the photometric assay or recombinant TcCTLTP-6C was folded incorrectly by the Hi5-cells leading to a nonfunctional peptidase. However, recombinant TcCTLTP-5C<sub>2</sub> showed a sevenfold increase of chymotrypsin activity (309  $\mu\text{mol}/\text{min}/\text{mg}$ ) after activation by trypsin when compared to the non activated control (42  $\mu\text{mol}/\text{min}/\text{mg}$ ; Fig. 43). Hence its activity was in the range of bovine chymotrypsin (297  $\mu\text{mol}/\text{min}/\text{mg}$ ). Chymotrypsin activity of TcCTLTP-5C<sub>2</sub> was not affected by the addition of trypsin as the control reactions showed only very low activity when trypsin was added before or after heat inactivation of the samples (Fig. 43).



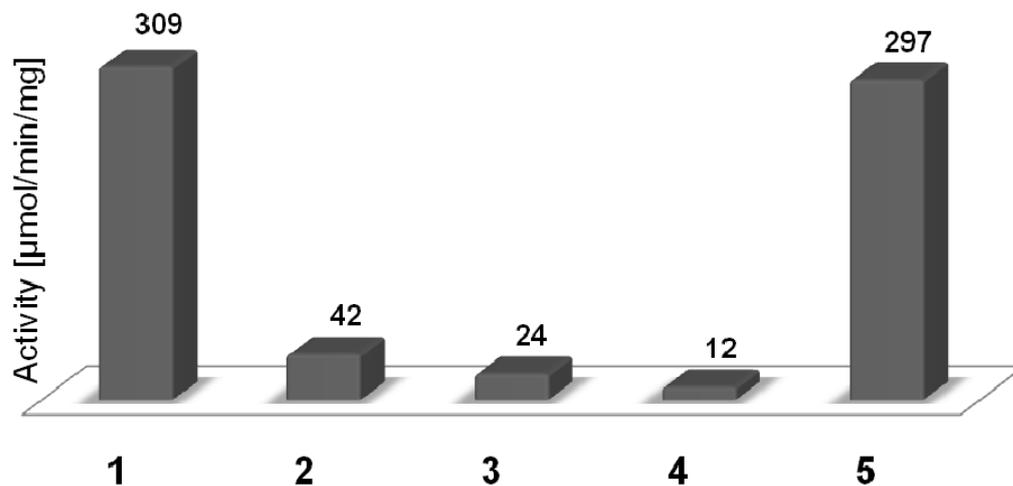
**Fig. 40.** Purification of two recombinant TcCTLTPs. (A) Hi5-cell cultures expressing recombinant TcCTLTP-5C<sub>2</sub> were separated in pellet (PE) and supernatant (SN) by centrifugation. Aliquots of both fractions were separated by SDS-Page and transferred to nitrocellulose membranes. The blot membranes were stained with the anti-His antibody. (B+C) cleared supernatants (SN) of Hi5-cell cultures expressing TcCTLTP-5C<sub>2</sub> (B) or TcCTLTP-6C (C) were submitted to a Ni-NTA column. After washing the column twice (W2) elution of bound proteins was accomplished by a step gradient of imidazole from 50 to 500 mM (50-500). FT, column flow trough fraction. Standard proteins are shown with molecular masses indicated in kDa (S).



**Fig. 41.** Anion exchange chromatography of TcCTLTP-5C<sub>2</sub>. Recombinant TcCTLTP-5C<sub>2</sub> was purified by Ni-NTA column and the 150-250 mM imidazole elution fractions were submitted to a Mono Q anion exchange chromatography. Aliquots of the elution fractions 1A7-1B1 were separated by SDS-Page and stained by Coomassie blue. As confirmed by SDS Page and the elution chromatogram (highlighted gray area), the fraction 1A9 contains the purified TcCTLTP-5C<sub>2</sub> without any contaminations. FT, column flow trough fraction; HS elution at high salt conditions (1 M NaCl). Standard proteins are shown with molecular masses indicated in kDa (S).



**Fig. 42.** Anti-TcCTLTP antibody specificity. To test the specificity of the anti-TcCTLTP-5C<sub>2</sub> and anti-TcCTLTP-6C antibodies, aliquots of supernatants from Hi5-cell cultures expressing TcCTLTP-5C<sub>2</sub> (1) and TcCTLTP-6C (2) as well as supernatants from untransfected Hi5 cells cultures (3) were separated by SDS Page. Following the transfer to nitrocellulose membranes, blot membranes were stained with the indicated antibodies.

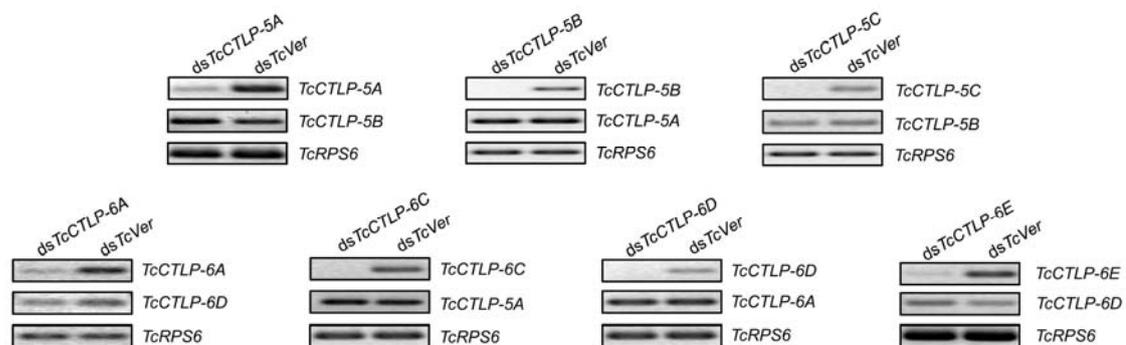


**Fig. 43.** TcCTLTP-5C<sub>2</sub> chymotrypsin activity assay. Recombinant TcCTLTP-5C<sub>2</sub> expressed by Hi5-cells was purified by Ni-NTA and anion exchange chromatography. The photometric chymotrypsin assay was performed using 0.1 μg/ml purified TcCTLTP-5C<sub>2</sub> and the substrate AAPF. 1, recombinant TcCTLTP-5C<sub>2</sub> activated by the addition of 0.01 μg recombinant trypsin; 2, recombinant TcCTLTP-5C<sub>2</sub> without the addition of trypsin; 3, recombinant TcCTLTP-5C<sub>2</sub> heat inactivated for 1 h followed by the addition of 0.01 μg recombinant trypsin; 4, recombinant TcCTLTP-5C<sub>2</sub> with the addition of 0.01 μg recombinant trypsin and heat inactivated for 1 h; 5, control using 0.1 μg/ml bovine chymotrypsin instead of recombinant TcCTLTP-5C<sub>2</sub>. Values above the bars indicate the activity of the respective sample ([μg/min/mg]).

### 3.13. RNAi of *Tribolium* CTLPs

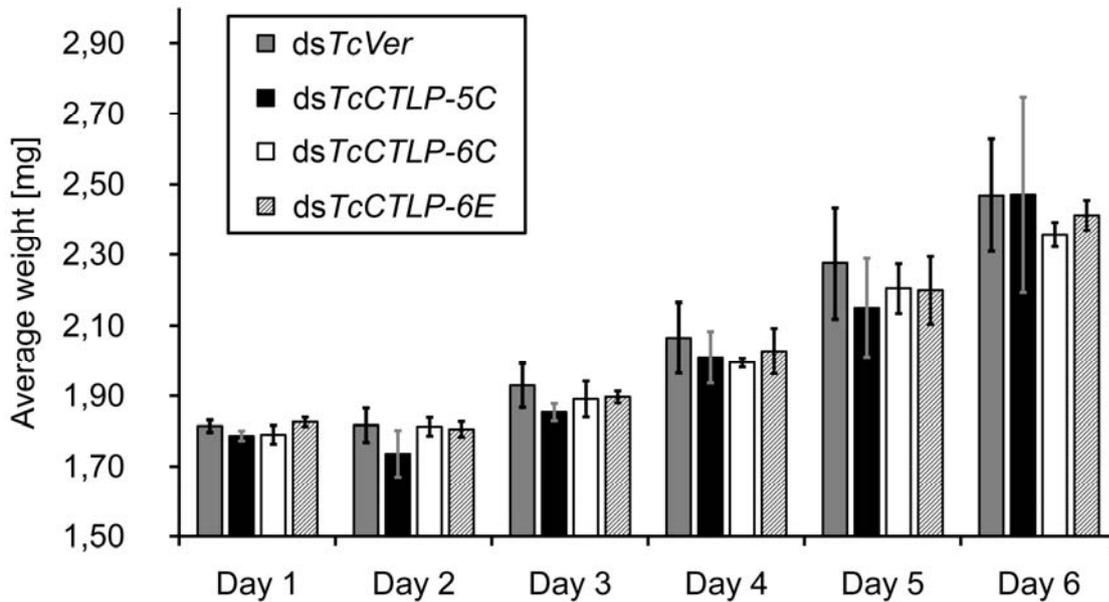
#### 3.13.1. Effects of dsRNA induced knockdown of *TcCTLP* transcripts

At least 30 penultimate instar larvae were injected with the respective *TcCTLP*-dsRNAs. Another set of control animals was injected with *TcVer*-dsRNA to knock down expression of the *tryptophan oxygenase* gene *vermilion*, leading to a white eye phenotype (Arakane *et al.*, 2008b; Lorenzen *et al.*, 2002). The effects of dsRNA injection on development and viability were monitored on a daily basis. RT-PCR with cDNAs from insects injected with *TcCTLP*-dsRNA into penultimate instar larvae was performed three days after injection. There was a significant reduction of transcript level when compared with animals that were injected with *TcVer* dsRNA as a control group (Fig. 44). To exclude any cross-reactions of dsRNA injections on the expression levels of the other *TcCTLP* genes, additional RT-PCRs were performed with primer pairs for closely related *TcCTLPs* (*TcCTLP-5B* for *TcCTLP-5A* and vice versa, nucleotide sequence identity: 47.7%; *TcCTLP-5B* for *TcCTLP-5C*, nucleotide sequence identity: 47.1%; *TcCTLP-6D* for *TcCTLP-6A* and vice versa, nucleotide sequence identity: 70.1%; *TcCTLP-5A* for *TcCTLP-6C*, identity: 53.9%; *TcCTLP-6D* for *TcCTLP-6E*, identity: 48.6% as determined using the ALIGN program; see also Table 3). None of these controls revealed any evidence of knockdown of non-target *TcCTLP* transcripts (Fig. 44).



**Fig. 44.** Effect of dsRNA injection to *TcCTLP-5A-C*, *TcCTLP-6A,C,D* and *TcCTLP-6E* on mRNA levels. dsRNA for *TcCTLP-5A-C*, *TcCTLP-6A,C,D* and *TcCTLP-6E* was injected into penultimate instar larvae and then total RNA was prepared three days after injections to the indicated gene (diagonal). The mRNA levels were determined by RT-PCR with primers to the indicated target gene (top panels). PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide. To test specificity of knockdowns, primers to closely related *TcCTLPs* were used to examine their mRNA levels by RT-PCR (middle panels). As a control RT-PCR with primers to the housekeeping gene *TcRSPS6* was performed on dsRNA injected animals (bottom panels).

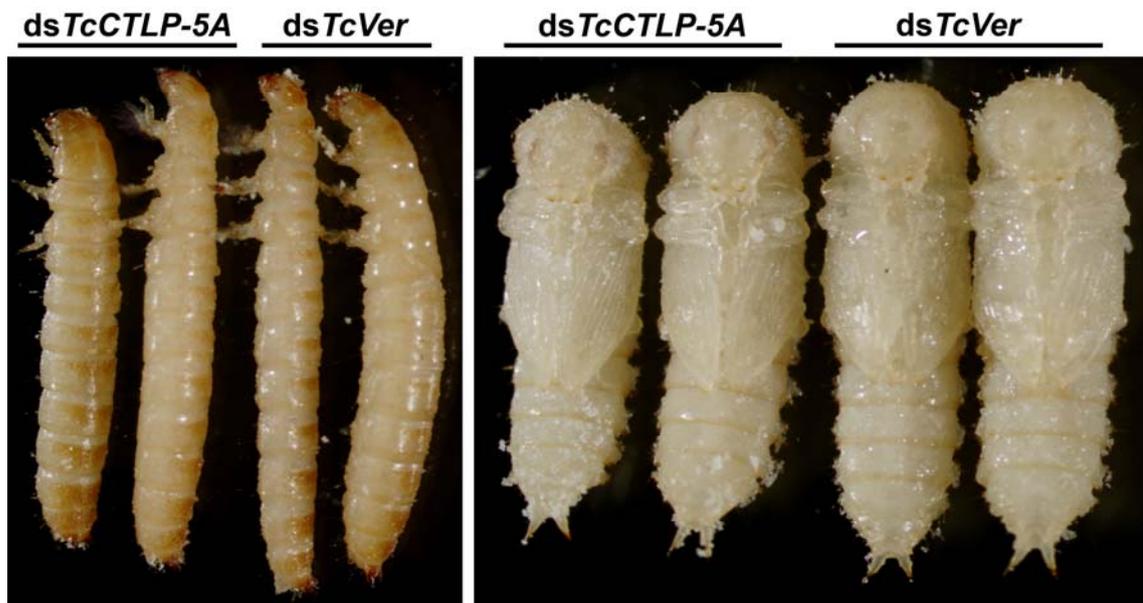
Additionally, the weights of the larvae were determined to test for possible effects on nutrient uptake. However, knockdown of any one of the seven *TcCTLPs* did not reveal any significant difference in the weight gain when compared with *TcVer*-dsRNA injected larvae (exemplarily shown for *TcCTLP-5C*, *TcCTLP-6C* and *TcCTLP-6E* dsRNA injections; Fig. 45).



**Fig. 45.** RNAi influence on body mass gain. At least 30 penultimate instar *Tribolium* larvae were injected with the indicated dsRNA. From the first to the sixth day post injection the body mass of the animals was determined by weighing each group of animals six times. Once the insects reached the prepupal stage they were omitted from the weighing process. Injection of dsRNA and weighing was repeated in independent experiments (mean values  $\pm$  S.E.,  $n = 5$ ).

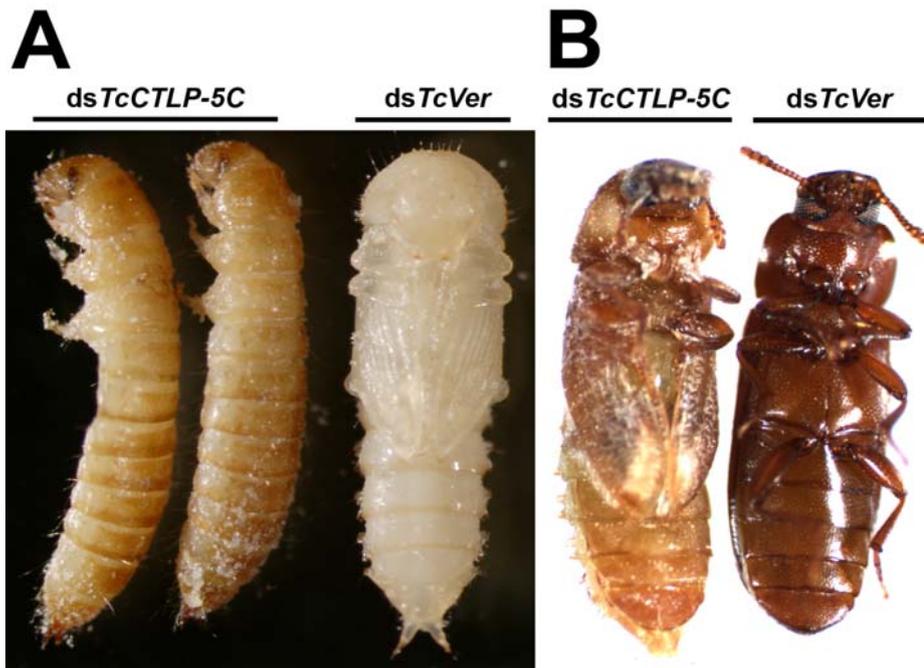
### 3.13.2. Knockdown of *TcCTLP-5C* transcripts affects larval-pupal and pupal-adult molt

Larvae injected with dsRNAs for *TcCTLP-5A* showed no abnormalities during pupal or adult development (Fig. 46). Similar results were obtained for injections of dsRNA for *TcCTLP-5B*, *TcCTLP-6A*, *TcCTLP-6D* and *TcCTLP-6E* (data not shown).



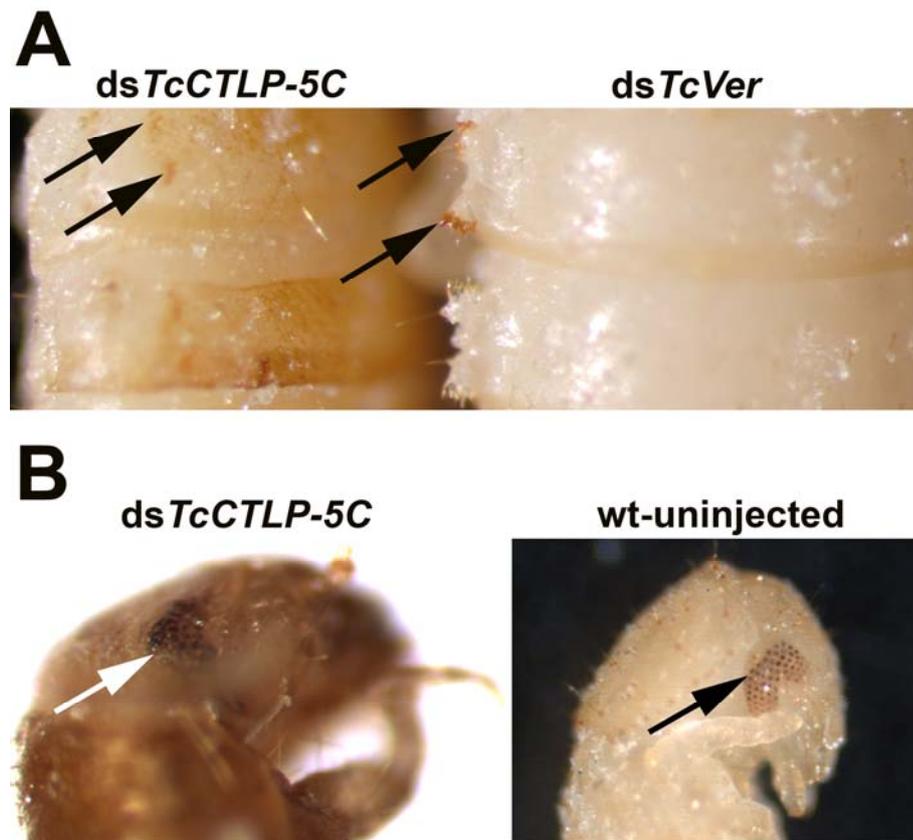
**Fig. 46.** Effects of injections of dsRNA for *TcCTLP-5A* on development. dsRNA for *TcCTLP-5A* and *TcVer* were injected into penultimate instar larvae and the phenotypes were monitored during development. *TcCTLP-5A*-dsRNA injected larvae developed normally into pupae and adults like the *TcVer*-dsRNA injected control animals, which exhibited, however, an expected no eye pigmented phenotype. Left panel: pharate pupae five days post dsRNA injection. Right panel: two days old pupae.

However, visual monitoring revealed distinct phenotypes for larvae that were injected with dsRNA for *TcCTLP-5C* and *TcCTLP-6C*. *TcCTLP-5C* dsRNA-injected animals developed normally until they reached the prepupal stage. Extensive pupal development proceeded under the old larval cuticle after ds*TcCTLP-5C* injection, but the insects died without molting (Figs. 47 A and 48). The expression data revealed that *TcCTLP-5C* transcripts are highly abundant at the P<sub>4</sub> and P<sub>5</sub> pupal stages just prior to the adult molt. Based on this observation it appeared likely that dsRNA injections at late prepupal and midpupal stages could also arrest the pupal-adult molt. To test this hypothesis, dsRNA for *TcCTLP-5C* was injected at these stages. About 50% of the dsRNA-*TcCTLP-5C*-injected prepupae and midpupae (two days old pupae) were indeed unable to continue development. They exhibited developmental arrest during the pupal-adult molt (Fig. 47 B).

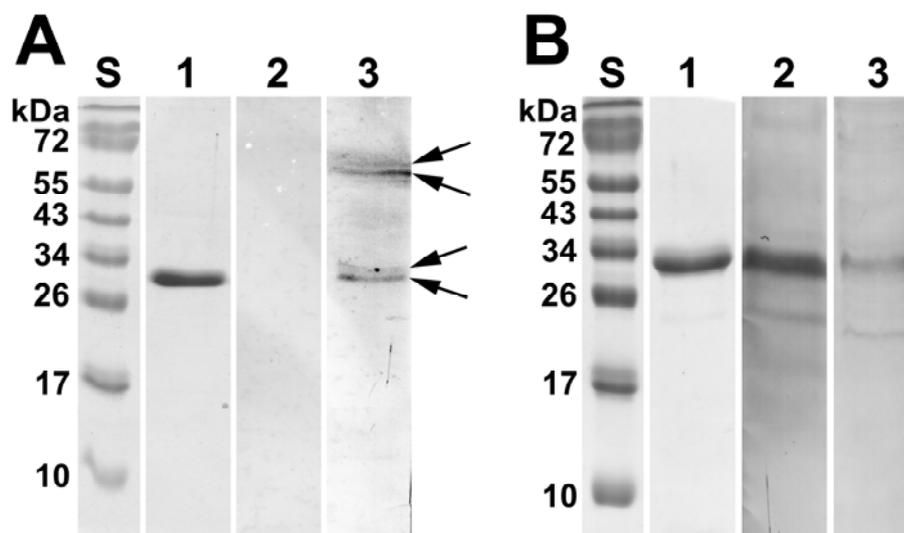


**Fig. 47.** Effects injections of dsRNA for *TcCTLP-5C* on development. dsRNA for *TcCTLP-5C* and *TcVer* were injected into penultimate instar larvae and midpupae (two days old pupae) and the phenotypes were monitored during development. (A) Injection of dsRNA for *TcCTLP-5C* into penultimate larvae led to a defective larval-pupal molt. (B) Injection of dsRNA for *TcCTLP-5C* into two days old pupae led to a defective pupal-adult molt. *TcVer* dsRNA injected into corresponding stages yielded normally developing animals except for the expected no eye pigmented phenotype.

As *TcCTLP-5C* is expressed in the carcass and its knockdown affects larval-pupal and pupal-adult molts, it apparently plays a role in molting. To support this hypothesis, immunoblots were performed with an anti-*TcCTLP-5C* antibody and proteins extracted from larval exuviae (Fig. 49 A, lane 3), to test whether expression of *TcCTLP-5C* is detectable in the integument. Indeed, the Western blot showed a *TcCTLP-5C*-specific signal with a protein of an apparent molecular mass of 30 kDa and a second signal with a protein of 60 kDa. Interestingly, the detected immunosignals appeared as double bands, presumably representing the two splice variants of *TcCTLP-5C*. In addition, immunoblots performed with proteins obtained from midgut preparations revealed no signal (Fig. 49 A, lane 2).



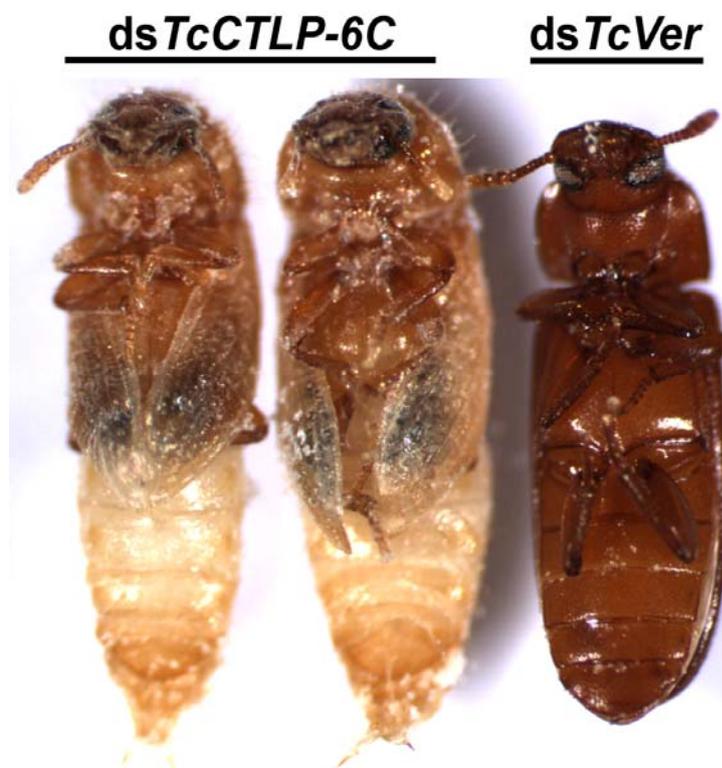
**Fig. 48.** Effects of injections of dsRNA for *dsTcCTLP-5C* on development. (A) Penultimate larvae injected with dsRNA for *TcCTLP-5C* developed pupal specific tissues such as the gin traps (black arrows) under the old larval cuticle (B) The old larval cuticle in the head region of the prepupa was removed exposing a complex eye (left panel, arrows), which normally is not fully developed before the pupal stage after same period of time in development (right panel, wildtype (wt)-uninjected).



**Fig. 49.** Immunoblots detecting *TcCTLP-5C<sub>1/2</sub>* and *TcCTLP-6C*. Proteins from midguts (lane 2) and larval exuviae (lane 3) were separated by SDS-PAGE, blotted onto nitrocellulose membranes and stained with antibodies to *TcCTLP-5C* (A) or *TcCTLP-6C* (B). Lane 1, recombinant *TcCTLP-5C<sub>2</sub>* (A) and *TcCTLP-6C* (B) used as antigens for antibody generation were separated by SDS-PAGE and stained with Coomassie blue. S, standard proteins with indicated molecular masses given in kDa.

### 3.13.3. Knockdown of *TcCTLP-6C* transcripts affects pupal-adult molting

Animals injected with *TcCTLP-6C* dsRNA at the larval stage developed normally through larval and early pupal stages. However, near the time of adult eclosion, almost all of the pharate adults failed to shed their old cuticle and consequently died (Fig. 50). As *TcCTLP-6C* also exhibits high expression in older adults (Fig. 30), it should be tested whether a knockdown of *TcCTLP-6C* affects the animals at the adult stage. Although RT-PCR proved that the knockdown in the beetles injected with *TcCTLP-6C* dsRNA shortly after adult eclosion had taken place, the insects showed no abnormal phenotypes and had the same life span as the *dsTcVer*-injected control animals (data not shown). The finding that *TcCTLP-6C* is highly expressed in the midgut suggests that it has a role in nutrient digestion. Additionally, the knockdown of *TcCTLP-6C* led to an arrest at the pupal-adult molt suggesting that it also functions in degrading the old endocuticle. Correspondingly, also lower transcript amounts could be detected in the carcass (Figs. 32 and 33). To detect the peptidase in the integument, larval exuviae were collected followed by extraction of the proteins to perform immunoblots with the anti-*TcCTLP-6C* antibodies. Indeed, an immunosignal for *TcCTLP-6C* was detected in this preparation (Fig. 49 B, lane 3). As expected from the expression profile obtained by RT-PCR (Figs. 32 ad 33), the signal intensity for the larval exuvial protein extract was significantly lower than that for the midgut protein extracts (Fig. 49 B, lane 2).



**Fig. 50.** Effect of injections of dsRNA for *TcCTLP-6C* on development. dsRNA for *TcCTLP-6C* and *TcVer* were injected into young larvae and the phenotypes were monitored during development. Injection of dsRNA for *TcCTLP-6C* led to a defective pupal-adult molt, while *TcVer*-dsRNA-injected control animals developed normally, except for the expected no eye pigmented phenotype.

## 4. Discussion

### 4.1. The structure function link

#### 4.1.1. Structure of *Manduca* chymotrypsin-like peptidases

Chymotrypsin like peptidases mainly exhibit S1-pocket specificity (Czapinska and Otlewski, 1999). The walls of this pocket are formed of three  $\beta$ -sheet regions (residues 189-192, 214-216, 225-228), which are connected by two surface loops (loop1: 185-188; loop2: 221-224) and the disulfide bond between Cys191 and Cys220 (Hedstrom, 2002; Hedstrom *et al.*, 1992; Perona and Craik, 1995; Perona and Craik, 1997; if not noted otherwise, chymotrypsin numbering according to Greer (1990) is used throughout). The specificity of the S1 pocket is primarily determined by the residue 189 at the pocket's bottom, whereas residues 216 and 226 are additional primary determinants located on one wall (Perona and Craik, 1997).

In the four *M. sexta* CTLPs identified in this work, the amino acids determining primary specificity are composed of Gly189, Gly216 and Asp226. This pattern was once considered as a rare variation of the Ser189, Gly216 and Gly216 motif typical for bovine chymotrypsin (Fig. 52; Hartley, 1964; Hartley and Kauffman, 1966) when it was first discovered in insects for the chymotrypsin C1 (GenBank|Q7SIG2.1) of the fire ant *Solenopsis invicta* (Botos *et al.*, 2000; Whitworth *et al.*, 1998). However, with increasing availability of insect ESTs and genomes, more and more insect CTLPs with S1 pocket residues matching this pattern have been discovered. In addition to the CTLPs from *M. sexta* and *S. invicta*, this includes serine peptidases from the orders hymenoptera (*Bombus ignitus* SP1; Choo *et al.*, 2007) diptera (*Anopheles gambiae* CT1; *A. gambiae* CT2; Vizioli *et al.*, 2001; *Aedes aegypti* CT2), siphonaptera (*Ctenocephalides felis* SP24; Gaines *et al.*, 1999), lepidoptera (*Spodoptera exigua* CT34; Herrero *et al.*, 2005) as well as several CTLPs from the coleopterans *T. castaneum* (TcCTLP-5A; TcCTLP-5B; TcCTLP-6C; this work) and *Rhyzopertha dominica* (Zhu and Baker, 2000; Fig. 52). Assays using synthetic substrates provided evidence of chymotrypsin activity for *S. invicta* chymotrypsin C1, *A. gambiae* CT1 and *S. exigua* CT34 (Herrero *et al.*, 2005; Vizioli *et al.*, 2001; Whitworth *et al.*, 1998).

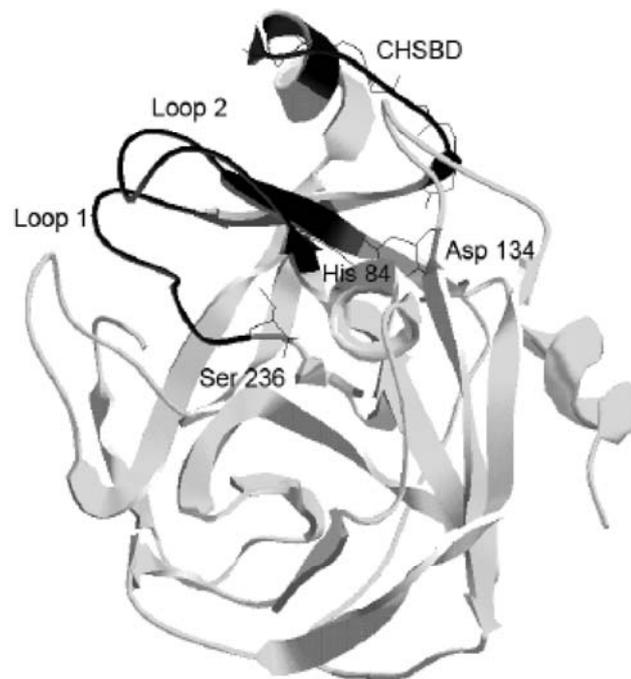
To estimate the substrate specificity of MsCTLP1-4, the residues accounting for chymotrypsin activity were compared to those of the CTLPs mentioned above. The S1 pockets of MsCTLP1-4 are composed strikingly similar, probably exhibiting comparable substrate specificities and seem to be closer related to other insect CTLPs than to bovine chymotrypsin (Fig 52). Compared to Ser189 found in bovine chymotrypsin, Gly189 at the S1 pocket's bottom of MsCTLP1-4 presumably allows a higher conformational flexibility (Betts and Russell, 2003) making the left-hand side of the S1 primary specificity site somewhat wider than in its bovine counterpart. Of the residues Val213, Gly216, Gly226, Ser190 and Tyr228, which define the shape and the size of S1 pocket's entrance in bovine chymotrypsin (Botos *et al.*, 2000), Val213 and Gly216 are found unchanged in MsCTLP1-4. The occurrence of Thr190 and Phe228 in MsCTLP1-4 can be considered as fairly neutral changes, as the former amino acid

introduces a methyl group in place of a hydrogen group found in serine while the latter one lacks the hydroxyl group in the *ortho* position on the benzene ring of tyrosine. The presence of Asp226 in one wall of the S1 site of MsCTLP1-4 probably has a much greater influence on the accessibility of the pocket's entrance. Thus the right-hand side of the S1-specificity pocket is most likely slightly bulkier, probably hindering large substrate side chains to access the pocket's bottom. However, the Asp226 of MsCTLP1-4 might also rotate aside allowing the S1 specificity pocket to accept the large aromatic residues of phenylalanin, tyrosin and tryptophan as revealed by the crystal structure for the *S. invicta* chymotrypsin C1 (Botos *et al.*, 2000).

Specificity is further determined by the interaction of surface loop1, loop2 and loop3 with the S1-site (Hedstrom, 2002; Perona and Craik, 1997; see 1.2). The loop1 and loop2 regions of MsCTLP1-4 show little homology to other insect CTLPs and differ in particular to those of bovine chymotrypsin. Additionally, loop1 is elongated in comparison to the corresponding region of the bovine equivalent (Fig. 52). Among the *Manduca* CTLPs, the surface loops show the highest variation between those of MsCTLP1 and MsCTLP2-4, whereas the latter are almost identical with the exception of those of MsCTLP2, which has an Arg186 instead of the Glu186 found in MsCTLP3-4. The amino acids Gly185 and Val188 present in loop1 of MsCTLP1 instead of the corresponding Ser185 and His188 residues found in MsCTLP2-4 most likely changes the loop's structure. Especially the latter replacement of a polar residue by an aliphatic, hydrophobic residue probably leads to an altered interaction of the loop with the S1-site, resulting in modified substrate specificity. This difference in interaction is probably further enhanced by the presence of Pro186, uniquely found in loop1 of MsCTLP1, as proline is known to introduce very tight turns into protein structures (Betts and Russell, 2003). Also loop2 differs notably between MsCTLP1 and MsCTLP2-4. Whereas loop2 of the latter enzymes consists of the residues Ala-Leu-Gly-Ala, the corresponding amino acids are Ala-Arg-Gly-Gly in MsCTLP1. Especially Arg222 changes the properties of loop2 of MsCTLP1 as it introduces a positive charge into the sequence, while loop2 of MsCTLP2-4 consists of non-polar amino acids. As the two loops connecting the S1-site do have a significant effect on enzyme activity and substrate specificity (Ma *et al.*, 2005), MsCTLP1 probably has a role in the midgut of *M. sexta* which is distinct from those of MsCTLP2-4. This assumption is supported by the finding that MsCTLP1 seems to be involved in the regulation of midgut specific chitin-synthase 2 (MsCHS2; see 4.2).

In this context, homology based 3D-structure modeling of MsCTLP1, calculated by using the crystal structures of *Sus scrofa* trypsin and *S. invicta* chymotrypsin C1, revealed an unusual surface structure (Fig. 51), possibly involved in the interaction with MsCHS2. This structure includes the surface loop3, which is elongated in the four MsCTLPs with little sequence homology to other chymotrypsin-like peptidases. Loop3 further lacks Trp172, which was demonstrated to be an important substrate specificity determinant (Hedstrom *et al.*, 1994). Because of the missing homologies, this loop could only be partially displayed in the 3D-structure model. While the sequence of loop3 is to some degree conserved among MsCTLP2-4, the corresponding amino acids differ in MsCTLP1, additionally exhibiting an accumulation of proline residues in the

sequence following the loop3. The finding that proline might play an important role in molecular recognition (Betts and Russell, 2003) becomes all the more interesting when one considers that MsCTLP1 co-localizes and binds to the extracellular, carboxy-terminal domain of zymogenic MsCHS2 (see 4.2). Together with the elongated loop region, this proline-rich stretch is possibly responsible for the binding to MsCHS2, and was hence suggested to be a chitin synthase binding domain (CHSBD; Figs. 51 and 52 highlighted region).



**Fig. 51.** Homology based model of the 3D structure of MsCTLP1. The structure-model was created using Swiss-Model, based on the crystal structures of chymotrypsin C1 of *Solenopsis invicta* (GenBank|Q7SIG2.1) and trypsin of *Sus scrofa* (GenBank|NP\_001156363.1). CHSBD (corresponding amino acids highlighted in Fig. 52), the potential chitin synthase binding domain, as well as loop1 and loop2 connecting the S1 specificity pocket are highlighted in black. His84, Asp134 and Ser236 (numbering according to the sequence of MsCTLP1), are amino acids of the catalytic triad. CHSBD is only partially displayed in the structure model because of missing homologies to other CTLPs thus allowing no precise prediction of the structure.

#### 4.1.2. Maintenance of catalysis at alkaline conditions

The exposure to an extremely alkaline milieu of the lepidopteran midgut of pH 10-12 (Dow, 1984) requires adaption mechanisms of the luminal serine peptidases in order to maintain enzyme's function. The alkaline conditions in the midgut do have a significant effect on the ionization states of Ile16 and His57 and hence on the functionality of the enzyme. At high pH, Ile16 (pKa = 9.6 when Ile16 is at the amino-terminus) is deprotonated, which results in the disruption of the salt bridge to Asp194 (pKa = 3.7) and finally shifts the conformational equilibrium toward an inactive zymogen-like conformation (Hedstrom, 2002). Above pH 9.5 His57 (pKa = 6.0) is deprotonated in bovine chymotrypsin (Rogez *et al.*, 1987) and thus affects proton transfer of the charge relay system. Maintenance of the catalysis at such a high alkaline

pH was previously suggested to be due to the larger number of arginine residues ( $pK_a = 12.5$ ) stabilizing the structure of lepidopteran chymotrypsins (Peterson *et al.*, 1995). Analysis of the amino acid content of lepidopteran chymotrypsin-like peptidases from different species showed that the arginine content varies between 2.10% and 9.00% (mean  $4.80 \pm 1.63$ ,  $n=19$ ) and the lysine content between 0% and 2.40% (mean  $0.53 \pm 0.72$ ,  $n=19$ ). However, a greater variability in the content of arginine and lysine residues can also be observed for dipteran chymotrypsin like peptidases, which exhibit lower pH values in the midgut. Here the arginine content varies between 0.40% and 5.30% (mean  $3.70 \pm 1.58$ ,  $n=11$ ) and the lysine content between 0.90% and 7.10% (mean  $3.90 \pm 1.93$ ,  $n=11$ ). Overall, the arginine content appears indeed slightly higher and the lysine content lower in lepidopteran than in dipteran chymotrypsin-like peptidases. However, as there are obvious exceptions, a general rule cannot be established, particularly because the arginine content of lepidopteran hemolymph peptidases is also high (mean  $4.55 \pm 1.30$ ,  $n=17$ ). Even more striking is that the lysine content of serine peptidases adapted to highly alkaline lepidopteran gut appears low, while it is high in lepidopteran hemolymph serine peptidases (mean  $5.79 \pm 0.66$ ,  $n=17$ ).

Preservation of proton transfer in the charge relay system at highly alkaline conditions is probably kept up by amino acid replacements in the neighborhood of His57 at the positions 59, 90, 103 and 213. Substitution of these residues, especially the occurrence of Trp59 in lepidopteran chymotrypsins, are thought to affect the  $pK_a$  value of His57 in such a way that it can adapt to the midgut's high pH (Lopes *et al.*, 2009). However, several lepidopteran peptidases are an exception to these observations, in particular MsCTLP1-4 as well as SeCT34 (Herrero *et al.*, 2005) as they lack Trp59 (Fig. 52). These peptidases probably have different ways of maintaining the enzyme's functionality at alkaline conditions as they exhibit Ile59, which, together with the similar amino acids leucine and valine, is most frequently found in insect CTLPs at this position (Fig. 52). In line with these considerations, sequence analysis revealed that MsCTLP1-4 as well as SeCT34 fit perfectly into a novel subgroup of insect CTLPs which is defined by sequence similarity and by the replacement of GP by SA in the highly conserved GDSGGP motif (Figs. 20 and 52). The GP→SA substitution appears to be one of the rare variations in this highly conserved region, which leads to a functional enzyme. Members of this CTLP subgroup have been suggested to exert specific functions rather than being involved in general digestion of dietary proteins (Herrero *et al.*, 2005). However, it also seems to be conceivable that the GP→SA substitution might affect Asp194 in such a way that it is able to maintain its salt bridge to Ile16 even at high alkaline conditions.

#### 4.1.3. Structure of *Tribolium* chymotrypsin-like peptidases

To identify putative CTLP genes in the *Tribolium* genome (Richards *et al.*, 2008), a Blast search was performed that yielded about 160 genes encoding S1A peptidases from which those peptidases were selected containing S1 pocket residues typically found in chymotrypsins (Perona and Craik, 1995). The search pattern was extended for Gly/Ser/Thr189, Gly216, Asp/Gly/Ala/Ser226, so as to include also the above

mentioned MsCTLTPs and other insect CTLTPs with non-classical S1 pocket residues. In doing so, 14 *TcCTLTP* genes were identified in the *Tribolium* genome from which seven were selected for further characterization. As reported for other insect genomes, genes encoding CTLTPs constitute only a small fraction of the total number of genes encoding S1A peptidases in *Tribolium*. This seems to be the case also in *M. sexta*, as genomic Southern Blots revealed up to eight gene copies encoding CTLTPs in the genome (Fig. 21). However, the possibility cannot be excluded that there are additional peptidases with chymotrypsin-like activities in *Tribolium* that were missed in the analysis due to the presence of non-canonical S1 pocket residues.

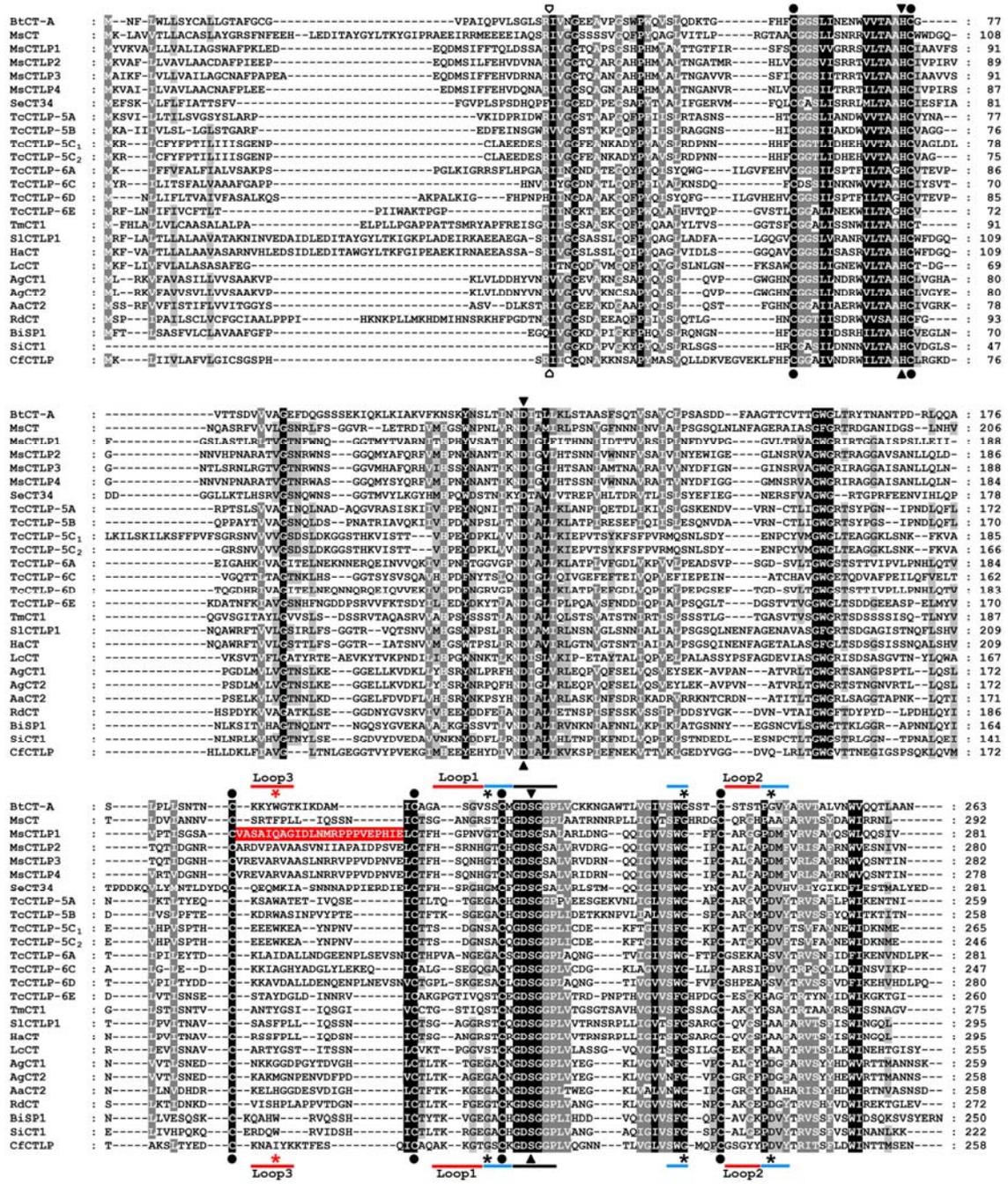
From the residues determining substrate selectivity and shaping the S1-site of the seven selected *TcCTLTPs* only Val213 and Gly216 are conserved throughout, whereas the composition of the remaining residues varies considerably (Table 2; Fig. 52). Thus *TcCTLTP-5C/-6D/-6E* exhibit Ser189 at the pocket's bottom as described for bovine chymotrypsin (Hartley, 1964; Hartley and Kauffman, 1966), whereas Gly189 is accounting for a relaxed specificity in *TcCTLTP-5A/-5B/-6A/-6C* (Botos *et al.*, 2000). Ala190 present in almost all seven *TcCTLTPs*, is slightly less voluminous than Ser190 found in bovine chymotrypsin and therefore probably widens the S1-site as described for chymotrypsin S1 of *S. invicta* (Botos *et al.*, 2000). Only *TcCTLTP-6E* deviates from this observation since it has Thr190 at this position, which was also observed for *Manduca* CTLTP1-4 and probably reflects a neutral exchange of the bovine motif. Tyrosine at position 228 present in bovine chymotrypsin was the same for most of the seven *TcCTLTPs* or was replaced in *TcCTLTP-5C* and *TcCTLTP-6E* by the very similar amino acid phenylalanine, which lacks the reactive hydroxyl group. Ala226 present in *TcCTLTP-6E* and Ser226 found in *TcCTLTP-6A* and *TcCTLTP-6D* occupy almost the same space and are slightly bulkier than Gly226 found in the bovine equivalent (Chothia, 1976; Zamyatin, 1972). Like in *Manduca* CTLTP1-4, this position is occupied by Asp226 in all *TcCTLTPs* of chromosome five as well as in *TcCTLTP-6C*. As discussed above, the aspartate residue may hinder large aromatic residues to enter the S1-site. However, the photometric assay using and the synthetic substrate AAPF, clearly showed chymotrypsin activity for recombinant *TcCTLTP-5C<sub>2</sub>* after activation by trypsin (see 3.12.4). This on the one hand proves that zymogen activation of *TcCTLTP-5C<sub>2</sub>* is indeed trypsin mediated. As *TcCTLTP-5C<sub>2</sub>* activity was in the range of bovine chymotrypsin, these findings on other hand demonstrate that an enzyme exhibiting a Ser189-Gly216-Asp226 configuration, a configuration which defines a pocket that is presumably narrower than that of bovine chymotrypsin and even that of *MsCTLTP1-4*, is able to cleave chymotrypsin substrates.

In *Tribolium*, the three surface loops involved in interaction with the S1-site (see 1.2) show little sequence homology among the seven *TcCTLTPs* as well as to other CTLTPs (Fig. 52) allowing only little interpretation about the structure-function relationship. However, loop2 is to some degree conserved throughout the *TcCTLTPs* of chromosome five, whereas the corresponding residues show a higher variation in the *TcCTLTPs* of chromosome 6, which additionally exhibit amino acid insertions in the case of *TcCTLTP-6A* and *TcCTLTP-6D*. Also loop1 regions of *TcCTLTP-6A*, *TcCTLTP-6D* and *TcCTLTP-6E* are significantly elongated. This finding appears unusual, as Ma *et al.*

(2005) demonstrated by sequence alignment that the probability that loop1 exhibits four or five residues is 54% or 46%, respectively. Furthermore, Ma and colleagues found loop2 regions to be exclusively composed of four residues in all cases investigated. These amino acid insertions in loop1 and loop2 most likely lead to heterogeneous interactions with the S1-site, possibly pointing to different functions these enzymes accomplish in *Tribolium*. Most notably, the loop3 regions of TcCTLP-5A, TcCTLP-5B and TcCTLP-5C exhibit Trp172, an important determinant of chymotrypsin specificity (Hedstrom *et al.*, 1994). Among other reasons this finding might explain, why recombinant expressed TcCTLP-5C<sub>2</sub> exhibited chymotrypsin activity, whereas TcCTLP-6C, which lacks the Trp172 in its sequence, showed no chymotrypsin activity with the same substrate (see 3.12.4). On the other hand, insect chymotrypsins were demonstrated to prefer smaller substrates like N-benzoyl-L-Tyr p-nitroanilide in contrast to bovine chymotrypsin, which favors larger substrates like the AAPF used in this work (Lopes *et al.*, 2009).

Additionally, loop3 of TcCTLP-5C shows an unusual accumulation of charged amino acids, especially glutamate, creating a net negative charge. Possibly this specific feature leads to a unique substrate selectivity which might be vital for TcCTLP-5C's specialized function in molting (see 4.4). In this context it seems to be conceivable that the occurrence of two splice variants for TcCTLP-5C might be related to the enzyme's role in ecdysis. However, a blastp search with the 19 amino acids sequence present in TcCTLP-5C<sub>1</sub> but missing in TcCTLP-5C<sub>2</sub> revealed no hit to any conserved motif hence making it difficult to draw any conclusions about its function.

Although chymotrypsin-like peptidases are one of the best characterized protein families, the interaction of the structural components defining specificity is still poorly understood, making purification or recombinant expression of the respective peptidases still the most reliable way to predict substrate selectivity. However this is quite laborious and time consuming and therefore data derived from protein analysis may help to create algorithms that allow predicting the behavior of amino acids within the polypeptide.



**Fig. 52.** ClustalW alignment of chymotrypsin-like peptidases. Conserved, highly conserved or identical amino acids are highlighted in light-gray, gray or black, respectively. Open arrow head, putative zymogen activation sites; black asterisks, conserved S1 pocket residues; red asterisks, important residue for substrate selectivity; closed circles, conserved cysteines; triangles, conserved residues of the catalytic triad; black line, highly conserved GDSGGP motif; red lines, surface loop regions; blue lines,  $\beta$ -sheet regions forming the S1-specificity pocket; sequence highlighted in red, putative chitin synthase binding region; CT, chymotrypsin; CTLP, chymotrypsinogen like peptidase; SP, serine peptidase; Bt, *Bos taurus* (GenBank|XP\_608091.3); Ms, *Manduca sexta* (MsCT (GenBank|AAA58743.1), MsCTLP1 (embl|CAL92020.1), MsCTLP2 (embl|CAM84317.1), MsCTLP3 (embl|CAM84318.1), MsCTLP4 (embl|CAM84319.1)); Se, *Spodoptera exigua* (GenBank|AAX35812.1); Tc, *Tribolium castaneum* (TcCTLP-5A (embl|CBC01177.1), TcCTLP-5B (embl|CBC01166.1), TcCTLP-5C<sub>1</sub> (embl|CBC01175.1), TcCTLP-5C<sub>2</sub> (embl|CBC01172.1), TcCTLP-6A (embl|CBC01171.1), TcCTLP-6C (embl|CBC01169.1), TcCTLP-6D (embl|CBC01179.1), TcCTLP-6E (embl|CBC01181.1)); Tm, *Tenebrio molitor* (GenBank|ABC88746.1); Sl, *Spodoptera litura* (GenBank|ACU00133.1); Ha, *Helicoverpa armigera* (GenBank|ACB54940.1); Lc, *Lucilia cuprina* (GenBank|AAA68986.1); Ag, *Anopheles gambiae* (AgCT1 (embl|CAA83568.1), AgCT2 (embl|CAA83567.1)); Aa, *Aedes aegypti* (GenBank|AAF43707.1); Rd, *Rhizopertha dominica* (GenBank|AAD31187.1); Bi, *Bombus ignitus* (GenBank|ABF39002.1); Si, *Solenopsis invicta* (GenBank|Q7SIG2.1); Cf, *Ctenocephalides felis* (gb|AAD21831.1).

## 4.2. Involvement in chitin synthesis

In insects, trypsin-mediated stimulation of chitin synthesis has been reported, in support of the enzyme's zymogenic nature (Cohen and Casida, 1980; Mayer *et al.*, 1980; Ward *et al.*, 1991). Chitin synthesis is also stimulated in the midgut of the tobacco hornworm upon trypsin treatment. However, trypsin does not directly act on chitin synthase but on a soluble factor, which in turn affects chitin synthase activity (Zimoch *et al.*, 2005). This finding implicates that trypsin-mediated activation of chitin synthesis involves either an inhibitor protein of chitin synthase that is inactivated, or an activator protein that is activated by tryptic cleavage.

Yeast two-hybrid screen on a *Manduca* midgut cDNA library identified a novel chymotrypsin-like peptidase (MsCTLP1) that interacts with the extracellular carboxy-terminal domain of MsCHS2 *in vitro* (Broehan *et al.*, 2007). Several lines of argument suggest that MsCTLP1 also interacts with MsCHS2 *in vivo*. Firstly, MsCTLP1 is expressed in the same tissues as MsCHS2, predominantly in the anterior midgut. Secondly, MsCTLP1 is evidently secreted into the gut lumen; thus, it can interact with the extracellular C7 domain of MsCHS2 (see Fig. 11). Thirdly, MsCTLP1 and MsCHS2 co-localize at the brush border microvilli of the columnar cells (Figs. 15 and 16 B+C). Fourthly and lastly, co-immunoprecipitation experiments demonstrated that the two proteins interact specifically in midgut cell lysates (Fig. 17).

Most important, however, is the fact that chymotrypsins are produced as zymogens, which are activated by tryptic cleavage (Kraut, 1977; Law *et al.*, 1977; Lehane *et al.*, 1996). The highly conserved trypsin cleavage site is also present in the MsCTLP1 amino acid sequence, suggesting that the precursor of MsCTLP1 is activated by trypsin (Figs. 12 and 13). Tryptic cleavage of MsCTLP1 is additionally supported by the immunological detection of the proteolytically processed MsCTLP1 in the gut contents, which exhibits a significantly smaller molecular mass than MsCTLP1 detected in extracts from midgut epithelial cells (Fig. 16 A). However, yet there is no direct evidence that MsCTLP1 exhibits a proteolytic activity that is induced by tryptic cleavage. Nevertheless, its binding to MsCHS2 may be the first clue for the participation of MsCTLP1 in the proteolytic stimulation of chitin synthesis (Zimoch *et al.*, 2005).

For control of the activity of midgut chitin synthase, a simplified and hypothetical model appears to be plausible in which trypsin activates the zymogenic form of MsCTLP1, being aware that the activation of MsCTLP1 might require further peptidases in addition to trypsin. However MsCTLP1 is secreted and proteolytically activated and the experimental data suggest that it binds to the extracellular C7 domain of MsCHS2. This interaction either could result in a direct conformational change inducing chitin synthase activity, or might be a crucial step in the processing of MsCHS2 involving the proteolytic activity of MsCTLP1.

In order to test whether MsCTLP1 directly cleaves MsCHS2, purification of zymogenic MsCTLP1 from the midgut tissue of *M. sexta* was performed (see 3.12.1). Since the purification proved to be difficult, MsCTLP1 was expressed as an active, recombinant enzyme in various systems (see 3.12.2). Although MsCTLP1 could be

purified from Sf21 cells expressing the recombinant peptidase, it has been so far not possible to refold it sufficiently to obtain an enzyme exhibiting significant proteolytic activity. However, previous results showed that chitin synthesis is stimulated in the presence of bovine chymotrypsin, suggesting that a chymotrypsin-like activity is involved in the activation of chitin synthesis (Broehan *et al.*, 2007). Chymotrypsin stimulated chitin synthesis comparably as well as trypsin, which was shown previously to stimulate chitin synthesis in *Manduca* midgut extracts (Zimoch *et al.*, 2005). Since Pefabloc SC, a known inhibitor of serine peptidases, impairs chitin synthesis, proteolytic activity appears to be necessary for chitin synthesis (Broehan *et al.*, 2007).

The hypothetical model for the proteolytic activation of chitin synthesis is attractive, for it couples the control of chitin synthase activity in the midgut to the nutritional state of the larvae. This is because trypsins are known to be secreted in response to dietary protein entering the midgut (Law *et al.*, 1977; Lehane *et al.*, 1996). Even though the precise mechanisms that lead to the activation of gut peptidases in response to nutrient uptake are not fully understood, particularly not in insects, it seems clear that trypsins initiate the activation of other gut peptidases such as chymotrypsins (Lehane *et al.*, 1996). Insect midgut peptidases may therefore act not only to digest nutrients but also to specifically modulate non-digestive gut enzymes such as chitin synthase, which is necessary for peritrophic matrix production. This notion may also be important for a better understanding of the structural and functional diversity of insect midgut peptidases, particularly observed in lepidopteran systems (Srinivasan *et al.*, 2006). Thus, peptidase inhibitors produced by plants as a defence against herbivorous insects could interfere not only with digestive gut peptidases but also with intestinal proteolytic signalling cascades controlling chitin synthesis and thus peritrophic matrix formation, which is necessary to protect the insect digestive tract from mechanical damage and infection by pathogens.

### 4.3. Digestion and beyond

A broad diversity of serine peptidases has been reported for different lepidopteran insect guts. For instance, *H. amigera*, a pest on numerous host plants, produces more than twenty different serine peptidases, of which seven have a chymotrypsin-like S1 specificity pocket. All of these peptidases have been reported to be active in the midgut accounting for 95% of total digestive activity (Srinivasan *et al.*, 2006). The legitimate question arises why so many serine peptidase isoforms are expressed at the same time. One plausible explanation may be the adaptation to (1) differences in the host-plants' nutrient composition and (2) to the exposure to different plant-derived peptidase inhibitors (Brito *et al.*, 2001; Jongsma *et al.*, 1995; Lopes *et al.*, 2004; Paulillo *et al.*, 2000; Srinivasan *et al.*, 2006). This explanation might be sufficient for *Helicoverpa*, which is a polyphagous insect facing different dietary compositions and peptidase inhibitors. Indeed, several studies showed that the composition of the peptidase cocktail in the gut changes depending on the type of plant fodder and the given peptidase inhibitors (Srinivasan *et al.*, 2006). In *Manduca* the situation is different, as the larvae

are oligophagous during their early life stages but monophagous at the later larval stages. Thus, the question arises what is the function of the complex proteolytic environment in the *Manduca* midgut. Expression of a larger set of digestive peptidases might be beneficial to monophagous insects in order to adapt to poorly digestible substrates as the host plants probably co-evolve in response to the herbivores. On the other hand, a great number of peptidases presumably allows at least some of them to differentiate into non-digestive functions.

In this work four *M. sexta* chymotrypsin-like serine peptidases have been identified (MsCTLP1–4), which are expressed by midgut epithelial cells of 5<sup>th</sup> instar larvae. Considering the previously described chymotrypsin (MsCT; Peterson *et al.*, 1995) as well as three trypsin isoforms (TRPA/B/C; Peterson *et al.*, 1994) and one trypsin-like peptidase, which has been cloned previously (embl|AM690450; Broehan *et al.*, 2008), at least nine different serine peptidases are expressed in the larval midgut of *M. sexta*. Genomic Southern blotting further revealed up to eight gene copies encoding MsCTLPs in the genome (Fig. 21). *Manduca* evidently produces a complex serine peptidase cocktail possibly expressing at least twelve different trypsin- and chymotrypsin-like peptidases. The proteolytic environment is complex although the animals face only one type of a well-defined, synthetic diet since years in the laboratory culture. Thus, changes in nutrient composition and peptidase inhibitors may be not sufficient in this case to explain the expression of as many different gut peptidases.

Contrary to expectations, the expression of MsCT, MsCTLP1, MsCTLP3 and MsCTLP4 in the midgut is down-regulated during molting but not during periods of starvation, while that of MsCTLP2 is down-regulated during molting and starvation. In addition, MsCT and MsCTLP2 mRNA amounts do not recover upon re-feeding while those of the other MsCTLPs reached almost the levels that they had before starvation (Fig. 24). These different types of responses may also point to non-redundant functions of the investigated MsCTLPs. Most notable, the larvae suspend secretion of MsCTLP1 during periods of molting and starvation (Figs. 25 and 26 A). Similar results were obtained for the larvae of *Mamestra configurata*, where the mRNA levels of two peptitrophic matrix associated trypsins (McSP1; GenBank|ACR15970.1 and McSP2; GenBank|ACR16004.1) and chymotrypsins (McSP25; GenBank|ACR15984.1 and McSP29; GenBank|ACR15980.1) showed no significant reduction in comparison to feeding animals when the larvae were starved. Unlike for the MsCTLPs, the four serine peptidases of *M. configurata* are not down-regulated upon molt. In line with this finding, the total serine peptidase activity of the of the midgut was measured to be the same in feeding and molting animals whereas starving larvae had approximately half the level of activity (Toprak *et al.*, 2009). As a common feedback to starvation both lepidopteran species do not down-regulate the respective serine peptidase mRNAs and MsCTLP secretion is furthermore suspended whereas the total serine peptidase activity of *M. configurata* is halved possibly because of the same reason. This analogy may be explained by the special biology of lepidopteran larvae. In contrast to blood sucking mosquitoes, for instance, whose trypsin and chymotrypsin mRNAs are up-regulated upon feeding and down-regulated during starvation (Borovsky, 2003), lepidopteran larvae usually do not face longer periods of starvation, except during molting, where

they actually down-regulate gene expression of all CTLPs. To forced periods of starvation they predominately respond by stopping secretion of CTLPs, obviously storing immature peptidases until nutrients are available again.

Functional diversity of chymotrypsin-like peptidases may also be deduced from the finding that Malpighian tubules exhibit significant expression of MsCT, MsCTLP1 and MsCTLP4 (Figs. 22 and 23). The Malpighian tubules actively transport solutes from the hemolymph into the lumen thereby generating an osmotic gradient that drives primary urine excretion (Beyenbach, 1995). Homeostasis of fluids is regulated by the action of diuretic and antidiuretic hormones (Gäde, 2004). Diuretic hormones are thought to be rapidly removed from the hemolymph by the cessation of their production in the brain and by proteolytic inactivation at their site of action, which is a G-protein coupled receptor at the basal membranes of the Malpighian tubule cells (Reagan, 1994; Reagan, 1995). In *Manduca*, the larger of the two diuretic hormones (Mas-DH) has been shown to be proteolytically cleaved between Leu29-Arg30 and Arg30-Ala31, presumably by two different peptidases (Li *et al.*, 1997). Since cleavage of Mas-DH was inhibited by the addition of EDTA, it was concluded that one of these enzymes is a metalloproteinase. The finding that chymotrypsin-like peptidases are expressed in the Malpighian tubules at the basal site raises the question, whether they are involved in inactivation of Mas-DH or other neuropeptides acting on the Malpighian tubule. Future studies will address this question.

In contrast to *Manduca* whose digestion is almost completely accomplished by serine peptidases (Srinivasan *et al.*, 2006), *Tribolium* for the most part relies on the digestion by cysteine peptidases (Vinokurov *et al.*, 2009; see 1.4.2). Thus the question arises, why *Tribolium* expresses a strikingly high number of chymotrypsin-like peptidases. From the 14 TcCTLPs identified in this work, the 7 TcCTLPs, which were the subject of further characterization, showed expression in the midgut (Fig. 32). Oppert *et al.* (2005) suggested one possible explanation, as the insects seem to shift from a cysteine- to a serine-based digestion when extensively taking up cysteine peptidase inhibitors. On the other hand, some of the newly discovered TcCTLPs may have functions, which differ from mere digestion. To get first hints about the functions of the seven *TcCTLPs* that were the focus of this study, their gene expression was profiled throughout various developmental stages of the animals. Some *TcCTLP* genes were expressed exclusively in feeding larval and adult stages (*TcCTLP-5A*, *TcCTLP-5B*, *TcCTLP-6A*), suggesting that these *TcCTLPs* might indeed play a role in digestion. Others were also found in embryonic and/or pupal stages (*TcCTLP-5C*, *TcCTLP-6C*, *TcCTLP-6D*, *TcCTLP-6E*; Figs. 30 and 31), indicating the functional diversity of these peptidases.

To evaluate the RT-PCR based *TcCTLP* expression data, the results were compared with microarray and proteomic data recently reported for the larval gut of *Tribolium* (Morris *et al.*, 2009). In that study, ranking scores between 1 and 100 were assigned to the transcripts of gut-expressed genes, which reflect the extent of gene expression in the larval gut based on a custom array hybridization experiment covering about 12000 GLEAN sequences of the *T. castaneum* genome. According to the data obtained in this work *TcCTLP-5A* and *TcCTLP-6A* were highly expressed in the midgut

as RT-PCR yielded high product amounts at low cycle numbers and significantly lower expression in the carcass (Fig. 32). This conclusion is consistent with the high gut ranking score of 39 obtained for both of these genes. According to the method used for the determination of midgut proteins with acidic and basic pIs by Vinokurov *et al.* (2009), TcCTLP-5A (pI: 6.47) and TcCTLP-6A (pI: 4.94) are anionic peptidases, probably belonging to one of the chymotrypsin fractions involved in digestion. In line with a gut ranking score of 1, which was assigned to *TcCTLP-5B*, *TcCTLP-5C* and *TcCTLP-6D*, was the finding that PCR products for these genes were only obtained at higher cycle numbers of  $\geq 30$  when using midgut cDNA as a template (Fig. 32). However, the high expression levels of *TcCTLP-6C* and *TcCTLP-6E* that were repeatedly detected in the midgut (Figs. 32 and 33) were not reflected by the corresponding gut ranking scores of 2 and 1, respectively. As each *Tribolium* gene was represented only by a single 35-mer oligomer on the custom array (Morris *et al.*, 2009), discrepancies between the data of this work and the custom array data could be due to sporadic chip hybridization artifacts. High gut ranking scores of 15, 27 and 26 were also detected for the so far uncharacterized enzymes *TcCTLP-3B*, *TcCTLP-6B* and *TcCTLP-7A*, respectively. As all three peptidases exhibit acidic pIs (4.32; 4.53; 4.52) they probably belong to one of the anionic chymotrypsin fractions involved in digestion (Vinokurov *et al.*, 2009). With 47, the highest gut ranking value of all TcCTLPs was assigned to TcCTLP-6F, making it a most likely candidate for being a digestive enzyme. However, at the pH conditions prevailing in the *Tribolium* midgut, TcCTLP-6F (GenBank|XP\_970603.1) is a cationic enzyme (pI: 9.35) presumably exhibiting chymotrypsin activity, which was reported to be found only in anionic midgut fractions (Vinokurov *et al.*, 2009). For this reason TcCTLP-6F might have different, yet uncharacterized functions. This might also apply for TcCTLP-8A (GenBank|XP\_001815065.1), another cationic TcCTLP (pI: 9.82), which is not listed in the gut ranking, thus making both peptidases an interesting target for RNAi studies.

An unexpected result was the failure to obtain any phenotypes from knockdown experiments involving any of the gut-specific TcCTLPs. Perhaps these peptidases are redundant such that after a knockdown of one, there could be some compensatory overproduction of another TcCTLP. Such a compensatory proteolytic effect has been reported previously in the *Tribolium* midgut in response to dietary peptidase inhibitors (Oppert *et al.*, 2005).

#### 4.4. Role in molt

Chymotrypsin-like peptidases that have been identified so far in insects are secreted into the gut lumen where they act as digestive enzymes in concert with various other types of peptidases (Law *et al.*, 1977; Srinivasan *et al.*, 2006; Terra *et al.*, 1996). However, CTLPs may not be restricted to the intestinal tract in insects and may have undergone functional differentiation as chymotrypsin-like activities have been reported also in the molting fluid (Brookhart and Kramer, 1990).

Initial indications that some of the seven *TcCTLPs* that were the focus of this study might have additional non-digestive functions during development were obtained

by RT-PCR, as *TcCTLP-5C*, *TcCTLP-6D*, *TcCTLP-6E* and *TcCTLP-6C* (Figs. 30 and 31) showed expression in embryonic and/or pupal stages. This assumption was supported by the fact that expression of *TcCTLP-5C* and *TcCTLP-6C* was also observed in a carcass preparation obtained from legs that were free of any potential contamination by midgut cells, because transcripts for *TcCTLP-6E* could not be amplified from this preparation (Fig. 33).

Functional differentiation with involvement in the insect's molt was observed for TcCTLP-5C, as both larval-pupal and pupal-adult ecdyses were affected when dsRNA for *TcCTLP-5C* was injected early enough before the respective molts (Fig. 47). This finding correlates with the results from RT-PCR showing that *TcCTLP-5C* reaches maximal expression at the prepupal as well as the P<sub>5</sub> late pupal stage (Fig 31). While the larval-pupal molt can be blocked almost completely by injection of dsRNA for *TcCTLP-5C* with the pupal development continuing under the old unshed larval cuticle, prepupal and midpupal stage injections of the same dsRNA led to only about 50% mortality. This observation may indicate that the amount of injected dsRNA was insufficient to knockdown TcCTLP-5C at late pupal stages where its expression is actually the highest (Fig 31). Similar results were obtained in *Locusta migratoria* for the RNAi-mediated knockdown of a trypsin-like serine peptidase, which is involved in molting (Wei *et al.*, 2007). Knockdown of this gene led to a developmental arrest at the adult molt and the phenotype was more severe when the dsRNA was injected earlier.

The conclusion that TcCTLP-5C has a function in molting is also supported by Western blots where TcCTLP-5C can be detected with an antibody to the recombinant peptidase in protein extracts from larval exuviae (Fig. 49 A). However, the apparent molecular mass of the protein resolved by SDS-PAGE was about 30 kDa and hence higher than the expected theoretical molecular mass of about 25 kDa for the activated peptidase. The decreased mobility might be explained by N-glycosylation, as TcCTLP-5C exhibits a putative N-glycosylation site, NLS<sub>D</sub>, at amino acid positions 137-140. Glycosylation of TcCTLP-5C seems conceivable due to the finding that also MsCTLP1 is glycosylated (see 3.2). A second protein detected by the anti-TcCTLP-5C antibody migrated as a protein of about 60 kDa and therefore could reflect the formation of dimers that are stable in SDS (Fig. 49 A).

Since the occurrence of alternatively spliced transcripts for TcCTLP-5C could be demonstrated, the question arises whether both or only one of the two splice variants are responsible for the molting phenotype after their knockdown by dsRNA injection. RT-PCR with primers that did not discriminate between the splice variants showed that *TcCTLP-5C* was expressed at a somewhat higher level in the carcass than in the midgut. To address the question whether this difference could be due to tissue specific expression of the splice variants, RT-PCR was performed with primers that allowed discrimination between the variants. However, *TcCTLP-5C<sub>1</sub>* and *TcCTLP-5C<sub>2</sub>* were both expressed in the midgut and the carcass (Fig. 33), suggesting that RNAi induced phenotypes are not mediated by the knockdown of any specific TcCTLP-5C variant. Interestingly, the immunoblot performed with antibodies to TcCTLP-5C detected double bands that differed only minimally in size (Fig. 49 A, lane 3). The protein bands presumably represent the respective protein forms of the splice variants with theoretical

molecular masses of 25.7 kDa for TcCTLP-5C<sub>1</sub> and 23.5 kDa for TcCTLP-5C<sub>2</sub>. In contrast to the results from RT-PCR where both splice variants could be detected in the midgut, the immunoblot did not reveal any protein staining in the midgut, which could be explained either by differential protein turnover in both tissues or by slight differences in the developmental stage of the larvae used for mRNA or protein preparations.

In contrast to *TcCTLP-5C*, *TcCTLP-6C* was assigned a high gut ranking score of 39 in the larval gut transcriptome of *Tribolium* (Morris *et al.*, 2009). This high score is supported by the results from RT-PCR, which show highest expression of TcCTLP-6C in the midgut and only low expression in the carcass (Figs. 32 and 33). Moreover, in immunoblots with antibodies to TcCTLP-6C, this peptidase was detected in high amounts in the midgut and in lower amounts in protein extracts from the larval exuviae (Fig. 49 B). The apparent molecular mass was again about 30 kDa and hence higher than the expected theoretical molecular mass of about 24 kDa for the activated peptidase. Also TcCTLP-6C exhibits putative N-glycosylation sites NATL at amino acid positions 30-33, NYTS at amino acid positions 102-105 and NATC at amino acid positions 136-139. Together with the finding that *TcCTLP-6C* is highly expressed at feeding stages (Fig. 30), these results imply that TcCTLP-6C is a gut digestive peptidase. Surprisingly, injection of ds*TcCTLP-6C* into *Tribolium* larvae led to an almost complete arrest of the pupal-adult molt (Fig. 50), indicating that TcCTLP-6C functions additionally as a molting fluid peptidase as well. This hypothesis is supported by the finding that *TcCTLP-6C* was detected at the mRNA and protein levels in the integument (Figs. 33 and 49 B). Although *TcCTLP-6C* expression is low throughout the pupal stage (Fig. 31), it is obviously essential for the pupal-adult molt. The phenotype obtained after ds*TcCTLP-6C* injection into larvae closely resembles that obtained for prepupal and midpupal ds*TcCTLP-5C* injections (Fig. 47 B). This may point to similar functions of both peptidases in pupal-adult molting. However, larval injection of ds*TcCTLP-6C* did not prevent the larval-larval and larval-pupal molt, as was the case for *TcCTLP-5C*. Possibly the animals escaped from molting arrest because the knockdown of *TcCTLP-6C* was insufficient at the relevant larval stage. However, this observation could also reflect a real functional difference between *TcCTLP-5C* and *TcCTLP-6C*.

Together, the results suggest that TcCTLP-5C and TcCTLP-6C are important cuticle matrix-digesting peptidases present in the molting fluid. TcCTLP-5C and TcCTLP-6C most likely digest cuticle proteins that form the matrix in which chitin microfibrils are embedded. However, they could also be required for the activation of other molting fluid peptidases or other cuticle-modifying enzymes as molting fluid peptidases have been reported to activate a pro-chitinase in the endocuticle (Law *et al.*, 1977). Furthermore, it has been suggested that zymogen activation of the chymotrypsin-like peptidase SeCT34 from *Spodoptera exigua* is mediated by chymotrypsin rather than trypsin (Herrero *et al.*, 2005). In any case, chymotrypsin-like peptidases apparently act in concert with other molting peptidases to degrade proteins in the old cuticle, which may involve also some of those chymotrypsin-like peptidases which have not been examined so far. However, TcCTLPs may also be involved in the activation of

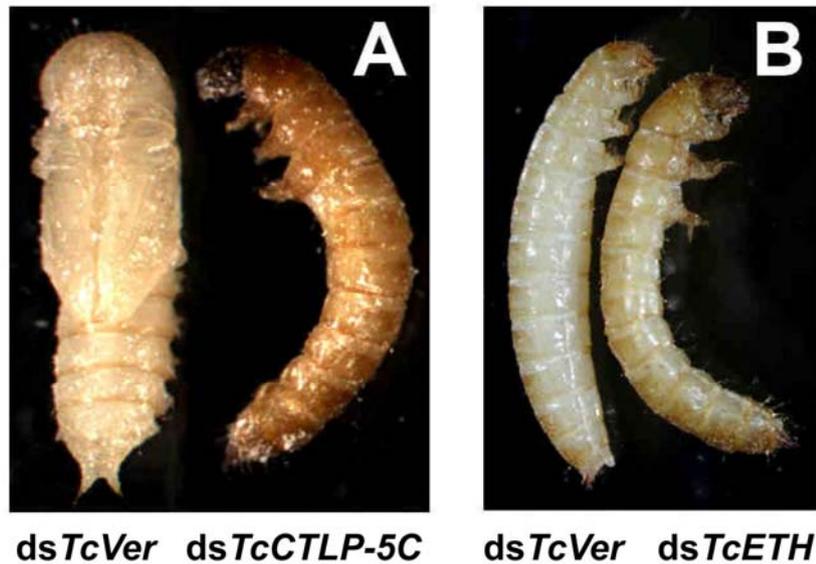
zymogenic chitin synthases like it has been suggested for MsCTLP1. In this context, preliminary results obtained by injection of dsRNA for *TcCTLP-6E* into penultimate larvae suggested a reduced production of peritrophic matrix in the midgut of *T. castaneum* (data not shown). The cuticle turnover process involving TcCTLP-5C and TcCTLP-6C is presumably tightly regulated by serine peptidase inhibitors that are present in the molting fluid as well (Samuels and Reynolds, 2000). Possibly cascades of proteolytic enzymes act in concert with peptidase inhibitors as described above for the immune responses in insects (see 1.4.1).

Regulation of the TcCTLPs which play a role in ecdysis might further involve several molting hormones. Molting is initiated by 20-hydroxy-ecdysone and coordinated by various peptide hormones (Arakane *et al.*, 2008a; Truman, 2005; Zitnan *et al.*, 2007). Some of the most important peptide hormones in this context are the eclosion hormone (EH), the ecdysis triggering hormone (ETH), the crustacean cardioactive peptide (CCAP) and bursicon.

Comparison of the RNAi induced knockdowns of these hormones with those for *TcCTLP-5C* and *TcCTLP-6C* showed high similarity of the phenotypes for *TcCTLP-5C* and *TcETH* dsRNA injections (Dr. Yasuyuki Arakane, personal communication). Injection of dsRNA for *TcETH* into penultimate larvae lead to an arrest of the larval-pupal molt as observed for ds*TcCTLP-5C*. Injected animals showed the same signs of pupal development under a larval cuticle without being able to molt into pupa (Fig. 53 A and B). Similar results were also obtained for dsRNA injection of *TcETH* and *TcCTLP-5C* into prepupal and midpupal stages since in both cases development was stopped at pre-ecdysis (Fig. 47 B and Fig. 4 in Arakane *et al.*, 2008a). Whether TcCTLP-5C also affects molting behavior as reported for TcETH (Arakane *et al.*, 2008a) needs to be addressed in future studies. The analogy between the ds*TcCTLP-5C* and ds*TcETH* phenotypes suggests a potential role of TcCTLP-5C in the TcETH pathway. To rule out the possibility that the *TcCTLP-5C* phenotype is caused by a cross-knockdown of *TcEH* or *TcETH*, RT-PCR was performed, using cDNA from ds*TcCTLP-5C* and ds*Ver* injected penultimate-larvae and primers specific for *TcCTLP-5C*, *TcEH* and *TcETH*. Whereas a knockdown effect for *TcCTLP-5C* could be shown, transcript levels for *TcEH* and *TcETH* stayed the same compared to the ds*Ver* injected control (data not shown).

However, involvement of TcCTLP-5C in the TcETH pathway is highly speculative, and the phenotypes could be similar by chance. Approximately two days before ecdysis, high ecdysteroid levels induce expression of ETH receptors in the CNS and increase ETH production in Inka cells (Zitnan *et al.*, 2007). Declining levels of ecdysteroids release EH from ventral medial neurons (Hewes and Truman, 1994) and enable the Inka cells to secrete ETH in response to EH. EH elicits cGMP-mediated ETH depletion from the Inka cells into the hemolymph and leads to pre-ecdysis behavior by binding of ETH to the ETH receptors A and B in central neurons (Zitnan *et al.*, 2007). Whereas TcETH is secreted by Inka cells into the hemolymph and acts on the CNS, TcCTLP-5C is probably predominantly secreted into the molting fluid by epidermal cells. This assumption was supported by the finding that Western blots using the anti-TcCTLP-5C antibody and proteins from larval exuviae showed a TcCTLP-5C-specific

signal (Fig. 49 A). This implies that TcETH and TcCTLP-5C most likely will not interact with each other under *in vivo* conditions due to differences in localization. However, involvement of TcCTLP-5C in the TcETH pathway cannot be fully excluded. To prove this hypothesis, it would be interesting if micro-injection of TcETH could restore the *TcCTLP-5C* dsRNA induced phenotype *in vivo*.



**Fig. 53.** Comparison of phenotypes induced by *TcCTLP-5C* and *TcETH* dsRNA injection. dsRNA for *TcCTLP-5C* (A) and *TcETH* (B) were injected into penultimate instar larvae and the phenotypes were monitored during development. In both cases the injection of the respective dsRNA led to a defective larval-pupal molt (The picture showing the *dsTcETH* phenotype is a kind gift of Dr. Yasuyuki Arakane). *TcVer* dsRNA injected into penultimate larvae yielded normally developing animals except for the expected no eye pigmented phenotype.

## 5. Summary

Digestion of proteins in the midgut of lepidopteran larvae relies on different types of peptidases, among the trypsins and chymotrypsins. In this work four chymotrypsin-like peptidases (MsCTLP1–4) were identified from the larval midgut of *M. sexta*, which are distantly related to another chymotrypsin (MsCT), a previously described peptidase present in the larval midgut of *M. sexta*. MsCTLP1–4 fit perfectly into a novel subgroup of insect CTLPs by sequence similarity and by the replacement of GP by SA in the highly conserved GDSGGP motif. Examination of *MsCTLP* expression in different tissues showed that most of the peptidases were predominantly expressed in the anterior and median midgut, while some were found in the Malpighian tubules. Expression analysis of *MsCTLPs* at different physiological states revealed that the mRNA amounts did not differ considerably in feeding and starving larvae except for *MsCTLP2*, whose mRNA dropped significantly upon starvation. During molting, however, the mRNA amounts of all *MsCTLPs* dropped significantly. Immunological determination of MsCTLP1 amounts showed that the mature peptidase was only detectable in the gut lumen of feeding and re-fed larvae, but not in that of starving or molting larvae, suggesting that MsCTLP1 secretion is suspended during starvation or molt. Differential regulation of transcript levels as well as their partial expression in Malpighian tubules might point to a role, which is distinct from digestion for at least some *MsCTLPs*. In line with this assumption, MsCTLP1 was shown to interact with the chitin synthase 2 (MsCHS2), necessary for chitin synthesis in the course of peritrophic matrix formation in the midgut of *M. sexta*. The occurrence of this interaction *in vivo* is supported by co-localization and co-immunoprecipitation. The data suggest that chitin synthesis is controlled by an intestinal proteolytic signaling cascade linking chitin synthase activity to the nutritional state of the larvae. As MsCTLP1 appears to be involved in such signaling cascades, other midgut peptidases could have other targets and may therefore regulate different activities.

To gain more insight into the functions of CTLPs, the gene family encoding these peptidases in the genome of the red flour beetle, *T. castaneum*, was analyzed. Using an extended search pattern, 14 *TcCTLP* genes were identified that encode peptidases with S1 specificity pocket residues typically found in chymotrypsin-like enzymes. Analysis of the expression patterns of seven *TcCTLP* genes at various developmental stages revealed that some *TcCTLP* genes were exclusively expressed in feeding larval and adult stages (*TcCTLP-5A/B*, *TcCTLP-6A*). Others were also detected in non-feeding embryonic (*TcCTLP-5C*, *TcCTLP-6D*) and pupal stages (*TcCTLP-5C*, *TcCTLP-6C/D/E*). *TcCTLP* genes were expressed predominantly in the midgut where they presumably function in digestion. However, *TcCTLP-5C* and *TcCTLP-6C* also showed considerable expression in the carcass. The latter two genes might therefore encode peptidases that act as molting fluid enzymes. To test this hypothesis, western blots were performed using protein extracts from larval exuviae. The extracts reacted with antibodies to *TcCTLP-5C* and *TcCTLP-6C* suggesting that the corresponding peptidases are secreted into the molting fluid. Finally, systemic RNAi experiments were performed. While injections of dsRNAs to *TcCTLP-5A/B* and *TcCTLP-6A/D/E* into penultimate larvae did not affect growth or development, injection of dsRNA for

*TcCTLP-5C* and *TcCTLP-6C* resulted in severe molting defects. Recombinant expressed *TcCTLP-5C*<sub>2</sub> was moreover activated by trypsin and was able to hydrolyze AAPF, hence making *TcCTLP-5C* the first described chymotrypsin-like peptidase ever to be involved in molting.

## 6. References

- Aggarwal, K. and Silverman, N.** (2008). Positive and negative regulation of the *Drosophila* immune response. *BMB Rep* **41**, 267-77.
- Arakane, Y., Dixit, R., Begum, K., Park, Y., Specht, C. A., Merzendorfer, H., Kramer, K. J., Muthukrishnan, S. and Beeman, R. W.** (2009). Analysis of functions of the chitin deacetylase gene family in *Tribolium castaneum*. *Insect Biochem Mol Biol* **39**, 355-65.
- Arakane, Y., Li, B., Muthukrishnan, S., Beeman, R. W., Kramer, K. J. and Park, Y.** (2008a). Functional analysis of four neuropeptides, EH, ETH, CCAP and bursicon, and their receptors in adult ecdysis behavior of the red flour beetle, *Tribolium castaneum*. *Mech Dev* **125**, 984-95.
- Arakane, Y., Specht, C. A., Kramer, K. J., Muthukrishnan, S. and Beeman, R. W.** (2008b). Chitin synthases are required for survival, fecundity and egg hatch in the red flour beetle, *Tribolium castaneum*. *Insect Biochem Mol Biol*.
- Asada, N. and Sezaki, H.** (1999). Properties of phenoloxidases generated from prophenoloxidase with 2-propanol and the natural activator in *Drosophila melanogaster*. *Biochem Genet* **37**, 149-58.
- Aso, Y., Kramer, K. J., Hopkins, T. L. and Lookhart, G. L.** (1985). Characterization of hemolymph protyrosinase and a cuticular activator from *Manduca sexta* (L.). *Insect Biochem.* **15**, 9-17.
- Aso, Y., Yamashita, T., Meno, K. and Murakami, M.** (1994). Inhibition of prophenoloxidase-activating enzyme from *Bombyx mori* by endogenous chymotrypsin inhibitors. *Biochem Mol Biol Int* **33**, 751-8.
- Bach, M.** (1870). Die Wunder der Insektenwelt. Das Insekt, sein Leben und Wirken in dem Haushalte der Natur, gemeinfaßlich dargestellt. Soest: Nasse'sche Verlagsbuchhandlung
- Baldwin, B. M. and Hakim, R. S.** (1991). Growth and differentiation of the larval midgut epithelium during moult in the moth, *Manduca sexta*. *Tissue & Cell* **23**, 411-422.
- Barillas-Mury, C. and Wells, M. A.** (1993). Cloning and sequencing of the blood meal-induced late trypsin gene from the mosquito *Aedes aegypti* and characterization of the upstream regulatory region. *Insect Mol Biol* **2**, 7-12.
- Beeman, R. W. and Stuart, J. J.** (1990). A gene for lindane + cyclodiene resistance in the red flour beetle (Coleoptera: Tenebrionidae). *J. Econ. Entomol.* **83**, 1745-1751.
- Bell, R. A. and Joachim, F. G.** (1974). Techniques for rearing laboratory colonies of tobacco hornworms and pink bollworms. *Ann Entomol Soc Am* **69**, 365-373.
- Bender, M. L. and Kézdy, F. J.** (1964). The Current Status of the  $\alpha$ -Chymotrypsin Mechanism. *J. Am. Chem. Soc.* **86** 3704-3714.
- Bendtsen, J. D., Nielsen, H., von Heijne, G. and Brunak, S.** (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**, 783-95.
- Berg, J. M., Tymoczko, J. L. and Stryer, L.** (2007). Biochemistry. New York: W. H. Freeman ; Basingstoke : Palgrave [distributor].
- Betts, M. J. and Russell, R. B.** (2003). Amino acid properties and consequences of substitutions. In *Bioinformatics for geneticists*, eds. M. R. Barnes and I. C. Gray. Chichester: Wiley.
- Beyenbach, K. W.** (1995). Mechanism and regulation of electrolyte transport in Malpighian tubules. *Journal of Insect Physiology* **41**, 197-207.

**Beyenbach, K. W. and Wieczorek, H.** (2006). The V-type H<sup>+</sup> ATPase: molecular structure and function, physiological roles and regulation. *J Exp Biol* **209**, 577-89.

**Birktoft, J. J. and Blow, D. M.** (1972). Structure of crystalline  $\alpha$ -chymotrypsin. V. The atomic structure of tosyl- $\alpha$ -chymotrypsin at 2 Å resolution. *J Mol Biol* **68**, 187-240.

**Bischoff, V., Vignal, C., Boneca, I. G., Michel, T., Hoffmann, J. A. and Royet, J.** (2004). Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. *Nat Immunol* **5**, 1175-80.

**Blow, D. M.** (1997). The tortuous story of Asp ... His ... Ser: structural analysis of alpha-chymotrypsin. *Trends Biochem Sci* **22**, 405-8.

**Bode, W., Meyer, E., Jr. and Powers, J. C.** (1989). Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism-based inhibitors. *Biochemistry* **28**, 1951-63.

**Bode, W., Schwager, P. and Huber, R.** (1978). The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of the bovine trypsinogen-pancreatic trypsin inhibitor complex and of its ternary complex with Ile-Val at 1.9 Å resolution. *J Mol Biol* **118**, 99-112.

**Bollag, D. M. and Edelman, S. J.** (1991). Protein Methods. New York: Wiley-Liss Inc.

**Borovsky, D.** (2003). Biosynthesis and control of mosquito gut proteases. *IUBMB Life* **55**, 435-41.

**Botos, I., Meyer, E., Nguyen, M., Swanson, S. M., Koomen, J. M., Russell, D. H. and Meyer, E. F.** (2000). The structure of an insect chymotrypsin. *J Mol Biol* **298**, 895-901.

**Bown, D. P., Wilkinson, H. S. and Gatehouse, J. A.** (1997). Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem Mol Biol* **27**, 625-38.

**Brito, L. O., Lopes, A. R., Parra, J. R., Terra, W. R. and Silva-Filho, M. C.** (2001). Adaptation of tobacco budworm *Heliothis virescens* to proteinase inhibitors may be mediated by the synthesis of new proteinases. *Comp Biochem Physiol B Biochem Mol Biol* **128**, 365-75.

**Broadway, R. M.** (1997). Dietary regulation of serine proteinases that are resistant to serine proteinase inhibitors. *J Insect Physiol* **43**, 855-874.

**Broehan, G., Arakane, Y., Beeman, R. W., Kramer, K. J., Muthukrishnan, S. and Merzendorfer, H.** (2010). Chymotrypsin-like peptidases from *Tribolium castaneum*: a role in molting revealed by RNA interference. *Insect Biochem Mol Biol* **40**, 274-83.

**Broehan, G., Kemper, M., Driemeier, D., Vogelpohl, I. and Merzendorfer, H.** (2008). Cloning and expression analysis of midgut chymotrypsin-like proteinases in the tobacco hornworm. *J Insect Physiol* **54**, 1243-52.

**Broehan, G., Zimoch, L., Wessels, A., Ertas, B. and Merzendorfer, H.** (2007). A chymotrypsin-like serine protease interacts with the chitin synthase from the midgut of the tobacco hornworm. *J Exp Biol* **210**, 3636-43.

**Brookhart, G. L. and Kramer, K. J.** (1990). Proteinases in the molting fluid of the tobacco hornworm, *Manduca sexta*. *Insect Biochemistry* **20**, 467-477.

**Burke, E. and Barik, S.** (2003). Megaprimer PCR: application in mutagenesis and gene fusion. *Methods Mol Biol* **226**, 525-32.

**Cerenius, L. and Söderhäll, K.** (2004). The prophenoloxidase-activating system in invertebrates. *Immunol Rev* **198**, 116-26.

- Chomczynski, P. and Sacchi, N.** (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**, 156-9.
- Choo, Y. M., Lee, K. S., Yoon, H. J., Lee, S. B., Kim, J. H., Sohn, H. D. and Jin, B. R.** (2007). A serine protease from the midgut of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae): cDNA cloning, gene structure, expression and enzyme activity. *Eur. J. Entomol* **104**, 1-7.
- Chothia, C.** (1976). The nature of the accessible and buried surfaces in proteins. *J Mol Biol* **105**, 1-12.
- Christensen, B. M., Li, J., Chen, C. C. and Nappi, A. J.** (2005). Melanization immune responses in mosquito vectors. *Trends Parasitol* **21**, 192-9.
- Cohen, E. and Casida, J. E.** (1980). Properties of *Tribolium* gut chitin synthetase. *Pestic Biochem Physiol* **13**, 121-128.
- Cooper, G. M. and Hausman, R. E.** (2009). The cell : a molecular approach. Sunderland, Mass.: Sinauer Associates.
- Czapinska, H. and Otlewski, J.** (1999). Structural and energetic determinants of the S1-site specificity in serine proteases. *Eur J Biochem* **260**, 571-95.
- Davis, C. A., Riddell, D. C., Higgins, M. J., Holden, J. J. and White, B. N.** (1985). A gene family in *Drosophila melanogaster* coding for trypsin-like enzymes. *Nucleic Acids Res* **13**, 6605-19.
- Davis, T. R., Trotter, K. M., Granados, R. R. and Wood, H. A.** (1992). Baculovirus expression of alkaline phosphatase as a reporter gene for evaluation of production, glycosylation and secretion. *Biotechnology (N Y)* **10**, 1148-50.
- De Gregorio, E., Spellman, P. T., Tzou, P., Rubin, G. M. and Lemaitre, B.** (2002). The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. *EMBO J* **21**, 2568-79.
- Di Cera, E.** (2009). Serine proteases. *IUBMB Life* **61**, 510-5.
- Dissing, M., Giordano, H. and DeLotto, R.** (2001). Autoproteolysis and feedback in a protease cascade directing *Drosophila* dorsal-ventral cell fate. *EMBO J* **20**, 2387-93.
- Dow, J. A.** (1984). Extremely high pH in biological systems: a model for carbonate transport. *Am J Physiol* **246**, R633-6.
- Dow, J. A.** (1992). pH GRADIENTS IN LEPIDOPTERAN MIDGUT. *J Exp Biol* **172**, 355-375.
- El Chamy, L., Leclerc, V., Caldelari, I. and Reichhart, J. M.** (2008). Sensing of 'danger signals' and pathogen-associated molecular patterns defines binary signaling pathways 'upstream' of Toll. *Nat Immunol* **9**, 1165-70.
- Emanuelsson, O., Brunak, S., von Heijne, G. and Nielsen, H.** (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**, 953-71.
- Epstein, D. M. and Abeles, R. H.** (1992). Role of serine 214 and tyrosine 171, components of the S2 subsite of alpha-lytic protease, in catalysis. *Biochemistry* **31**, 11216-23.
- Ewer, J.** (2005). Behavioral actions of neuropeptides in invertebrates: insights from *Drosophila*. *Horm Behav* **48**, 418-29.
- Fukamizo, T. and Kramer, K. J.** (1985). Mechanism of chitin hydrolysis by the binary chitinase system in insect moulting fluid. *Insect Biochemistry* **15**, 141-145.
- Gäde, G.** (2004). Regulation of intermediary metabolism and water balance of insects by neuropeptides. *Annu Rev Entomol* **49**, 93-113.

- Gaines, P. J., Sampson, C. M., Rushlow, K. E. and Stiegler, G. L.** (1999). Cloning of a family of serine protease genes from the cat flea *Ctenocephalides felis*. *Insect Mol Biol* **8**, 11-22.
- Gatehouse, L. N., Shannon, A. L., Burgess, E. P. and Christeller, J. T.** (1997). Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. *Insect Biochem Mol Biol* **27**, 929-44.
- Gobert, V., Gottar, M., Matskevich, A. A., Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J. A. and Ferrandon, D.** (2003). Dual activation of the *Drosophila* toll pathway by two pattern recognition receptors. *Science* **302**, 2126-30.
- Gottar, M., Gobert, V., Matskevich, A. A., Reichhart, J. M., Wang, C., Butt, T. M., Belvin, M., Hoffmann, J. A. and Ferrandon, D.** (2006). Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. *Cell* **127**, 1425-37.
- Gräf, R., Lepier, A., Harvey, W. R. and Wieczorek, H.** (1994). A novel 14-kDa V-ATPase subunit in the tobacco hornworm midgut. *J Biol Chem* **269**, 3767-74.
- Greer, J.** (1990). Comparative modeling methods: application to the family of the mammalian serine proteases. *Proteins* **7**, 317-34.
- Halicak, J. P. and Beeman, R. W.** (1983). Status of malathion resistance in five genera of beetles infesting farm-stored corn, wheat and oats in the United States. *J. Econ. Entomol.* **76**, 717-722.
- Hall, M., Scott, T., Sugumaran, M., Soderhall, K. and Law, J. H.** (1995). Proenzyme of *Manduca sexta* phenol oxidase: purification, activation, substrate specificity of the active enzyme, and molecular cloning. *Proc Natl Acad Sci U S A* **92**, 7764-8.
- Hanahan, D.** (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-80.
- Hartley, B. S.** (1964). Amino-Acid Sequence of Bovine Chymotrypsinogen-A. *Nature* **201**, 1284-7.
- Hartley, B. S. and Kauffman, D. L.** (1966). Corrections to the amino acid sequence of bovine chymotrypsinogen A. *Biochem J* **101**, 229-31.
- He, N., Aso, Y., Fujii, H., Banno, Y. and Yamamoto, K.** (2004). In vivo and in vitro interactions of the *Bombyx mori* chymotrypsin inhibitor b1 with *Escherichia coli*. *Biosci Biotechnol Biochem* **68**, 835-40.
- Hedstrom, L.** (2002). Serine protease mechanism and specificity. *Chem Rev* **102**, 4501-24.
- Hedstrom, L., Perona, J. J. and Rutter, W. J.** (1994). Converting trypsin to chymotrypsin: residue 172 is a substrate specificity determinant. *Biochemistry* **33**, 8757-63.
- Hedstrom, L., Szilagyi, L. and Rutter, W. J.** (1992). Converting trypsin to chymotrypsin: the role of surface loops. *Science* **255**, 1249-53.
- Henderson, R.** (1970). Structure of crystalline alpha-chymotrypsin. IV. The structure of indoleacryloyl-alpha-chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme. *J Mol Biol* **54**, 341-54.
- Herrero, S., Combes, E., Van Oers, M. M., Vlak, J. M., de Maagd, R. A. and Beekwilder, J.** (2005). Identification and recombinant expression of a novel chymotrypsin from *Spodoptera exigua*. *Insect Biochem Mol Biol* **35**, 1073-82.
- Hewes, R. S. and Truman, J. W.** (1994). Steroid regulation of excitability in identified insect neurosecretory cells. *J Neurosci* **14**, 1812-9.
- Higgins, D. G.** (1994). CLUSTAL V: multiple alignment of DNA and protein sequences. *Methods Mol Biol* **25**, 307-18.

- Hung, S. H. and Hedstrom, L.** (1998). Converting trypsin to elastase: substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity. *Protein Eng* **11**, 669-73.
- Ishimoto, M. and Chrispeels, M. J.** (1996). Protective mechanism of the Mexican bean weevil against high levels of alpha-amylase inhibitor in the common bean. *Plant Physiol* **111**, 393-401.
- Jiang, H. and Kanost, M. R.** (2000). The clip-domain family of serine proteinases in arthropods. *Insect Biochem Mol Biol* **30**, 95-105.
- Jiang, H., Wang, Y. and Kanost, M. R.** (1998). Pro-phenol oxidase activating proteinase from an insect, *Manduca sexta*: a bacteria-inducible protein similar to *Drosophila* easter. *Proc Natl Acad Sci U S A* **95**, 12220-5.
- Jiang, H., Wang, Y., Yu, X. Q. and Kanost, M. R.** (2003a). Prophenoloxidase-activating proteinase-2 from hemolymph of *Manduca sexta*. A bacteria-inducible serine proteinase containing two clip domains. *J Biol Chem* **278**, 3552-61.
- Jiang, H., Wang, Y., Yu, X. Q., Zhu, Y. and Kanost, M.** (2003b). Prophenoloxidase-activating proteinase-3 (PAP-3) from *Manduca sexta* hemolymph: a clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs. *Insect Biochem Mol Biol* **33**, 1049-60.
- Jongsma, M. A., Bakker, P. L., Peters, J., Bosch, D. and Stiekema, W. J.** (1995). Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc Natl Acad Sci U S A* **92**, 8041-5.
- Kambris, Z., Brun, S., Jang, I. H., Nam, H. J., Romeo, Y., Takahashi, K., Lee, W. J., Ueda, R. and Lemaitre, B.** (2006). *Drosophila* immunity: a large-scale in vivo RNAi screen identifies five serine proteases required for Toll activation. *Curr Biol* **16**, 808-13.
- Kanost, M. R. and Gorman, M. J.** (2008). Phenoloxidases in insect immunity. In *Insect immunology*, (ed. N. E. Beckage), pp. 69-96. London: Academic.
- Katzenellenbogen, B. S. and Kafatos, F. C.** (1970). Some properties of silkworm moulting gel and moulting fluid *Journal of Insect Physiology* **16**, 2241-2256.
- Katzenellenbogen, B. S. and Kafatos, F. C.** (1971a). Inactive proteinases in silkworm moulting gel *Journal of Insect Physiology* **17**, 823-832.
- Katzenellenbogen, B. S. and Kafatos, F. C.** (1971b). Proteinases of silkworm moulting fluid: Physical and catalytic properties. *Journal of Insect Physiology* **17**, 775-800.
- Kim, M. S., Baek, M. J., Lee, M. H., Park, J. W., Lee, S. Y., Soderhall, K. and Lee, B. L.** (2002). A new easter-type serine protease cleaves a masquerade-like protein during prophenoloxidase activation in *Holotrichia diomphalia* larvae. *J Biol Chem* **277**, 39999-40004.
- Klein, U., Löffelmann, G. and Wieczorek, H.** (1991). The Midgut as a Model System for Insect K<sup>+</sup>-Transporting Epithelia: Immunocytochemical Localization of a Vacuolar-Type H<sup>+</sup> Pump *J Exp Biol* **161**, 61-75.
- Kraut, J.** (1977). Serine proteases: structure and mechanism of catalysis. *Annu Rev Biochem* **46**, 331-58.
- Kwon, T. H., Kim, M. S., Choi, H. W., Joo, C. H., Cho, M. Y. and Lee, B. L.** (2000). A masquerade-like serine proteinase homologue is necessary for phenoloxidase activity in the coleopteran insect, *Holotrichia diomphalia* larvae. *Eur J Biochem* **267**, 6188-96.
- Kyhse-Andersen, J.** (1984). Electrophoretic transfer of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J Biochem Biophys Methods* **10**, 203-9.

- Kyte, J. and Doolittle, R. F.** (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32.
- Laemmli, U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.
- Law, J. H., Dunn, P. E. and Kramer, K. J.** (1977). Insect proteases and peptidases. *Adv Enzymol Relat Areas Mol Biol* **45**, 389-425.
- Lee, K. Y., Zhang, R., Kim, M. S., Park, J. W., Park, H. Y., Kawabata, S. and Lee, B. L.** (2002). A zymogen form of masquerade-like serine proteinase homologue is cleaved during pro-phenoloxidase activation by  $\text{Ca}^{2+}$  in coleopteran and *Tenebrio molitor* larvae. *Eur J Biochem* **269**, 4375-83.
- Lee, S. Y., Cho, M. Y., Hyun, J. H., Lee, K. M., Homma, K. I., Natori, S., Kawabata, S. I., Iwanaga, S. and Lee, B. L.** (1998). Molecular cloning of cDNA for pro-phenol-oxidase-activating factor I, a serine protease is induced by lipopolysaccharide or 1,3-beta-glucan in coleopteran insect, *Holotrichia diomphalia* larvae. *Eur J Biochem* **257**, 615-21.
- Lehane, M. J. and Billingsley, P. F.** (1996). *Biology of the insect midgut*. London: Chapman & Hall.
- Lehane, M. J., Müller, H. M. and Crisanti, A.** (1996). Mechanisms controlling the synthesis and secretion of digestive enzymes in insects. In *Biology of the insect midgut*, eds. M. J. Lehane and P. F. Billingsley, pp. 195-205. London: Chapman & Hall.
- LeMosy, E. K., Tan, Y. Q. and Hashimoto, C.** (2001). Activation of a protease cascade involved in patterning the *Drosophila* embryo. *Proc Natl Acad Sci U S A* **98**, 5055-60.
- Li, H., Wang, H., Schegg, K. M. and Schooley, D. A.** (1997). Metabolism of an insect diuretic hormone by Malpighian tubules studied by liquid chromatography coupled with electrospray ionization mass spectrometry. *Proc Natl Acad Sci U S A* **94**, 13463-8.
- Ligoxygakis, P., Roth, S. and Reichhart, J. M.** (2003). A serpin regulates dorsal-ventral axis formation in the *Drosophila* embryo. *Curr Biol* **13**, 2097-102.
- Liu, J., Shi, G. P., Zhang, W. Q., Zhang, G. R. and Xu, W. H.** (2006). Cathepsin L function in insect moulting: molecular cloning and functional analysis in cotton bollworm, *Helicoverpa armigera*. *Insect Mol Biol* **15**, 823-34.
- Liu, Y., Sui, Y. P., Wang, J. X. and Zhao, X. F.** (2009). Characterization of the trypsin-like protease (Ha-TLP2) constitutively expressed in the integument of the cotton bollworm, *Helicoverpa armigera*. *Arch Insect Biochem Physiol*.
- Lopes, A. R., Juliano, M. A., Juliano, L. and Terra, W. R.** (2004). Coevolution of insect trypsins and inhibitors. *Arch Insect Biochem Physiol* **55**, 140-52.
- Lopes, A. R., Sato, P. M. and Terra, W. R.** (2009). Insect chymotrypsins: chloromethyl ketone inactivation and substrate specificity relative to possible coevolutional adaptation of insects and plants. *Arch Insect Biochem Physiol* **70**, 188-203.
- Lorenzen, M. D., Brown, S. J., Denell, R. E. and Beeman, R. W.** (2002). Cloning and characterization of the *Tribolium castaneum* eye-color genes encoding tryptophan oxygenase and kynurenine 3-monooxygenase. *Genetics* **160**, 225-34.
- Lu, D., Futterer, K., Korolev, S., Zheng, X., Tan, K., Waksman, G. and Sadler, J. E.** (1999). Crystal structure of enteropeptidase light chain complexed with an analog of the trypsinogen activation peptide. *J Mol Biol* **292**, 361-73.
- Ma, W., Tang, C. and Lai, L.** (2005). Specificity of trypsin and chymotrypsin: loop-motion-controlled dynamic correlation as a determinant. *Biophys J* **89**, 1183-93.

- Maruyama, K. and Sugano, S.** (1994). Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**, 171-4.
- Mayer, R. T., Chen, A. C. and DeLoach, J. R.** (1980). Characterization of a chitin synthase from the stable fly, *Stomoxys calcitrans* (L.). *Insect Biochem.* **10**, 549-556.
- Mazumdar-Leighton, S. and Broadway, R. M.** (2001a). Identification of six chymotrypsin cDNAs from larval midguts of *Helicoverpa zea* and *Agrotis ipsilon* feeding on the soybean (Kunitz) trypsin inhibitor. *Insect Biochem Mol Biol* **31**, 633-44.
- Mazumdar-Leighton, S. and Broadway, R. M.** (2001b). Transcriptional induction of diverse midgut trypsins in larval *Agrotis ipsilon* and *Helicoverpa zea* feeding on the soybean trypsin inhibitor. *Insect Biochem Mol Biol* **31**, 645-57.
- McGrath, M. E., Vasquez, J. R., Craik, C. S., Yang, A. S., Honig, B. and Fletterick, R. J.** (1992). Perturbing the polar environment of Asp102 in trypsin: consequences of replacing conserved Ser214. *Biochemistry* **31**, 3059-64.
- Merzendorfer, H.** (2006). Insect chitin synthases: a review. *J Comp Physiol [B]* **176**, 1-15.
- Merzendorfer, H., Huss, M., Schmid, R., Harvey, W. R. and Wieczorek, H.** (1999). A novel insect V-ATPase subunit M9.7 is glycosylated extensively. *J Biol Chem* **274**, 17372-8.
- Merzendorfer, H., Reineke, S., Zhao, X. F., Jacobmeier, B., Harvey, W. R. and Wieczorek, H.** (2000). The multigene family of the tobacco hornworm V-ATPase: novel subunits a, C, D, H, and putative isoforms. *Biochim Biophys Acta* **1467**, 369-79.
- Michel, T., Reichhart, J. M., Hoffmann, J. A. and Royet, J.** (2001). *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* **414**, 756-9.
- Morisato, D. and Anderson, K. V.** (1995). Signaling pathways that establish the dorsal-ventral pattern of the *Drosophila* embryo. *Annu Rev Genet* **29**, 371-99.
- Morris, K., Lorenzen, M. D., Hiromasa, Y., Tomich, J. M., Oppert, C., Elpidina, E. N., Vinokurov, K., Jurat-Fuentes, J. L., Fabrick, J. and Oppert, B.** (2009). *Tribolium castaneum* larval gut transcriptome and proteome: A resource for the study of the coleopteran gut. *J Proteome Res* **8**, 3889-98.
- Muller, H. M., Catteruccia, F., Vizioli, J., della Torre, A. and Crisanti, A.** (1995). Constitutive and blood meal-induced trypsin genes in *Anopheles gambiae*. *Exp Parasitol* **81**, 371-85.
- Neurath, H.** (1984). Evolution of proteolytic enzymes. *Science* **224**, 350-7.
- Neurath, H. and Dixon, G. H.** (1957). Structure and activation of trypsinogen and chymotrypsinogen. *Fed Proc* **16**, 791-801.
- Onishi, E., Dohke, K. and Ashida, M.** (1970). Activation of prephenoloxidase. II. Activation by alpha-chymotrypsin. *Arch Biochem Biophys* **139**, 143-8.
- Oppert, B., Morgan, T. D., Hartzer, K. and Kramer, K. J.** (2005). Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Comp Biochem Physiol C Toxicol Pharmacol* **140**, 53-8.
- Oppert, B., Walters, P. and Zuercher, M.** (2006). Digestive proteinases of the larger black flour beetle, *Cynaues angustus* (Coleoptera: Tenebrionidae). *Bull Entomol Res* **96**, 167-72.
- Ote, M., Mita, K., Kawasaki, H., Daimon, T., Kobayashi, M. and Shimada, T.** (2005a). Identification of molting fluid carboxypeptidase A (MF-CPA) in *Bombyx mori*. *Comp Biochem Physiol B Biochem Mol Biol* **141**, 314-22.

- Ote, M., Mita, K., Kawasaki, H., Kobayashi, M. and Shimada, T. (2005b). Characteristics of two genes encoding proteins with an ADAM-type metalloprotease domain, which are induced during the molting periods in *Bombyx mori*. *Arch Insect Biochem Physiol* **59**, 91-8.
- Page, M. J. and Di Cera, E. (2008). Serine peptidases: classification, structure and function. *Cell Mol Life Sci* **65**, 1220-36.
- Patankar, A. G., Giri, A. P., Harsulkar, A. M., Sainani, M. N., Deshpande, V. V., Ranjekar, P. K. and Gupta, V. S. (2001). Complexity in specificities and expression of *Helicoverpa armigera* gut proteinases explains polyphagous nature of the insect pest. *Insect Biochem Mol Biol* **31**, 453-64.
- Pau, R. N. and Kelly, C. (1975). The hydroxylation of tyrosine by an enzyme from third-instar larvae of the blowfly *Calliphora erythrocephala*. *Biochem J* **147**, 565-73.
- Paulillo, L. C., Lopes, A. R., Cristofolletti, P. T., Parra, J. R., Terra, W. R. and Silva-Filho, M. C. (2000). Changes in midgut endopeptidase activity of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) are responsible for adaptation to soybean proteinase inhibitors. *J Econ Entomol* **93**, 892-6.
- Perona, J. J. and Craik, C. S. (1995). Structural basis of substrate specificity in the serine proteases. *Protein Sci* **4**, 337-60.
- Perona, J. J. and Craik, C. S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J Biol Chem* **272**, 29987-90.
- Perona, J. J., Hedstrom, L., Rutter, W. J. and Fletterick, R. J. (1995). Structural origins of substrate discrimination in trypsin and chymotrypsin. *Biochemistry* **34**, 1489-99.
- Peterson, A. M., Barillas-Mury, C. V. and Wells, M. A. (1994). Sequence of three cDNAs encoding an alkaline midgut trypsin from *Manduca sexta*. *Insect Biochem Mol Biol* **24**, 463-71.
- Peterson, A. M., Fernando, G. J. and Wells, M. A. (1995). Purification, characterization and cDNA sequence of an alkaline chymotrypsin from the midgut of *Manduca sexta*. *Insect Biochem Mol Biol* **25**, 765-74.
- Piao, S., Song, Y. L., Kim, J. H., Park, S. Y., Park, J. W., Lee, B. L., Oh, B. H. and Ha, N. C. (2005). Crystal structure of a clip-domain serine protease and functional roles of the clip domains. *EMBO J* **24**, 4404-14.
- Pili-Floury, S., Leulier, F., Takahashi, K., Saigo, K., Samain, E., Ueda, R. and Lemaitre, B. (2004). In vivo RNA interference analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in *Drosophila* adults. *J Biol Chem* **279**, 12848-53.
- Polanowski, A. and Wilusz, T. (1996). Serine proteinase inhibitors from insect hemolymph. *Acta Biochim Pol* **43**, 445-53.
- Polgar, L. (2005). The catalytic triad of serine peptidases. *Cell Mol Life Sci* **62**, 2161-72.
- Rabossi, A., Stoka, V., Puizdar, V., Turk, V. and Quesada-Allue, L. A. (2008). Purification and characterization of two cysteine peptidases of the Mediterranean fruit fly *Ceratitidis capitata* during metamorphosis. *Arch Insect Biochem Physiol* **68**, 1-13.
- Rawlings, N. D. and Barrett, A. J. (1994). Families of serine peptidases. *Methods Enzymol* **244**, 19-61.
- Rawlings, N. D., Barrett, A. J. and Bateman, A. (2010). MEROPS: the peptidase database. *Nucleic Acids Res* **38**, D227-33.

- Reagan, J. D.** (1994). Expression cloning of an insect diuretic hormone receptor. A member of the calcitonin/secretin receptor family. *J Biol Chem* **269**, 9-12.
- Reagan, J. D.** (1995). Functional expression of a diuretic hormone receptor in baculovirus-infected insect cells: evidence suggesting that the N-terminal region of diuretic hormone is associated with receptor activation. *Insect Biochem Mol Biol* **25**, 535-9.
- Reynolds, S. E., Nottingham, S. F. and Stephensa, A. E.** (1985). Food and water economy and its relation to growth in fifth-instar larvae of the tobacco hornworm, *Manduca sexta*. *Journal of Insect Physiology* **31**, 119-127
- Reynolds, S. E. and Samuels, R. I.** (1996). Physiology and Biochemistry of Insect Moulting Fluid. In *Advances in Insect Physiology*, vol. 26 (ed. P. D. Evans). London: Academic Press.
- Richards, S. Gibbs, R. A. Weinstock, G. M. Brown, S. J. Denell, R. Beeman, R. W. Gibbs, R. Bucher, G. Friedrich, M. Grimmelikhuijzen, C. J. et al.** (2008). The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**, 949-55.
- Rogez, D., Cerf, R., Andrianjara, R., Salehi, S. T. and Fouladgar, H.** (1987). Ultrasonic studies of proton-transfer reactions at the catalytic site of alpha-chymotrypsin. *FEBS Lett* **219**, 22-6.
- Roh, K. B., Kim, C. H., Lee, H., Kwon, H. M., Park, J. W., Ryu, J. H., Kurokawa, K., Ha, N. C., Lee, W. J., Lemaitre, B. et al.** (2009). Proteolytic cascade for the activation of the insect toll pathway induced by the fungal cell wall component. *J Biol Chem* **284**, 19474-81.
- Ross, J., Jiang, H., Kanost, M. R. and Wang, Y.** (2003). Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships. *Gene* **304**, 117-31.
- Sambrook, J., Fritsch, E. F. and Maniatis, T.** (1989). Molecular cloning : a laboratory manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Samuels, R. I., Charnley, A. K. and Reynolds, S. E.** (1993a). An aminopeptidase from the moulting fluid of the tobacco hornworm, *Manduca sexta*. *Insect Biochem Mol Biol* **23**, 615-620
- Samuels, R. I., Charnley, A. K. and Reynolds, S. E.** (1993b). A cuticle-degrading proteinase from the moulting fluid of the tobacco hornworm, *Manduca sexta*. *Insect Biochem Mol Biol* **23**, 607-14.
- Samuels, R. I. and Reynolds, S. E.** (1993). Molting fluid enzymes of the tobacco hornworm, *Manduca sexta*: Timing of proteolytic and chitinolytic activity in relation to pre-ecdysial development. *Archives of Insect Biochemistry and Physiology* **24**, 33-44.
- Samuels, R. I. and Reynolds, S. E.** (2000). Proteinase inhibitors from the molting fluid of the pharate adult tobacco hornworm, *Manduca sexta*. *Arch Insect Biochem Physiol* **43**, 33-43.
- Sato, P. M., Lopes, A. R., Juliano, L., Juliano, M. A. and Terra, W. R.** (2008). Subsite substrate specificity of midgut insect chymotrypsins. *Insect Biochem Mol Biol* **38**, 628-33.
- Satoh, D., Horii, A., Ochiai, M. and Ashida, M.** (1999). Prophenoloxidase-activating enzyme of the silkworm, *Bombyx mori*. Purification, characterization, and cDNA cloning. *J Biol Chem* **274**, 7441-53.
- Schaefer, B. C.** (1995). Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Anal Biochem* **227**, 255-73.
- Schechter, I. and Berger, A.** (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157-62.

- Schellenberger, V., Braune, K., Hofmann, H. J. and Jakubke, H. D.** (1991). The specificity of chymotrypsin. A statistical analysis of hydrolysis data. *Eur J Biochem* **199**, 623-36.
- Srinivasan, A., Giri, A. P. and Gupta, V. S.** (2006). Structural and functional diversities in lepidopteran serine proteases. *Cell Mol Biol Lett* **11**, 132-54.
- Stein, R. L., Strimpler, A. M., Hori, H. and Powers, J. C.** (1987). Catalysis by human leukocyte elastase: mechanistic insights into specificity requirements. *Biochemistry* **26**, 1301-5.
- Sui, Y. P., Liu, X. B., Chai, L. Q., Wang, J. X. and Zhao, X. F.** (2009a). Characterization and influences of classical insect hormones on the expression profiles of a molting carboxypeptidase A from the cotton bollworm (*Helicoverpa armigera*). *Insect Mol Biol* **18**, 353-63.
- Sui, Y. P., Wang, J. X. and Zhao, X. F.** (2009b). The impacts of classical insect hormones on the expression profiles of a new digestive trypsin-like protease (TLP) from the cotton bollworm, *Helicoverpa armigera*. *Insect Mol Biol* **18**, 443-52.
- Tamura, K., Dudley, J., Nei, M. and Kumar, S.** (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596-9.
- Telang, M. A., Giri, A. P., Sainani, M. N. and Gupta, V. S.** (2005). Characterization of two midgut proteinases of *Helicoverpa armigera* and their interaction with proteinase inhibitors. *J Insect Physiol* **51**, 513-22.
- Terra, W. R., Ferreira, C., Jordao, B. P. and Dillon, R. J.** (1996). Digestive enzymes. In *Biology of the insect midgut*, eds. M. J. Lehane and P. F. Billingsley, pp. 153-194. London: Chapman & Hall.
- Tomoyasu, Y. and Denell, R. E.** (2004). Larval RNAi in *Tribolium* (Coleoptera) for analyzing adult development. *Dev Genes Evol* **214**, 575-8.
- Toprak, U., Baldwin, D., Erlandson, M., Gillott, C. and Hegedus, D. D.** (2009). Insect intestinal mucins and serine proteases associated with the peritrophic matrix from feeding, starved and moulting *Mamestra configurata* larvae. *Insect Mol Biol*.
- Towbin, H., Staehelin, T. and Gordon, J.** (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4.
- Truman, J. W.** (2005). Hormonal control of insect ecdysis: endocrine cascades for coordinating behavior with physiology. *Vitam Horm* **73**, 1-30.
- Tsukada, H. and Blow, D. M.** (1985). Structure of alpha-chymotrypsin refined at 1.68 Å resolution. *J Mol Biol* **184**, 703-11.
- Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. and McCawley, P.** (1977). The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro* **13**, 213-7.
- Venekei, I., Szilagy, L., Graf, L. and Rutter, W. J.** (1996). Attempts to convert chymotrypsin to trypsin. *FEBS Lett* **383**, 143-7.
- Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A.** (2006). Diversity of digestive proteinases in *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae. *Comp Biochem Physiol B Biochem Mol Biol* **145**, 126-37.
- Vinokurov, K. S., Elpidina, E. N., Zhuzhikov, D. P., Oppert, B., Kodrik, D. and Sehnal, F.** (2009). Digestive proteolysis organization in two closely related Tenebrionid beetles: red flour beetle (*Tribolium castaneum*) and confused flour beetle (*Tribolium confusum*). *Arch Insect Biochem Physiol* **70**, 254-79.

- Vizioli, J., Catteruccia, F., della Torre, A., Reckmann, I. and Muller, H. M. (2001). Blood digestion in the malaria mosquito *Anopheles gambiae*: molecular cloning and biochemical characterization of two inducible chymotrypsins. *Eur J Biochem* **268**, 4027-35.
- Volpicella, M., Ceci, L. R., Cordewener, J., America, T., Gallerani, R., Bode, W., Jongtsma, M. A. and Beekwilder, J. (2003). Properties of purified gut trypsin from *Helicoverpa zea*, adapted to proteinase inhibitors. *Eur J Biochem* **270**, 10-9.
- Wang, R., Lee, S. Y., Cerenius, L. and Soderhall, K. (2001). Properties of the prophenoloxidase activating enzyme of the freshwater crayfish, *Pacifastacus leniusculus*. *Eur J Biochem* **268**, 895-902.
- Wang, Y. and Jiang, H. (2004). Prophenoloxidase (proPO) activation in *Manduca sexta*: an analysis of molecular interactions among proPO, proPO-activating proteinase-3, and a cofactor. *Insect Biochem Mol Biol* **34**, 731-42.
- Ward, G. B., Mayer, R. T., Feldlaufer, M. F. and Svoboda, J. A. (1991). Gut chitin synthase and sterols from larvae of *Diaprepes abbreviatus* (Coleoptera, Curculionidae). *Arch. Insect Biochem. Physiol.* **18**, 105-117.
- Wei, Z., Yin, Y., Zhang, B., Wang, Z., Peng, G., Cao, Y. and Xia, Y. (2007). Cloning of a novel protease required for the molting of *Locusta migratoria manilensis*. *Dev Growth Differ* **49**, 611-21.
- Whitworth, S. T., Blum, M. S. and Travis, J. (1998). Proteolytic enzymes from larvae of the fire ant, *Solenopsis invicta*. Isolation and characterization of four serine endopeptidases. *J Biol Chem* **273**, 14430-4.
- Wieczorek, H., Cioffi, M., Klein, U., Harvey, W. R., Schweikl, H. and Wolfersberger, M. G. (1990). Isolation of goblet cell apical membrane from tobacco hornworm midgut and purification of its vacuolar-type ATPase. *Methods Enzymol* **192**, 608-16.
- Yan, J., Cheng, Q., Li, C. B. and Aksoy, S. (2001). Molecular characterization of two serine proteases expressed in gut tissue of the African trypanosome vector, *Glossina morsitans morsitans*. *Insect Mol Biol* **10**, 47-56.
- Yu, X. Q., Jiang, H., Wang, Y. and Kanost, M. R. (2003). Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. *Insect Biochem Mol Biol* **33**, 197-208.
- Zamyatnin, A. A. (1972). Protein volume in solution. *Prog Biophys Mol Biol* **24**, 107-23.
- Zhao, Q. L., Shen, X. J., Zhu, L. J., Yi, Y. Z., Tang, S. M., Zhang, G. Z. and Guo, X. J. (2007). Characterization of Clb1 gene promoter from silkworm, *Bombyx mori*. *Z Naturforsch C* **62**, 875-80.
- Zhu, Y. C. and Baker, J. E. (2000). Molecular cloning and characterization of a midgut chymotrypsin-like enzyme from the lesser grain borer, *Rhyzopertha dominica*. *Arch Insect Biochem Physiol* **43**, 173-84.
- Zimoch, L., Hogenkamp, D. G., Kramer, K. J., Muthukrishnan, S. and Merzendorfer, H. (2005). Regulation of chitin synthesis in the larval midgut of *Manduca sexta*. *Insect Biochem Mol Biol* **35**, 515-27.
- Zimoch, L. and Merzendorfer, H. (2002). Immunolocalization of chitin synthase in the tobacco hornworm. *Cell Tissue Res* **308**, 287-97.
- Zitnan, D., Kim, Y. J., Zitnanova, I., Roller, L. and Adams, M. E. (2007). Complex steroid-peptide-receptor cascade controls insect ecdysis. *Gen Comp Endocrinol* **153**, 88-96.

**Zitnan, D., Ross, L. S., Zitnanova, I., Hermesman, J. L., Gill, S. S. and Adams, M. E.** (1999). Steroid induction of a peptide hormone gene leads to orchestration of a defined behavioral sequence. *Neuron* **23**, 523-35.

**Zou, Z., Lopez, D. L., Kanost, M. R., Evans, J. D. and Jiang, H.** (2006). Comparative analysis of serine protease-related genes in the honey bee genome: possible involvement in embryonic development and innate immunity. *Insect Mol Biol* **15**, 603-14.

## 7. Appendix

### 7.1. Abbreviations

AAPF	N-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide
BAPNA	N <sub>α</sub> -benzoyl-D,L-arginine 4-nitroanilide hydrochloride
bp	Base pair
BSA	Bovine serum albumin
cDNA	complementary Deoxyribonucleic acid
CHS	Chitin synthase
CNS	Central nervous system
CTLP	Chymotrypsin-like peptidase
Da	Dalton
DEPC	Diethylpyrocarbonate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dsRNA	Double stranded RNA
EDTA	ethylenediaminetetraacetic acid
GuHCL	Guanidinium hydrochloride
HP	Hemolymph peptidase
IgG	Immunoglobulin G
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger RNA
Ms	<i>Manduca sexta</i>
NLS	Sodium lauroyl sarcosinate
PAP	Prophenoloxidase-activating protein
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PGRP	Peptidoglycan recognition protein
PO	Phenoloxidase
PPAE	Prophenoloxidase-activating enzyme
PPAF	Prophenoloxidase-activating factor

---

proPO	Prophenoloxidase
RACE-PCR	Rapid amplification of cDNA ends-PCR
RNA	Ribonucleic acid
RNAi	RNA interference
rpm	Rounds per minute
rRNA	Ribosomal RNA
SAE	Spätzle-processing enzyme-activating enzyme
SDS	Sodium dodecyl sulfate
Serpin	Serine peptidase inhibitor
SPE	Spätzle-processing enzyme
SPH	Serine peptidase homologue
TBS	TRIS buffered saline
Tc	<i>Tribolium castaneum</i>
TLP	Trypsin like peptidase
TRIS	tris(hydroxymethyl)aminomethane
U	Unit
Ver	Tryptophan oxygenase
v/v	volume/volume percentage
w/v	weight/volume percentage
x g	-times gravitational acceleration
X-Gal	bromo-chloro-indolyl-galactopyranoside

## 7.2. Bacterial strains and insect cell lines

Table 4. Bacterial strains and insect cell lines

Strain	Genotype/Origin	Application
<b>DH5<math>\alpha</math></b>	F <sup>-</sup> , $\Phi$ 80 <i>dlacZ</i> $\Delta$ M15, $\Delta$ ( <i>lacZYA-argF</i> )U169, <i>endA1</i> , <i>recA1</i> , <i>hs-dR17</i> ( <i>r<sub>K</sub>m<sub>K</sub></i> ), <i>deoR</i> , <i>thi-1</i> , <i>supE44</i> , $\lambda$ <sup>-</sup> <i>gyrA96</i> , <i>relA1</i>	Amplification of plasmid DNA
<b>Rosetta (DE3)pLysS</b>	F <sup>-</sup> <i>ompT hsdSB</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>gal dcm</i> (DE3) pLysSRARE (Cam <sup>R</sup> )	Expression of recombinant proteins
<b>DH10Bac</b>	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 recA1 endA1 araD139 <math>\Delta</math>(<i>ara, leu</i>)7697 <i>galU galK</i> <math>\lambda</math><sup>-</sup> <i>rpsL nupG</i> /pMON14272 / pMON7124</i>	Amplification of baculovirus
<b>Sf21</b>	Pupal ovarian tissue of <i>Spodoptera frugiperda</i> (Vaughn <i>et al.</i> , 1977)	Expression of recombinant proteins
<b>High-Five</b>	Ovarian cells of <i>Trichoplusia ni</i> (Davis <i>et al.</i> , 1992)	Expression of recombinant proteins

## 7.3. Antibodies

Table 5. Primary antibodies

Antibody	p/m	Host	Dilution	Reference
<b>Anti-MsCHS</b>	polyclonal	Rabbit	1:1000	(Zimoch and Merzendorfer, 2002)
<b>Anti-MsCTLP1</b>	polyclonal	Rabbit	1:1000	(Broehan <i>et al.</i> , 2007)
<b>Anti-MsV<sub>1</sub>A (221-9)</b>	monoclonal	Mouse	1:10	(Klein <i>et al.</i> , 1991)
<b>Anti-TcCTLP-5C<sub>2</sub></b>	polyclonal	Mouse	1:100	(Broehan <i>et al.</i> , 2010)
<b>Anti-TcCTLP-6C</b>	polyclonal	Rabbit	1:10000	(Broehan <i>et al.</i> , 2010)
<b>Anti-Penta-His</b>	polyclonal	Mouse	1:1000	Qiagen
<b>Anti-DIG-Gold</b>	Gold	Sheep	1:10000	Roche

Table 6. Secondary antibodies

Antibody	Conjugate	Host	Dilution	Manufacturer
<b>Anti-rabbit</b>	AP	Goat	1:10000	Sigma
<b>Anti-rabbit</b>	Cy3	Sheep	1:10000	Sigma
<b>Anti-mouse</b>	AP	Goat	1:30000	Sigma
<b>Anti-DIG (F<sub>AB</sub>)</b>	AP	Sheep	1:10000	Roche

## 7.4. Oligonucleotides

Table 7. Oligonucleotides

Primers	Forward primer (5' to 3')	Reverse primer (5' to 3')
<b>TcCTLP-5A</b>	CGCAGATGCACAAGGTGTTTC	GTACACATCAGGCACTCCTC
<b>TcCTLP-5B</b>	ATCAACAGGTGCTCGTTTCG	ATGAGTGGACCTCCAGAATC
<b>TcCTLP-5C</b>	GGTCATTTCCACGACGGTTC	ACGCCTCCTTCCATTCTTCC
<b>TcCTLP-5C<sub>1</sub></b>	GACCTCCTCAAATCTTATC	
<b>TcCTLP-5C<sub>2</sub></b>	GCAAACAAAGCCGACTACCC	
<b>TcCTLP-6A</b>	CTAAACCCTCTCCGGTCTG	CACGGGCTTACCAAATCGC
<b>TcCTLP-6C</b>	TCTGTAACAGTCGGCCAAAC	GCTCCTTCTCCAAGTAAAGC
<b>TcCTLP-6D</b>	GAATCTTGGGCGTCCACGAG	GGTTCAGGCAACCCGATAGG
<b>TcCTLP-6E</b>	TCTAACGGCCGGTCAATTGCG	GCGTGAAACCGGCAGGTTTG
<b>TeRPS6</b>	AGATATATGGAAGCATCATGAAGC	CGTCGTCTTCTTTGCTCAAATG
<b>TcCTLP-5C-5'-Out</b>		GGATTGACAAAAATATGGAA
<b>TcCTLP-5C-5'-In</b>		ACGCCTCCTTCCATTCTTCC
<b>TcCTLP-5C-3'-Out</b>	ATGAAGCGTTTGTGTTTTTAC	
<b>TcCTLP-5C-3'-In</b>	CCTTAATCGACCACGAACAC	
<b>TcCTLP-5C<sub>2</sub>-BH</b>	TACTCAGGATCCATGAAGCGTTTGTGTTTTTAC	TACTCAAAGCTTCTAATGGTGATGATGGTGATGGTGATGTCCATATTTTGTCAATCC
<b>TcCTLP-6C-BH</b>	TACTCAGGATCCATGTACCGAATTTAATCAC	TACTCAAAGCTTCTAATGGTGATGATGGTGATGGTGATGCGGTTTTATTACCGAATTAA
<b>TcCTLP-5A-T7</b>	TAATACGACTCACTATAGGAACTCTCACTTATG	TAATACGACTCACTATAGGATCAGGCACTCCTCG
<b>TcCTLP-5B-T7</b>	TAATACGACTCACTATAGGACCAGCTTATACTGT	TAATACGACTCACTATAGGATTACGAACCGCATC
<b>TcCTLP-5C-T7</b>	TAATACGACTCACTATAGGCGTTTTTCCCGTTT	TAATACGACTCACTATAGGCCCCCGCTTCCGTT
<b>TcCTLP-6A-T7</b>	TAATACGACTCACTATAGGCACTGGAGGTGTAGG	TAATACGACTCACTATAGGCCGTATATTCCAAGA
<b>TcCTLP-6C-T7</b>	TAATACGACTCACTATAGGAAATCGTTGGCGAAT	TAATACGACTCACTATAGGGTCCCTCAGACCCCA
<b>TcCTLP-6D-T7</b>	TAATACGACTCACTATAGGGGCACGCGTCACCG	TAATACGACTCACTATAGGTGTGAATTCAGAGCC
<b>TcCTLP-6E-T7</b>	TAATACGACTCACTATAGGAGTATACTCTAGCTA	TAATACGACTCACTATAGGGTCAAGTCCGTCATA
<b>TcCTLP-5A-ORF</b>	AAAAGTGTAATATTGTTGAC	AATGTTTGTGTTTTCTTTAA
<b>TcCTLP-5B-ORF</b>	AAAGCAATAATAATTGTATT	GTTTTGTAATGGTTTTTGTAA
<b>TcCTLP-5C-ORF</b>	ATGAAGCGTTTGTGTTTTTAC	TCATTCCATATTTTGTCAATCC
<b>TcCTLP-6A-ORF</b>	AAACTATTTTTCGTGTTTGC	TTTGGCAAGTCGTTTACGT
<b>TcCTLP-6C-ORF</b>	TACCGAATTTAATCACTAC	CGGTTTTATTACCGAATTAA
<b>TcCTLP-6D-ORF</b>	AACCTACTCATCTTTCTAAC	TTGGGGCAAATCGTGGACAT
<b>TcCTLP-6E-ORF</b>	CGTTTTCTCAATTTAATTTT	AATACCAGTTTTTCCTTTAA

Primers	Forward primer (5' to 3')	Reverse primer (5' to 3')
MsCTLP1-5'-In		GTAGGCACAATGATCTCCAG
MsCTLP1-5'-Out		GAAGGTGCACAGCTCGATGT
MsCTLP2-DIG	CTGGAGATCATTTGTGCCTAC	ATACCACACCGATCTGCTGG
MsCTLP2-5'-In		CTAAGTGCCTCATAGTAGCA
MsCTLP2-5'-Out		TTTGAACGCGTACATCTGA
MsCTLP3-Deg	ATGTTCCCTNGGCGTGGCCACNGTNAAC	TCCATNGCNTCGCGCATCGTNAG
MsCTLP3-5'-In		TGTGTTACCCTAACGATGC
MsCTLP3-5'-Out		CACTGTTCCAGCGATTGGTA
MsCTLP3-3'-In	TGCGAACCTGCTTCAACTGA	
MsCTLP3-3'-Out	ACTCCAGAGTTGCTGGATGG	
MsCTLP3-DIG	TGACGGTAACACACTCAGCA	CGGAGTCACCGTTGCAAGTT
MsCT-P	TTGCGCTTCTCTGCCTACG	GTCCTGGTTTCGAGTCTGA
MsCTLP1-SP	CCTTCCCAAAGCTCGAAGAT	GAGAACACCACCGGCACAT
MsCTLP2-SP	CCTTCCCTATTGAGGAGCCG	GTTCAACCTTCTCCGATCC
MsCTLP3-SP	CCTTCCCAGCTCCTGAAGCA	GCGCGGACAGCATTGGTCAT
MsCTLP4-P	TTCTTCGAGCATGTGGACCAA	AGTTCATCCCACCTCCGATAA
MsCTLP1-NP	TTCGATGAAATACTTCGGGC	CGAGCGACGGTGTACATGGT
MsCTLP2-NP	GCCGCTCTAGAACTAGTGGA	TTTGAACGCGTACATCTGA
MsCTLP3-NP	TAGCGGTAACACACTCAGCA	CGGAGTCACCGTTGCAAGTT
MsCT	CTTGGATCTAACCCTTGTT	TAGTACGACC GAAACCGGAA
MsCTLP1	TCTTCACGCAGCTCGATTGG	GCGAGGGAACCGAACTGAA
MsCTLP2	TCTTCGAGCATGTGGACGTC	GGTGAACATTGTTGCCGACT
MsCTLP3	TCTTCGAGCATGTGGACCGC	CTGAGTGTGTTACCCTGAC
MsCTLP4	TCGGCACTAACCGCTGGGCAT	AGTTCATCCCACCTCCGATAA
MsCHS2	TTGGACCGCATAGGGACCTC	CGCAAGGGTCTGCATTTCTC
MsRPS6	TACGCTGAGAAAGTTGCC	CATGGACTTGGCTCTCTG
MsCTLP1-BH	TACTCAGGATCCATGTACGTGAAAGTAGCACT	TACTCAAAGCTTCTAATGGTGATGATGGTGATGA TGGTGACGATGCTCTGCTG
MsCTLP1-Z-Bip	TACTCAACTAGTTTCCCAAAGCTCGAAGATGA	
MsCTLP1-A-Bip	TACTCAACTAGTATCGTGGGTGGTACCCAGGC	
MsCTLP1/3		GATAGCACCACCGCCCTGATCCTGCCCA
MsCTLP3/1		AGGTGGTCTCATGTTGAGGCTGGCGGCGAC
MsCT3		AGGCGCTGAGCCTTACGAAC
MsCTLP3-BP	TACTCAGGATCCATGGCGATTAAATTTTAGT	TACTCACTGCAGTTAGTGGTGGTGGTGGTGGTGG TGGTGGTTGATGGTGTGCTCTGGA
MsRPS3	TACGCTGAGAAAGTTGCC	CATGGACTTGGCTCTCTG



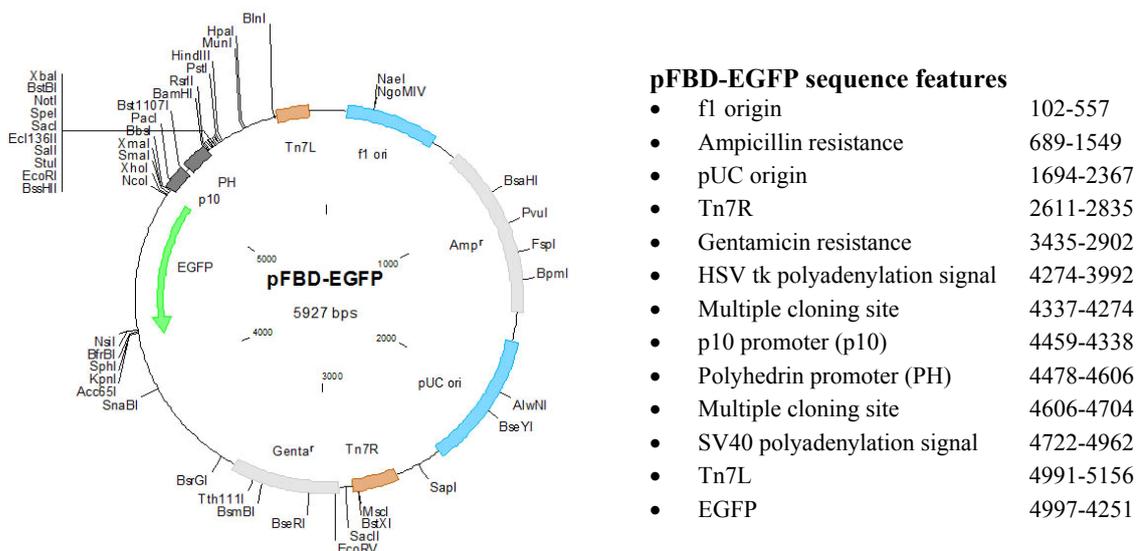


Fig. 56. The pFastBac Dual - EGFP vector map and sequence features

## 7.6. Kits

The following kits were used according to the manufacturer's recommendations:

Table 8. Kits

Kit	Application	Manufacturer
RNA-easy Mini Kit	Total RNA isolation	Qiagen
SuperScript III First-Strand Synthesis System	First strand cDNA synthesis	Invitrogen
ThermoScript RT-PCR System	First strand cDNA synthesis	Invitrogen
FirstChoice RLM-RACE Kit	Amplification of unknown 3' and 5' mRNA regions	Ambion
QIAquick Gel Extraction Kit	Purification of DNA from agarose-gels	Qiagen
QIAprep Spin Miniprep Kit	Plasmid preparations	Qiagen
DIG DNA Labeling Mix	DNA-Probe synthesis	Roche
DIG RNA Labeling Kit (Sp6/T7)	RNA-Probe synthesis	Roche
AmpliScribe T7-Flash transcription kit	Synthesis of dsRNA	Epicentre
IntenSETM silver enhancement kit	<i>In situ</i> hybridization	GE Amersham
Protein G Immunoprecipitation kit	Co-immunoprecipitation of proteins	Sigma
DIG Glycan Differentiation Kit	Detection of Glycoproteins	Roche
Affi-Gel Hz Immunoaffinity Kit	Purification of native proteins	Bio-Rad

## 7.7. Curriculum vitae

### Personal Data

Name: Gunnar Bröhan  
Address: Kommenderiestraße 106, 49080 Osnabrück  
Contact: 0541/4074975;  
g.broehan@gmx.net  
Day of birth: 29.07.1977  
Birthplace: Stade  
Marital status: unmarried

### Studies and education

2006-2010 PhD student at the University of Osnabrück  
Stipendiary of the Graduate-College 612  
Follow-up support by the SFB 431  
Thesis title: „Chymotrypsin-like peptidases in insects”

1998-2005 Studies in biology at the University of Osnabrück,  
Diploma in biology  
Thesis title: „Zur Rolle der NHE-Familie im  
Insektendarm.“

1989-1997 Vincent-Lübeck Gymnasium Stade,  
High school graduation

## 7.8. Publications

**Broehan, G., Arakane, Y., Beeman, R. W., Kramer, K. J., Muthukrishnan, S. and Merzendorfer, H.** (2010). Chymotrypsin-like peptidases from *Tribolium castaneum*: a role in molting revealed by RNA interference. *Insect Biochem Mol Biol* **40**, 274-83.

**Broehan, G., Kemper, M., Driemeier, D., Vogelpohl, I. and Merzendorfer, H.** (2008). Cloning and expression analysis of midgut chymotrypsin-like proteinases in the tobacco hornworm. *J Insect Physiol* **54**, 1243-52.

**Broehan, G., Zimoch, L., Wessels, A., Ertas, B. and Merzendorfer, H.** (2007). A chymotrypsin-like serine protease interacts with the chitin synthase from the midgut of the tobacco hornworm. *J Exp Biol* **210**, 3636-43.

## 7.9. Erklärung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich / unentgeltlich geholfen.

1. MsCTLP1 als potentieller Interaktionspartner von MsCHS2 wurde mittels eines Yeast-two-hybrid Screenings von Anton Wessels während seiner Diplomarbeit identifiziert. Im Rahmen weiterer Diplomarbeiten von Michael Kemper und Beyhan Ertas wurden Teilsequenzen bzw. vollständige Sequenzen von *MsCTLP1*, *MsCTLP2* und *MsCTLP4* kloniert.

Weitere Personen waren an der inhaltlichen materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

.....  
(Ort, Datum)

.....  
(Unterschrift)

## 8. Acknowledgements

First of all I would like to thank Prof. Dr Hans Merzendorfer not only for excellent supervision but also for giving me the opportunity to do research also on topics of my own choice. I also have to thank him for all the inspiring discussions and for always having an open ear for new ideas or problems. Not to forget about the opportunity to visit the Kansas State University even twice during my PhD thesis. I profited very much from the stay, not only scientifically. And which biologist would not agree that seeing a real bison herd would have already been worth the trip!?

I would like to thank Prof Dr. Helmut Wiczorek for giving me the opportunity to work on my dissertation project in his department. Furthermore I am grateful that he agreed to be a member of my examination board.

I thank Prof. Dr. Achim Paululat for writing the second opinion on my dissertation.

Dr. Vera Linke I thank for unhesitatingly agreeing to be a member of my examination board.

The Graduate College 612 I thank for funding me throughout my stipendiary. All the members of the Graduate College I thank for the good cooperation and for three nice conferences.

Many thanks to all the members of the “Tierphysiologie” department at the University of Osnabrück. In the first place I would like to thank Derek Meissner for his friendship, all the inspiring discussions and for bringing color into everyday-lab-life. But also Christin Nardmann, Svenja Bockelmann, Felix Tiburcy and Jothini Odman Naresh shall not be forgotten. I had a great time working with all of them.

A special thanks goes to Margret Düvel and Ulla Mädler whose help in maintaining the cell culture and in doing cryosections is highly appreciated.

I also like to thank Christian Kaufmann, our guest scientist from Switzerland. He became a good friend and it was a pleasure to work with him. He has the rare gift to inspire people for science.

I owe very special thanks to Asha Muthukrishnan and Prof. Dr. Subbaratnam Muthukrishnan who took me up in their household like a son of their own. Furthermore I have to thank Dr. Krishnan for giving me the opportunity to work in his department, for all the discussions about biology and for sharing his expertise with me. I have to thank Asha for all the idlis and dosas and the supply with all the other wonderful Indian food. I rarely met people as kind as they are; those two made me feel at home in a foreign country.

Of course, I have to thank all the other people, which I have worked with at the Kansas State University. I thank Dr. Karl Kramer for margaritas, salsa and the good Mexican food and for showing me a new perspective on science. Dr. Yasuyuki Arakane, Mr. RNAi himself, I thank for teaching me the RNAi injection

technique and for always having an open ear for my questions. Dr. Dick Beeman I thank for interesting discussions on our Thursday meetings. I profited a lot from their expertise and I really appreciate it.

My lab mates at Kansas State University, Sujata Chaudhari and Sinu Jasrapuria, I thank for the good collaboration. I really had a wonderful time working with them.

I also have to thank Dr. Norbert for always putting me in a good mood and thus helping me to finish this work.

My deepest and sincere thanks go to my girlfriend Katharina Richter for her love and patience, even if this meant another weekend we could not meet. She continuously encouraged me to do my best and she was always there for me even when I was far away. And I have to thank her for not being too annoyed with me when I confessed to her that I will leave for a further stay at the Kansas State University.

Last but not least, I owe my deepest gratitude to my parents for all their patience, support, encouragement and their love. I have to thank them for always believing in me. Without them, this would not have been possible.