CHITIN METABOLISM IN INSECTS: CHITIN SYNTHASES AND *BETA*-*N*-ACETYLGLUCOSAMINIDASES

by

DAVID GEORGE HOGENKAMP

B.S., Brock University, Canada, 1998

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Graduate Biochemistry Group
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2006

ABSTRACT

Chitin, a linear homopolymer of *beta*-1,4-linked *N*-acetylglucosamine, is the second most abundant biopolymer next to cellulose. It is the major structural polysaccharide in the insect's exoskeleton and gut lining. An extensive study of two of the major genes encoding enzymes involved in chitin metabolism, chitin synthases (CHSs) and *beta*-*N*-acetylglucosaminidases (NAGs), was undertaken. CHS genes from the tobacco hornworm, *Manduca sexta*, and NAG genes from the red flour beetle, *Tribolium castaneum*, were identified and characterized.

In general, chitin deposition occurs in two major extracellular structures of insects, the cuticle that overlays the epidermis, and the peritrophic membrane (PM) that lines the midgut. Only two CHS genes were identified in *M. sexta* using Southern blot analysis. Extensive expression studies of both *M. sexta* CHS genes, *MsCHS1* and *MsCHS2*, suggest a strict functional specialization of these two genes for the synthesis of epidermal and PM-associated chitin, respectively. Furthermore, two alternatively spliced transcripts of *MsCHS1*, *MsCHS1a* and *MsCHS1b*, were identified. Analysis of the levels of these transcripts in different tissues and stages of development indicated that the *MsCHS1a* transcript predominates in the integument during the feeding and pupal stages, whereas the *MsCHS1b* transcript is more abundantly present in the tracheae, foregut, and hindgut during all developmental stages tested.

Four genes encoding putative NAGs (*TcNAG1*, *TcNAG2*, *TcNAG3*, and *TcNAG4*) were identified by searching the *Tribolium* genomic database. The full-length cDNAs for all four NAGs were cloned and sequenced, and the exon-intron organizations were determined. Studies on developmental expression patterns of each gene indicated that they are expressed during most developmental stages with *TcNAG1* being the predominant one. The function of each NAG was assessed by down regulating the level of each transcript at various developmental stages using RNA interference. Selective knock down of each transcript, without significant reduction in the expression levels of the other NAG transcripts, was verified and the resulting phenotypes were documented. Knockdown of *TcNAG1* interrupted larval-larval, larval-pupal, and pupal-adult molting, and the insects were unable to completely shed their old cuticles.

CHITIN METABOLISM IN INSECTS: CHITIN SYNTHASES AND *BETA-N*-ACETYLGLUCOSAMINIDASES

by

DAVID GEORGE HOGENKAMP

B.S., Brock University, Canada, 1998

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Graduate Biochemistry Group
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2006

Approved by:	Approved by:
Co-Major Professor	Co-Major Professor
Subbaratnam Muthukrishnan	Karl J. Kramer

COPYRIGHT

ABSTRACT

Chitin, a linear homopolymer of *beta*-1,4-linked *N*-acetylglucosamine, is the second most abundant biopolymer next to cellulose. It is the major structural polysaccharide in the insect's exoskeleton and gut lining. An extensive study of two of the major genes encoding enzymes involved in chitin metabolism, chitin synthases (CHSs) and *beta*-*N*-acetylglucosaminidases (NAGs), was undertaken. CHS genes from the tobacco hornworm, *Manduca sexta*, and NAG genes from the red flour beetle, *Tribolium castaneum*, were identified and characterized.

In general, chitin deposition occurs in two major extracellular structures of insects, the cuticle that overlays the epidermis, and the peritrophic membrane (PM) that lines the midgut. Only two CHS genes were identified in *M. sexta* using Southern blot analysis. Extensive expression studies of both *M. sexta* CHS genes, *MsCHS1* and *MsCHS2*, suggest a strict functional specialization of these two genes for the synthesis of epidermal and PM-associated chitin, respectively. Furthermore, two alternatively spliced transcripts of *MsCHS1*, *MsCHS1a* and *MsCHS1b*, were identified. Analysis of the levels of these transcripts in different tissues and stages of development indicated that the *MsCHS1a* transcript predominates in the integument during the feeding and pupal stages, whereas the *MsCHS1b* transcript is more abundantly present in the tracheae, foregut, and hindgut during all developmental stages tested.

Four genes encoding putative NAGs (*TcNAG1*, *TcNAG2*, *TcNAG3*, and *TcNAG4*) were identified by searching the *Tribolium* genomic database. The full-length cDNAs for all four NAGs were cloned and sequenced, and the exon-intron organizations were determined. Studies on developmental expression patterns of each gene indicated that they are expressed during most developmental stages with *TcNAG1* being the predominant one. The function of each NAG was assessed by down regulating the level of each transcript at various developmental stages using RNA interference. Selective knock down of each transcript, without significant reduction in the expression levels of the other NAG transcripts, was verified and the resulting phenotypes were documented. Knockdown of *TcNAG1* interrupted larval-larval, larval-pupal, and pupal-adult molting, and the insects were unable to completely shed their old cuticles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations	Name
Aa	*Aedes aegypti*
A. aegypti	*Aedes aegypti*
Ag	*Anopheles gambiae*
A. gambiae	*Anopheles gambiae*
CHI	chitinase
CHS	chitin synthase
Dm	*Drosophila melanogaster*
D. melanogaster	*Drosophila melanogaster*
dsRNA	double-stranded RNA
ENG	endo-*β*-*N*-acetylglucosaminidase
FDL	*fused lobes*
GH	glycosylhydrolase
GlcNAc	*N*-acetylglucosamine
GT	glycosyltransferase
Lc	*Lucilia cuprina*
L. cuprina	*Lucilia cuprina*
Ms	*Manduca sexta*
M. sexta	*Manduca sexta*
MW	molecular weight
NAG	*β*-*N*-acetylglucosaminidase
ORF	open reading frame
PCR	polymerase chain reaction
PM	peritrophic membrane
RNAi	RNA interference
RpS3	ribosomal protein S3 (*Manduca sexta*)
RpS6	ribosomal protein S6 (*Tribolium castaneum*)
RT	reverse transcription
Tc	*Tribolium castaneum*

T. castaneum	*Tribolium castaneum*
TMS	transmembrane segment
UDP	uridine diphosphate
UDP-GlcNAc	uridine diphosphate-*N*-acetylglucosamine
UPGMA	unweighted pair group method with arithmetic mean
UTR	untranslated region

ACKNOWLEDGEMENTS

I would like to sincerely thank my major professors, Prof. Subbaratnam Muthukrishnan and Dr. Karl J. Kramer, for their support, guidance, and encouragement. I greatly appreciate everything they have done throughout my studies. I would also like to sincerely thank Dr. Richard W. Beeman for his valuable assistance throughout my studies and for serving on my dissertation committee. Thanks also to Prof. Ramaswamy Krishnamoorthi for serving on my committee and Prof. Barbara Valent for serving as the outside chairperson on my committee. Special thanks also to Prof. Michael R. Kanost, the head of the Biochemistry Department, and Prof. Lawrence C. Davis, the chair of the Graduate Biochemistry Group, for their advice and assistance.

I am deeply indebted to Dr. Yasuyuki Arakane, a Research Assistant Professor in our laboratory, for teaching me virtually everything I learned in the laboratory, his expert advice, and valuable friendship. I would like to thank all of the current and past members of our laboratory as well as those in other labs with whom I have worked closely including Dr. Qingsong Zhu, Dr. Renata Bolognesi, Dr. Marcé Lorenzen, Radhika Dixit, Debarshi Banerjee, Peter K. Lundquist, Kendra S. Siebert, Merrilee S. Haas, and Janna Voorhees . To everyone at the Grain Marketing Production and Research Center and in the Department of Biochemistry; professors, students, and staff, thank you all.

DEDICATION

I would like to dedicate this dissertation to my wife, Theresa, and my sons, Jacob and Tyler, for their love and support.

INTRODUCTION

RATIONALE

Within the animal kingdom, insects (Insecta) are undoubtedly the most successful class. There are approximately one million species of insects and there are considerably more that have yet to be identified. For every person, there are many orders of magnitude more individual insects, which makes us grossly outnumbered by these highly successful arthropods. Insects have successfully populated virtually all terrestrial and freshwater habitats and there have even been some marine species identified.

The success of insects can be attributed to their unique characteristics, which have been acquired through several hundred million years of evolution. One such feature of insects is the presence of an exoskeleton, which provides extensive protection from the environment. A major component of the insect exoskeleton is chitin, a β-1,4-linked linear homopolymer of *N*-acetylglucosamine. Chitin is the second most abundant biopolymer on earth next to cellulose. In addition to the exoskeleton, chitin is found in the peritrophic membranes (PM) and cocoons of insects. Chitin is also widespread in other arthropods and fungi (Muzzarelli, 1977). Furthermore, chitin is found in the eggshells and pharynx of nematodes (Veronico et al., 2001). This ubiquitous biopolymer is even found in the cyst wall of *Entamoeba histolytica* (Van Dellen et al., 2006). Chitin has been implicated in a wide variety of applications including health beneficial and antimicrobial uses (Yalpani et al., 1992; Howling et al., 2001; Shahidi & Abuzaytoun, 2005). It has been estimated that 10 gigatons of chitin are synthesized and degraded in the biosphere annually (Muzzarelli, 1999). Because of its absence in higher plants and mammals, and its biological necessity in insects, chitin metabolism represents a rather selective target for insect control agents (Kramer & Muthukrishnan, 1997; 2005).

STRUCTURE OF CHITIN

Chitin is a linear polymer of β-(1,4)-linked 2-acetamido-2-deoxy- β-D-glucopyranoside. Chitin is typically present as microfibrils of varying length and diameter (Merzendorfer, 2006).

The individual chitin chains within the microfibrils are linked by hydrogen bonds between the carbonyl and amino groups of *N*-acetylglucosamine. In nature, chitin occurs in three forms known as α-, β-, and γ-chitin (Kramer & Koga, 1986). In α-chitin, the adjacent chains are arranged in an antiparallel orientation, whereas in β-chitin the chains are arranged in a parallel manner. In γ-chitin, the chains are grouped in sets of three strands where two chains form a parallel orientation and one antiparallel (Rudall, 1965). The chitin microfibrils of α-chitin are tightly packed and this form is stabilized by a high number of inter-chain hydrogen bonds. The most thermodynamically stable form is α-chitin, which is also the most abundant form (Kameda et al., 2005). Insect cuticles are primarily composed of α-chitin, which confers the highest degree of strength and stability (Giraud-Guille & Bouligand, 1986).

In contrast to α-chitin, β- and γ-chitins are less tightly packed and form a greater number of hydrogen bonds with water resulting in a greater degree of hydration relative to α-chitin (Merzendorfer, 2006). Because of its open structure, β-chitin is more facile to enzymatic modifications (Gardner and Blackwell, 1975; Saito et al., 2002). Both β-chitin and γ-chitin therefore form a more flexible and soft chitinous structure such as the PM and cocoon (Kenchington, 1973; Peters, 1992).

OCCURRENCE OF CHITIN IN INSECTS

Chitin is widespread throughout the Insecta, which utilize this remarkable biopolymer in various anatomical structures. The two major extracellular structures where chitin deposition occurs in insects are the cuticle that overlays the epidermis and the peritrophic membrane that lines the midgut. In addition, significant chitin deposition occurs in the tracheae, which form spiral thickenings within the tracheal trunk known as taenidia. The lining of the foregut and hindgut, known as intima, is also composed of chitin. Chitin is also found in the salivary glands and mouthparts of some insects (Kramer & Muthukrishnan, 2005). The chitin content of the exuvial dry mass may be as high as 40% but varies considerably depending on the insect species and cuticle type (Kramer et al., 1995).

THE INSECT EXOSKELETON

The insect cuticle (exoskeleton) is a remarkable structure and its mechanical properties are attributable in part to its α-chitin component (Vincent & Wegst, 2004). The cuticle is composed of several layers, which vary in composition, mechanical properties, and function. The cement layer is the outermost cuticular layer of the epicuticle. The waxy layer immediately underneath the cement layer is composed primarily of lipids and functions in waterproofing. Underneath the epicuticle are three distinct layers: the exocuticle, the mesocuticle, and the endocuticle, which are initially secreted as the procuticle. Varying degrees of sclerotization results in these three different forms of procuticle. Sclerotization is a chemical process where the cuticle is irreversibly transformed into a stiffer and harder structure, which is characterized by the decreased extractability of cuticular proteins and increased resistance towards enzymatic degradation (Anderson, 2005). This process involves the oxidative conjugation of catechols with cuticular proteins (Kramer et al., 2001). During sclerotization, the cuticle may appear black or various shades of brown in a process known as tanning. In the red flour beetle, *Tribolium castaneum*, *Laccase 2* is the phenoloxidase gene required for cuticle tanning (Arakane et al., 2005). The exocuticle is the most highly sclerotized structure and generally is very stiff and hard. The mesocuticle, however, is only partially sclerotized. The endocuticle, which overlays the epidermis, is composed of soft flexible protein and chitin layers, and is partially degraded prior to ecdysis. The onset of molting is characterized by apolysis, which involves the separation of the endocuticle from the epidermal cell layer (Locke & Huie, 1979). This separation results in the formation of the exuvial space where the new cuticle is formed. Enzymes necessary for the partial degradation of the endocuticle are secreted into the molting fluid, which is separated from the newly synthesized procuticle by an ecdysial membrane. Epidermal cell division precedes the deposition of newly synthesized cuticle at ecdysis (Locke, 1970). In order to assist in the shedding of the exuvia, the insect ingests air, which results in the expansion of the exoskeleton. Following this cuticle expansion, sclerotization occurs resulting in the characteristic rigid exoskeleton of the intermolt.

THE INSECT PERITROPHIC MEMBRANE

The insect peritrophic membrane (PM) lines the midgut and separates the food bolus from the midgut epithelium. The PM occurs in most but not all insects (Lehane, 1997). Formation of the PM results in the compartmentalization of the midgut, creating an endoperitrophic space, where the food bolus is located, and an ectoperitrophic space, located between the midgut epithelium and the PM. The PM functions to protect the insect against food abrasion and microbial invasion, but there are also numerous other functions it provides (Terra, 2001). Some of these functions include prevention of non-specific binding of undigested material onto the midgut cell surface, increased efficiency in digestion, and enzyme immobilization (Bolognesi et al., 2001). Furthermore, the PM facilitates endo-ectoperitrophic circulation. Here, food (inside the PM) flows from the anterior midgut to the posterior midgut, whereas water (outside the PM) flows from the posterior midgut to the caeca (pouches) located at the anterior midgut where reabsorption takes place. This endo-ectoperitrophic circulation prevents enzyme excretion and therefore digestive enzymes can be recycled.

PM component proteins are divided into one of four classes based on their varying degrees of extractability (Tellam et al., 1999). The fourth class of PM proteins are those that are not solubilized by strong denaturants. Presumably, these proteins are covalently cross-linked, thus preventing their extraction. Insect PM proteins typically contain chitin-binding domains, which bind to the chitin fibril (Wang & Granados, 2001). Peritrophins, which are important PM proteins, strongly interact with chitin presumably through their chitin-binding domains. One class of peritrophins, the mucin-like peritrophins, probably function in the lubrication of the passage of the food through the gut, in addition to protective roles. In all, the PM is a complex extracellular structure that functions in numerous vital processes of insect digestion.

CHITIN SYNTHESIS IN INSECTS

Despite its biological significance, relatively little information is known about the chitin biosynthetic pathway in insects or other invertebrates (Cohen, 2001). The chitin biosynthetic pathway begins with the production of glucose from the cleavage of trehalose by trehalase (EC 3.2.1.28, Merzendorfer & Zimoch, 2003). The last step in the chitin biosynthetic pathway is

catalyzed by chitin synthase (CHS, UDP-GlcNAc: chitin 4-β-N-acetylglucosaminyltransferase, EC 2.4.1.16), which catalyzes the polymerization of chitin from activated UDP-N-acetylglucosamine (UDP-GlcNAc) monomers. Accordingly, RNAi-mediated down-regulation of *T. castaneum* CHS genes results in the reduction of chitin content (Arakane et al., 2005). One of the members of the group of Halloween mutants in *Drosophila* is *mummy* (*mmy*). The *mmy* gene encodes the *Drosophila* UDP-*N*-acetylglucosamine-pyrophosphorylase (EC 2.7.7.23), which catalyzes the step immediately preceding CHS in the chitin biosynthetic pathway (Tonning et al., 2005a). As expected, *mmy* mutants exhibited severe cuticle defects. Similarly, mutations in the *cystic* allele, which also encodes UDP-*N*-acetylglucosamine pyrophosphorylase in *Drosophila*, results in severe tracheal abnormalities implicating its role of chitin synthesis in epithelial tube morphogenesis (Devine et al., 2005). Furthermore, chitin is required for uniform tracheal tube expansion and loss of tracheal chitin causes tubular constrictions in *Drosophila* (Tonning et al., 2005b). Similarly, chitin is required for cuticle assembly in *Drosophila* (Moussian et al., 2005).

CHSs are large enzymes located in the plasma membrane, enabling the newly synthesized chitin to be extruded from the cell into extracellular locations (Cohen, 1991). Typically, CHSs contain several transmembrane spans, which are particularly abundant at the N-terminus of the protein. CHSs belong to the GT2 family of glycosyltransferases, which also includes the closely related cellulose synthases (Coutinho et al., 2003). The GT2 family of glycosyltransferases employs an inverting mechanism of catalysis and contains a characteristic GT-A fold. The GT-A fold is characterized by two intimately associated β/α/β domains, which are composed of at least eight β-strands. The alternating orientation of the GlcNAc residues in the chitin polymer has lead to the hypothesis that CHSs contain two active sites. Evidence for the two-active site mechanism for CHS was obtained using simple uridine-derived dimeric inhibitors (Yeager & Finney, 2003). These dimeric inhibitors exhibited 10-fold greater inhibition than a monomeric inhibitor control, consistent with the two active site model. Assays using several other uridine dimers reinforced these conclusions (Yeager & Finney, 2004).

CHSs have been extensively studied in fungi (Valdivieso et al., 1999). Fungal CHSs are encoded by a large family of genes and as many as eight different CHS genes have been identified in a single fungal species (Munro and Gow, 2001). Various fungal CHSs have been found to have different roles including sporulation and cell division, and their expression varies

throughout different developmental stages (Specht et al., 1996; Merz et al., 1999; Valdivieso et al., 1999; Munro and Gow, 2001; Roncero, 2002). However, in contrast to fungi, nematodes appear to have fewer CHS genes. Following the genome sequencing of the nematode, *Caenorhabditis elegans*, only two CHS genes were identified (Veronico et al., 2001). Furthermore, one or two CHS genes have been reported in several other nematode species including *Brugia malayi*, *Meloidogyne artiellia*, and *Dirofilaria immitis* (Harris et al., 2000; Veronico et al., 2001; Harris and Fuhrman, 2002).

In contrast to fungi, insect CHSs have been characterized only recently. The first cDNA sequence was reported for a CHS from the sheep blowfly, *Lucilia cuprina*, just over five years ago (Tellam et al., 2000). Since then, CHS cDNAs/genes from several other insect species including *Drosophila melanogaster, Anopheles gambiae, Aedes aegypti, Tribolium castaneum*, and *Manduca sexta* have been characterized (reviewed in Kramer & Muthukrishnan, 2005). Fungal CHSs are divided into six classes, based on their amino acid sequence similarities (Munro and Gow, 2001). Insect CHSs, however, are divided into only two classes, classes A and B, based on the limited number of insect CHS amino acid sequences available (Fig. 1). After completion of sequencing of the genomes of *D. melanogaster* and *A. gambiae*, only two putative CHS genes were found in each species, one belonging to each class. In addition, only two putative CHS genes, one in class A and one in class B, were found in *T. castaneum* by screening a BAC library of genomic DNA and no additional CHS genes were identified by Southern blot analysis of genomic DNA (Arakane et al., 2004). Furthermore, no additional CHS genes have been identified following the recent genome sequencing of this beetle species. The characterization of CHS genes from several different insect species representing three orders (Diptera, Coleoptera, and Lepidoptera) suggests that most insects encode only two CHS genes, one belonging to each class.

In general, chitin deposition in insects occurs in two major extracellular structures. This insoluble polysaccharide is used in the assembly of the cuticle and the PM of the midgut. The division of insect CHSs into two classes may be of functional relevance. Indeed, functional specialization has been reported in fungal CHSs (Martín-Udíroz et al., 2004). In *L. cuprina*, studies of tissue specificity showed that the class A CHS is expressed in the epidermis and tracheal epithelial cells but not in the midgut (Tellam et al., 2000). Furthermore, the presence of an alternate exon has been identified in genes encoding several class A CHSs, but to date no

alternate exon has been identified in those encoding class B CHSs (Arakane et al., 2004). *TcCHS2*, the class B CHS gene in *T. castaneum*, is expressed during periods when the insect is actively feeding (Arakane et al., 2004). Furthermore, the class B CHS gene in *A. aegypti* is expressed in the midgut, and the level of expression increases following blood feeding in females (Ibrahim et al., 2000). Finally, chitin staining of the PM of the fall armyworm, *Spodoptera frugiperda*, indicated the presence of chitin only during periods when the class B CHS of this insect is expressed (Bolognesi et al., 2005). Therefore, it appears that class A CHSs specialize in cuticle chitin production, whereas class B CHSs specialize in chitin production for the PM.

CHITIN DEGRADATION IN INSECTS

The two primary enzymes responsible for the degradation of chitin in insects are chitinases (CHIs, EC 3.2.1.14) and *β*-*N*-acetylglucosaminidases (NAGs, EC 3.2.1.30). Both of these enzymes have been studied in a number of model insect species. NAGs belong to the CAZy glycoside hydrolase family 20, GH20, whereas CHIs that employ a retaining mechanism of hydrolysis (which is typical of these enzymes in insects) belong to the CAZy glycoside hydrolase family 18, GH18. Insect CHI proteins are encoded by a large and complex gene family (Zhu, 2005). CHIs catalyze the endohydrolysis of chitin at random internal positions within the chitin polymer. Exochitinases, which cleave the exohydrolysis of the chitin polymer from the nonreducing end, are divided into two subcategories: chitobiosidases (EC 3.2.1.29) and NAGs (Dahiya et al., 2005). Chitobiosidases catalyze the progressive release of diacetylchitobiose, whereas NAGs generate monomers of GlcNAc (Sahai & Manocha, 1993). In the molting fluid of *M. sexta*, insect chitin hydrolysis is carried out by a tandem binary enzyme system composed of CHI and NAG (Fukamizo & Kramer, 1985b). Presence of these two enzymes together results in a synergistic effect on chitin catabolism and the rate of hydrolysis is up to six times higher than the sum of the rates observed with either enzyme alone. This enhancement of catalytic activity by the binary enzyme chitinase system is highly sensitive to the concentration ratio of CHI to NAG. This finding implies that the regulation of both enzymes is under strict hormonal and/or developmental control because disruptions in the metabolic flux of these enzymes could potentially have severe detrimental effects on the insect.

CHI is unable to efficiently convert its chitin substrate into individual GlcNAc monomers. Therefore, the enzyme primarily responsible for the production of GlcNAc monomers is NAG. Kinetic studies on the mechanism of chitin oligosaccharide hydrolysis have revealed that chitin oligomers are produced by the action of CHI and these chitin oligomers are the substrate for NAG (Fukamizo et al., 1985a). These short chitin fragments can then be cleaved by NAG, resulting in the release of GlcNAc monomers (Sahai & Manocha, 1993). Kinetics studies with *M. sexta* CHI have revealed that this enzyme is subject to substrate inhibition (Koga et al., 1982; Koga et al., 1983). Therefore, a potential function of NAG may involve the prevention of the accumulation of high concentrations of chitin oligomers that can prevent the efficient degradation of chitin.

It is unknown if a recycling mechanism exists for GlcNAc in insects (Fig. 2). A recycling pathway for GlcNAc has been characterized in *E. coli* (Park et al., 2001). One of the components of this pathway is the enzyme *N*-acetylglucosamine kinase (EC 2.7.1.59), which catalyzes the transfer of a phosphate group from ATP to GlcNAc, producing GlcNAc-6-P (Asensio & Ruiz-Amil, 1966). Murine and human *N*-acetylglucosamine kinase homologs have been identified (Berger et al., 2002; Hinderlich et al., 2000). If such an enzyme exists in insects, it is conceivable that GlcNAc could re-enter the chitin synthesis pathway as *N*-acetylglucosamine-6-phosphate.

M. sexta has been utilized as a model species for the study of CHIs and NAGs. A NAG gene from *M. sexta* was found to be expressed primarily in the epidermal and midgut tissues (Zen et al., 1996). The level of expression in these tissues was highest at six to seven days following the fifth instar molt. This enzyme likely functions in the degradation of chitin and its expression has been shown to be under the control of the hormone 20-hydroxyecdysone. NAGs have also been purified from another lepidopteran, the silkworm, *Bombyx mori* (Nagamatsu et al., 1995). Furthermore, NAGs have been identified in several other insect species including the yellow fever mosquito, *A. aegypti* (Filho et al., 2002). This enzyme was found to be present in the gut of unfed mosquitoes, but the NAG activity increased rapidly upon feeding on an artificial protein-free diet or blood.

Some NAGs that have been identified in a variety of insect species have been implicated in various roles other than chitin catabolism. An unusual membrane-bound NAG has been identified in the lepidopteran insect cell line, Sf21 (Altmann, et al., 1995). Through

ultracentrifugation experiments, it was determined that this enzyme exists in both a soluble and membrane-bound form. The membrane-bound form accounted for two-thirds of the total NAG activity. Studies of the specificity of this NAG indicated that the enzyme exclusively removed the terminal GlcNAc residue attached to the α1,3-linked mannose of the pentasaccharide core of *N*-glycans. The specificity of this NAG therefore implies a function in the processing of protein *N*-glycans. However, enzyme assays indicated that this NAG could also hydrolyze the chitin oligomer, chitotriose. Therefore, this unusual NAG may have diverse functions such as *N*-glycan processing as well as chitin oligosaccharide catabolism.

Another NAG was recently identified in *D. melanogaster* and its corresponding gene has been termed the *fused lobes* (*fdl*) gene because homozygous mutants showed a fusion of the β-lobes in the adult brain (Boquet et al., 2000). When the recombinant *Drosophila fdl* (Dmfdl) was expressed in *Pichia pastoris*, the purified enzyme exhibited a similar specificity for *N*-glycans as the unique NAG from Sf21 cells (Léonard et al., 2006). Dmfdl exclusively hydrolyzed the terminal GlcNAc residue attached to the α1,3-linked mannose of the core pentasaccharide of *N*-glycans. In contrast to the Sf21 NAG, the recombinant enzyme did not cleave chitooligosaccharides. However, when *Drosophila* HEXO1 and HEXO2 were expressed in *P. pastoris*, both recombinant enzymes cleaved chitooligosaccharides but not *N*-glycans. Finally, a NAG has been identified in *Caenorhabditis elegans*, which exhibits a specificity for *N*-glycans similar to that of Dmfdl, indicating that this NAG is perhaps not as unique as previously thought (Zhang et al., 2003).

Another class of NAGs, known as endo-β-N-acetylglucosaminidases (mannosyl-glycoprotein endo- *β* -*N*-acetylglucosaminidase, ENG, EC 3.2.1.96) belongs to the CAZy glycoside hydrolase family 18 (GH18), the same family as CHIs. In contrast, the chitinolytic NAGs and Dmfdl homologs belong to the CAZy glycoside hydrolase family 20 (GH20). ENG catalyzes the endohydrolysis of the di-N-acetylchitobiosyl unit in high-mannose glycopeptides and glycoproteins containing the [Man $(GlcNAc)_2$]Asn-structure. Following this cleavage reaction, one *N*-acetyl-D-glucosamine residue remains attached to the protein and the rest of the oligosaccharide is released intact. Akin to Dmfdl, ENG also functions in the processing of *N*-glycans as well as free chitooligosaccharides in the cytosol (Suzuki et al., 2002).

Although the crystal structure of an insect NAG is yet to be completed, structures are currently available for the closely related human hexosaminidase β-chain, *HEXB* (Maier et al.,

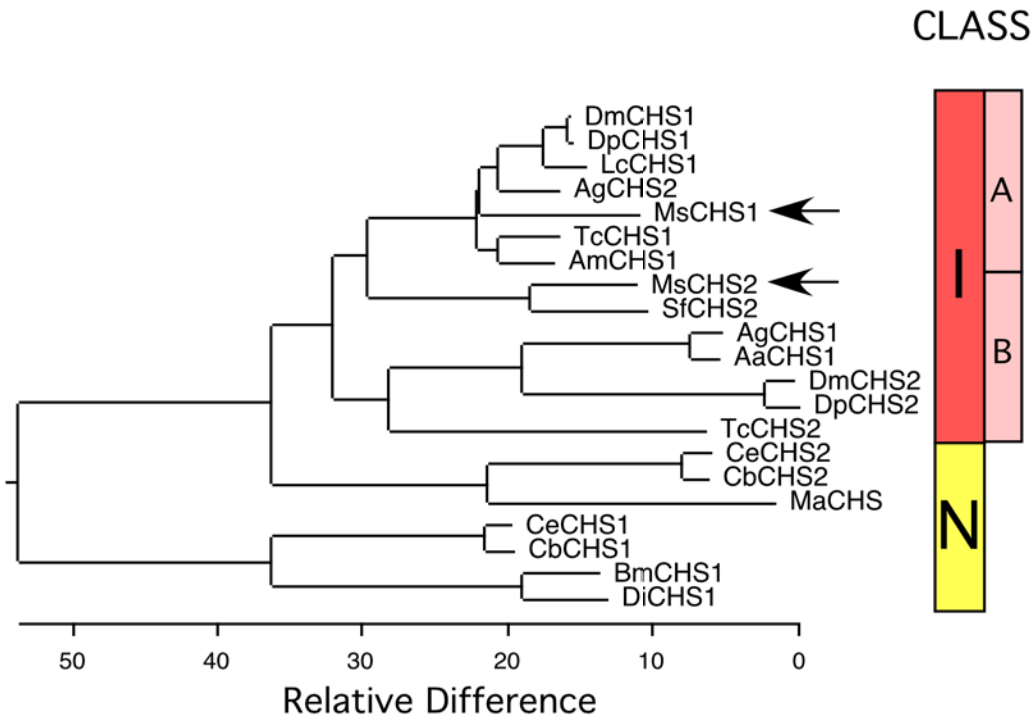
2003). The human β-chain of *HEXB* is one of two lysosomal β-hexosaminidases and mutations in this gene results in Sandhoff disease, a fatal neurodegenerative disorder. The *HEXB* β-chain forms a heterodimer with the α-chain hexosaminidase, which is encoded by the gene *HEXA*. Homodimeric hexosaminidases are also formed from two *HEXB* β-chains or two *HEXA* α-chains. Dimerization of hexosaminidases is required for the formation of a catalytically active enzyme. Mutations in the human *HEXA* gene are responsible for the lysosomal storage disorder, Tay-Sachs disease. Evidence for multimeric NAGs exists in a limited number of insects. A NAG purified from *T. castaneum* exhibited a dimeric structure with subunits of different molecular weights (Kramer & Aoki, 1987). Similarly, a NAG from *B. mori* exhibited a dimeric structure, although it is probably composed of identical subunits (Kimura, 1974).

IMPACT

In the present study, a characterization of all CHS genes in *M. sexta* and NAG genes in *T. castaneum* was undertaken. The number of CHS and NAG genes was determined in these species and the transcripts of each of the genes were fully characterized. Extensive developmental expression studies of these genes were conducted using multiple methods of analysis. The findings from these studies suggest a specialization of these genes for the synthesis/degradation of cuticular and PM-associated chitin. Furthermore, RNAi-mediated down-regulation studies of NAG genes demonstrates specialized functions of these enzymes at various molts. This is one of only a limited number of studies employing reverse genetics, RNAi, to study an important class of chitin metabolism enzymes.

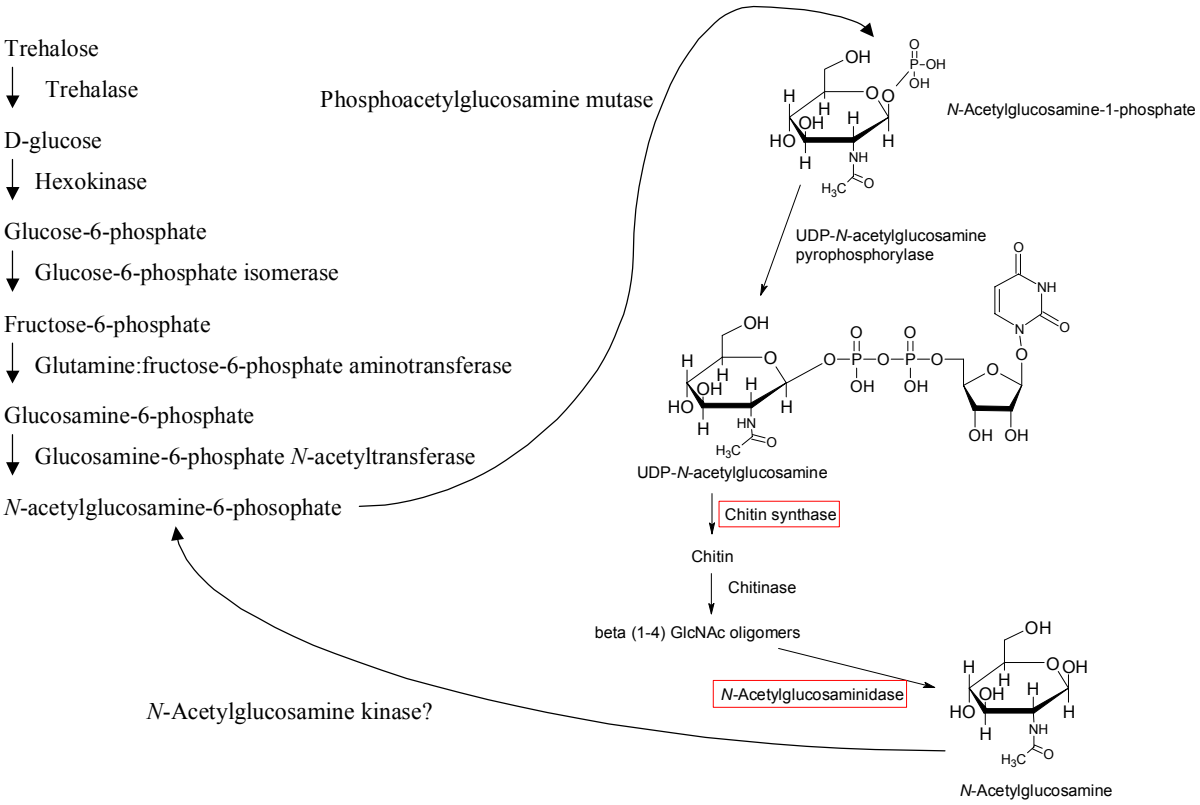
Although chitin is vital for insect survival, it is absent in animals and higher plants. Therefore, the enzymes involved in chitin metabolism are attractive targets for insect growth regulators (IGRs). Disruption of the chitin synthesis pathway or the chitin degradation pathway is deleterious to the insect. Therefore, IGRs that are capable of targeting CHSs or NAGs would have significant uses in insect control. Knockdown studies of NAG genes in *T. castaneum*, a major pest of stored products, demonstrated the essentiality of these enzymes. Therefore, IGRs targeting these enzymes are likely to be effective. Furthermore, studies on the developmental pattern of expression of these chitin metabolism genes may provide insight into the optimal time to deploy the IGRs, thereby maximizing their effectiveness.

Figure 1. Classification of insect (I) and nematode (N) CHSs based on amino acid sequence similarities.



Highly conserved sequences of approximately 500 amino acids from CATMWHET to WGTRE (two signature motifs of CHSs) were aligned by ClustalW (PAM250) to generate the phylogenetic tree. This sequence begins at the start of the catalytic domain and extends 164 amino acids downstream from the C terminus of the catalytic box. CHSs were from the insects (I) *Drosophila melanogaster* (Dm), *Drosophila pseudoobscura* (Dp), *Lucilia cuprina* (Lc), *Anopheles gambiae* (Ag), *Manduca sexta* (Ms), *Tribolium castaneum* (Tc), *Apis mellifera* (Am), *Spodoptera frugiperda* (Sf), *Aedes aegypti* (Aa), and the nematodes (N) *Caenorhabditis elegans* (Ce), *Caenorhabditis briggsae* (Cb), *Meloidogyne artiellia* (Ma), *Brugia malayi* (Bm), and *Dirofilaria immitis* (Di). Genbank Accession Numbers (DNA) are as follows: *DmCHS1* (NM_206430), *DmCHS2* (NM_079485), *DpCHS1* (AADE01001241), *DpCHS2* (AADE01000881), *LcCHS1* (AF221067), *AgCHS1* (AY056833), *AgCHS2* (XM_321336), *MsCHS1* (AY062175), *MsCHS2* (AY821560), *TcCHS1* (AY291475), *TcCHS2* (AY291477), *AmCHS1* (XM_395677), *SfCHS2* (AY525599), *AaCHS1* (AF223577), *CeCHS1* (T25G3.2), *CeCHS2* (F48A11.1), *CbCHS1* (CAE66574), *CbCHS2* (CAE62792), *MaCHS* (AY013285), *BmCHS1* (AF274311), and *DiCHS1* (AF288618). Insect CHSs fall within one of two classes, A and B. Arrows indicate the two *Manduca sexta* CHSs, one belonging to class A and the other belonging to class B. (Reprinted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs

Figure 2. Chitin metabolism pathway.



The enzymes involved in the insect chitin metabolism pathway are indicated next to the arrows corresponding to the biochemical reaction that they catalyze. The enzyme chitin synthase, which catalyzes the final step in the chitin synthesis pathway, and N-acetylglucosaminidase, which catalyzes the final step in the chitin degradation pathway are boxed in red. It is unkown if a recycling mechanism exists in insects where *N*-acetylglucosamine kinase may phosphorylate *N*-acetylglucosamine to produce *N*-acetylglucosamine-6-phosphate.

CHAPTER 1:
CHITIN SYNTHASES IN *Manduca sexta*

INTRODUCTION

In insects, as well as nematodes and fungi, the synthesis of chitin is catalyzed by the enzyme chitin synthase (UDP-GlcNAc: chitin 4-β-N-acetylglucosaminyltransferase, EC 2.4.1.16) from the biosynthetic precursor UDP-N-acetylglucosamine (UDP-GlcNAc). Chitin synthase (CHS) is a large enzyme that is partially located in the plasma membrane, which ensures that the newly synthesized chitin is extruded from the cell into the extracellular matrix. Although CHS has been extensively studied in fungi, only recently have insect CHSs begun to be characterized. In fact, just over five years ago, the first CHS gene was characterized in *Lucilia cuprina* (Tellam et al., 2000). Since then, several other insect CHSs have been characterized, although no more than two CHSs have been identified in a single insect species. Even after the genome sequencing of several insects belonging to various orders including Diptera, Lepidoptera, and Coleoptera, no more than two CHSs were identified in a single species. Generally speaking, chitin deposition in insects occurs in only two major extracellular structures. Chitin is used in the assembly of the cuticle (exoskeleton) and gut lining (peritrophic membrane). Because insects appear to have only two CHSs, perhaps one gene is used in the synthesis of cuticular chitin, while the other may specialize in the synthesis of peritrophic membrane-associated chitin in the midgut. A gene encoding a CHS (*MsCHS1*) was previously identified in the tobacco hornworm, *M. sexta* (Zhu et al., 2002). In this chapter, I describe the characterization of the other CHS gene in *M. sexta*, *MsCHS2*. Also, the presence of two alternatively spliced transcripts of *MsCHS1* are characterized and the levels of these transcripts in different tissues and stages of development were studied. These expression studies suggest a functional specialization of *M. sexta* CHSs and the alternately spliced isoforms. The large size and well-characterized life cycle of *M. sexta* makes it an excellent model organism for the study of all CHS isoforms in this insect.

MATERIALS AND METHODS

Insect cultures

M. sexta larvae were reared at 27°C using an artificial diet as described previously (Bell and Joachim, 1976).

Dissection of tissues from *Manduca sexta*

Tissue dissections were carried out on *M. sexta* at various developmental stages. Immediately following dissection, tissues were submerged in approximately ten volumes of RNA*later*™ RNA Stabilization Reagent (Qiagen) on ice for approximately 30 min. Tissues were then frozen at -80°C for later use. Alternatively, RNA was isolated from some tissue samples immediately following dissection. Midguts of *M. sexta* were dissected following the removal of malpighian tubules and tracheae as far as possible. Inevitably, some visceral tracheae remained in the midgut preparations. Likewise, the integument preparations also contained some visceral tracheae. However, the majority of fat body and muscle tissue was removed in these preparations. Abdominal tracheae, located between the midgut and the lateral longitudinal tracheal trunk, were dissected from various developmental stages for analysis of *MsCHS* expression in isolated tracheae. For most experiments, a representative portion of the *M. sexta* integument was utilized, which contained a part of the head capsule and spiracles, as well as the leg. For the determination of the relative expression of both alternatively spliced transcripts of *MsCHS1* in the integument, these regions of the integument were separately analyzed in addition to the dorsal integument (the region of the integument that overlays the dorsal longitudinal tracheal trunk). For analysis of *MsCHS* expression throughout the digestive tract, the dissected midgut was divided into three equal sections corresponding to the anterior, medial, and posterior portions of the midgut. The hindgut was also divided into two segments corresponding to the anterior hindgut (intestine) and posterior hindgut (rectum). Throughout all dissections, meticulous care was taken in order to avoid contamination from other tissues.

Isolation of total RNA

The RNeasy® Protect Mini Kit (Qiagen) was used to isolate total RNA from dissected *M. sexta* tissue samples according to the manufacturer's instructions. Briefly, approximately 30 mg of freshly dissected tissues or frozen samples in RNA*later*™ RNA Stabilization Reagent were ground after freezing in liquid nitrogen. Lysis buffer containing beta-mercaptoethanol was then added to the homogenate and extracted by passing through a 20-gauge needle. Following centrifugation at 15,000 g, the total RNA in the supernatant was purified using the RNeasy® mini spin column.

Isolation of genomic DNA

Genomic DNA was extracted from dissected epidermal tissue from 20 fourth instar larvae after homogenization in 1.7 mM PIPES buffer (pH 6.5). Following phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v) and chloroform/isoamyl alcohol (24:1, v/v) extractions, the DNA was ethanol precipitated in 0.3 M sodium acetate overnight and redissolved in water.

Synthesis of first-strand cDNA

The SuperScript™ III First-Strand Synthesis System for RT-PCR (Invitrogen) was used to synthesize the first-strand cDNA according to the manufacturer's instructions. Briefly, an oligo$(dT)_{20}$ primer was used for reverse transcription and 5 μg of total RNA was used as template. Reverse transcription was carried out at 45°C for 50 min. Template RNA was removed following reverse transcription by treatment with RNase H.

PCR amplification

RT-PCR was performed using a PTC-200 (MJ Research, Waltham, MA) thermal cycler using extension times and annealing temperatures appropriate for the individual primer pairs. For the PCR amplification of *MsCHS2*, the location and sequence of the individual primers used may be found in Fig. 3. For PCR amplification of overlapping genomic DNA fragments of *MsCHS1* and *MsCHS2*, pairs of gene-specific primers were designed from the corresponding cDNA sequences and used with genomic DNA as a template.

Cloning

Cloning of PCR products was carried out using the TOPO® TA Cloning Kit (Invitrogen) according to the manufacturer's instructions. Briefly, the PCR products were cloned directly from the PCR solution into the pCR®4-TOPO® vector and used for transformation of One Shot® TOP10 chemically competent *E. coli*. Following growth under ampicillin selection, clones containing the desired insert were identified by PCR and grown up in LB medium overnight. Plasmid isolation was carried out using the QIAprep® Spin Miniprep Kit according to manufacturer's instructions.

DNA sequencing

DNA sequencing was conducted at the DNA sequencing facility at Kansas State University using an automated sequencer (ABI Prism 3700).

3'-RACE of *MsCHS2*

The 3'-end of *MsCHS2* was obtained using 3'- Rapid Amplification of cDNA Ends (RACE). An adapter-oligo$(dT)_{17}$ primer (5'-GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT TTT T-3') was used in the first strand cDNA synthesis reaction. Following reverse transcription, a *MsCHS2* gene-specific primer, 3RACEGSP1 (5'-GTT TGT GAT CGT CAG TAT CGT CC-3') was solely used for linear PCR in order to overcome problems encountered with non-specific amplification by the adapter primer. Following linear PCR, another nested gene-specific primer, 3RACEGSP2 (5'-ATG GTT GGC TAG CGC CAT CTA G-3') was used in conjunction with the adapter-oligo$(dT)_{17}$ primer in the second PCR reaction. Finally, a third PCR reaction was carried out using a third nested gene-specific primer, 3RACEGSP3 (5'-CAC TGC GGC AAC ATT TGG-3') and the adapter primer (5'-GGC CAC GCG TCG ACT AGT AC-3'). The resulting PCR product was cloned and sequenced as described above.

5'-RACE of *MsCHS2*

The 5'-end of *MsCHS2* was obtained using the 5'-RACE System for Rapid Amplification of cDNA Ends (Invitrogen). To ensure the production of a full-length cDNA and to aid in the specificity of the reaction, a gene-specific primer, 5RACEGSP1 (5'-GTA TAA GAA TTT TGC

AAG C-3'), was used instead of an oligo-$(dT)_{17}$ primer in the reverse transcription reaction. Due to catalytic limitations of the reverse transcriptase enzyme, the reaction is typically carried out at a low temperature and therefore the T_m of the 5RACEGSP1 primer was designed to be around 42°C. Furthermore, the primer was designed to have a high AT content in order to maximize the length of the primer and therefore the specificity of the reaction. The resulting cDNA produced after reverse transcription with the 5RACEGSP1 primer was purified using the S.N.A.P.™ column (Invitrogen) according to the manufacturer's instructions in order to remove residual dNTPs. The purified cDNA was then dC-tailed by solely adding dCTP to the terminal deoxynucleotidyl transferase tailing reaction. The 5'-RACE Abridged Anchor Primer (5'-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3') was then used in conjunction with a second nested *MsCHS2* gene-specific primer, 5RACEGSP2 (5'-TGT AGT ACC GAG TCC TCT TGA GG-3'), in the first PCR reaction. Finally, the Abridged Universal Amplification Primer, AUAP (5'-GGC CAC GCG TCG ACT AGT AC-3'), was used in conjunction with the third nested *MsCHS2* gene-specific primer, 5RACEGSP3 (5'-CGT AGT TCC TGC AGC ACC G-3'), and the resulting PCR product was cloned and sequenced as described above.

DNA and protein sequence analyses

The amino acid sequence of MsCHS2 was deduced following translation of the ORF in the corresponding cDNA sequence using the translation tool at the ExPASy Proteomics website (http://us.expasy.org/tools/dna.html). Other protein sequence analysis tools used in this study, including those for prediction of MW and pI were obtained from the ExPASy Proteomics website (http://us.expasy.org/). TMHMM (v2.0) software (www.cbs.dtu.dk/services/TMHMM-2.0/) was used to predict transmembrane helices in the putative protein sequences. Potential coiled-coil domains were identified using the Paircoil Program (http://paircoil.lcs.mit.edu/cgi-bin/paircoil, Berger et al., 1995). Signal peptide and signal anchor prediction was conducted using the SignalP 3.0 server (www.cbs.dtu.dk/services/SignalP/). The deduced amino acid sequences of CHSs from *Anopheles* (XM_321951), *Drosophila* (NM_079485), and *Tribolium* (AY291477) were obtained as described previously (Arakane et al., 2004). The sequence of *MsCHS1* was obtained from the published data (Zhu et al., 2002). Multiple sequence alignments of deduced amino acid sequences were made using Multiple Alignment software (http://npsa-

pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html). Alignments of nucleotide sequences were made using ClustalW software (PAM250).

RT-PCR analysis of *MsCHS* expression

The developmental pattern of *MsCHS* expression was analyzed using RT-PCR (Hogenkamp et al., 2005). Total RNA was isolated from dissected epidermal and anterior midgut tissues from various stages of development using the RNeasy® Protect Mini Kit (Qiagen) according to the manufacturer's instructions. Two micrograms of total RNA were utilized as templates for cDNA synthesis using an oligo-(dT) primer. This cDNA then served as a template for the subsequent PCR reactions. To minimize variations in primer annealing in the *MsCHS1*/*MsCHS2* expression analysis, a common forward primer was used that was designed from a region where the sequences of both transcripts are identical. The sequence of this primer was 5'-GAA AGG CGC TCA TGG ACG-3', which spans positions 2432–2449 in *MsCHS1* and 2398–2415 in *MsCHS2*. To allow for the simultaneous analysis of *MsCHS1*/*MsCHS2* expression, two gene-specific reverse primers were designed to produce different size PCR products when used in conjunction with the common forward primer. The sequences of these primers were 5'-TGA AGG AAG CCC AAG AGA G-3' (spanning positions 3275–3293) for *MsCHS1*, and 5'-ACG TTG TTC AAA TTG CAT AGG-3' (spanning positions 3120–3140) for *MsCHS2*. These primers were designed to have similar melting temperatures, enabling both PCR reactions to be carried out in the same vessel. The same strategy was used to examine the alternate exon usage of *MsCHS1*. Here, a common reverse primer was designed downstream of the alternate exon with the sequence 5'-TTC GTT ATT AGC ACC TAG GG-3' (spanning positions 4466–4485). A forward primer specific for the *MsCHS1a* alternate exon with the sequence 5'-TGA AAG AAT TGA GAG ACT CG-3' (located near the 5'-end of the alternate exon and spanning positions 3788–3807) and another forward primer specific for the *MsCHS1b* alternate exon with the sequence 5'-ATT ACC TAC ATC GAG GAG AC-3' (located near the 3'-end of the alternate exon and spanning positions 3919–3938) were used. For controls, primers for the constitutively expressed housekeeping gene RpS3 were utilized. A series of PCR reactions of different cycle numbers ranging from 15 to 30 were carried out in order to determine the appropriate cycle number to be used in the analysis. All PCR reactions used in the

expression analysis were conducted using the following conditions: denaturation at 94°C for 30 sec., annealing at 58°C for 45 sec. and polymerization at 72°C for 1 min. for 20 cycles.

Northern blot analysis of *MsCHS* expression

The developmental expression pattern of *MsCHS* transcripts was determined by northern blot analysis (Hogenkamp et al., 2005). Total RNA was isolated from dissected midgut and epidermal tissues from various stages using the RNeasy® Protect Mini Kit (Qiagen) according to the manufacturer's instructions. Following isopropanol precipitation using ammonium acetate, the concentrated RNA samples were redissolved in water. Approximately 10 μg of total RNA from each sample was fractionated in a 1.5% formaldehyde/agarose gel. The RNA was transferred in 20 X SSC (SSC = 0.15 M NaCl and 0.015 M sodium citrate pH 8.0) to a nylon membrane (Hybond™-N^+, Amersham Biosciences). Probes specific for *MsCHS1* and *MsCHS2* were designed from dissimilar sequences corresponding to nucleotide positions 4333–4588 and 4241–4477, respectively. The DNA probes corresponding to these regions were prepared using Ready-To-Go™ DNA Labeling Beads (dCTP, Amersham) and radiolabeled with (α-^{32}P) dCTP (Perkin Elmer). Hybridization and washing were performed under high stringency conditions: 0.5 M phosphate buffer, pH 7.2 with 7% SDS, and 1 mM EDTA (Church and Gilbert, 1984), at 62°C overnight, followed by washing with 1 X SSC, 0.1% SDS at 65°C. Following detection of the *MsCHS2* transcript by autoradiography, the membrane was stripped by boiling in 0.1% SDS and then re-probed with a mixture of the *MsCHS1* probe and a probe specific for the constitutively expressed housekeeping gene, ribosomal protein S3, RpS3 (Jiang et al., 1996).

Real-time RT-PCR analysis of *MsCHS* and *MsCHI* expression

Analysis of *MsCHS* and *MsCHI* expression was carried out using real-time RT-PCR. Gene-specific primers for *MsRpS3*, *MsCHI*, *MsCHS2*, and both alternatively spliced isoforms of *MsCHS1* were designed using Beacon Designer 2.0 software (Premier Biosoft International). For determination of the number of *MsCHS2* transcripts, the forward primer, RTMsCHS2F1 (5'-GAG GAG GGA CTC ACT GCT CGC-3'), and reverse primer, RTMsCHS2R1 (5'-GCC GCT CAT GTC GTT CAC CG-3'), were used. For analysis of the alternately spliced *MsCHS1a* transcript, the primers RTMsCHS1aF1 (5'-AGA GAT CTG AAA GAA TTG AGA GAC TCG-3') and RTMsCHS1aR1 (5'-AAC TCC GAA GGG CCA TTT AAC G-3') were designed to

anneal within the alternate exon of *MsCHS1a*. Likewise, the primers used for the analysis of *MsCHS1b* transcripts, RTMsCHS1bF1 (5'-GCT CGC ATT GCA GGC GAT CTG C-3'), and RTMsCHS1bR1 (5'-TCC CAA AGG CCA AAC GAC GTG G-3'), were designed to anneal within the alternate exon of *MsCHS1b*. The primers RTMsCHIF1 (5'-ACA CCA CAC CGA CTC CTG AA-3') and RTMsCHIR1 (5'-TGG TAG TAG TCT TCA CGG TCG T-3') were used in the analysis of chitinase transcripts. Finally, the primers corresponding to the constitutively expressed housekeeping gene, RpS3, RTRpS3F1 (5'-AGA AAG TTG CCA CCC GTG-3') and RTRpS3R1 (5'-CGC TCC AGA CTC CAT GAT G-3'), were used as an internal control.

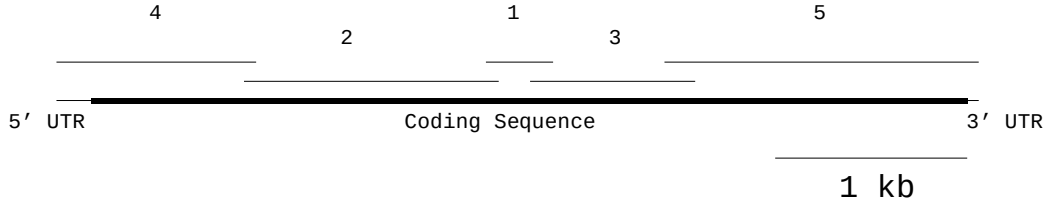
The PCR products for each of the primer pairs used in the quantitative PCR experiments were cloned into the pCR®4-TOPO® vector (Invitrogen) and after transformation of *E. coli*, the plasmid was purified from an overnight culture. The DNA concentration of each construct was determined using a VersaFluor™ Fluorometer (Bio-Rad) and the DNA intercalating dye Hoechst 33258. Sets of serial dilutions were then made from each construct and the dilutions were analyzed by real-time PCR using the nucleic acid stain SYBR green (Bio-Rad). The Bio-Rad iCycler iQ® real-time PCR detection system at the Kansas State University Gene Expression Facility was used in the quantitative PCR experiments. The protocol used in all experiments is outlined in Figure 3. Standard curves for each gene of interest were constructed using serial dilutions of each construct. At the end of each quantitative PCR experiment, a melt curve was generated to rule out the possibility of primer dimer formation. Using the standard curves, the transcript copy number was determined for each gene of interest along with the copy number of the internal housekeeping gene, RpS3. All experiments were repeated in triplicate using cDNAs prepared from different insects.

Southern blot analysis of *MsCHS* genes

MsCHS genes were studied using *M. sexta* genomic DNA and Southern blot analysis (Hogenkamp et al., 2005). Three different restriction enzymes, *Eco*RV, *Eco*RI, and *Sal*I, were used to digest 10 μg of *M. sexta* genomic DNA using one endonuclease in each reaction. The resulting digests were subjected to electrophoresis in a 0.9% agarose gel and transferred to a nylon membrane (Hybond™-N^+, Amersham Biosciences) in 0.4 M NaOH. The *MsCHS1* probe was designed from a highly conserved region of the gene that is 88% identical to the *MsCHS2* gene, allowing for the simultaneous detection of both CHSs. This probe was radiolabeled with

α^{32}P-dCTP (Ready-To-Go™ DNA Labeling Beads, Amersham) and corresponded to nucleotide positions 2315–2515 of the *MsCHS1* cDNA. To demonstrate the ability of the probe to detect even distantly related CHS genes from another species, digestions (10 μg each) of genomic DNA from *Drosophila* were carried out in parallel using the restriction enzymes *Hin*dIII and *Bam*HI. For controls, non-radioactive PCR fragments (200 pg each) corresponding to the *MsCHS1* and *MsCHS2* probes were also analyzed in parallel. Hybridization was carried out at 50°C followed by washing at 37°C with 1 X SSC/0.1% SDS. The membrane was exposed to X-ray film for 7 d.

Figure 3. MsCHS2 cDNA cloning strategy.



PCR Fragment	Size (bp)	*Primer*		
		Direction[1]	Type[2]	Sequence (5′-3′)
1	359	F	D	tggwsncargtnatgtayatg
		R	D	tcrtccatnarngcrttncc
2	1333	F	D	ggntggtgggaraa
		R	G	gatccacggatattggcagg
3	869	F	G	cctggtaccagatgttcgag
		R	D	acytcnckngtncccca
4	1049	F	5′-RACE Adapter	ggccacgcgtcgactagtac
		R	G	cgtagttcctgcagcaccg
5	1648	F	G	cactgcggcaacatttgg
		R	3′-RACE Adapter	gaccacgcgtatcgatgtcga

[1] F: Forward, R: Reverse
[2] D: degenerate primer, G: gene-specific primer

The full-length *MsCHS2* cDNA sequence was determined by sequencing five overlapping PCR fragments using midgut cDNA as a template. PCR fragments 4 and 5 were obtained using 5'- and 3'-RACE, respectively (Reprinted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development, 529-540, Copyright 2005, with permission from Elsevier)

Figure 4. Protocol used for real-time RT-PCR experiments.

```
Cycle 1:       (1X)
     Step 1:            95°C for 3:00
Cycle 2:       (45X)
     Step 1:            95°C for 0:20
     Step 2:            60°C for 0:45
     Data collection and real-time analysis enabled
Cycle 3:       (1X)
     Step 1:            95°C for 1:00
Cycle 4:       (1X)
     Step 1:            55°C for 1:00
Cycle 5:       (80X)
     Step 1:            55°C for 0:10
     Increase setpoint temperature after cycle 2 by 0.5°C
     Melt curve data collection and analysis enabled
```

Data analysis steps are outlined in cycles 1-2 and the steps for melt curve construction are outlined in cycles 3-5.

RESULTS

Sequence analysis of *MsCHS2* cDNA

The sequence of the entire cDNA (Genbank accession number AY821560) corresponding to the *MsCHS2* gene was obtained by sequencing DNA fragments obtained by PCR reactions using cDNA prepared from midgut tissue as the template and degenerate primers covering most of the open reading frame (ORF) as well as by 5'- and 3'-RACE (Fig. 3). In total, 4821 nucleotides of this transcript were sequenced, which contained an open reading frame of 4575 nucleotides. The encoded protein is 1524 amino acids in length with a predicted MW of approximately 175 kDa. The predicted pI of MsCHS2 is pH 6.06, which is slightly more acidic than the predicted pI of pH 6.56 for MsCHS1. Using the topology prediction software, TMHMM (v2.0) (www.cbs.dtu.dk/services/TMHMM-2.0/), MsCHS2 is predicted to be an integral membrane protein with 16 transmembrane helical spans (Fig. 5). The catalytic domain of CHS is probably located on the cytoplasmic side of the membrane (Tellam et al., 2000). Alignment of MsCHS2 with other class B insect CHSs shows a high degree of conservation, particularly in the middle of the putative catalytic domain (Fig. 6). The superior alignment of MsCHS2 with other class B insect CHSs, as opposed to class A CHSs, demonstrates that this protein is very probably an enzyme of the B class. Located near the C-terminal end of the catalytic domain of MsCHS2 is the QRRRW 'signature sequence' of CHSs (Nagahashi et al., 1995). The CHS 'signature sequence', QRRRW, may correspond to the QXXRW motif found in cellulose synthases (Saxena et al., 2001). Five transmembrane segments (5-TMS) are predicted to be located immediately following the putative central catalytic domain. This topology is similar to that of cellulose synthases and has been predicted to constitute a pore in the membrane through which the newly synthesized carbohydrate polymer may be extruded (Richmond, 2000). Class A CHSs have two unique characteristics, a transmembrane segment encoded by an exon that is alternatively spliced and a coiled-coil domain (Kramer and Muthukrishnan, 2005). Genes encoding class B CHSs, however, do not contain an alternate exon and the coiled-coil domain is typically not predicted in these proteins. Consistent with other class B CHSs, MsCHS2 is not predicted to have a coiled-coil domain immediately following the cluster of five transmembrane segments (data not shown).

Sequencing of *MsCHS2* and *MsCHS1* genes

The *MsCHS2* gene (AY821561) was sequenced using several sets of primers that were designed from the corresponding cDNA sequence to PCR amplify overlapping genomic DNA fragments (Fig. 7). Genomic DNA encoding *MsCHS2* is 18,525 bp including the 5'- and 3'-UTRs and contains 24 short exons ranging in size from 98 to 331 nucleotides, which are separated by 23 introns of varying lengths. In contrast to the gene sequences of class B CHSs from other insects, the *MsCHS2* gene is considerably larger and interspersed with a greater number of introns. This observation may be a general characteristic of the *M. sexta* genome, as a pattern of numerous introns interspersed in the gene is also observed in *MsCHS1*. The 5'-UTR of *MsCHS2* is interrupted by a single intron, whereas the short 3'-UTR contains no introns. A careful analysis of the *MsCHS2* gene sequence and comparison with the corresponding cDNA sequence indicate that no alternate exon is present in this gene, as in the case of other class B CHS genes (Arakane et al., 2004). However, following the completion of the *MsCHS1* gene sequence, an alternate exon was identified that corresponds to exon 20 in the *MsCHS2* gene. Within this region of the *MsCHS1* gene sequence (equivalent to the region between exons 19 and 21 of *MsCHS2*), the sizes and positions of the exons in *MsCHS1* and *MsCHS2* perfectly match, with differences observed only in the intron size and in the presence of the alternate exon (corresponding to exon 20 of *MsCHS2*) in *MsCHS1*. Furthermore, the intron positions are conserved between the two CHS genes in 10 additional positions for a total of 14 identical positions (Fig. 6). The alternate exons of *MsCHS1* are equal in length (177 nucleotides) and encode 59-amino acid-long segments composed of a putative TMS located in the middle of the exon sequence and two short flanking sequences, one facing the cytosolic side and the other the extracellular side. The corresponding exon in *MsCHS2*, exon 20, also encodes a topologically similar peptide sequence. In all, the presence of an alternate exon in the *MsCHS1* gene demonstrates a higher degree of complexity of organization compared to the *MsCHS2* gene.

Analysis of expression of *MsCHS1a*/*MsCHS1b* and *MsCHS2* by RT-PCR

The developmental pattern of expression of *MsCHS1* and *MsCHS2* was studied by RT-PCR (Fig. 8). Consistent with the expression pattern reported for the previously characterized transcript of *MsCHS1* (Zhu et al., 2002), expression in the integument is observed during the feeding and pupal stages. Transcript levels rapidly decline at the onset of the wandering stage

and gradually reappear prior to pupation. In the tracheae, however, *MsCHS1* transcripts are virtually undetectable in all developmental stages analyzed except for wandering days two and four. The same is true in anterior midgut preparations (which are invariably contaminated with tracheae) where *MsCHS1* is detected during periods that coincide with tracheal *MsCHS1* expression. In contrast, *MsCHS2* is only detectable in the midgut and no expression was observed in any of the integument or tracheal preparations. Furthermore, expression of this transcript was not observed in several other tissues including malpighian tubules, hemocytes, and fat body at various developmental stages (data not shown). This midgut-specific CHS is expressed during periods when the larvae are actively feeding. However, after the second day of the wandering stage, no significant quantity of *MsCHS2* mRNA can be detected.

The utilization of the two alternate exons of *MsCHS1* for the generation of alternately spliced transcripts was also investigated by RT-PCR (Fig. 8). At all time points when *MsCHS1* is expressed in epidermal tissues, both splice variants are present. However, the relative amounts of the two transcripts vary substantially during development. The relative abundance of the *MsCHS1a* transcript is greater than that of *MsCHS1b* during the larval feeding and pupal stages in the epidermis. However, in the tracheae, there is a significant accumulation of *MsCHS1b* mRNA relative to *MsCHS1a* during the prepupal stage (W-4). This excess of *MsCHS1b* is also evident in the anterior midgut at the same developmental stage. At this stage, *MsCHS1* expression in the tracheae associated with the midgut is significantly increased and the *MsCHS1b* transcripts predominate, although both splice variants are expressed. By extrapolation, it is possible that some of the *MsCHS1b* transcripts that appear in the integument preparations (which is invariably contaminated with tracheae) may have originated from tracheae.

Expression of *MsCHS2* in the midgut by northern blot analysis

Northern blot analysis of midguts from the feeding stage, which were divided into three equal segments corresponding to the anterior, medial, and posterior sections of the midgut, show the highest level of *MsCHS2* expression in the anterior midgut with expression levels tapering off in the medial and posterior midguts (Fig. 9, right). The level of *MsCHS2* expression in the anterior midgut is significantly higher compared to the expression levels in the medial and posterior midgut samples, indicating that *MsCHS2* is expressed primarily in the anterior midgut.

In agreement with the RT-PCR expression data, *MsCHS2* is expressed only during the feeding stages and at the onset of the wandering stage. Following the decline of *MsCHS2* transcripts, trace amounts of *MsCHS1* transcripts are observed in the midgut preparations during days two and four of the wandering stage when tracheal *MsCHS1* expression is known to occur (Zimoch et al., 2005).

Developmental pattern of expression of *MsCHI* in the integument, tracheae, and midgut by real-time RT-PCR

The developmental expression pattern of the previously characterized CHI gene (Kramer et al., 1993) from *M. sexta*, MsCHI, was analyzed by real-time RT-PCR (Fig. 10). Standard curves were generated from serial dilutions of cloned *MsCHI* and *RpS3* PCR products. Both standard curves showed a correlation coefficient >0.998 and PCR efficiency between 98% and 103%. At the end of each quantitative PCR experiment, a melt curve was generated and all samples displayed only a single peak indicating that no primer dimers were formed.

High levels of *MsCHI* expression are seen at the time of the fifth instar molt and pupal molt in all tissues examined: the integument, tracheae, and midgut. Significant levels of expression were also observed on the second day of the wandering stage, particularly in the integument and tracheae.

Developmental pattern of expression of *MsCHS2* in the midgut by real-time RT-PCR

The developmental expression pattern of *MsCHS2* in larvae starting from the fourth instar to three-day-old pupae was analyzed by real-time RT-PCR (Fig. 11). Standard curves were generated from serial dilutions of cloned *MsCHS2* and *RpS3* PCR products. Both standard curves showed a correlation coefficient >0.998 and PCR efficiency between 100% and 105%. At the end of each quantitative PCR experiment, a melt curve was generated and all samples displayed only a single peak indicating that no primer dimers were formed.

These studies indicate that *MsCHS2* is expressed during periods when the larvae are actively feeding such as the fourth instar and fifth instar. However, just prior to the fourth to fifth instar molt, the level of transcripts drops significantly. This drop in expression coincides with the time point where the larvae stop feeding. However, during periods of active feeding, *MsCHS2* expression is consistently seen. Transcript levels persist at the onset of the wandering

stage but rapidly decline to virtually undetectable levels by the fourth day of the wandering stage.

Developmental pattern of expression of *MsCHS1a*/*MsCHS1b* in the integument, tracheae, and midgut by real-time RT-PCR

The developmental expression pattern of *MsCHS1a* and *MsCHS1b* starting from fourth instar larvae to three-day-old pupae was analyzed by real-time RT-PCR (Fig. 12). Standard curves were generated from serial dilutions of cloned *MsCHS1a*, *MsCHS1b*, and *RpS3* PCR products. All standard curves showed a correlation coefficient >0.99 and PCR efficiency between 90% and 105%. At the end of each quantitative PCR experiment, a melt curve was generated and all samples displayed only a single peak indicating that no primer dimers were formed.

Quantitative analysis of *MsCHS1* transcripts in the integument, tracheae, and midgut clearly show that this gene is primarily expressed in the integument and tracheae with only minor levels of expression observed in the midgut preparations. The pattern of expression of *MsCHS1* in the tracheae closely resembles that of the midgut preparations except that the level of transcripts is more than twice that of the midgut preparations. This is consistent with the premise that the *MsCHS1* transcripts observed in the midgut originate from a contaminating tissue, the tracheae. Furthermore, the ratio of *MsCHS1b* transcripts verses *MsCHS1a* transcripts is similar in both tissues where *MsCHS1b* transcripts predominate. On the contrary, in the integument, it is the *MsCHS1a* transcripts that predominate at all developmental stages with the exception of day four of the wandering stage. It is at this stage when *MsCHS1b* expression in the tracheae is maximal and, therefore, the transcripts corresponding to this alternately spliced mRNA in the integument might have originated from contaminating tracheae in the integument preparations. The expression of *MsCHS1a* is most pronounced during molting periods leading to the fifth instar as well as the pupal molt (prepupa to pupa), consistent with its proposed role in the synthesis of cuticular chitin. *MsCHS1b* transcripts, however, are primarily expressed during periods where tracheae are actively proliferating, that is during the period around the fifth instar molt and wandering days two and four (Zimoch et al., 2005).

Expression of *MsCHS1* and differential accumulation of the alternatively spliced transcripts *MsCHS1a* and *MsCHS1b* in different regions of the integument by real-time RT-PCR

The expression of *MsCHS1a* and *MsCHS1b* from day two larvae in the fifth instar was analyzed in various regions of the integument in order to determine if the relative expression of the two alternately spliced transcripts is consistent throughout the integument (Fig. 13). Levels of transcripts for *MsCHS1a*, *MsCHS1b*, and *RpS3* were determined from standard curves generated from serial dilutions of each of these three cloned cDNAs. All standard curves showed a correlation coefficient >0.99 and PCR efficiency between 90% and 105%. At the end of each quantitative PCR experiment, a melt curve was generated and all samples displayed only a single peak indicating that no primer dimers were formed.

Analysis of the expression level of both alternatively spliced transcripts indicated no significant difference between the four different regions of the integument including the spiracle, dorsal integument, head capsule, and leg during the fifth instar, day 2. At this developmental stage, significant levels of *MsCHS1b* transcripts were not observed in isolated tracheae and therefore regions of the integument with a greater concentration of tracheae, such as the spiracles, did not show a larger proportion of the *MsCHS1b* transcript. However, when a different developmental stage was analyzed, wandering day 4 (when tracheal *MsCHS1* expression is maximal), a higher proportion of *MsCHS1b* transcripts was observed in the spiracle region of the integument.

***Manduca sexta* CHS expression along the length of the digestive tract by real-time RT-PCR**

The expression of *MsCHS2* and the alternatively spliced transcripts *MsCHS1a* and *MsCHS1b* were examined in different segments of the digestive tract of fifth instar *Manduca* using real-time RT-PCR (Fig. 14). The quantity of each of the three transcripts was determined using standard curves made from each of the cloned cDNAs. The standard curves for all transcripts analyzed showed a PCR efficiency between 90% and 105% and the correlation coefficient was >0.99 for all. A melt curve was generated for all samples, which showed only a single peak indicating that no primer dimers were formed during the course of the reaction.

During the fifth instar (days 0-2), transcripts for *MsCHS2* were only detected in the anterior, medial, and posterior regions of midgut with no quantifiable expression in foregut or

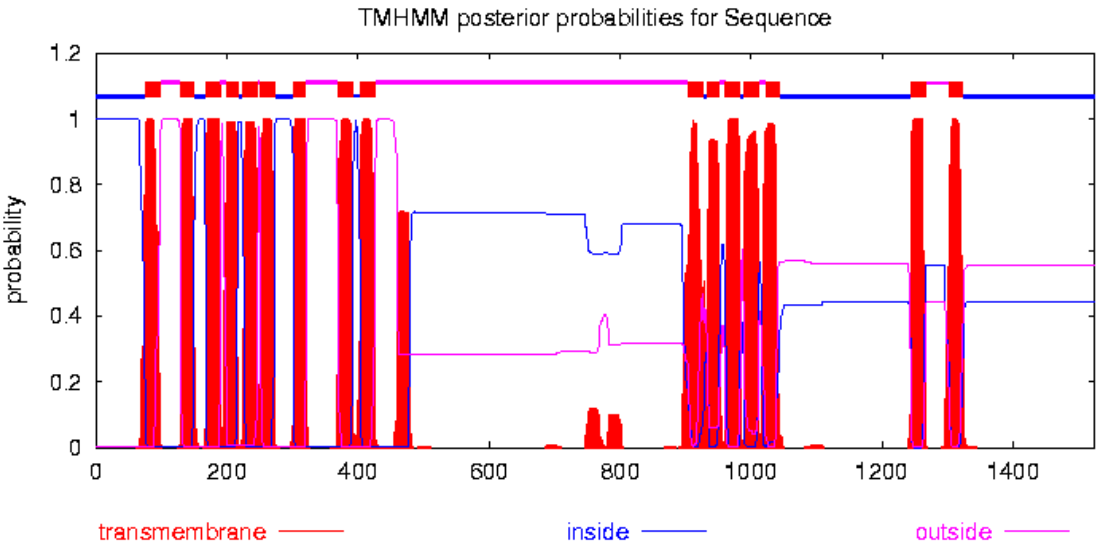
hindgut (Fig. 14). A gradient of decreasing expression of *MsCHS2* was also observed from the anterior to posterior portions of the midgut with the majority of transcripts in the anterior midgut. The alternately spliced transcript primarily expressed in the tracheae, *MsCHS1b*, showed significantly higher levels of expression than *MsCHS1a* in the foregut and anterior as well as posterior regions of the hindgut. Although these tissues are invariably contaminated with tracheae expressing *MsCHS1b*, the level of expression of this transcript in the tracheae is minimal at this developmental stage (Figs. 6 & 10). Therefore, MsCHS1b primarily carries out the synthesis of chitin in the intima of the foregut and hindgut.

Southern blot analysis of CHS genes in *Manduca sexta*

Southern blot analysis of genomic DNA was carried out to determine the copy number of CHS genes in *M. sexta* (Fig. 15). A probe designed to be specific for CHS genes was designed from the *MsCHS1* gene around a conserved region where its sequence is 88% identical to the *MsCHS2* gene. This probe was capable of hybridizing to both CHS genes in *M. sexta* as shown by cross-hybridization of the radioactive probe with unlabeled probe DNAs containing either of the CHS gene sequences. Furthermore, the ability of this probe to detect other CHS genes was demonstrated by its ability to recognize CHS gene sequences in digests of genomic DNA from a distantly related insect species, *D. melanogaster*. The CHS probe was able to recognize genes encoding both class A and class B CHSs in *Drosophila*, namely *DmCHS1* and *DmCHS2*, respectively. Because the probe was able to recognize both CHS genes from a distantly related insect, it is likely that this probe would be able to recognize all CHS gene(s) in *M. sexta*. The restriction enzymes were selected, after analysis of the gene sequences, based on of their ability to produce differently sized DNA fragments from the two CHS genes. Only two bands are visible in the *Eco*RI and *Sal*I digestions and the sizes of these fragments correspond to the predicted sizes of the digestion products of *MsCHS1* and *MsCHS2* (expected to be detected by this probe) as deduced from the genomic DNA sequences. The *Eco*RV digestion lane shows a single band of approximately 3 kb (corresponding to the *MsCHS1* gene) and two closely spaced bands of approximately 7 kb. The sizes of the two larger bands correspond to two possible fragments (6933 and 7831 bp) in *MsCHS2*, which would be produced as a result of either incomplete digestion and/or methylation of the restriction enzyme cleavage sites. We conclude

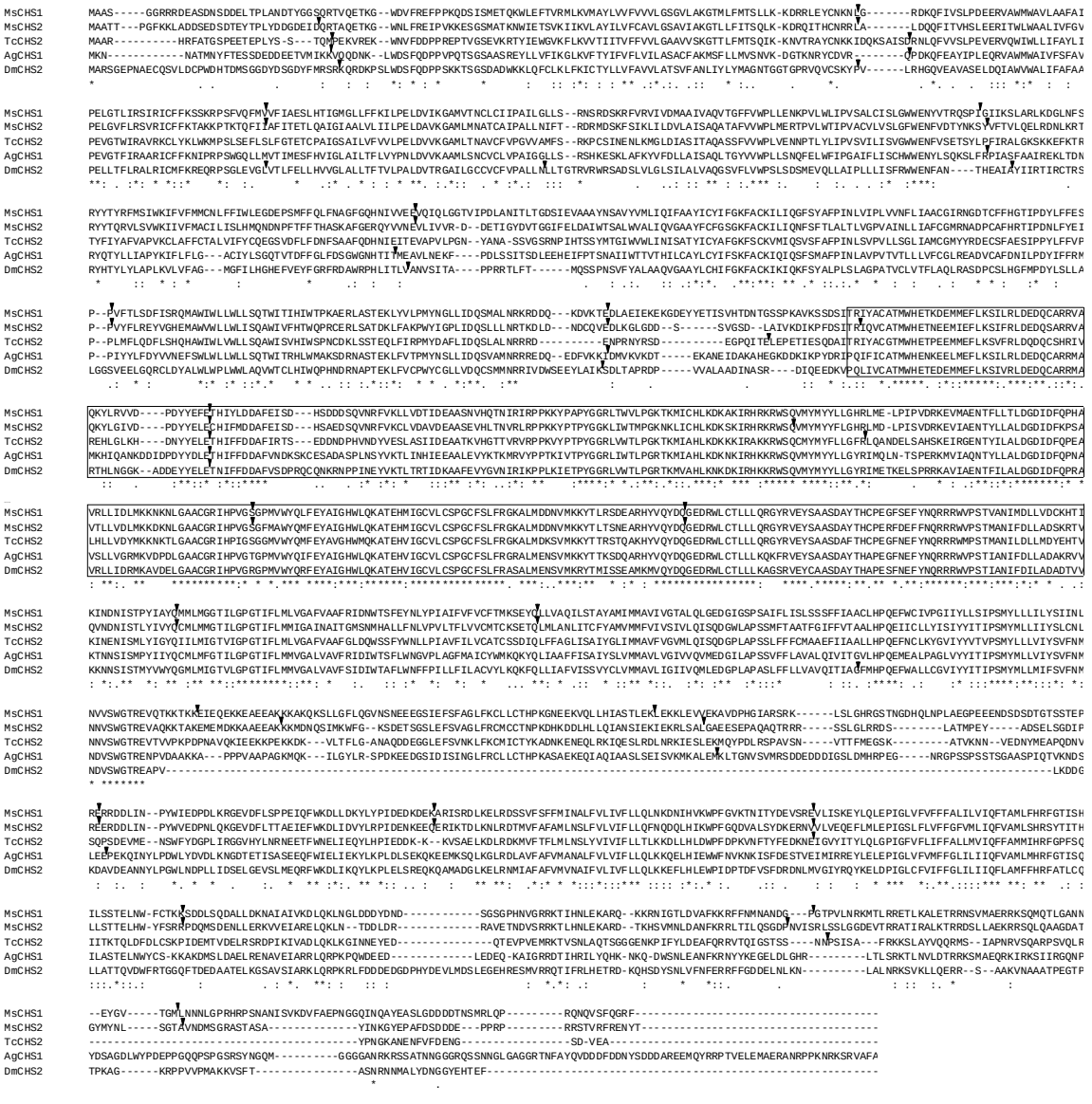
that like other insects, *M. sexta* very probably has only two CHS genes, namely *MsCHS1* and *MsCHS2*, as well as three CHS enzymes, MsCHS1a, MsCHS1b, and MsCHS2.

Figure 5. Topology prediction of MsCHS2.



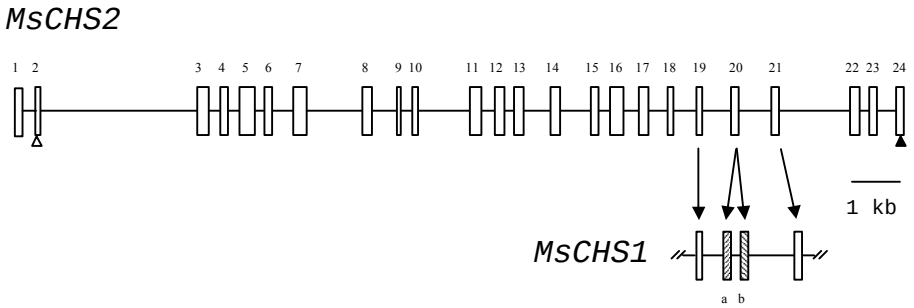
TMHMM (v2.0) software (www.cbs.dtu.dk/services/TMHMM-2.0/) was used to predict transmembrane segments in the deduced protein sequence. A total of 16 TMSs are predicted with probability >0.8. The central putative catalytic domain is probably intracellular (Tellam et al., 2000).

Figure 6. Alignment of the deduced amino acid sequences of MsCHS1, MsCHS2, TcCHS2, AgCHS1, and DmCHS2 using multiple alignment software.



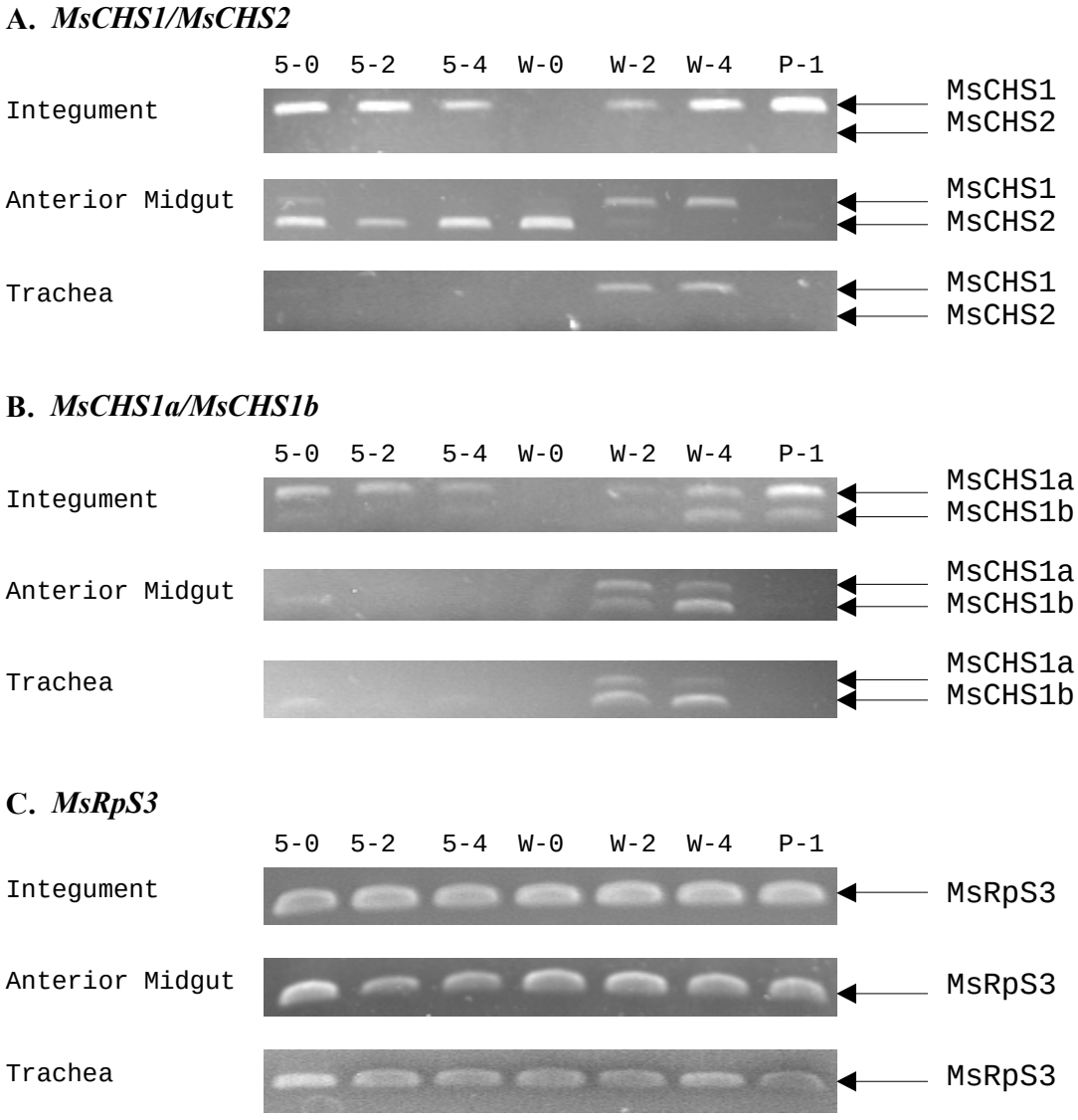
The putative catalytic domain is boxed. The positions in the protein sequences where coding regions are interrupted by introns are indicated by shaded arrowheads. Symbols below the aligned amino acid sequences indicate identical (*), highly conserved (:), and conserved residues (.).(Reprinted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development, 529-540, Copyright 2005, with permission from Elsevier)

Figure 7. Schematic diagram of the exon-intron organization of the *MsCHS2* gene and a portion of the *MsCHS1* gene.



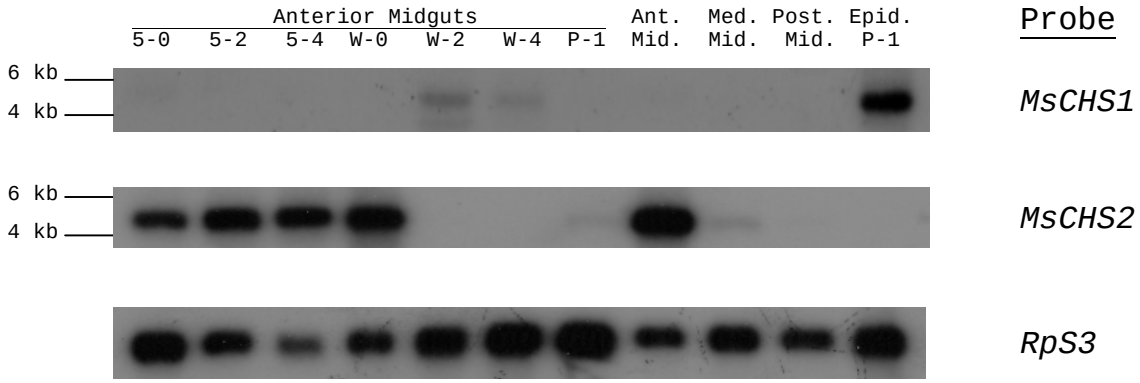
Boxes indicate exons and lines indicate introns. The locations of the ATG start codon and TAA stop codon are represented by open and shaded triangles, respectively. The crosshatched exons, a and b, represent the alternative exons in the *MsCHS1* gene that correspond to exon 20 in the *MsCHS2* gene. The majority of the *MsCHS1* gene organization at the 5'-end was previously determined (Zhu et al., 2002). (Reprinted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development, 529-540, Copyright 2005, with permission from Elsevier)

Figure 8. RT-PCR analysis of the expression of CHS genes in the integument, anterior midgut, and tracheae of *M. sexta* during development.



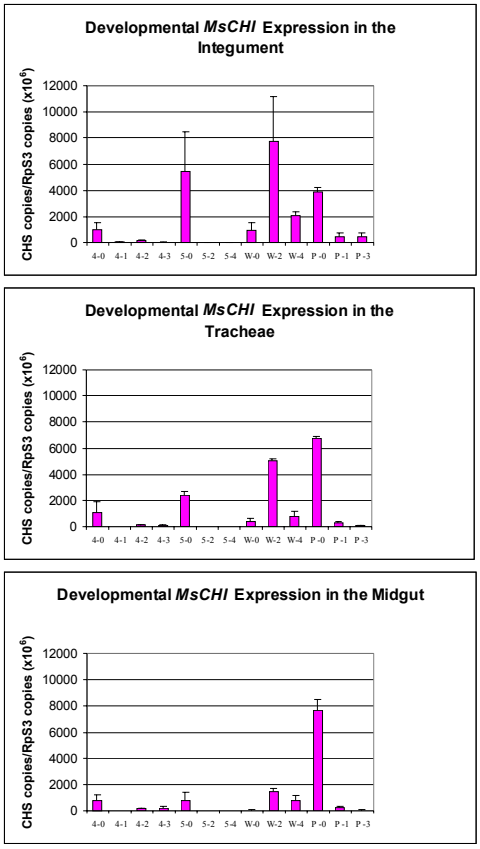
The cDNAs used for PCR analysis were prepared from total RNA extracted from dissected epidermal, anterior midgut, and tracheal tissues at various stages during development: Fifth instar days 0, 2, 4; Wandering days 0, 2, 4; Pupal day 1. Two sets of gene-specific primers were used to study the relative expression of *MsCHS1*/*MsCHS2* in panel A, and *MsCHS1a*/*MsCHS1b* in panel B. These primers were designed to produce different size PCR products to allow for simultaneous analysis. For controls, primers for the constitutively expressed housekeeping gene, *MsRpS3*, were used in panel C. (Adapted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development, 529-540, Copyright 2005, and Insect Biochemistry and Molecular Biology, 35, Zimoch et al., Regulation of chitin synthesis in the larval midgut of *Manduca sexta*. 515-527, Copyright 2005, with permission from Elsevier)

Figure 9. Northern blot analysis of chitin synthase transcripts in different sections of midgut during development.



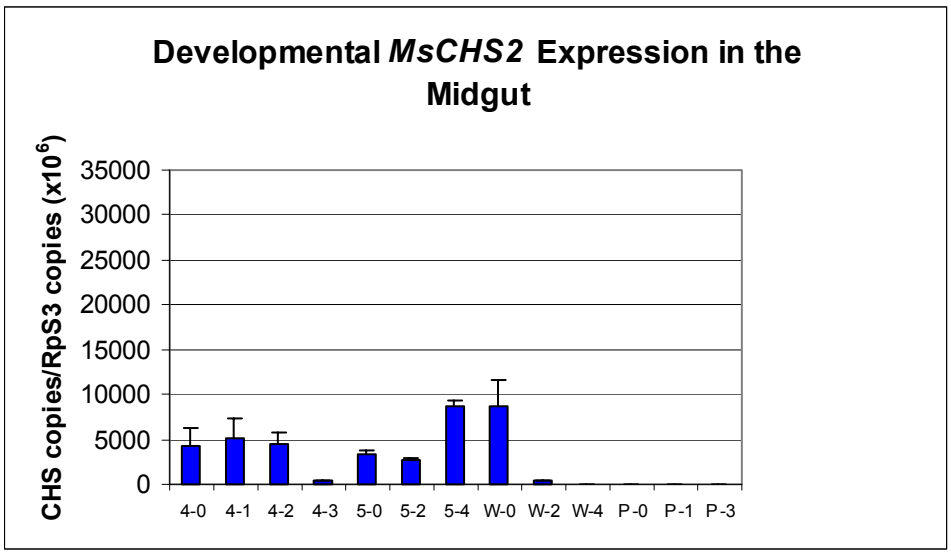
Total RNA was extracted from anterior midguts at various stages during development: fifth instar days 0, 2, 4 (lanes 5-0, 5-2, and 5-4); Wandering days 0, 2, 4 (lanes W-0, W-2, and W-4); Pupa day 1 (lane P-1). Probes specific for *MsCHS1*, *MsCHS2*, and *RpS3* were radiolabeled using (α-^{32}P) dCTP. Following hybridization and detection by autoradiography using the *MsCHS2* probe, the membrane was stripped of the probe in boiling 0.1% SDS. The membrane was then re-probed with both the *MsCHS1* and *RpS3* probes simultaneously. Abbreviations are as follows: Ant. Mid., anterior midgut; Med. Mid., medial midgut; Post. Mid., posterior midgut; Epid. P-1, epidermis pupa day 1. Anterior, medial, and posterior midgut samples were obtained from dissected midgut tissues during the fifth larval instar, day 2. (Reprinted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development, 529-540, Copyright 2005, with permission from Elsevier)

Figure 10. Developmental pattern of expression of *Manduca sexta* chitinase in the integument, tracheae, and midgut by real-time RT-PCR.



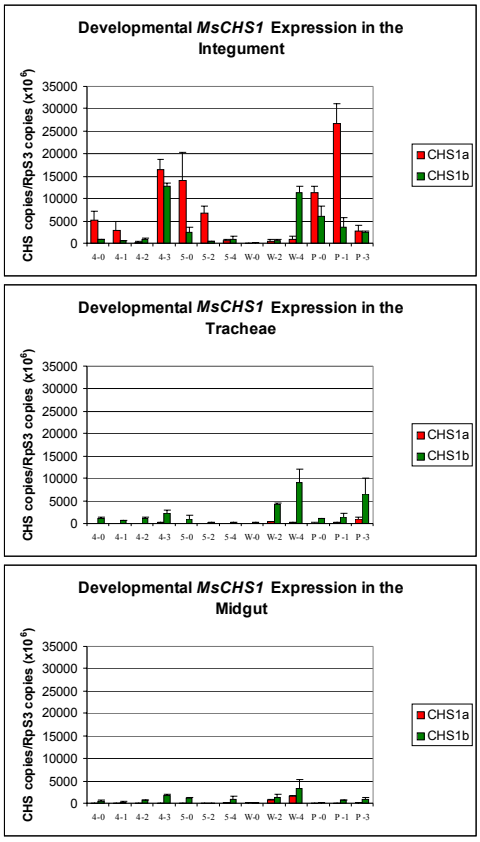
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from dissected epidermal, tracheal, and midgut tissues at various stages during development: Fourth instar days 0, 1, 2, 3 (lanes 4-0, 4-1, 4-2, and 4-3); Fifth instar days 0, 2, 4 (lanes 5-0, 5-2, and 5-4); Wandering days 0, 2, 4 (lanes W-0, W-2, and W-4); Pupal days 0, 1, 3 (lanes P-0, P-1, and P-3). Transcript copies were determined from standard curves constructed from cloned CHI and RpS3 constructs. CHI expression is presented relative to the internal housekeeping gene, RpS3. The specificity of the primers used in this analysis was not determined for the multiple chitinases present in *M. sexta*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 11. Developmental pattern of expression of *MsCHS2* in the midgut by real-time RT-PCR.



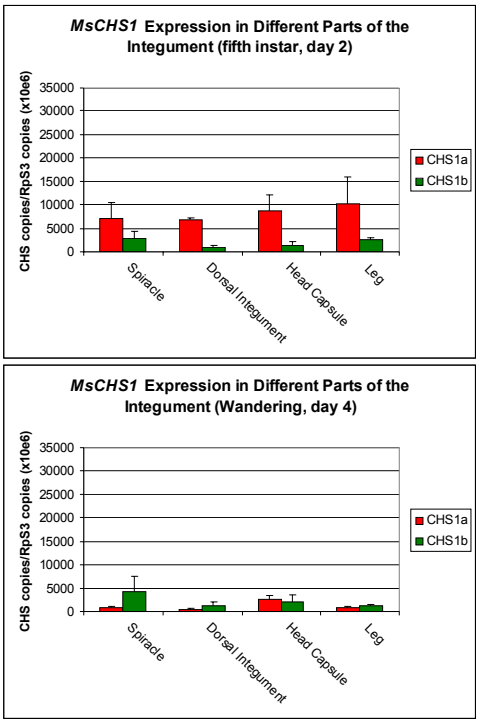
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from dissected midgut tissues at various stages during development: Fourth instar days 0, 1, 2, 3 (lanes 4-0, 4-1, 4-2, and 4-3); Fifth instar days 0, 2, 4 (lanes 5-0, 5-2, and 5-4); Wandering days 0, 2, 4 (lanes W-0, W-2, and W-4); Pupal days 0, 1, 3 (lanes P-0, P-1, and P-3). Transcript copies were determined from standard curves constructed from cloned *MsCHS2* and *RpS3* constructs. *MsCHS2* expression is presented relative to the internal housekeeping gene, *RpS3*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 12. Developmental pattern of expression of *MsCHS1a*/*MsCHS1b* in the integument, tracheae, and midgut by real-time RT-PCR.



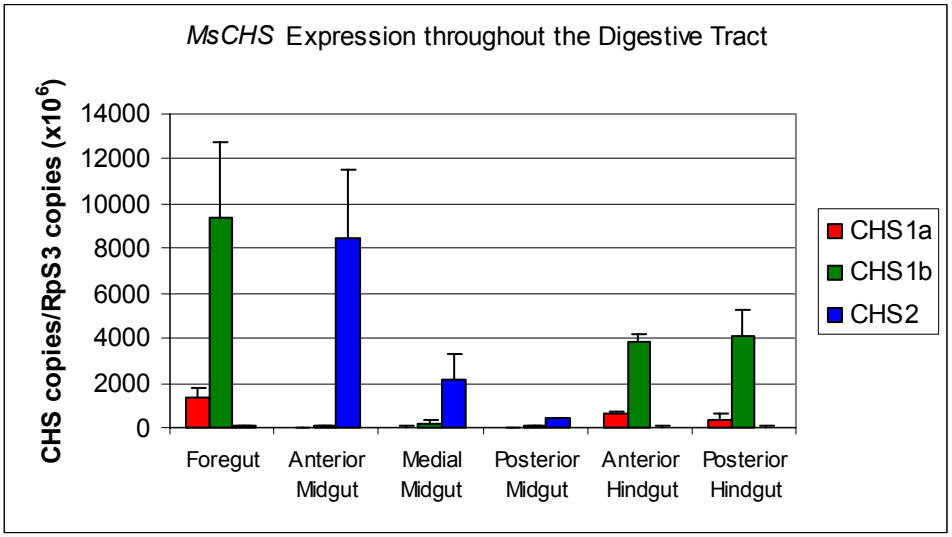
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from dissected integument, tracheal, and midgut tissues at various stages during development: Fourth instar days 0, 1, 2, 3 (lanes 4-0, 4-1, 4-2, and 4-3); Fifth instar days 0, 2, 4 (lanes 5-0, 5-2, and 5-4); Wandering days 0, 2, 4 (lanes W-0, W-2, and W-4); Pupal days 0, 1, 3 (lanes P-0, P-1, and P-3). Transcript copies were determined from standard curves constructed from cloned *MsCHS1a*, *MsCHS1b*, and *RpS3* constructs. *MsCHS1a* and *MsCHS1b* expression is presented relative to the internal housekeeping gene, *RpS3*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 13. Expression of *MsCHS1* in different regions of the integument.



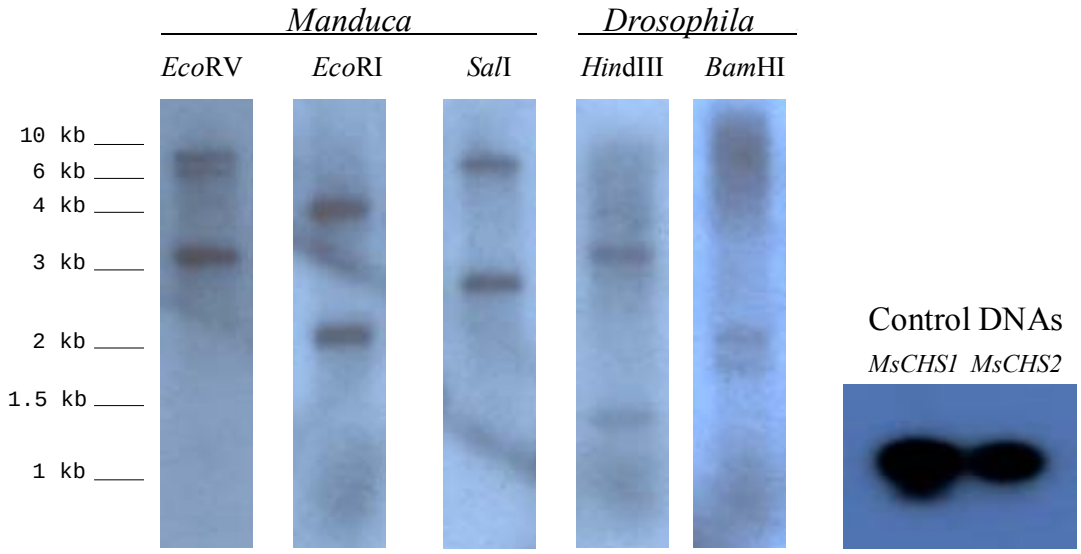
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from dissected integument obtained from different regions of the fifth instar (day 2) or wandering (day 4) larvae. Transcript copies were determined from standard curves constructed from cloned *MsCHS1a*, *MsCHS1b*, and *RpS3* constructs. *MsCHS1a* and *MsCHS1b* expression is presented relative to the internal housekeeping gene, *RpS3*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 14. Expression of CHSs along the length of the digestive tract of *Manduca sexta*.



The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from dissected regions of the digestive tract of *M. sexta* fifth instar larvae, days 0 to 2. Transcript copies were determined from standard curves constructed from cloned *MsCHS1a*, *MsCHS1b*, *MsCHS2*, and *RpS3* constructs. *MsCHS1a*, *MsCHS1b*, and *MsCHS2* expression is presented relative to the internal housekeeping gene, *RpS3*. Data are presented as the average of three or more separate experiments with error bars showing the standard deviation from the mean.

Figure 15. Southern blot analysis of *Manduca sexta* and *Drosophila melanogaster* genomic DNAs.



M. sexta DNA (10 µg per lane) was digested with the following restriction enzymes: *Eco*RV, *Eco*RI, and *Sal*I. *D. melanogaster* DNA (10 µg per lane) was digested with *Hin*dIII and *Bam*HI restriction enzymes. A ^{32}P-radiolabeled probe corresponding to the highly conserved nucleotide positions 2315 to 2515 of *MsCHS1* was used in the hybridization. For controls, non-radioactive PCR fragments (200 pg each) corresponding to *MsCHS1* and *MsCHS2* probes were also assayed in parallel. (Reprinted from Insect Biochemistry and Molecular Biology, 35, Hogenkamp et al., Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development, 529-540, Copyright 2005, with permission from Elsevier)

DISCUSSION

The cloning and partial characterization of the first *M. sexta* CHS gene, *MsCHS1*, was reported previously (Zhu et al., 2002). The previous work described a single CHS transcript and part of the corresponding genomic sequence. In the present work, a second *M. sexta* CHS gene, *MsCHS2*, has been fully characterized and the missing genomic sequences representing the 3'-end of *MsCHS1* gene have been completed. Southern blot analyses suggest the existence of only these two CHS genes in the *M. sexta* genome (Fig. 15). The presence of only two distinct CHS genes in the lepidopteran *M. sexta*, the coleopteran *T. castaneum*, and the dipterans *D. melanogaster* and *A. gambiae* (Arakane et al., 2004) has led to the hypothesis that these enzymes may have distinct functions and may be responsible for deposition of chitin associated in different chitinous structures within insects. Since the major chitin-containing structures found in insects are the cuticle and the PM that lines the midgut, it has been hypothesized that one CHS is utilized for the synthesis of chitin in the cuticle, whereas the other enzyme functions to synthesize chitin in the PM. However, until now, this hypothesis has never been tested in a single insect species by analyzing the expression of both CHS genes in several tissues at different developmental stages. The large size and well-characterized life cycle of *M. sexta* make it a suitable model organism for this study. Because of its size, dissection of epidermal, tracheal, and midgut tissues with minimal contamination from other tissues was possible in *M. sexta*, allowing for straightforward analysis of CHS expression in various tissues. Furthermore, sections of anterior, medial, and posterior midgut samples were easily obtained to further study the differential expression of CHS genes throughout the length of the midgut.

Using various methods of gene expression analysis including RT-PCR, northern blots, and real-time RT-PCR, it is clear that *MsCHS1* is expressed in the epidermis and tracheae, whereas *MsCHS2* is expressed in the midgut. The low amounts of *MsCHS1* transcripts that are detected in midgut tissues (at times when *MsCHS2* expression is low or undetectable) are attributable to tracheal contamination. Results from different methods of quantifying the level of expression, including northern blots and quantitative real-time RT-PCR, have indicated that the relative level of expression of *MsCHS1* transcripts in the midgut is at least two fold lower than that of isolated tracheae. This observation is consistent with the notion that these transcripts arise from a contaminating tissue, most likely the tracheae. Furthermore, the developmental

expression pattern of *MsCHS1* in the tracheae closely resembles that of the midgut. Finally, the relative amounts of the alternately spliced *MsCHS1*a/*MsCHS1b* transcripts are similar in the isolated tracheae and midgut tissues, whereas this ratio is reversed in epidermal preparations.

The extensive data studying the expression of CHS genes in *Manduca* indicates that there is strict tissue specificity of expression of the two CHS genes. The observed difference in the developmental patterns of expression of the two CHS genes also supports their distinct biological roles. *MsCHS1* expression is observed during periods of growth such as during the larval-larval and larval-pupal intermolts when the insect is actively synthesizing new cuticle and forming new tracheae, which requires an increase in the amount of exoskeletal and tracheal chitin, respectively. On the other hand, *MsCHS2*, is expressed only during larval feeding periods when synthesis and elaboration of a new PM occurs (Fig. 8). In *Tribolium*, this gut-specific transcript is also expressed during adult feeding periods but at a reduced level (Arakane et al., 2004). In *M. sexta*, expression of this gut-specific CHS is observed throughout the feeding stage in the fifth instar and transcripts are still seen at the point when the insect enters the wandering stage. However, only trace amounts of *MsCHS2* transcripts are detected on the second day of the wandering stage and beyond, suggesting a rapid down-regulation of synthesis and/or turnover of these transcripts during the late wandering stages (Fig. 11). These abrupt changes in levels of *MsCHS2* transcripts may reflect changes in the rate of synthesis/turnover of this transcript at this stage when rapid changes in hormonal levels are known to occur (Riddiford, 1994). These observations are in agreement with the hypothesis that MsCHS2 synthesizes PM-associated chitin and is required only when the PM is being synthesized. The absence of *MsCHS2* transcripts in other tissues also supports this hypothesis.

As discussed above, MsCHS2 is gut-specific and functions to synthesize PM-associated chitin in the midgut. However, trace amounts of *MsCHS1* transcripts were observed in midgut preparations at various developmental stages. Nevertheless, MsCHS1 has no role in the synthesis of PM for the following reasons. A comparison of the developmental profile of expression of the two CHS genes in the midgut is clearly different, peaking at different times. In fact, on day zero of the wandering stage, when *MsCHS2* expression is at a high level, there is no detectable *MsCHS1* expression in the midgut (Fig. 12). Conversely, on day four of the wandering stage, when *MsCHS1* levels in the midgut are maximal, there are no detectable *MsCHS2* transcripts. These results suggest that the *MsCHS1* expression in the midgut is under a

different control mechanism from that of *MsCHS2* and transcripts for *MsCHS1* found in the midgut originate from a cell type other than the midgut columnar cells, which express only *MsCHS2* (Zimoch et al., 2005). The possibility of significant epidermal contamination of the midgut is low because of the size of the insect at the time of dissection and the meticulous care taken during the isolation of the midgut. Therefore, the source of *MsCHS1* transcripts in the midgut must be from a non-epidermal tissue that is intimately associated with the midgut. We have ruled out the possibility that contamination of our midgut preparations with foregut or hindgut tissues was the source of *MsCHS1* transcripts by carefully dissecting the midgut to avoid such contamination. The most likely candidate therefore is tracheal tissue, which is closely associated with the midgut columnar cells. Previous immunolocalization studies have shown that tracheae indeed have CHS protein (Zimoch & Merzendorfer, 2002). However, the nature of the CHS gene that was expressed in the tracheae was not established in that study. In poly-A RNA isolated directly from tracheae (free of other midgut tissue), using gene-specific probes, only transcripts for *MsCHS1*, but not those for *MsCHS2*, were detected in tracheae (Hogenkamp et al., 2005). Therefore, tracheae express *MsCHS1* exclusively with no detectable levels of *MsCHS2* expression. Direct evidence for the expression of *MsCHS1* by tracheae has also been previously obtained (Zimoch et al., 2005). The trace amounts of this transcript detected in midgut preparations during the feeding stage and somewhat larger amounts during the wandering stage are, therefore, probably due to tracheae being a minor component of the midgut tissue and the different timing of expression of the CHS genes in these two tissues. Further analyses including *in situ* hybridization with gene-specific probes or immunolocalization studies using antibodies specific for each CHS are required to more definitively address this issue.

In the *MsCHS1* gene, the presence of two alternate exons that correspond to exon 20 in the *MsCHS2* gene was revealed following completion of sequencing of both genes (Fig. 7). This region of the gene was not sequenced in the study that initially characterized the *MsCHS1* cDNA (Zhu et al., 2002). Classification of the *MsCHS1* and *MsCHS2* genes based on amino acid sequence similarities of the encoded proteins indicates that *MsCHS1* encodes a class A CHS, whereas *MsCHS2* codes for a class B CHS (Fig. 1). The presence of alternate exons in genes for class A CHSs has been reported in several other insect species including *Tribolium*, *Drosophila*, and *Anopheles* (Arakane et al., 2004). However, an alternate exon has not been identified in genes encoding any insect class B CHS sequenced to date, including *MsCHS2*. Although the

genes encoding class A CHSs of most insect species are designated as CHS1 genes, the naming of *A. gambiae* CHS genes are reversed, with *AgCHS2* encoding the class A CHS and *AgCHS1* encoding the class B CHS. Similarly, the naming of *Aedes aegypti* CHSs are also reversed. Thus, all CHS genes characterized so far encoding class A CHSs from insects belonging to the lepidopteran, coleopteran, and dipteran orders have two alternate exons for a region near the C-terminus that includes a TMS. The presence of these alternate exons appears to be a distinguishing feature of class A CHSs in all insect species.

The tissue specificity of expression of the alternately spliced transcripts, *MsCHS1a* and *MsCHS1b*, has provided some interesting insights into the biological importance of the two isoforms of the enzymes translated from alternatively spliced transcripts. In the epidermis, the *MsCHS1a* transcript preferentially accumulates over the *MsCHS1b* transcript, particularly during the feeding and pupal stages (Figs. 6 & 10). However, the preferential accumulation of the *MsCHS1b* transcript in tracheae originating from the midgut (prepupal stage) suggests an important role for the encoded enzyme in tracheal development (Fig. 12). Although we have not analyzed levels of these transcripts in isolated tracheae from epidermal tissue, it is likely that tracheae embedded in the epidermis will have the same preference for the *MsCHS1b* isoform as the tracheae associated with midgut tissue. In fact, tracheal contamination probably contributes to some (or possibly all) of the *MsCHS1b* transcript levels observed in the epidermis, and therefore the preference for the *MsCHS1a* isoform in epidermal tissue may be even greater. In an attempt to address this issue, dissected epidermal tissues were taken from different regions corresponding to the dorsal integument, head capsule, leg, and the region surrounding and including the spiracle, which are known to vary in their tracheal content. When the relative expression level of both isoforms was analyzed in these preparations during day two of the fifth instar, no significant difference in the *MsCHS1a*/*MsCHS1b* ratio was observed (Fig. 13). During this developmental stage, the level of CHS expression in the tracheae was negligible. However, when these same tissue preparations were taken at a developmental stage when tracheal CHS expression is maximal (day four of the wandering stage), a significantly higher level of *MsCHS1b* expression was observed in the spiracle preparations, in which the local tracheal concentration is greater than in preparations of other regions of the integument. This result provides further evidence for the preferential expression of the *MsCHS1b* isoform in the tracheae and of the *MsCHS1a* isoform in the integument.

A comparison of the deduced amino acid sequences of the segments encoded by the two alternate exons of *MsCHS1* reveals a high degree of sequence similarity. However, a significant difference between the two isoforms is the presence of a potential *N*-glycosylation site, NKSV, in *MsCHS1b*, which is absent in *MsCHS1a* (replaced by the sequence DSSV). This potential *N*-glycosylation site is also present in the class B CHSs of several insects including *T. castaneum*, *A. gambiae*, and *D. melanogaster* (Arakane et al., 2004). Perhaps the function of the alternate exon is to provide a site for *N*-glycosylation in the *MsCHS1b* isoform. This isoform could, therefore, have different enzymatic properties and/or be differentially regulated, and serve a slightly different function from its *MsCHS1a* counterpart. An analysis of the alternate exon usage of *MsCHS1* in the midgut indicates that the *MsCHS1b* isoform is preferentially expressed over *MsCHS1a*, particularly during day four of the wandering stage (Fig. 12). At this stage, the tracheae undergo a major structural rearrangement and actively synthesize chitin (Zimoch et al., 2005). It would be interesting to determine whether the presence of the *MsCHS1b* isoform has any influence on the rate of chitin synthesis and/or chitin deposition in the tracheae.

The expression of CHS along the length of the digestive tract of *M. sexta* was also studied by real-time RT-PCR (Fig. 14). Consistent with the northern blot analysis (Fig. 9), a gradient in the level of expression of *MsCHS2* is observed starting from the anterior section of the midgut and tapering off at the posterior section of the midgut. The level of this midgut-specific CHS is maximal in the anterior midgut, but the level of expression in the medial midgut is reduced by approximately 20%. A similar level of reduction in the number of transcripts is observed between the medial and posterior midgut where the level of *MsCHS2* transcripts in the posterior midgut is less than 10% than that of the anterior midgut. Lepidopterans secrete a type-I PM, which is present only when the larvae are feeding (Terra, 2001). Therefore, at the cessation of feeding, this type of PM (including its constituent chitin) is no longer needed and is degraded. Indeed, the expression of chitinases has been previously observed to be up-regulated during this transitional period (Kramer et al., 1993). Furthermore, quantitative PCR analysis of chitinase expression revealed an increased level of expression during the wandering stage as well as the molting periods in not just the midgut but also the integument and tracheae (Fig. 10). This correlation of down-regulation of CHS and up-regulation of chitinase during the wandering stage suggests that the expression of both genes is under strict reciprocal hormonal and/or developmental control. This period during the wandering stage coincides with a large peak in

ecdysteriod titers (Riddiford et al., 2003). In fact, 20-hydroxyecdysone has been shown to upregulate the expression of both CHI and NAG in *M. sexta* (Kramer et al., 1993; Zen et al., 1996). In *D. melanogaster*, CHS transcripts are barely detected during the period where the late-larval ecdysone pulse occurs (Gagou et al., 2002). Therefore, the timing of expression of CHS, CHI, and NAG may be under hormonal control. Generally speaking, CHI and NAG expression preceeds CHS expression where the chitinolytic enzymes are needed for chitin degradation just prior to the synthesis of new chitin by CHS.

Some insects that have a type-I PM produce this structure from the secretory epithelium located throughout the entire midgut (Terra, 2001). On the other hand, the type-II PM is secreted by only a few rows of epithelial cells known as the cardia, which are located at the entrance of the midgut. The finding of a gradient of *MsCHS2* expression, with the highest amounts in the anterior midgut and progressively lower amounts in the medial and posterior midgut sections, suggests that the PM is secreted by epithelial cells throughout the midgut and not by a select few cells, as is the case in insects with a type-I PM. This gradient of CHS expression has been observed previously in immunolabeling studies of anterior, medial, and posterior midgut cells (Zimoch & Merzendorfer, 2002). These studies also demonstrated that CHS is localized in the apical brush borders of midgut cells. Although a gradient of CHS expression exists along the length of the midgut of *M. sexta*, the entire midgut epithelium is secreting PM, which is in agreement with its classification as a type-I PM that is present in most lepidopterans (Peters, 1992).

Chitin-containing structures are found not only in the midgut of the insect digestive system but also in the foregut and hindgut. These latter two tissues are lined with a cuticular layer known as the intima. Real-time RT-PCR was used to follow CHS expression in these tissues to determine which isoform is responsible for the synthesis of chitin in the intima (Fig. 14). No significant levels of *MsCHS2* expression were observed in the foregut and hindgut preparations. Therefore, *MsCHS2* is midgut-specific and not expressed in other parts of the digestive system. Both the foregut and hindgut preparations showed significantly greater levels of *MsCHS1b* expression over *MsCHS1a* during the fifth instar (days zero to two), indicating that the synthesis of intima-associated chitin is primarily due to the *MsCHS1b* isoform. Although the foregut and hindgut preparations are invariably contaminated with tracheae, the level of CHS expression in the tracheae at this developmental stage is minimal (Fig. 12). Therefore, the

MsCHS1b isoform is primarily responsible for the synthesis of intima-associated chitin in the foregut and hindgut. Finally, no significant difference in the CHS expression level was observed between the anterior (intestine) and posterior (rectum) portions of the hindgut.

The existence of at least two CHS genes in *Manduca* was previously determined, but the exact number of genes remained unclear (Zimoch & Merzendorfer, 2002). Using Southern blot analysis, only two CHS genes were identified in *M. sexta* (Fig. 15). To date, no more than two CHS genes have been identified in any insect species. Analysis of the completed genome sequences of *D. melanogaster* and *A. gambiae* indicates that only two putative CHS genes are present in each species. The *Tribolium* genome sequencing is complete and only two CHS genes have been identified so far. It is unlikely that additional CHS genes will be found since Southern blot analysis and extensive BAC library screening also indicated the presence of only two CHS genes (Arakane et al., 2004). Even though it appears that there are only two CHS genes, three isoforms of CHS are possible because of the utilization of the two alternate exons of *MsCHS1*, a common characteristic of all insect genes encoding class A CHSs analyzed thus far. The function of *MsCHS2*, which is made by the midgut, is clearly the formation of PM by the midgut cells as shown by numerous experiments studying the expression of CHSs in *M. sexta* in which the expression of both genes was followed during different developmental stages (Figs. 6, 7, & 10). Furthermore, gene knockdown studies using RNA interference in *T. castaneum* has illustrated that the chitin synthase genes *TcCHS1* and *TcCHS2* are specialized for the synthesis of epidermal cuticle and the midgut peritrophic membrane, respectively (Arakane et al., 2005). In that study, the specific knockdown of *TcCHS1* disrupted the larval-larval, larval-pupal, and pupal-adult molts and significantly reduced the whole-body chitin content. Specific knockdown of *TcCHS2*, on the other hand, led to the cessation of feeding, a dramatic reduction in the larval size, and reduced chitin content in the midgut.

In this work, a second CHS gene in the model insect, *M. sexta*, was characterized. Southern blot analysis indicated that only two CHS genes are present in this species. An extensive study of both genes was carried out that included the developmental expression and tissue-specificity of these genes. Through these studies, it is clear that there is strict tissue-specific expression of both genes during all developmental stages analyzed. These results have led to the conclusion that MsCHS1 is utilized for the synthesis of cuticular, tracheal, and intima-associated chitin, whereas MsCHS2 is utilized for the synthesis of PM-associated chitin.

Furthermore, the alternative splicing of *MsCHS1* was revealed in this study. Two alternate exons were identified in *MsCHS1*, *MsCHS1a* and *MsCHS1b*, with the latter containing a potential *N*-glycosylation site. Studies of the tissue-specificity of expression of both alternately spliced transcripts have shown that the MsCHS1b isoform is the primary enzyme that is used for the synthesis of chitin in the tracheae and intima of the foregut and hindgut. In contrast, the MsCHS1a isoform is primarily responsible for the synthesis of cuticular chitin. The characterization of all CHS genes in a single model insect, *M. sexta*, as well as the developmental expression studies in specific tissues have greatly increased our understanding of chitin synthesis.

CHAPTER 2:

beta-N-ACETYLGLUCOSAMINIDASES IN *Tribolium castaneum*

INTRODUCTION

The degradation of chitin in insects is carried out by the synergistic action of two families of chitinolytic enzymes, chitinases (CHIs) and *β-N*-acetylglucosaminidases (Fukamizo & Kramer, 1985a, b). The insoluble chitin polymer is first hydrolyzed by the action of chitinase proteins. These enzymes typically degrade the chitin polymer into various small oligomers but are unable to efficiently hydrolyze the chitin polymer into the monomeric *N*-acetylglucosamine (GlcNAc) residues. The cleavage of chitin oligomers from the nonreducing end into the GlcNAc monomers is carried out by *β-N*-acetylglucosaminidases (NAGs). Several NAGs have been identified in various insect species including *M. sexta* (Zen et al., 1995) and *B. mori* (Nagamatsu et al., 1995). These NAGs were present in the molting fluid and their timing of expression suggests a function in the turnover of chitin during molting. Furthermore, a *M. sexta* NAG was found to be under hormonal control. Both *M. sexta* NAG and CHI were found to be up-regulated by the injection of 20-hydroxyecdysone into isolated abdomens (Zen et al., 1996; Kramer et al., 1993). However, the topical application of the juvenile hormone analog, fenoxycarb, suppressed the inductive effect of molting hormone on NAG and CHI expression. A combination of these two enzymes is much more efficient in hydrolysis of chitin than either enzyme alone (Fukamizo & Kramer, 1985a, b).

NAGs in other species have been implicated in various cellular processes including *N*-glycan processing (Léonard et al., 2006) and in the catabolism of free oligosaccharides in the cytosol (Suzuki et al., 2002). Additionally, deficiencies of the closely related hexosaminidases (CAZy family GH20) have been implicated in a number of degenerative neural diseases including Tay-Sachs and Sandhoff diseases (Neufeld, 1989).

In contrast to CHIs, less work has been conducted on the study of NAGs in insects. Furthermore, a systematic study of all NAG genes in a single insect species has not been conducted. In fact, the establishment of the number of NAGs has not been determined even in the popular model insects such as *A. gambiae* and *T. castaneum*. To date, there has been no complete functional characterization of all NAGs in any single insect species.

The recent sequencing of the *T. castaneum* genome and amenability of this insect to RNAi (Arakane et al., 2005) make it an excellent model species to study NAGs. A functional

analysis of putative NAGs identified from searching the genome database can be conducted by selectively knocking down specific transcripts through RNAi. Therefore, *T. castaneum* serves as an excellent model insect for the complete functional characterization of all NAG genes in this species.

MATERIALS AND METHODS

Insect cultures

The GA-1 strain (Haliscak and Beeman, 1983) of *T. castaneum* was used in all experiments. Insects were reared at 30°C under standard conditions as described previously (Beeman & Stuart, 1990).

Isolation of total RNA

The RNeasy® Protect Mini Kit (Qiagen) was used to isolate total RNA from *Tribolium* tissue samples according to the manufacturer's instructions. Briefly, approximately 10 mg of *Tribolium* insect samples were ground after freezing in liquid nitrogen. Lysis buffer containing *β*-mercaptoethanol was then added to the homogenate and extracted by passing through a 20-gauge needle. Following centrifugation at 15,000 g, total RNA in the supernatant was purified using the RNeasy® mini spin column. Total RNA samples were treated with RNase-free DNase I for 30 minutes at 37°C to remove contaminating genomic DNA.

Synthesis of first-strand cDNA

The SuperScript™ III First-Strand Synthesis System for RT-PCR (Invitrogen) was used to synthesize the first-strand cDNA according to the manufacturer's instructions. Briefly, an oligo$(dT)_{20}$ primer was used for reverse transcription and 5 µg of total RNA was used as template. Reverse transcription was carried out at 45°C for 50 min. Template RNA was removed following reverse transcription by treatment with RNase H.

Identification of putative *Tribolium* NAGs, ENG, and hexosaminidases

The amino acid sequence of a previously characterized *β*-*N*-acetylglucosaminidase (Zen et al., 1996) from the insect, *M. sexta*, was used to identify potential *T. castaneum* NAGs by using the *tblastn* program in the *Tribolium* genome database (http://www.bioinformatics.ksu.edu/BeetleBase/). Similarly, for the identification of potential *T. castaneum* hexosaminidases and ENG, the human hexosaminidase (HEXA) and human ENG protein sequences were used as query sequences in the blast searches, respectively.

PCR amplification

For PCR amplification of *T. castaneum* NAG cDNAs, using the blast results from the *Tribolium* genome database (http://www.bioinformatics.ksu.edu/BeetleBase/), pairs of gene-specific primers were designed from conserved regions of insect NAGs to PCR amplify a part of each cDNA. Following 3'- and 5'-RACE, pairs of gene-specific primers were designed for each TcNAG and the full-length cDNA was amplified using these primers and larval or pupal cDNA as template.

Cloning

Cloning of PCR products was carried out using the TOPO® TA Cloning Kit (Invitrogen) according to the manufacturer's instructions. Briefly, the PCR products were ligated into the pCR®4-TOPO® vector and the ligation mixture was used to transform One Shot® TOP10 chemically competent *E. coli*. Following growth under ampicillin selection, clones containing the desired insert were identified by PCR and grown up in LB medium overnight. Plasmid isolation was carried out using the QIAprep® Spin Miniprep Kit according to the manufacturer's instructions.

DNA sequencing

DNA sequencing was conducted at the DNA sequencing facility at Kansas State University using an automated sequencer (ABI Prism 3700).

Phylogenetic analysis of NAGs and hexosaminidases

ClusalW software (www.ebi.ac.uk/clustalw/) was used to perform multiple sequence alignment prior to phylogenetic analysis. MEGA 3.0 (Kumar et al., 2004) was used to construct the consensus phylogenetic tree using UPGMA. To evaluate the branch strength of the phylogenetic tree, bootstrap analysis of 5000 replications were performed. The deduced amino acid sequences of the three *T. castaneum* hexosaminidases: TcHEX1, TcHEX2, and TcHEX3 are shown in appendix A, B, and C, respectively. The deduced amino acid sequence of *T. castaneum* ENG is shown in appendix D.

Cloning of putative *Tribolium* NAGs

Following the identification of potential *T. castaneum* NAG genes, pairs of gene-specific primers were designed for each of the four potential NAGs. Following PCR amplification using a cDNA obtained from a mixture of *T. castaneum* larvae or pupae as template, the products were cloned and sequenced as described above. Cloning of the four full-length NAGs was carried out after designing pairs of primers from the sequences obtained from 5'- and 3'-RACE.

3'-RACE of *Tribolium* NAGs

The 3'-ends of four putative *T. castaneum* NAGs were obtained using 3'-Rapid Amplification of cDNA Ends (RACE). An adapter-oligo$(dT)_{17}$ primer (5'-GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT TTT T-3') was used in the first strand cDNA synthesis reactions. Following reverse transcription, a *T. castaneum* NAG gene-specific primer was used in conjunction with the adapter-oligo$(dT)_{17}$ primer in the first PCR reaction (see Table 1 for primer sequences). A second round of PCR was then carried out using the adapter primer (5'-GGC CAC GCG TCG ACT AGT AC-3') and a second nested *T. castaneum* NAG gene-specific primer (Table 1). The resulting PCR product was cloned and sequenced as described above.

5'-RACE of *Tribolium* NAGs

The 5'-ends of the putative *T. castaneum* NAGs were obtained using the 5'-RACE System for Rapid Amplification of cDNA Ends (Invitrogen). To ensure the production of a full-length cDNA and to aid in the specificity of the reaction, *T. castaneum* NAG gene-specific primers (see Table 2 for primer sequences) were used instead of an oligo-$(dT)_{17}$ primer in the reverse transcription reaction. Due to catalytic limitations of the reverse transcriptase enzyme, the reaction is typically carried out at a low temperature and, therefore, the T_m of these gene-specific primers were designed to be around 50°C. Furthermore, the primer was designed to have a high AT content in order to maximize the length of the primer and the specificity of the reaction. The resulting cDNAs produced after reverse transcription were purified using the S.N.A.P.™ column (Invitrogen) according to the manufacturer's instructions in order to remove residual dNTPs. The purified cDNAs were then dC-tailed by including only dCTP during the

terminal deoxynucleotidyl transferase tailing reaction. The 5'-RACE Abridged Anchor Primer (5'-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3') was then used in conjunction with a second nested *T. castaneum* NAG gene-specific primer (Table 2) in the first PCR reaction. Finally, the Abridged Universal Amplification Primer, AUAP (5'-GGC CAC GCG TCG ACT AGT AC-3'), was used in conjunction with a third nested *T. castaneum* NAG gene-specific primer (Table 2) and the resulting PCR product was cloned and sequenced as described above.

DNA and protein sequence analyses

Protein sequence analysis tools used in this study, including translation, MW, pI, and topology predictions, were obtained from the ExPASy Proteomics website (http://us.expasy.org/). TMHMM (v2.0) software (www.cbs.dtu.dk/services/TMHMM-2.0/) was used to predict transmembrane segments in the putative protein sequences. Potential coiled-coil domains were identified using the Paircoil Program (http://paircoil.lcs.mit.edu/cgi-bin/paircoil, Berger et al., 1995). Signal peptide and signal anchor prediction was conducted using the SignalP 3.0 server (www.cbs.dtu.dk/services/SignalP/). Sequence alignments of multiple proteins were carried out using the MultAlin website (http://prodes.toulouse.inra.fr/multalin/multalin.html, Corpet, 1988).

RT-PCR analysis of *TcNAG* expression

The developmental pattern of expression of all four putative *T. castaneum* NAGs was determined as well as their expression in dissected larval midguts by RT-PCR. Total RNA was isolated from beetles at various developmental stages and dissected midguts of late larvae using the RNeasy® Protect Mini Kit (Qiagen) according to the manufacturer's instructions. Two micrograms of total RNA were utilized as templates for cDNA synthesis using an oligo-(dT) primer. This cDNA then served as a template for the subsequent PCR reactions. To allow for the simultaneous analysis of expression of all four putative *T. castaneum* NAGs, the pairs of gene-specific primers used in the analysis (Table 3) were designed to produce different size PCR products. All primers were designed to have similar melting temperatures, enabling all PCR reactions to be carried out in the same vessel. For controls, primers for the constitutively expressed housekeeping gene *Ubiquitin* were also utilized. A series of PCR reactions of

different cycle numbers was carried out in order to determine the appropriate cycle number to be used in the analysis. All PCR reactions used in the expression analysis were conducted using the following conditions: denaturation at 94°C for 30 sec, annealing at 58°C for 45 sec and polymerization at 72°C for 1 min for 25 cycles.

Real-time RT-PCR analysis of *TcNAG* expression

Analysis of the expression of all four putative *T. castaneum* NAGs along with the internal control, *TcRpS6*, was carried out using real-time RT-PCR. Beacon Designer 2.0 software (Premier Biosoft International) was used to design all gene-specific primers (Table 4) for quantitative PCR analysis.

The PCR products for each of the primer pairs used in the quantitative PCR experiment were cloned into the pCR®4-TOPO® vector and after transformation of *E. coli*, the plasmid was purified from an overnight culture. The DNA concentration of each construct was determined using a VersaFluor™ Fluorometer (Bio-Rad) and the DNA intercalating dye Hoechst 33258. Sets of serial dilutions were then made from each construct and the dilutions were analyzed by real-time RT-PCR using the nucleic acid stain SYBR green (Bio-Rad). The Bio-Rad iCycler iQ® real-time RT-PCR detection system at the Kansas State University Gene Expression Facility was used in the quantitative PCR experiments. The protocol used in all experiments is outlined in Figure 3. Standard curves for each gene of interest were constructed using the serial dilutions of each construct. At the end of each quantitative PCR experiment, a melt curve was generated to rule out the possibility of primer dimer formation. Using the standard curves, the transcript copy number was determined for each gene of interest along with the copy number of the internal housekeeping gene, *TcRpS6*.

Double-stranded RNA synthesis

The most dissimilar regions of the *T. castaneum* NAG genes were chosen to synthesize double-stranded RNA (dsRNA) in order to minimize the likelihood of silencing other NAGs that share similar nucleotide sequences (Table 5). Following cloning and sequencing of the four putative NAGs, forward and reverse primers were designed corresponding to the dsRNA regions in Table 5. All primer sets contained the T7 RNA polymerase promoter sequence (5'-TAA TAC GAC TCA CTA TAG-3') at the 5'-end of the forward and reverse primers. Therefore, T7 RNA

polymerase could catalyze the synthesis of an RNA molecule from both strands of the dsDNA to produce a dsRNA. Following PCR amplification with the *T. castaneum* NAG primers with the T7 sequence at the 5'-end, the resulting products were excised from a 1.5% agarose gel, extracted in phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v), ethanol precipitated in 0.3 M sodium acetate overnight, and redissolved in DEPC-treated water. The resulting dsDNAs with flanking T7 sequences were then used as templates for the *in vitro* transcription of the dsRNAs. Double-stranded RNAs were synthesized using the AmpliScribe T7-Flash Transcription Kit (Epicentre) according to the manufacturer's instructions. The resulting dsRNAs were then extracted in phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v), isopropanol precipitated in 2 M ammonium acetate overnight, and redissolved in DEPC-treated water.

Injection of double-stranded RNA into *Tribolium*

Double-stranded RNAs corresponding to all four putative NAGs were injected into *T. castaneum* young larvae (3^{rd} to 4^{th} instar) and late larvae (penultimate instar larvae and last instar larvae) using a microinjection system equipped with a glass needle and dissection stereomicroscope according to established protocols (Tomoyasu and Denell, 2004). Briefly, larvae were anaesthetized with ether for 3.5 min and then placed on a piece of double-stick tape on top of a glass slide. Approximately 0.2 µg of dsRNA (~0.2µl of 1 µg/µl dsRNA) was injected into the dorsal side of the first or second abdomen segment of each *T. castaneum* larva. Following injection, the larvae were removed from the glass slide and incubated at room temperature for two hours before being transferred to whole-wheat flour (with 5% yeast) and incubated at 30°C until the appropriate stage for analysis.

Homology modeling of putative *Tribolium* NAGs

The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology models for the four putative *T. castaneum* NAGs. TcNAG1 and TcNAG3 homology models were obtained using human β-hexosaminidase B (pdb 1NOU, Mark et al., 2003) as the template. The X-ray crystal structure of human β-hexosaminidase B (pdb 1NOU) was obtained for the mature protein purified from human placenta and crystallized in its native mature (lysosomal) form. TcNAG2 and TcNAG4 homology models were obtained using human β-hexosaminidase B (pdb 1o7a, Maier et al., 2003) as the template. The X-ray crystal

structure of human β-hexosaminidase B (pdb 1o7a) was obtained for the recombinantly expressed protein in insect cells. The human β-hexosaminidase B enzymes, 1o7a (Appendix E) and 1NOU (Appendix F), are isoforms with 1o7a being slightly shorter than 1NOU. The two isoforms, 1o7a and 1NOU, exhibit over 95% identities.

Table 1. Primers used in 3'-RACE of putative *Tribolium* NAGs.

Gene	Primer Name	Primer Sequence (5'-3')
TcNAG1	3RACETcNAG1F1	TGGTGCAAAGGGAAAGACTG
	3RACETcNAG1F2	CGGACGCTCTTGAACCCGAATG
TcNAG2	3RACETcNAG2F1	CATGAACTACACCGACTTGCAC
	3RACETcNAG2F2	GGTGATCATCTCGCACGTGGAC
TcNAG3	3RACETcNAG3F1	CACCTATCATTCTATGGACGAGTC
	3RACETcNAG3F2	GTGGATGACAGTTCCGTGGAGTC
TcNAG4	3RACETcNAG4F1	CTAGACGTGAAGGTAACAGTGC
	3RACETcNAG4F2	GGGTGTGCGACCCATACACC

Table 2. Primers used in 5'-RACE of putative *Tribolium* NAGs.

Gene	Primer Name	Primer Sequence (5’-3’)
TcNAG1	5RACETcNAG1R1	CTCGTCCACAAAATGACAG
	5RACETcNAG1R2	TACTGATACGCCGGCTCGTC
	5RACETcNAG1R3	GGCGACCGTCAAAGCGTACG
TcNAG2	5RACETcNAG2R1	CGAGGATTTCGTAAACGTTG
	5RACETcNAG2R2	CGACCAAAGCCTTGACGTCTTC
	5RACETcNAG2R3	GGCCATTCCGTCCAACACTCGTC
TcNAG3	5RACETcNAG3R1	CCTAGAAGATTGGTAATATCTTC
	5RACETcNAG3R2	CGAGCAACAATCCTCTATGTGG
	5RACETcNAG3R3	GGCGTCAGAAATTGACGCTTCGTC
TcNAG4	5RACETcNAG4R1	GTCAGATGCAAATGCAGGAC
	5RACETcNAG4R2	GCCATTGCCATTCCATCAACAAC
	5RACETcNAG4R3	GGGCGGTGTCAATCATCACTCC

Table 3. Primers used in the RT-PCR analysis of putative *Tribolium* NAGs.

Gene	Primer Name	Primer Sequence (5'-3')	Amplicon Size (bp)
TcNAG1	TcNAG1ExpF1	CGGGCTGGAGACTCTGAACC	1006
	TcNAG1ExpR1	CTTTTTATCGCCGGCGATTTTGGC	
TcNAG2	TcNAG2ExpF1	AATGTCTTCCACTGGCACCTCAC	894
	TcNAG2ExpR1	CCTCGTATCGAGCGACGTCTC	
TcNAG3	TcNAG3ExpF1	GGATTGTTGCTCGATACTGCCAG	730
	TcNAG3ExpR1	CCCAGGTTTGTATAACGTATCTGG	
TcNAG4	TcNAG4ExpF1	CGGTTTGGTGCTTACGGACCTG	815
	TcNAG4ExpR1	GCTGAAGGGTCGCTCCAGATG	
Ubiquitin	UbiqF1	GACCGGCAAGACCATCACT	192
	UbiqR1	CGCAGACGCAAAACTAAATGAAGG	

Table 4. Primers used in the quantitative real-time RT-PCR analysis of putative *Tribolium* NAGs.

Gene	Primer Name	Primer Sequence (5′-3′)
TcNAG1	RTTcNAG1F1	GGCGACGAGGTCCATTTGGG
	RTTcNAG1R1	ATGAAATCTCCCTCAGAGCGGC
TcNAG2	RTTcNAG2F1	AATGTCTTCCACTGGCACCTCACC
	RTTcNAG2R1	CGGCGTGTAGATCATGTCGGGTC
TcNAG3	RTTcNAG3F1	TCCACTGGCACATAACTGACTC
	RTTcNAG3R1	AACTCCTCGCAATTTAGCATAGC
TcNAG4	RTTcNAG4F1	AAGTTGTTGATGGAATGGCAATGG
	RTTcNAG4R1	TGTGTAGATCATATCAGGTCCGTAAG
TcRpS6	RTTcRpS6F1	GACGAATGGAAGGGCTATGTTC
	RTTcRpS6R1	CGCCTTTACGCACGATAACG

Table 5. Summary of *Tribolium* NAG dsRNAs.

Gene	Nucleotide Position	Length (bp)
TcNAG1	1116-1380	264
TcNAG2	1227-1485	258
TcNAG3	1158-1425	267
TcNAG4	1134-1353	219

RESULTS

Identification of putative *Tribolium* NAGs

In order to identify possible NAGs in *T. castaneum*, the amino acid sequence from a previously characterized NAG (CAZy family GH20) from *M. sexta* (Zen et al., 1996) was used to blast the *Tribolium* genome database using the *tblastn* program (Table 6). A total of seven genes were identified as possible hits. Using the cutoff of 50% similarity to *M. sexta* NAG, a total of four putative *T. castaneum* NAGs were identified: *TcNAG1*, *TcNAG2*, *TcNAG3*, and *TcNAG4*. The additional three hits showed less than 50% similarity to *M. sexta* NAG and instead more closely resembled the human hexosaminidase, *HEXB*, with >50% similarity to *HEXB* (CAZy family GH20). These three hits were, therefore, annotated as hexosaminidases and named *TcHEX1*, *TcHEX2*, and *TcHEX3*. Furthermore, the four *T. castaneum* NAGs showed less than 50% similarity to the Human *HEXB*.

Cloning of putative *Tribolium* NAG cDNAs

The full-length cDNAs of each of the four putative *T. castaneum* NAGs were cloned and sequenced. Firstly, forward and reverse primers were designed from the genomic sequence obtained from the *Tribolium* genome database. Following PCR amplification from a larval or pupal cDNA, the resulting product was cloned and sequenced. The 5'- and 3'-ends were determined using 5'- and 3'-RACE, respectively. Forward and reverse primers specific for each of the four NAGs were then used to PCR amplify the full-length cDNA of each, which were then cloned and sequenced.

Sequence analysis of *Tribolium* NAGs

The amino acid sequences of each of the four putative *T. castaneum* NAGs were analyzed for the presence of a signal anchor or signal peptide (Table 7). Three of the NAGs, TcNAG1, TcNAG3, and TcNAG4, were predicted to have a signal peptide and are probably secreted proteins. One NAG, TcNAG2, was predicted to have a signal anchor and the presence of a single TMS (residues 9-26) was verified using TMHMM (v2.0) software. Furthermore, the predicted pI of this NAG was slightly higher (pH 6.61) than the other NAGs, all of which had a

predicted pI of less than pH 6.0. Finally, the MW of TcNAG2 was slightly higher (~63 kDa) than the other NAGs, which were all less than 60 kDa.

Gene organization of *Tribolium* NAGs

Following the cloning of the four full-length putative NAG cDNAs from *T. castaneum*, and comparison with the genome sequences, the exon-intron organization of each gene was determined (Fig. 16). The untranslated regions of all genes are relatively short (less than 200 nucleotides). A large intron is found adjacent to the putative ATG start codons of *TcNAG1*, *TcNAG2*, and *TcNAG3*, but not in *TcNAG4*. The first intron of *TcNAG2* is by far the largest. In fact, this intron is larger than any of the genes encoding the other NAGs. *TcNAG4*, which consists of only two exons, contains only one small intron located just after the putative translation start site. The gene organizations of *TcNAG1* and *TcNAG3* are remarkably similar, whereas *TcNAG2* and *TcNAG4* are quite unique.

Phylogenetic analysis of NAGs, hexosaminidases, and ENGs

Phylogenetic analysis of the putative *T. castaneum* NAGs, hexosaminidases, and ENG was carried out with the orthologs of other insect and metazoan species (Fig. 17). To evaluate the branch strength of the tree, a bootstrap analysis of 5000 replications was performed. Furthermore, in addition to the Unweighted Pair Group Method with Arithmetic mean (UPGMA), another distance-based method, neighbor-joining, was employed as well as maximum parsimony and minimum evolution. All four trees were consistent and grouped the proteins into five major classes: Chitinolytic NAGs Group I, Chitinolytic NAGs Group II, *N*-glycan processing NAGs, Hexosaminidases, and ENGs as previously classified (Léonard et al., 2006). TcNAG1 was grouped with the well-characterized NAGs including *M. sexta* NAG in Group I of the Chitinolytic NAGs. TcNAG3 was grouped with DmHEXO2 in the Chitinolytic NAGs, Group II. TcNAG2 was grouped with the *D. melanogaster fused lobes* (DmFDL) along with the *fused lobes* (*fdl*) homologs of other insect species. This result was consistent with the prediction of the fused lobes proteins having a transmembrane anchor and TcNAG2 is also predicted to have a signal anchor (Table 7). Curiously, TcNAG4 did not fall into any obvious group and appeared on a separate branch. The *T. castaneum* hexosaminidases, however, were clearly grouped with the metazoan hexosaminidases. Similarly, the *T. castaneum* ENG was

grouped with the human ENG with a bootstrap value of 100. The amino acid sequence alignment of the four putative *T. castaneum* NAGs with the well-characterized *M. sexta* NAG shows a high degree of conservation, reinforcing the classification of these proteins as NAGs (Fig. 18).

Developmental pattern of expression of *Tribolium* NAGs and expression in the larval midgut by RT-PCR

The expression pattern during insect development and in the larval midgut of all four *T. castaneum* NAGs was investigated by RT-PCR (Fig. 19). All four NAGs were expressed at the early and late larval stages. *TcNAG4* was not expressed at a detectable level following the late larval stage in contrast to the other three NAGs. The expression levels of *TcNAG2* and *TcNAG3* appeared to decrease in the middle of the pupal stage. All four NAGs were expressed in the larval midgut. However, the relative level of expression of *TcNAG2* was somewhat lower in comparison to the other *T. castaneum* NAGs.

Developmental pattern of expression of *Tribolium* NAGs by real-time RT-PCR

Using real-time RT-PCR, the developmental patterns of expression of all four *T. castaneum* NAGs were determined (Fig. 30). *TcNAG1* is expressed at the highest level relative to *RpS6*. The number of *TcNAG4* transcripts is significantly lower than those of the other NAGs and significant expression is observed only in the larval stages. In contrast, the other three NAGs are expressed at all developmental stages with the highest levels at the larval and late pupal stages. A slight decrease in the level of expression of *TcNAG1*, *TcNAG2*, and *TcNAG3* is observed at the late prepupal and mid-pupal stages as well as young adults. Overall, the pattern of expression of these three NAGs is quite similar.

Expression of *Tribolium* NAGs in the larval midgut and carcass by real-time RT-PCR

In order to study the tissue-specificity of the four putative *T. castaneum* NAGs, real-time RT-PCR was used to study their expression in dissected larval midguts and carcasses (larval abdomen with midgut removed). The expression level of *TcNAG1* is comparable in both the midgut and carcass indicating no strict tissue specific expression in these body parts (Fig. 31). Conversely, both *TcNAG3* and *TcNAG4* are expressed at a significantly higher levels in the

midgut than the carcass, indicating that these NAGs may function in the breakdown of PM-associated chitin. Both NAGs, however, do show limited expression in the carcass and therefore are not midgut-specific. *TcNAG2*, on the other hand, is expressed at a higher level in the carcass than the midgut with the expression in the carcass being approximately double than that in the midgut. Therefore, all four putative *T. castaneum* NAGs are differentially expressed in the midgut and surrounding tissues.

Determination of extent of knockdown of *Tribolium* NAGs by real-time RT-PCR

Real-time RT-PCR was used to determine the extent of knockdown of each *T. castaneum* NAG following dsRNA injections (Fig. 32). The expression levels in larvae were determined three days after injection with dsRNAs corresponding to unique regions of each *T. castaneum* NAG and the Enhanced Green Fluorescent Protein (*EGFP*) as a control. Injection of a dsRNA corresponding to one NAG did not significantly change the level of expression of the other three NAGs. The extent of knockdown ranged between 80-95% for *TcNAG1*, *TcNAG2*, and *TcNAG3* relative to *RpS6*. *TcNAG4*, which is expressed at a significantly lower level than the other three NAGs, showed an extent of knockdown greater than 50%. Furthermore, specific knockdown of *TcNAG1*, *TcNAG2*, and *TcNAG3* was also observed in *T. castaneum* pharate adults (data not shown).

Phenotypes resulting from specific knockdown of *Tribolium* NAGs

Double-stranded RNA corresponding to each of the four putative *T. castaneum* NAGs was injected into pharate pupae to study the effect of knockdown on pupal-adult molting. Specific knockdown of three of the NAGs; *TcNAG1*, *TcNAG2*, and *TcNAG3*, resulted in the arrest of the pupal-adult molt (Fig. 33). In each cohort, the synthesis of adult cuticle proceeded without any obvious defect, but the insects were unable to completely shed their old pupal cuticle. Over 75% of the insects exhibited such a phenotype when injected with the dsRNAs corresponding to each of these three NAGs (Table 8). Only a small number of similar phenotypes (approximately 10%) were observed in *TcNAG4* dsRNA treated beetles (data not shown).

Administration of *TcNAG1* dsRNA into larvae also affected the larval-larval and larval-pupal molts (Figs. 22 and 23, respectively). The insects were unable to shed their old larval

cuticle during these molts, resulting in a lethal phenotype. Similar phenotypes were also observed in *TcNAG2*, *TcNAG3*, and *TcNAG4* dsRNA-treated beetles, but less than 25% of the injected individuals exhibited an obvious phenotype (data not shown).

Homology modeling of putative *Tribolium* NAGs

X-ray crystal structures of two different proteins were used as templates for the homology modeling of putative *T. castaneum* NAGs. The X-ray crystal structure of the human β-hexosaminidase B (pdb 1NOU, Mark et al., 2003) determined for the mature protein purified from human placenta (crystallized in its native mature, lysosomal, form) is available in the database and this structure served as a template for the TcNAG1 and TcNAG3 homology models. The X-ray crystal structure of the human β-hexosaminidase B (pdb 1o7a, Maier et al., 2003) was determined also for the recombinantly expressed protein in insect cells and this structure served as a template for the TcNAG2 and TcNAG4 homology models. The two human β-hexosaminidase B proteins were highly similar and shared 97% identity, with the difference between the two proteins being only in their length (1o7a is slightly shorter than 1NOU). The human β-hexosaminidase B with the highest similarity to the NAGs was selected as a template for the homology modeling. The human β-hexosaminidase B, 1NOU, shared 33% identities with both TcNAG1 and TcNAG3. The human β-hexosaminidase B, 1o7a, shared 35% and 29% identities with TcNAG2 and TcNAG4, respectively.

The amino acid residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of human β-hexosaminidase B are catalytically important (Maier et al., 2003). Accordingly, these nine residues are highly conserved in TcNAG1 and TcNAG3 (Fig. 36) as well as in TcNAG2 and TcNAG4 (Fig. 37). Furthermore, the homology models of TcNAG1 (Fig. 38), TcNAG2 (Fig. 39), TcNAG3 (Fig. 40), and TcNAG4 (Fig. 41) indicate that these residues are located closely together, which is consistent with their proposed role in catalysis.

Both human β-hexosaminidase B proteins, 1NOU and 1o7a, contain a centrally located $(\beta,\alpha)_8$-barrel domain. Accordingly, a central $(\beta,\alpha)_8$-barrel domain is present in the homology models of all four *T. castaneum* NAGs (Fig. 42). This $(\beta,\alpha)_8$-barrel domain is easily discernible in each of the *T. castaneum* NAG homology models.

Table 6. Results of blast searches to identify putative NAG genes from the *Tribolium* genome database.

Annotation	Hit #	Contig #	Linkage Group	Scaffold	Scaffold nt start-nt stop	Percent Similarity with *Manduca* NAGase	Percent Similarity with Human HEXB
TcNAG1	1	2762	7	Contig4580_ Contig7715	794742-790319	69.7	40.3
TcNAG2	2	740	7	Contig4580_ Contig7715	1220316-1215143	54.7	38.7
TcNAG3	3	872	1	Contig2645_ Contig5355	55073-58142	56.1	40.8
TcNAG4	4	856	2	Contig4302_ Contig1329	1274669-1272919	50.0	39.9
TcHEX1	5	525	7	Contig3587_ Contig1062	3278306-327630	41.0	54.5
TcHEX2	6	525	7	Contig3587_ Contig1062	3283951-3285426	40.8	54.7
TcHEX3	7	525	7	Contig3587_ Contig1062	3282024-3280406	40.1	53.9

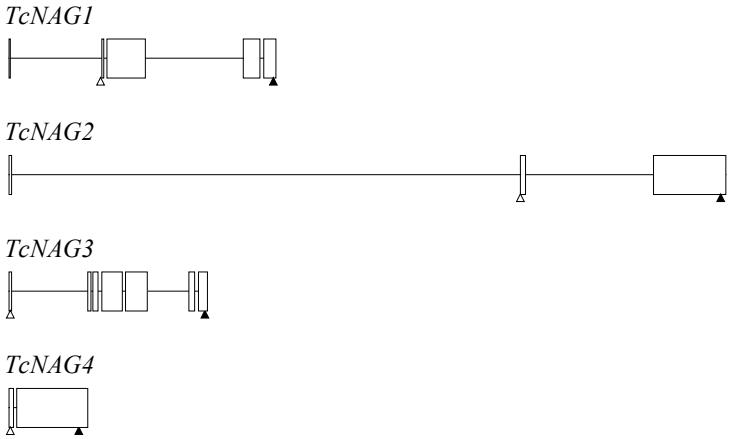
Table 7. Summary of properties of putative *Tribolium* NAG proteins.

Gene	Predicted amino acid length	Predicted MW (kDa)	Predicted pI	Predicted signal anchor	Predicted signal peptide
TcNAG1	598	67.3	5.18	No	Yes
TcNAG2	630	72.6	6.61	Yes	No
TcNAG3	593	67.4	5.47	No	Yes
TcNAG4	582	66.5	5.62	No	Yes

Table 8. Summary of the percentage of lethal phenotypes observed at various molts following *TcNAG1*, *TcNAG2*, *TcNAG3*, and *TcNAG4* knockdowns.

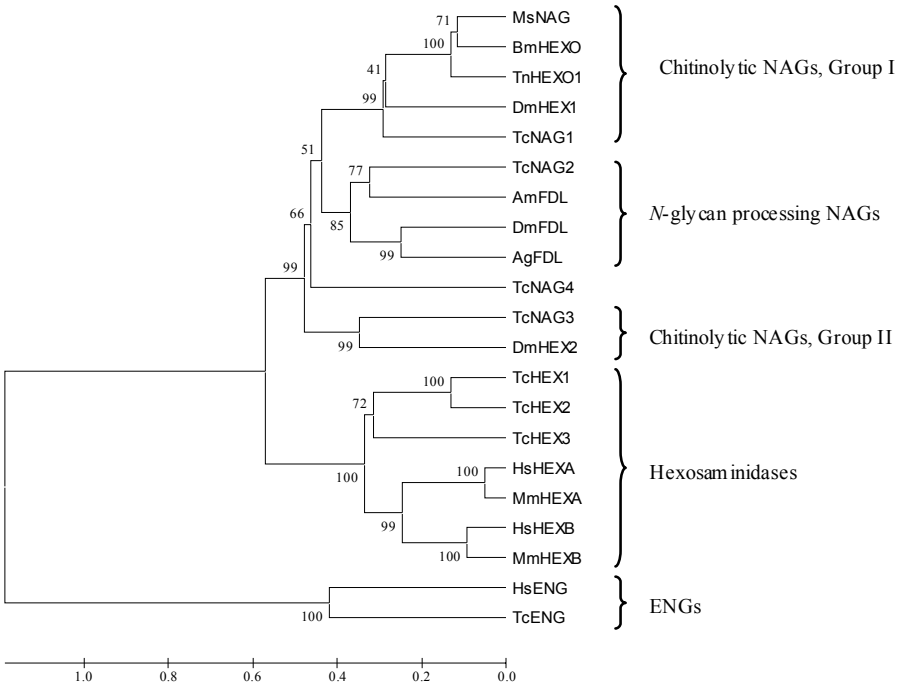
dsRNA Injected	% of individuals with the lethal phenotype at the molt		
	Larval-Larval Molt	Larval-Pupal Molt	Pupal-Adult Molt
TcNAG1	78%	79%	91%
TcNAG2	18%	10%	77%
TcNAG3	23%	14%	83%
TcNAG4	22%	9%	10%

Figure 16. Schematic diagram of the exon-intron organizations of the putative *Tribolium* NAG genes.



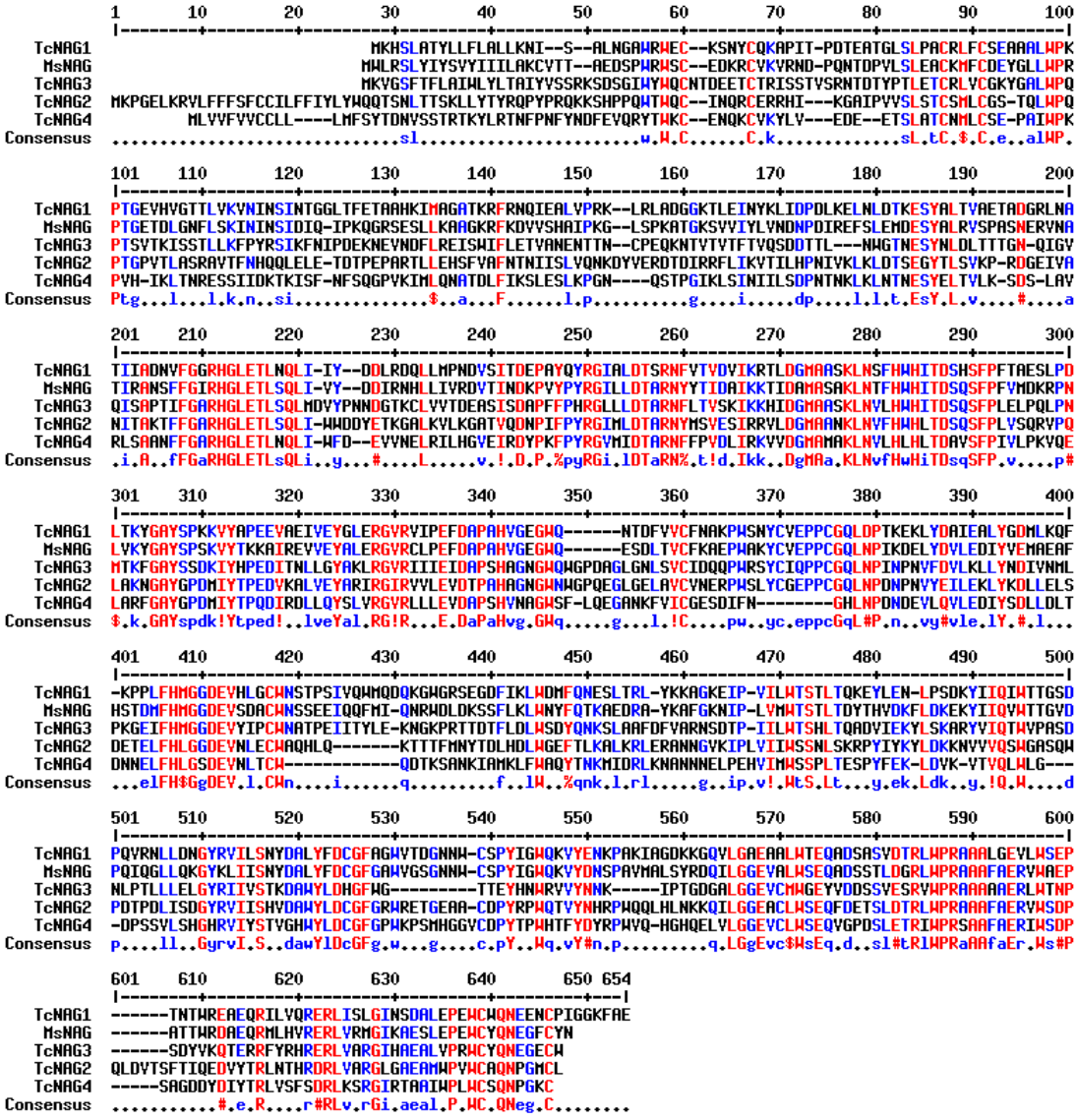
Boxes indicate exons and lines indicate introns. The location of the start codons and stop codons are represented by open and shaded triangles, respectively.

Figure 17. Phylogenetic analysis of NAGs, hexosaminidases, and ENGs in *Tribolium*, other insects, and metazoans.



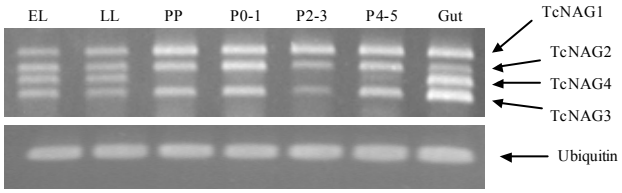
MEGA 3.0 (Kumar et al., 2004) was used to construct the consensus phylogenetic tree using UPGMA. Bootstrap analysis of 5000 replications are shown. Protein sequences extracted from GenBank include: MsNAG, *Manduca sexta* (AY368703); BmHEXO, *Bombyx mori* (S77548); TnHEXO, *Trichoplusia ni* (AY078172); DmHEXO1, *Drosophila melanogaster* (NM_079200); AmFDL, *Apis mellifera* (XP_394963.1); DmFDL, *Drosophila melanogaster* (NP_725178); AgFDL, *Anopheles gambiae* (XP_308677.2); Dm HEXO2, *Drosophila melanogaster* (NM_080342); HsHEXA, *Homo sapiens* (NM_000520); MmHEXA, *Mus musculus* (BC010755); HsHEXB, *Homo sapiens* (NP_000512); MmHEXB, *Mus musculus* (NM_010422); HsENG, *Homo sapiens* (AAM80487). *Tribolium castaneum* (Tc) NAG, HEX, and ENG sequences were deduced from their corresponding cDNA sequences.

Figure 18. Alignment of amino acid sequences of the deduced *Tribolium castaneum* and *Manduca sexta* NAGs.



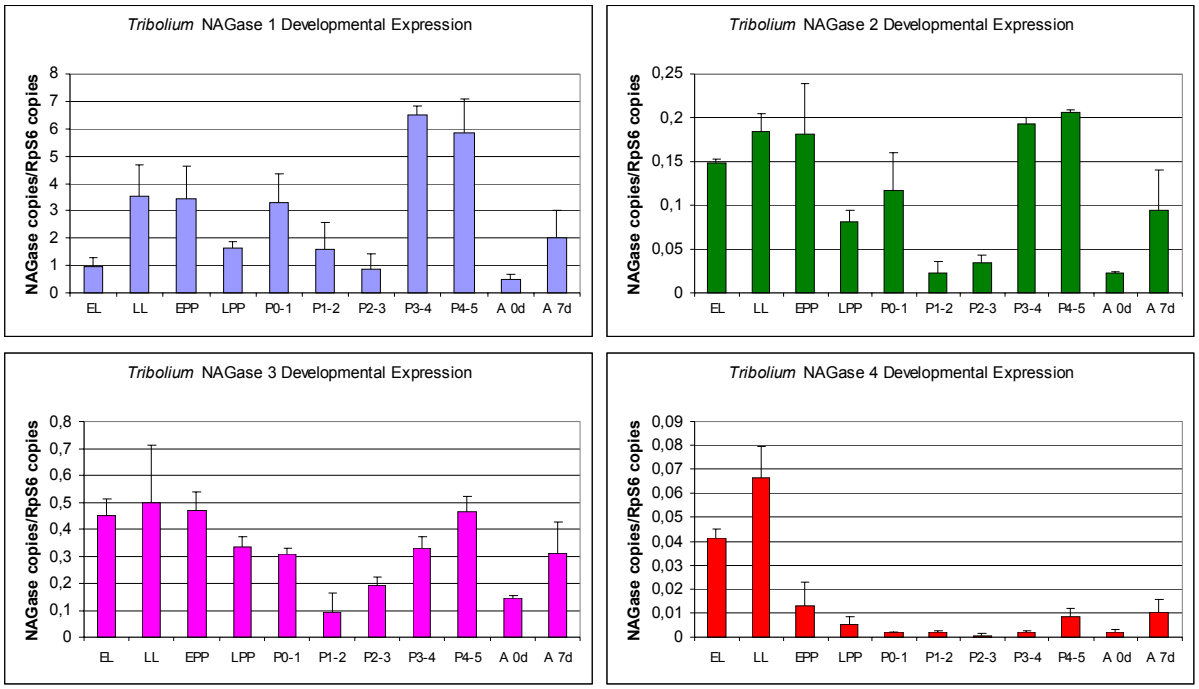
Amino acid sequences were aligned using Multalin (http://prodes.toulouse.inra.fr/multalin/multalin.html, Corpet, 1988). Highly conserved residues are shown in red and conserved residues in blue. Gaps in the alignment of sequences are denoted as dashes.

Figure 19. Developmental pattern of expression of *Tribolium* NAGs and expression in the larval midgut by RT-PCR.



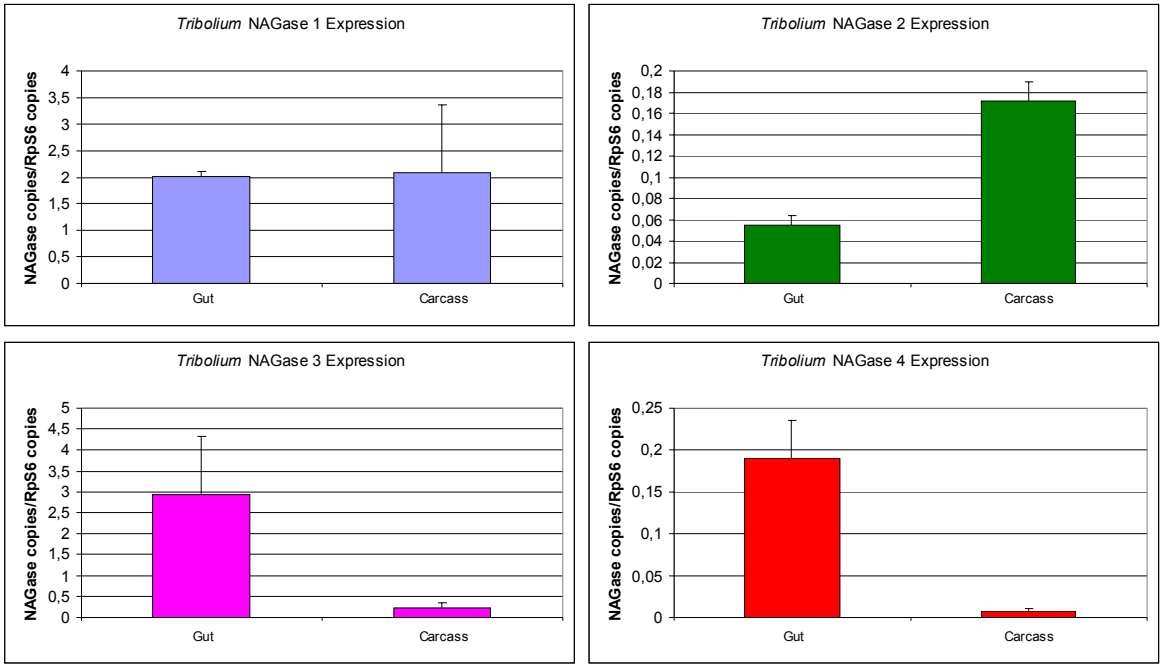
The developmental pattern of expression and the larval midgut expression of the four putative *Tribolium* NAGs was investigated by RT-PCR. The cDNAs used for PCR analysis were prepared from total RNA extracted from whole beetles at different developmental stages: EL, early larvae; LL, late larvae; PP, prepupae; P0-1, pupae days 0-1; P2-3, pupae days 2-3; P4-5, pupae days 4-5; or from the dissected midguts of late larvae. Four sets of gene-specific primers were used to study the relative expression of *TcNAG1*, *TcNAG2*, *TcNAG3*, and *TcNAG4*. These primers were designed to produce different size PCR products to allow for simultaneous analysis. For controls, primers for the constitutively expressed housekeeping gene, *Ubiquitin* were used.

Figure 20. Developmental pattern of expression of *Tribolium* NAGs by real-time RT-PCR.



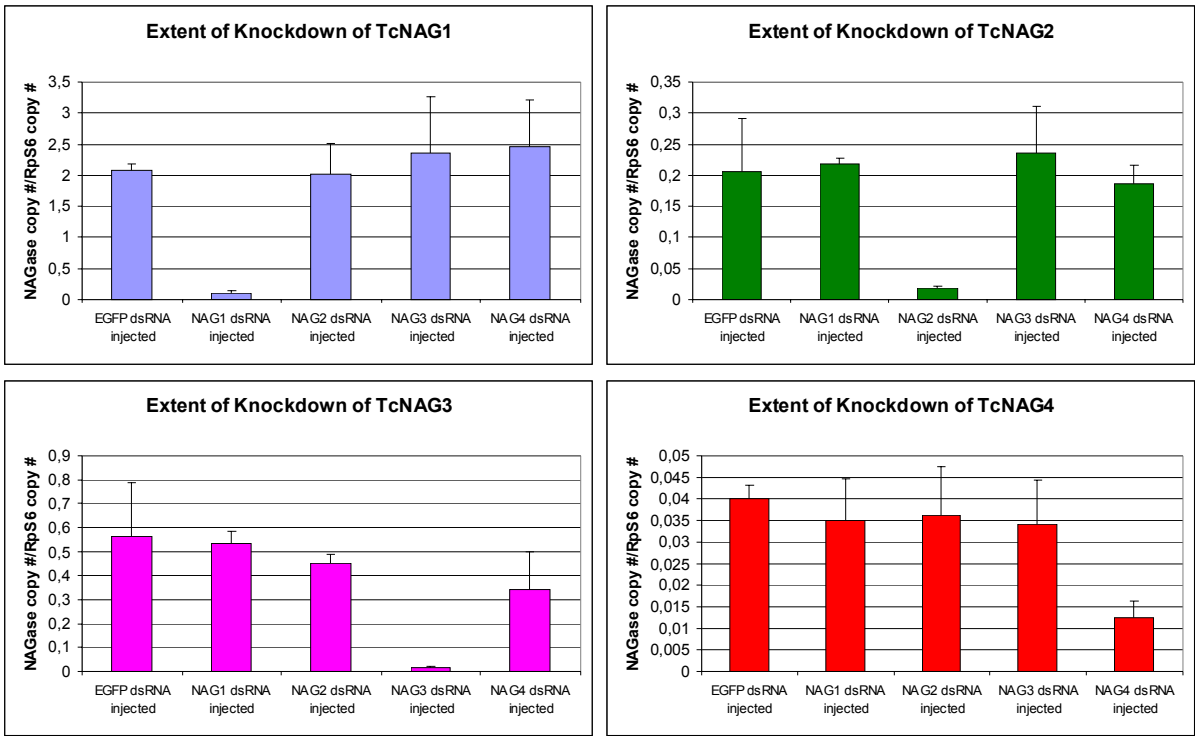
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from beetles at various developmental stages: EL, early larvae; LL, late larvae; EPP early prepupae; LPP, late prepupae; P0-1, pupae days 0-1; P1-2, pupae days 1-2; P2-3, pupae days 2-3; P3-4, pupae days 3-4; P4-5, pupae days 4-5; A 0d, adults day 0; A 7d, adults day 7. Transcript copies were determined from standard curves constructed from cloned NAG and *RpS6* constructs. NAG expression is presented relative to the internal housekeeping gene, *RpS6*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 21. Expression of *Tribolium* NAG genes in the larval midgut and carcass.



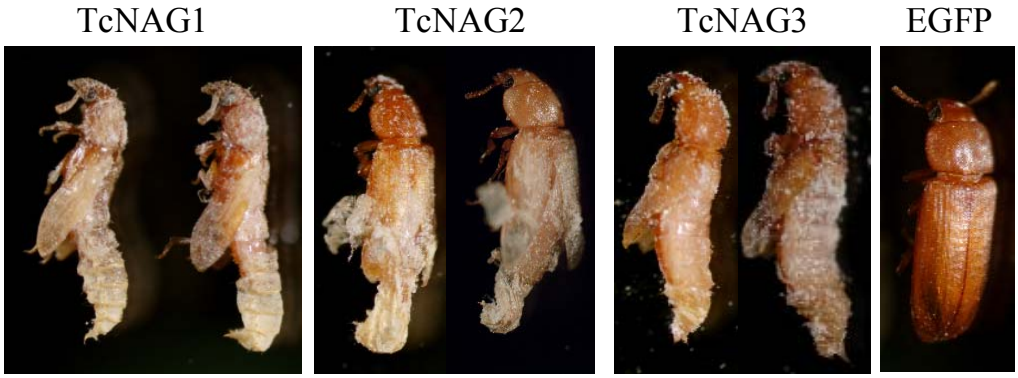
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from dissected larval midguts and carcasses with the midguts removed. Transcript copies were determined from standard curves constructed from cloned NAG and *RpS6* constructs. NAG expression is presented relative to the internal housekeeping gene, *RpS6*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 22. Extent of knockdown of *Tribolium* NAGs by RNAi.



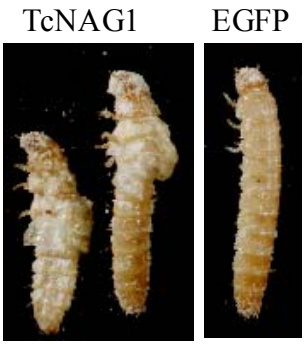
The cDNAs used for real-time RT-PCR analysis were prepared from total RNA extracted from larvae three days after injection with dsRNAs corresponding to the four putative *T. castaneum* NAGs and the Enhanced Green Fluorescent Protein (*EGFP*) for controls. Transcript copies were determined from standard curves constructed from cloned NAG and *RpS6* constructs. NAG expression is presented relative to the internal housekeeping gene, *RpS6*. Data are presented as the average of three separate experiments with error bars showing the standard deviation from the mean.

Figure 23. Phenotypes from treatment of *TcNAG1*, *TcNAG2*, and *TcNAG3* dsRNA on pupal-adult molting.



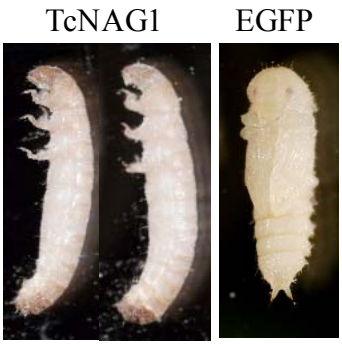
Double-stranded RNA (approximately 0.2 µg per insect) corresponding to *TcNAG1*, *TcNAG2*, or *TcNAG3* was injected into pharate pupae and EGFP dsRNA was injected as a control. The pupal to adult molt was arrested following specific knockdown of each NAG. The synthesis of adult cuticle proceeded without any obvious defect, but the insects were unable to shed their old pupal cuticle. Typical phenotypes obtained for over 75% of treated insects are depicted for each NAG.

Figure 24. Phenotypes from treatment of *TcNAG1* dsRNA on larval-larval molting.



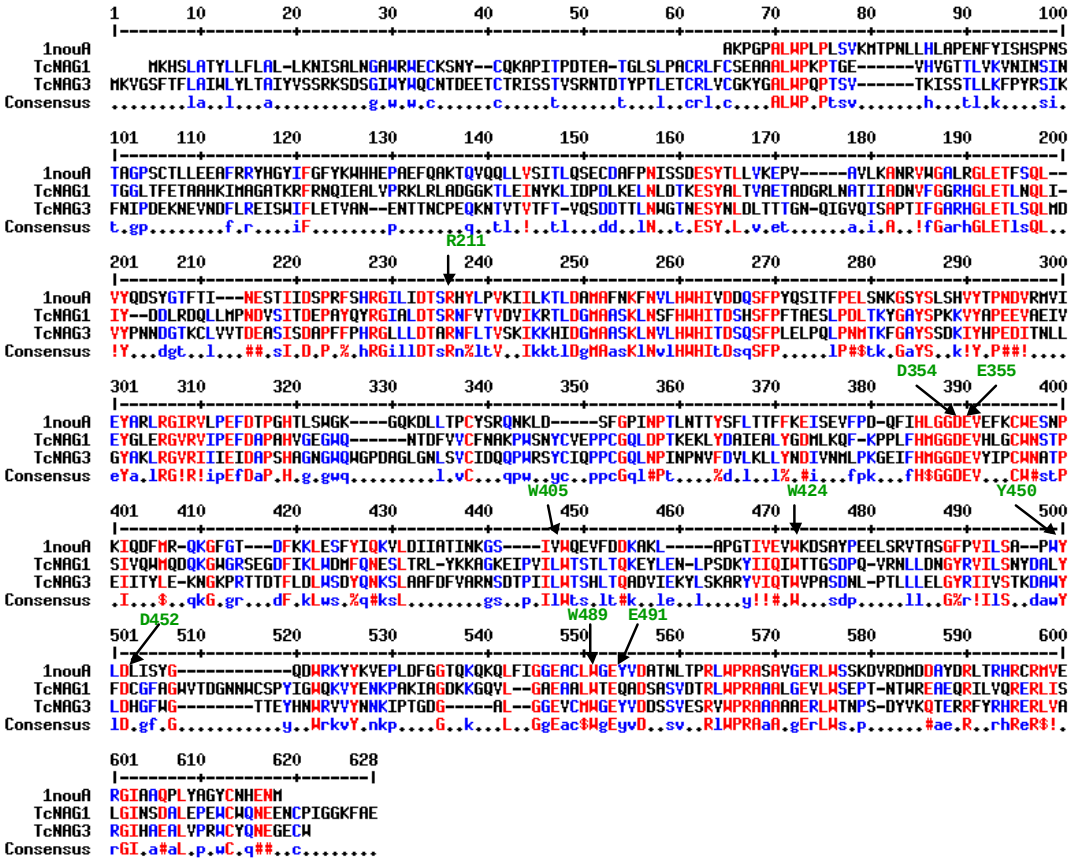
Double-stranded RNA (approximately 0.2 μg per insect) corresponding to *TcNAG1* was injected into late larvae as well as EGFP as a control. The larval to larval molt was arrested following specific knockdown of *TcNAG1*. The synthesis of new larval cuticle proceeded without any obvious defect, but the insects were unable to shed their old larval cuticle. Typical phenotypes for over 75% of *TcNAG1* dsRNA-treated beetles are shown.

Figure 25. Phenotypes from treatment of *TcNAG1* dsRNA on larval-pupal molting.



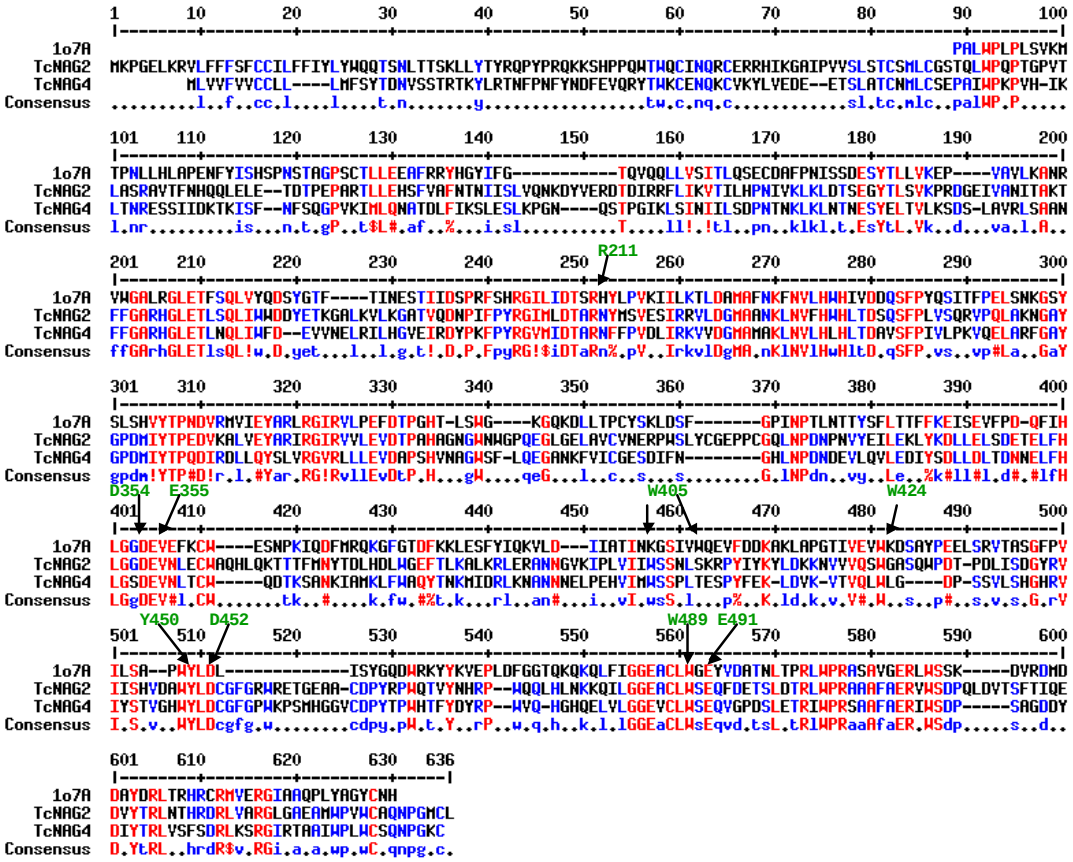
Double-stranded RNA (approximately 0.2 μg per insect) corresponding to *TcNAG1* was injected into last instar larvae. Double-stranded RNA for *EGFP* served as a negative control. The larval to pupal molt was arrested following specific knockdown of *TcNAG1* transcripts. The synthesis of new pupal cuticle proceeded underneath without any obvious defect, but the insects were unable to shed their exterior larval cuticle. Typical phenotypes obtained for >75% of insects treated with *TcNAG1* dsRNA are depicted.

Figure 26. Alignment of a human β-hexosaminidase B isoform (pdb 1NOU) with TcNAG1 and TcNAG3.



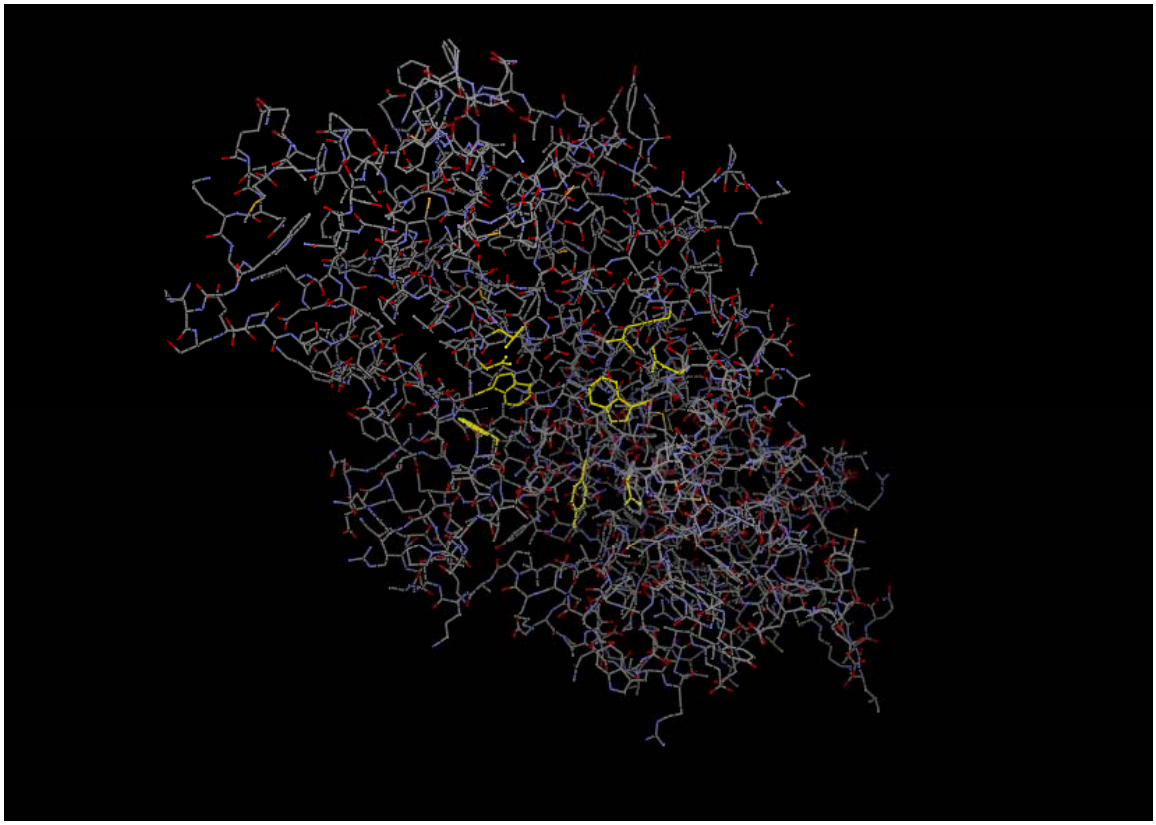
Amino acid sequences were aligned using Multalin (http://prodes.toulouse.inra.fr/multalin/multalin.html; Corpet, 1988). Highly conserved residues are shown in red and conserved residues in blue. Gaps in the alignment of sequences are denoted as dashes. The catalytically important amino acid residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of 1NOU are indicated in green at conserved positions in the alignment with TcNAG1 and TcNAG3.

Figure 27. Alignment of a human β-hexosaminidase B isoform (pdb 1NOU) with TcNAG2 and TcNAG4.



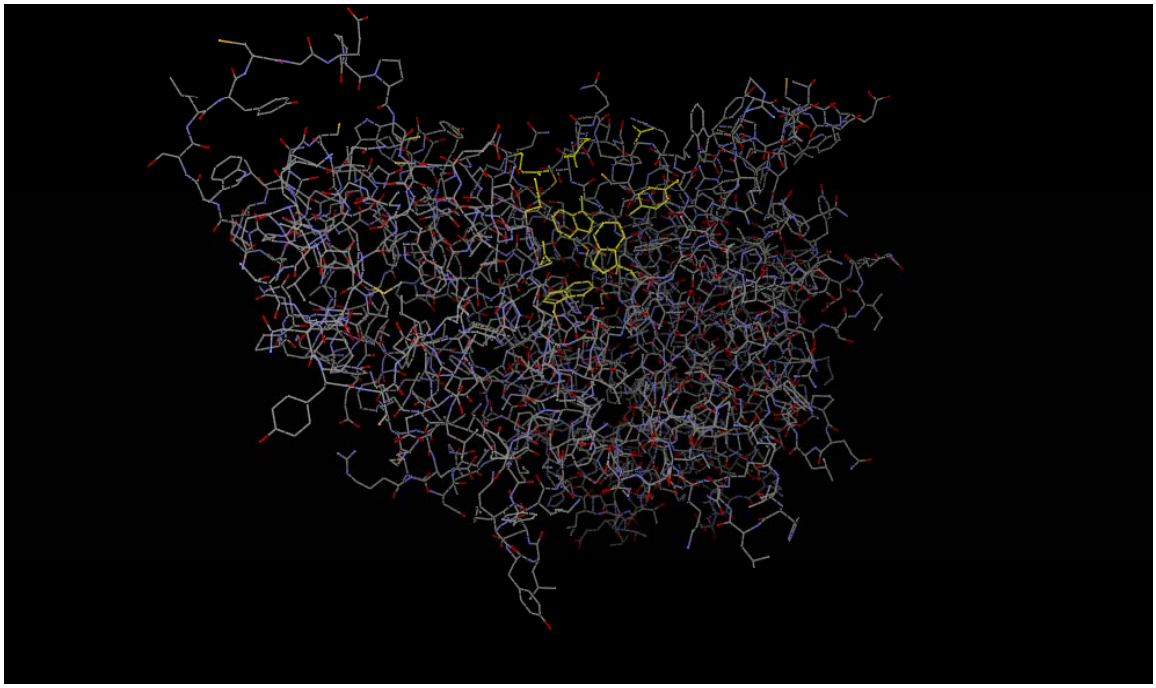
Amino acid sequences were aligned using Multalin (http://prodes.toulouse.inra.fr/multalin/multalin.html, Corpet, 1988). Highly conserved residues are shown in red and conserved residues in blue. Gaps in the alignment of sequences are denoted as dashes. The catalytically important amino acid residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of human *β*-hexosaminidase 1o7A are indicated in green at conserved positions in the alignment with TcNAG2 and TcNAG4.

Figure 28. Homology stick model of TcNAG1.



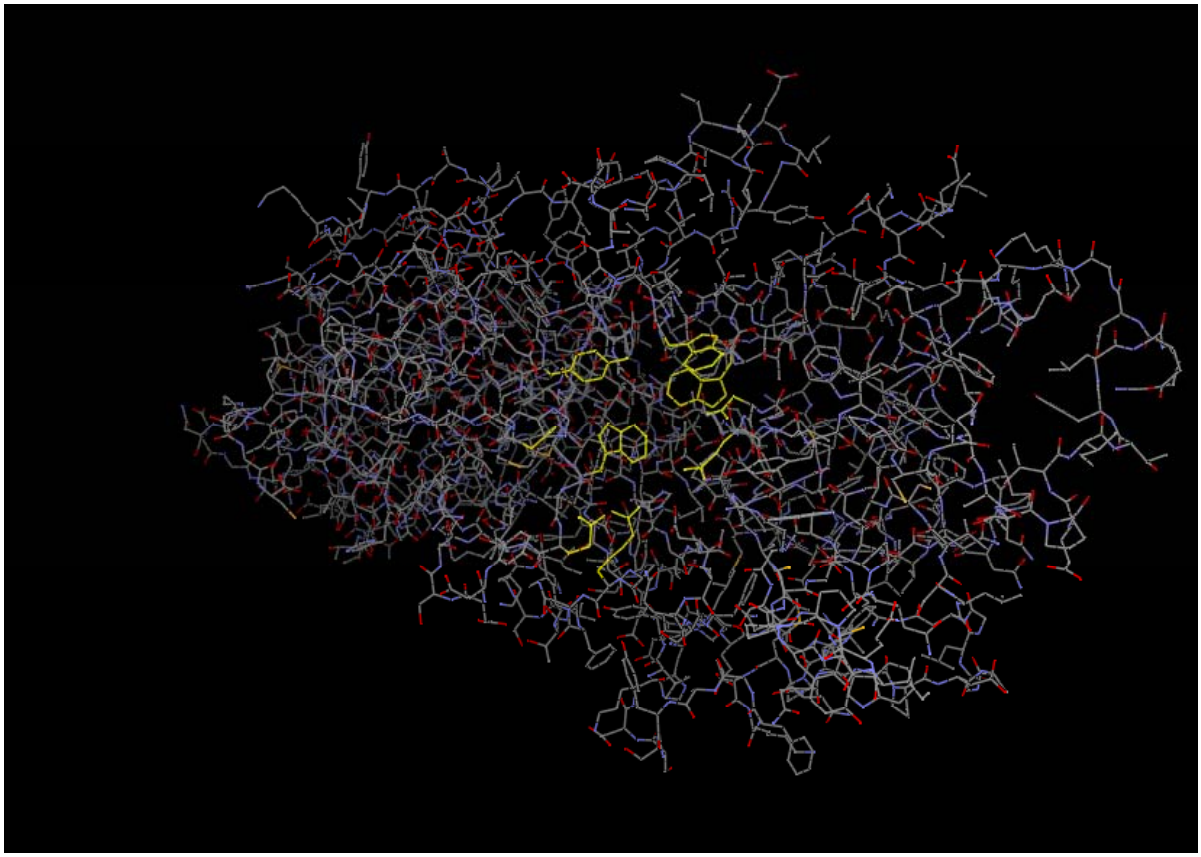
The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology model of TcNAG1 using the crystal structure of human β-hexosaminidase B as a template (pdb 1NOU, Mark et al., 2003). The catalytically important residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of human *β*-hexosaminidase 1NOU are highlighted in yellow.

Figure 29. Homology stick model of TcNAG2.



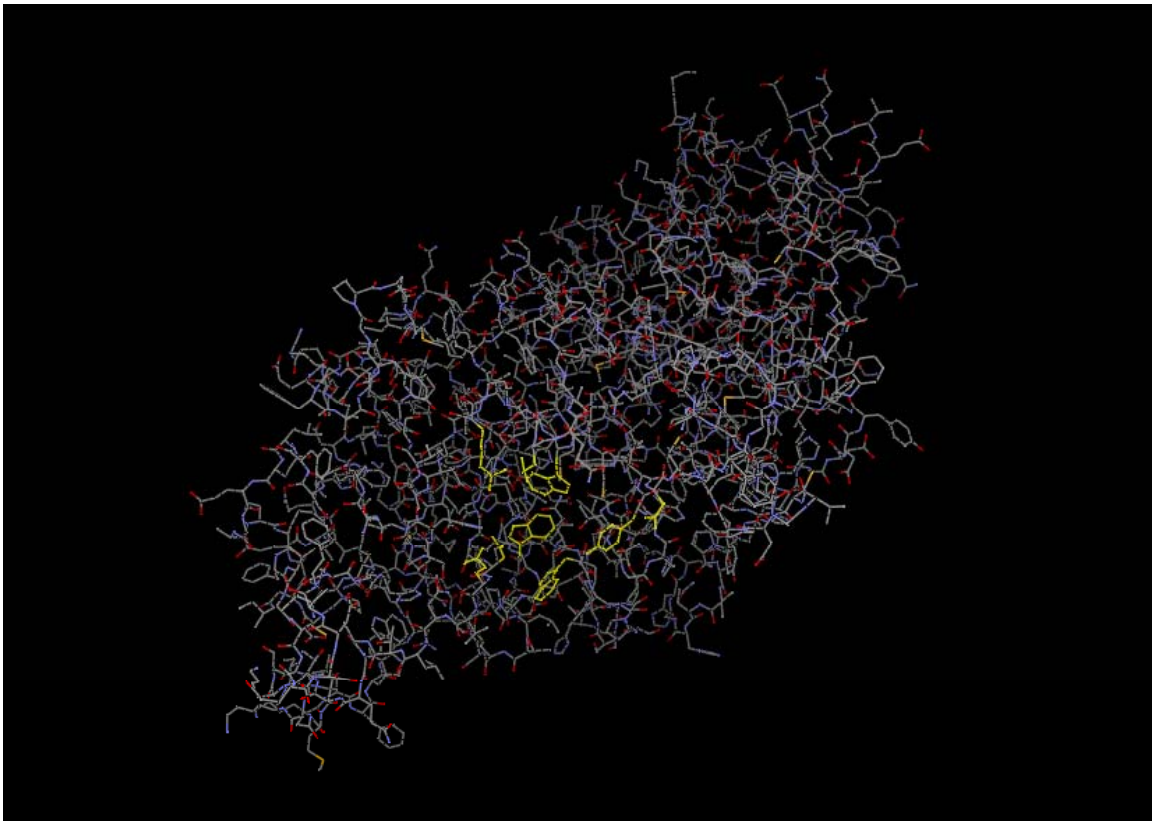
The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology model of TcNAG2 using the crystal structure of human β-hexosaminidase B as a template (pdb 1o7a, Maier et al., 2003). The catalytically important residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of human *β*-hexosaminidase 1o7a are highlighted in yellow.

Figure 30. Homology stick model of TcNAG3.



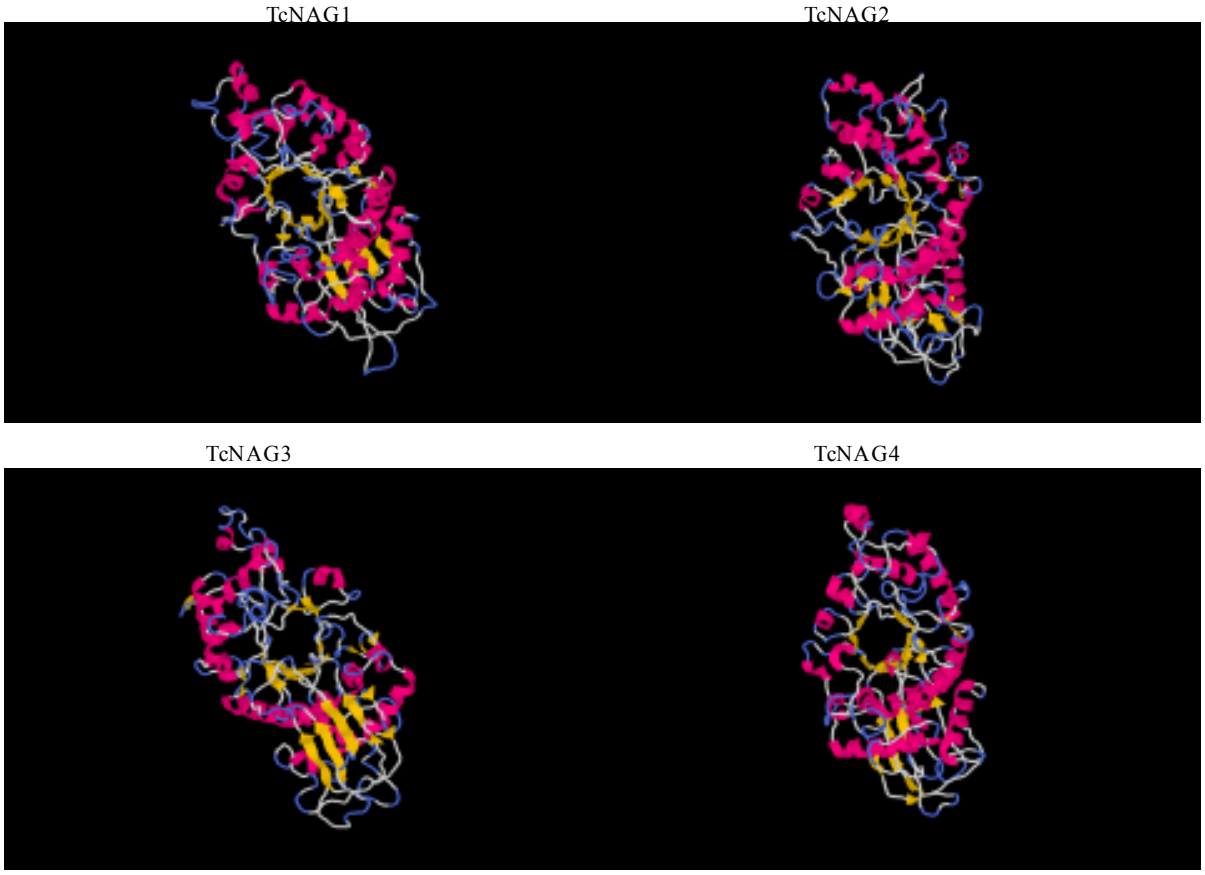
The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology model of TcNAG3 using the crystal structure of human β-hexosaminidase B as a template (pdb 1NOU, Mark et al., 2003). The catalytically important residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of human *β*-hexosaminidase 1NOU are highlighted in yellow.

Figure 31. Homology stick model of TcNAG4.



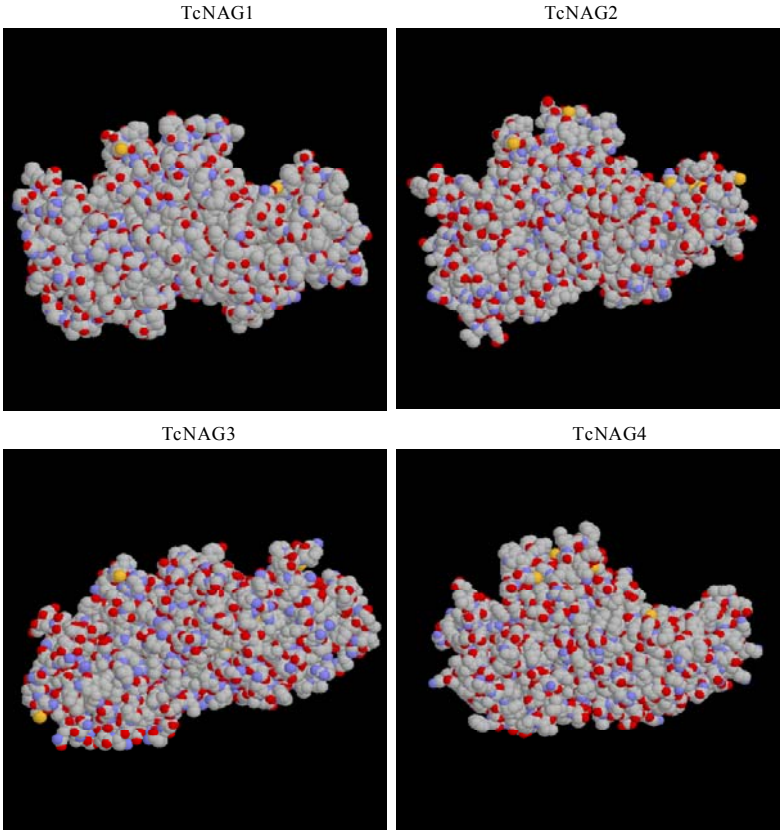
The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology model of TcNAG4 using the crystal structure of human β-hexosaminidase B as a template (pdb 1o7a, Maier et al., 2003). The catalytically important residues R211, D354, E355, W405, W424, Y450, D452, E491, and W489 in the active site of 1o7a are highlighted in yellow.

Figure 32. Homology models of the $(\beta,\alpha)_8$-barrel domains of *Tribolium* NAGs.



The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology model of the four *T. castaneum* NAGs. The crystal structure of human β-hexosaminidase B, 1NOU, was used as a template for TcNAG1 and TcNAG (Mark et al., 2003). The crystal structure of human β-hexosaminidase B, 1o7a, was used as a template for TcNAG2 and TcNAG4 (Maier et al., 2003). The $(\beta,\alpha)_8$-barrel is located near the center of each *T. castaneum* NAG.

Figure 33. Space filling homology models of TcNAG1, TcNAG2, TcNAG3, and TcNAG4 illustrating the clefts formed around the catalytic domain.



The Geno3D modeling program (http://geno3d-pbil.ibcp.fr, Combet et al., 2002) was used to generate the homology model of the four *T. castaneum* NAGs. The crystal structure of human β-hexosaminidase B, 1NOU, was used as a template for TcNAG1 and TcNAG (Mark et al., 2003). The crystal structure of human β-hexosaminidase B, 1o7a, was used as a template for TcNAG2 and TcNAG4 (Maier et al., 2003). The clefts formed around the catalytic domains in each of the four NAGs is located in the top right corner of the models.

DISCUSSION

An extensive study of *β-N*-acetylglucosaminidases (NAGs) of *Tribolium castaneum*, the red flour beetle, has been carried out. A comprehensive search of the *Tribolium* genome database led to the identification of four putative NAGs and three closely related *β*-hexosaminidases in addition to a single endo- *β-N*-acetylglucosaminidase (ENG). The three genes encoding proteins from the closely related CAZy (Carbohydrate-Active enzymes) glycoside hydrolase family 20 (Family GH20) *β*-hexosaminidases were identified through the expansion of the genome database search to include sequences with lower similarity. The human orthologs of these proteins have been extensively studied and deficiencies in these enzymes have been implicated in Tay-Sachs disease (Filho & Shapiro, 2004) and the less prevalent Sandhoff's disease (Sandhoff et al., 1968). The identification of other closely related Family GH20 *β*-hexosaminidases makes it unlikely that the search criteria employed were too narrow as to preclude the identification other potential NAGs in the *Tribolium* genome. Furthermore, because of the high genome coverage of the database (7.5X represented in the database), it is unlikely that regions of the genome containing other *T. castaneum* NAGs were not sequenced. Therefore, *T. castaneum* probably contains a total of four NAG genes.

Following the identification of the four putative NAGs in *T. castaneum*, preliminary gene annotation was carried out using sequence comparisons with previously characterized NAGs from other insect species as a guide. The possible splice sites and likely exon regions were readily identified in regions where a high degree of conservation exists between insect NAGs. Following the *in silico* determination of potential exons, pairs of gene-specific primers were designed from the exon regions of each putative NAG. In order to maximize the likelihood of successful prediction of exon sequences, these primers were designed around highly conserved regions of insect NAGs. Therefore, the likelihood of successful amplification of cDNAs by PCR was maximized. Consequently, cDNA clones encoding portions of all four putative NAGs were successfully amplified, cloned, and sequenced. Subsequently, 5'- and 3'-RACE (Rapid Amplification of cDNA Ends) was carried out in order to obtain the 5'- and 3'-ends of each *T. castaneum* NAG. Following cloning and sequencing of each of the 5'- and 3'-RACE PCR products, sets of gene-specific primers located at the 5'- and 3'-ends were designed for the amplification of the full-length cDNA of each NAG. Determination of the exon/intron

organization of the genes was then carried out by comparing the genomic sequences with those of the full-length cDNA sequences of each *T. castaneum* NAG (Fig. 16).

The amino acid sequences were deduced from the open reading frame of each full-length cDNA. These predicted protein sequences were then subjected to a battery of analyses including the prediction of signal peptides, transmembrane domains, as well as MWs and pIs (Table 7). In addition, several other prediction software programs were employed including coiled-coil domain prediction. However, none of the NAGs were predicted to have coiled-coil domains (data not shown). Based on SignalP 3.0 software, only one of the *T. castaneum* NAGs, TcNAG2, was predicted to contain a transmembrane domain, which was located near the N-terminus of the protein. The other three, TcNAG1, TcNAG3, and TcNAG4, were predicted to contain a signal peptide and are, therefore, probably secreted proteins. Analysis of the predicted properties of each NAG protein clearly shows that TcNAG2 stands out from the other NAGs. Firstly, it is the largest with a deduced amino acid sequence of over 600 residues and MW greater than 70 kDa. The other three NAGs are all less than 600 amino acid residues in length and have MWs of less than 70 kDa. Furthermore, the pI of TcNAG2 is greater than pH 6.0 in contrast to the other NAGs whose predicted pI's are all less than pH 6.0. Finally, the predicted signal anchor in TcNAG2 sets this protein apart from the other NAGs, making it a membrane-associated NAG as opposed to the others, which are predicted to be secreted proteins.

The active site of human β-hexosaminidase B contains several catalytically important residues including R211, D354, E355, W405, W424, Y450, D452, E491, and W489 (Maier et al., 2003). Accordingly, these nine residues are highly conserved in all four putative *Tribolium* NAGs. Furthermore, in the homology models of each *Tribolium* NAG, all nine of these residues are located in close proximity to each other. This is as expected from their proposed role in catalysis. The $(\beta,\alpha)_8$-barrel domain is a remarkably common fold in several structurally characterized enzymes. This domain was first reported for triose phosphate isomerase and was therefore called the TIM-barrel. The $(\beta,\alpha)_8$-barrel domain is well characterized in both human β-hexosaminidase B proteins, which were used as a template for the homology models of *Tribolium* NAGs. Accordingly, the $(\beta,\alpha)_8$-barrel domain was present in the homology models of all four *Tribolium* NAGs (Fig. 42).

Initial attempts to characterize the newly identified NAGs in *T. castaneum* included a phylogenetic analysis of these proteins with other previously characterized NAGs from insects

and other metazoa (Fig. 17). In addition, the three β-hexosaminidases and ENG identified in *T. castaneum* were included in the phylogenetic analysis. Using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) method, a phylogenetic tree was constructed. In order to evaluate the branch strength of the constructed tree, a bootstrap analysis with 5000 replications was performed. Furthermore, a number of different methods of inferring a phylogenetic tree were employed including neighbor-joining, maximum parsimony, and minimum evolution. The resulting phylogenetic trees were all consistent, reinforcing the robustness of the tree.

A total of five major groups were identified in the phylogenetic tree of NAGs using the UPGMA method: chitinolytic NAGs group I, chitinolytic NAGs Group II, *N*-glycan processing NAGs, hexosaminidases, and ENGs. Within each group, several subgroups can be categorized including both hexosaminidases, A and B, of *Homo sapiens* and *Mus musculus*. The *T. castaneum* hexosaminidases clearly group with these previously characterized metazoan hexosaminidases. Furthermore, the *T. castaneum* ENG is grouped with the human ENG with a bootstrap value of 100.

TcNAG1 falls within the chitinolytic NAG group I, which contains the well-characterized NAG from *M. sexta*, which was used in the blast search to identify potential NAGs in *T. castaneum*. All of the group I chitinolytic NAGs are predicted (SignalP 3.0) to be secreted proteins with a signal peptide, which is consistent with their proposed role in chitin degradation. In fact, MsNAG was found in the molting fluid of pharate pupae (Zen et al., 1996). Similarly, BmHEXO protein was purified from the integument tissue during the larval to pupal transformation (Nagamatsu et al., 1995). Furthermore, functional analysis of the purified enzymes from these lepidopterans has firmly established their role in degradation of chitin oligosaccharides. Finally, following the expression of another group I chitinolytic NAG, DmHEXO1, in *P. pastoris*, the purified recombinant enzyme displayed chitotriose-degrading activity (Léonard et al., 2006). Taken together, TcNAG1 is most probably the principal chitinolytic NAG in *T. castaneum*.

The expression pattern of *TcNAG1* is consistent with its role as a chitinolytic NAG. The level of expression of this putative NAG peaks during the larval-pupal and pupal-adult molts. The number of transcripts drops in the middle of the pupal stage when the need for this enzyme is expected to subside. The expression data using both RT-PCR (Fig. 19) and quantitative PCR

(Fig. 30) show similar patterns of expression. Furthermore, the relative level of expression of this NAG is significantly higher than those of the other NAGs (up to 100 fold), indicating that this is the primary NAG in *T. castaneum*.

The most compelling evidence suggesting that *TcNAG1* is important in chitin catabolism during the molt is that lethal phenotypes are obtained following specific knockdown of this transcript by RNA interference (RNAi) in *T. castaneum* larvae, pharate pupae, and pharate adults. Administration of dsRNA corresponding to a unique region of the *TcNAG1* transcript resulted in the selective knockdown of this transcript by greater than 90% in late larvae (Fig. 32). Similar decreased expression levels were also observed in dsRNA-injected pupae (data not shown). However, dsRNAs for the other *T. castaneum* NAGs did not significantly reduce transcript levels for *TcNAG*1, indicating highly specific knockdown of this gene. The administration of *TcNAG1* dsRNA at developmental stages prior to the larval-larval, larval-pupal, and pupal-adult molts resulted in the arrest and subsequent death of the insects at all three molts. Therefore, TcNAG1 is essential for all three of these molts and knockdown of this transcript alone is sufficient to arrest ecdysis and cause mortality. Furthermore, analysis of the terminal phenotypes resulting for *TcNAG1* knockdown is consistent with it role in chitin turnover in the cuticle. During the larval-larval molt, the insects were unable to completely shed their exoskeleton. Old cuticle can be seen still attached to the larvae (Fig. 34). In the case of larval-pupal molting, the pupae were unable to shed the old larval cuticle. The phenotypes observed in pharate adults were clearly the most obvious (Fig. 33). The cuticle of these adult insects was synthesized without any obvious defect, but they were unable to completely shed their old pupal cuticle. The terminal phenotypes of these insects show a large quantity of incompletely digested pupal cuticle over the entire body of the insect.

The *TcNAG1* knockdown phenotype is strikingly similar to that obtained from the knockdown of *T. castaneum* chitinase 5 (*TcCHI5*), in which the animals failed to shed their old cuticles with the new cuticles visible underneath (Zhu, 2005). In *M. sexta*, CHI and NAG act in concert in the degradation of chitin and chitin oligosaccharides (Fukamizo & Kramer, 1985a, b). Combination of NAG with CHI resulted in a synergistic effect on the degradation of chitin. The rate of chitinolysis was increased as much as six times compared to that of the individual enzymes alone. This synergistic effect was also shown to be highly dependent on the concentration ratio of these two enzymes, suggesting that alterations in the proportions of these

two enzymes could lead to inefficient chitin degradation. Indeed, the net effect of knocking down *TcNAG1* was a disruption in the binary enzyme system, resulting in inefficient chitin degradation. Furthermore, it has been shown that in *M. sexta*, CHI is susceptible to substrate oligosaccharide inhibition (Koga et al., 1982; Koga et al., 1983). Knockdown of NAG may result in the accumulation of high levels of chitin oligomers in the molting fluid. Consequently, the knockdown of NAG may inhibit CHI activity, resulting in a phenotype similar to the CHI knockdown beetles. Based on the high level of expression, the phylogenetic relationship, and the phenotypic effect of knocking down *TcNAG1*, we conclude that this NAG is the primary enzyme responsible for the degradation of cuticular chitin in concert with CHI in *T. castaneum*.

Although TcNAG1 is the principal chitinolytic NAG, the other three putative NAGs identified in *T. castaneum* also appear to play important roles in cuticle turnover. Phylogenetic analysis groups TcNAG3 with DmHEXO2 as part of the Group II chitinolytic NAGs. Recombinant DmHEXO2 showed chitotriose-degrading activity, indicating that this NAG is probably involved in chitin degradation (Léonard et al., 2006). The bootstrap value between TcNAG3 and DmHEXO2 was 99, indicating a very close phylogenetic relationship between the two proteins (Fig. 17). Further evidence that TcNAG3 plays a role in chitin degradation was obtained in transcript knockdown studies using RNAi. Following the administration of dsRNA, terminal phenotypes were obtained in >75% of treated insects with the pharate adults unable to fully eclose from the pupal cuticle (Fig. 33). This observed phenotype is highly similar to that obtained for TcNAG1 knockdown studies, which indicate that new adult cuticle was formed underneath the incompletely digested old pupal cuticle. Real-time RT-PCR analysis of NAG expression indicated that the level of *TcNAG1* expression was unaffected in the *TcNAG3* knockdown beetles, indicating that the phenotype observed after dsRNA for TcNAG3 injection was not a result of *TcNAG1* expression being affected. Furthermore, larval-larval and larval-pupal molts were also arrested but less than 25% of the treated insects exhibited a phenotype (data not shown). Perhaps the insects are less sensitive to reductions in the levels of TcNAG3 than to the reduction in the levels of TcNAG1 during earlier developmental stages. The developmental expression pattern of the two enzymes, however, is remarkably similar with high levels of expression around the time of the pupal and adult molts as well as larval stages. Relative to *TcRpS6*, the level of *TcNAG3* expression is approximately 10-fold lower than that of *TcNAG1*, corroborating the hypothesis that *TcNAG1* is the principal chitinolytic NAG in *T.*

castaneum. Another interesting finding from the expression studies of these two NAGs involved their expression in different tissues including the midgut. When the relative levels of expression of both genes were analyzed in dissected larval midguts and the carcass (larval abdomen with midgut removed), a significantly higher ratio of expression in the midgut to carcass tissues was observed for *TcNAG3* (Fig. 31). In contrast, no significant difference in the expression levels of *TcNAG1* was observed in these two tissue preparations. Perhaps *TcNAG3* also plays a critical role in the degradation of chitin in the peritrophic membrane (PM). However, given the high level of *TcNAG1* expression in the midgut, the role of this NAG in PM-associated chitin cannot be dismissed.

Similar to *TcNAG3*, *TcNAG4* was also expressed at a significantly higher level in the midgut than the larval carcass (Fig. 32). Furthermore, an analysis of the developmental pattern of expression of *TcNAG4* indicated that it is primarily expressed during the larval stages (Fig. 19). Unlike the other three putative NAGs, down-regulation of *TcNAG4* transcripts did not consistently result in lethal phenotypes and the majority of dsRNA-injected insects survived to adulthood with no visible ill effects on growth or metamorphosis. However, a small number of individuals (approximately 20%) did exhibit a lethal larval phenotype similar to that of *TcNAG1* knockdown (data not shown). In addition, a small number (approximately 10%) of *TcNAG4* knockdown insects exhibited a lethal pharate adult phenotype, and these insects were unable to fully shed their old pupal cuticle, just as in *TcNAG1* and *TcNAG3* knockdown beetles (data not shown). Analysis of the developmental pattern of expression of *TcNAG4* by quantitative PCR showed a slight increase in the number of transcripts close to the time of pupal-adult molting (Fig. 30). Therefore, this NAG may play a role, albeit not as significant as the other NAGs, in the degradation of cuticular chitin. The relative expression of *TcNAG4* is approximately 100-fold less than that of *TcNAG1*, which is consistent with *TcNAG4* not being as vital as *TcNAG1* for molting. The phylogenetic analysis of TcNAG4 failed to group this protein with any of the other insect NAGs (Fig. 17). This may reflect a lack of an adequate number of other insect NAGs available for the analysis. In *D. melanogaster*, only three putative NAGs have been identified. Because four NAGs have been identified in *T. castaneum*, it is difficult to assign the fourth NAG to the corresponding ortholog in *D. melanogaster*. Therefore, *TcNAG4* appears to be a unique NAG and its relatively high expression in the midgut vs. the carcass suggests that it may be more specialized in the turnover of PM-associated chitin than cuticular chitin.

Through this work, it is clear that *TcNAG2* also appears to be a unique NAG in several respects. TcNAG2 is composed of a larger number of amino acids (630) than the other NAGs, which are all less than 600 residues long. The predicted MW of TcNAG2, is 72.6 kDa in contrast to the other NAGs, which are all less than 68 kDa. Furthermore, the pI of TcNAG2 is predicted to be over pH 6.5 whereas the other three NAGs are predicted to have a pI around pH 5.5. However, the most significant difference in TcNAG2 is the predicted signal anchor at the N-terminus of the protein. The other three NAGs are all predicted to have a signal peptide and are therefore probably secreted. The developmental expression profile of *TcNAG2* is very similar to that of *TcNAG1*, although the level of expression relative to *RpS6* is approximately 30-fold lower than *TcNAG1* (Fig. 30). Furthermore, the size of the *TcNAG2* gene is considerably larger than the other three NAGs with an extremely large intron (>13 kb) located just upstream of the putative translation start site (Fig. 16).

TcNAG2 is unique in this respect and the phylogenetic analysis indicates that it is most closely related to the homologs of the *D. melanogaster fused lobes* protein, DmFDL. Both proteins are also predicted to be a membrane-anchored with a single TMS located near the *N*-terminus of the polypeptide. Furthermore, ultracentrifugation experiments on a lepidopteran *fused lobes* homolog from Sf21 cells indicated that the majority of NAG activity resided in the membrane fraction although some activity was observed in the soluble fraction as well (Altmann et al., 1995). This lepidopteran NAG was capable of effectively hydrolyzing chitotriose-PA (pyridylamino). In contrast, the recombinantly expressed DmFDL was unable to digest chitotriose (Léonard et al., 2006). Instead, the enzyme hydrolyzed only the GlcNAc residue attached to the α 1,3-linked mannose of the core pentasaccharide of *N*-glycans. No cleavage activity of any other GlcNAc residues was observed including the GlcNAc residue attached to the α 1,6-linked mannose of the core pentasaccharide. Furthermore, DmFDL did not catalyze the endohydrolysis of the *N*, *N'*-diacetylchitobiosyl unit in the high-mannose pentasaccharide core, which is catalyzed by ENGs. Similar *N*-glycan substrate specificity for the terminal GlcNAc attached to the α 1,3-linked mannose was observed in the lepidopteran FDL from several cell lines including Sf21, Bm-N, and Mb-0503 (Altmann et al., 1995). Given the similarity of TcNAG2 to the other previously characterized homologs of the *D. melanogaster fused lobes* protein, TcNAG2 may serve a similar function in *N*-glycan processing. However, since the lepidopteran homolog of DmFDL catalyzed the hydrolysis of chitotriose, TcNAG2 may also be

involved in chitin catabolism as well. Indeed, when TcNAG2 transcripts were selectively knocked down by RNAi, the resulting pharate adults were unable to completely shed their old pupal cuticle, particularly at the posterior end (Fig. 33). These phenotypes imply a possible involvement of TcNAG2 in chitin degradation. Unlike RNAi involving *TcNAG1* dsRNA, which resulted in lethal phenotypes in most of the treated insects, only a small percentage (<20%) exhibited lethal terminal phenotypes similar to those observed for *TcNAG1* knockdown insects at the larval-larval and larval-pupal molts (data not shown). The reason for this lower number of phenotypic individuals is unknown.

Analysis of the expression of *TcNAG2* by real-time RT-PCR indicated that this transcript is expressed at a lower level in the midgut compared to the carcass (Fig. 31). Therefore, the lethal phenotype at the pharate adult stage may be a direct result of the knockdown of this transcript in the cuticular epidermal cells. At ecdysis, an ecdysial membrane is formed which separates the newly synthesized cuticle and epidermal cells from the molting fluid (Locke & Krishnan, 1973). If TcNAG2 does in fact play a role in chitin turnover in the cuticle, then the protein should be secreted and not be membrane-bound as predicted by SignalP 3.0. One possible explanation is that this protein is cleaved by an unknown protease, which releases the NAG into the molting fluid.

The observed phenotypes resulting from the knockdown of *TcNAG2*, as well as the other NAGs, may not be a result of a disruption of chitin catabolism at all but instead a disruption of another cellular process which manifests itself during the molt and produces the observed phenotypes. It was interesting to find that TcNAG1 was unable to compensate for the other NAGs when they were knocked down by injection of dsRNA. One possible explanation for this is the formation of multimeric structures such as heterodimers. In the case of human hexosaminidases, heterodimers are known to form between HEXA and HEXB and this dimerization is required for catalytic activity (Maier et al., 2003). Dimerization of NAG in *B. mori* is known to occur (Kimura, 1976). Furthermore, a NAG dimer was previously characterized in *T. castaneum* and it is likely that this enzyme is composed of nonidentical subunits (Kramer & Aoki, 1987). Perhaps the formation of *T. castaneum* NAG heterodimers is required for catalytic activity, as is the case in human hexosaminidases. Another possibility is that the *T. castaneum* NAGs do not serve redundant functions. There may be a degree of specificity for the four NAGs for different chitooligosaccharides substrates. One may be

particularly catalytically active on hydrolyzing chitotriose, whereas another may specialize in the degradation of the tetramer and so on. Therefore, a knockdown of one of the NAGs results in the accumulation of a particular chitooligosaccharide. This could in turn lead to substrate inhibition of CHI in the binary enzyme chitinase system. Further research in this area is required to more definitively address this issue.

CONCLUSIONS

An extensive study of two of the major genes involved in chitin metabolism, chitin synthases (CHSs) and *β-N*-acetylglusosaminidases (NAGs), has been carried out. Two model insect systems have been used in these studies; the tobacco hornworm, *M. sexta*; and the red flour beetle, *T. castaneum*. The use of each model species provided unique sets of advantages for their use in the study of these two classes of chitin metabolism enzymes. The large size and well-characterized life cycle of *M. sexta* allowed for the straightforward analysis of CHS gene expression at various developmental stages. Furthermore, the shear mass of this insect facilitated the dissection of various tissues for tissue-specific CHS expression analysis. *T. castaneum*, however, provided distinct advantages over *M. sexta* for the study of NAGs. *T. castaneum* is rapidly becoming a sophisticated model species for coleopteran representing the largest order in insects, Coleoptera, as a result of the recent genome sequencing. This significant achievement has enabled the identification of all potential NAG genes in the *T. castaneum* genome. Furthermore, the exquisite amenability of this insect to knockdown studies using RNA interference (RNAi) has allowed for the functional analysis of each gene at various developmental stages. In this work, the unique characteristics of *M. sexta* and *T. castaneum* have been exploited in the extensive study of CHS and NAG genes, respectively.

The characterization of the two CHS genes in *M. sexta* has led to several important findings. Firstly, Southern blot analysis clearly demonstrates, in accord with other insect species, that there are only two CHS genes in *M. sexta*, one belonging to class A and one to class B. Secondly, using various methods of gene expression analysis including RT-PCR, northern blots, and quantitative PCR, it was determined that the class B CHS (encoded by *MsCHS2*) is expressed in an endodermal-derived tissue, the midgut. In contrast, the class A CHS (encoded by *MsCHS1*) is expressed in various ectodermal-derived tissues including the integument, tracheae, foregut, and hindgut. These findings suggest that the two CHS genes in *M. sexta*, *MsCHS2* and *MsCHS1*, are specialized for the synthesis of chitin in the midgut peritrophic membrane and epidermal cuticle, respectively. Furthermore, these studies have identified the presence of two isoforms of *MsCHS1*, *MsCHS1a* and *MsCHS1b*, which arise from the alternative splicing of transcripts of this class A CHS gene. Analysis of the relative levels of expression of

both alternatively spliced transcripts indicates a tissue-specific preferential accumulation of the *MsCHS1a* transcript in the integument during the feeding and pupal stages, and an excess of the *MsCHS1b* transcript in the tracheae and along the length of the digestive system during all developmental stages tested. The characterization of CHS genes in *M. sexta*, accompanied by the extensive expression studies has significantly advanced our understanding of chitin synthesis in this model insect.

All putative NAGs in *T. castaneum* have been identified and extensively studied. Four putative NAG genes have been fully characterized and their developmental pattern of expression has been analyzed. Furthermore, the functional analysis of each NAG has been studied by knocking down each transcript at various developmental stages through the injection of gene-specific dsRNAs. Selective knockdown of each transcript, without significant disruption in the expression level of the other NAGs, was confirmed and the resulting phenotypes were documented. From these studies, it is clear that *TcNAG1* is certainly the most prominent in terms of the relative level of expression and functional importance. Overall, this NAG is expressed at the highest level throughout all developmental stages. Accordingly, selective knockdown of this transcript results in the most severe lethal phenotypic effects, which are characterized by an incomplete digestion of the cuticle and developmental arrest at the time of molting. The other three NAGs, *TcNAG2*, *TcNAG3*, and *TcNAG4*, appear to play less important biochemical roles because knockdown of these transcripts results in the disruption of the larval-larval and larval-pupal molts in only a minority of dsRNA injected animals. Therefore, the investigation of all putative NAGs in *T. castaneum* has led to several important findings in terms of the expression and functional relevance of these genes.

In conclusion, the examination of all putative CHS and NAG genes in two model insects has lead to an increased understanding of the biochemical processes behind the synthesis and breakdown of chitin in insects.

SUGGESTIONS FOR FUTURE WORK

1. *In situ* hybridization and immunolocalization studies of CHSs in *M. sexta*. These experiments would more definitively address the tissue-specific expression of CHSs in this insect. Using probes specific for each alternately spliced transcript of *MsCHS1*, or monoclonal antibodies specific for MsCHS1a and MsCHS1b, determination if epidermal cells in the integument utilize MsCHS1b could be carried out.

2. Expression and purification of *M. sexta* CHSs and determine if they are enzymatically active. A comparison of the catalytic activities and the effect of various inhibitors could be determined. Furthermore, analysis of the different properties of the two isoforms, MsCHS1a and MsCHS1b, could be determined including protein-protein interactions.

3. Expression and purification of *T. castaneum* NAGs and determination if they are enzymatically active. The properties of each enzyme could be studied and the ability to form heterodimers could be determined. Furthermore, the substrate specificity and catalytic efficiencies of each NAG may provide valuable insight into the role these enzymes play in chitin degradation.

4. Tissue-specific expression studies of the NAGs in *T. castaneum*. This work could provide insight into the functional specialization of these enzymes.

5. Characterize the other enzymes involved in the chitin metabolism pathway. Study their expression throughout development and in various tissues.

6. Use RNAi to selectively knock down the other genes involved in the chitin metabolism pathway. Document the resulting phenotypes from the RNAi experiments.

7. Conduct a more detailed pathology of the phenotypes resulting from the knock down of *T. castaneum* NAGs. Study possible effects on various tissues including the midgut.

8. Characterize the ENG and hexosaminidases identified in *T. castaneum*. Study their developmental expression as well as their expression in various tissues. Use RNAi to knock down these genes and document the resulting phenotypes.

REFERENCES

Altmann, F., Schwihla, H., Staudacher, E., Glössl, J., & März, L. (1995). Insect cells contain an unusual, membrane-bound β-*N*-acetylglucosaminidase probably involved in the processing of protein *N*-glycans. The Journal of Biological Chemistry, 270, 17344-17349.

Anderson, S.O. (2005). Cuticular sclerotization and tanning. In: Gilbert, L., I., Iatrou, K., & and Gill, S. eds., Comprehensive Molecular Insect Science. Vol. 4, Biochemistry and Molecular Biology, Chapter 4. Elsevier Press, Oxford, UK., 145-170.

Arakane, Y., Hogenkamp, D. G., Zhu, Y. C., Kramer, K. J., Specht, C. A., Beeman, R. W., Kanost, M. R., & Muthukrishnan, S. (2004). Characterization of two chitin synthase genes of the red flour beetle, *Tribolium castaneum*, and alternate exon usage in one of the genes during development. Insect Biochemistry and Molecular Biology, 34, 291-304.

Arakane, Y., Muthukrishnan, S., Beeman, R. W., Kanost, M. R., & Kramer, K. J. (2005). *Laccase 2* is the phenoloxidase gene required for beetle cuticle tanning. Proceedings of the National Academy of Sciences (USA), 102, 1137-11342.

Arakane, Y., Muthukrishnan, S., Kramer, K. J., Specht, C. A., Tomoyasu, Y., Lorenzen, M. D., Kanost, M. R., & Beeman, R. W. (2005). The *Tribolium* chitin synthase genes *TcCHS1* and *TcCHS2* are specialized for synthesis of epidermal cuticle and midgut peritrophic matrix. Insect Molecular Biology, 14, 453-463.

Beeman, R.W., & Stuart, J. J. (1990). A gene for lindane + cyclodiene resistance in the red flour beetle (Coleopetera: Tenebrionidae). Journal of Economic Entomology, 83, 1745-1751.

Bell, R. A., & Joachim, F. G. (1976). Techniques for rearing laboratory colonies of tobacco hornworms and pink bollworms. Annals of the Entomological Society of America, 69, 365-373.

Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., Kim, P. S. (1995). Predicting coiled coils by use of pairwise residue correlations. Proceedings of the National Academy of Sciences (USA), 92, 8259-8263.

Berger, M., Chen, H., Reutter, W., & Hinderlich, S. (2002). Structure and function of *N*-acetylglucosamine kinase. European Journal of Biochemistry, 269, 4212-4218.

Bolognesi, R., Arakane, Y., Muthukrishnan, S., Kramer, K. J., Terra, W. R., Ferreira, C. (2005). Sequences of cDNAs and expression of genes encoding chitin synthases and chitinase in the midgut of *Spodoptera frugiperda*. Insect Biochemistry and Molecular Biology, 35, 1249-1259.

Bolognesi, R., Ribeiro, A.F., Terra, W.R., & Ferreira, C. (2001). The peritrophic membrane of *Spodoptera frugiperda*: Secretion of peritrophins and role in immobilization and recycling digestive enzymes. Archives of Insect Biochemistry and Physiology, 47, 62-75.

Boquet, I., Hitier, R., Dumas, M., Chaminade, M., & Préat, T. (2000). Central brain postembryonic development in *Drosophila*: Implication of genes expressed at the interhemispheric junction. Journal of neurobiology, 42, 33-48.

Church, G. M., & Gilbert, W. (1984). Genomic sequence. Proceedings of the National Academy of Sciences (USA), 81, 1991-1995.

Combet, C., Jambon, M., Deleage, G., & Geourjon, C. (2002). Geno3D: automatic comparative molecular modeling of protein. Bioinformatics, 18, 213-214.

Cohen, E. (1991). Chitin biochemistry. In: K. Binnington and A. Retnakaran, eds., Physiology of the Insect Epidermis. CSIRO Publications, Australia. 94-112.

Cohen, E. (2001). Chitin synthesis and inhibition: a revisit. Pest Management Science, 57, 946-950.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Research, 16, 10881-10890.

Coutinho, P. M., Deleury, E., Davies, G. J., & Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. Journal of Molecular Biology, 328, 307-317.

Dahiya, N., Tewari, R., Hoondal, G.S. (2005). Biotechnological aspects of chitinolytic enzymes: a review. Applied Microbiology and Biotechnology [On-line], 1-10.

Devine, W.P., Lubarsky, B., Shaw, K., Luschnig, S., Messina, L., & Krasnow, M.A. (2005). Requirement for chitin biosynthesis in epithelial tube morphogenesis. Proceedings of the National Academy of Sciences (USA), 102, 17014-17019.

Filho, B. P. D., Lemos, F. J. A., Secundino, N. F. C., Páscoa, V., Pereira, S. T., & Pimenta, P. F. P. (2002). Presence of chitinase and *beta-N*-acetylglucosaminidase in the *Aedes aegypti* A chitinolytic system involving peritrophic matrix formation and degradation. Insect Biochemistry and Molecular Biology, 32, 1723-1729.

Filho, J. A. F., Shapiro, B. E. (2004). Tay-Sachs Disease. Archives of Neurology, 61, 1466-1468.

Fukamizo, T. & Kramer, K. J. (1985a). Mechanism of chitin oligosaccharides hydrolysis by the binary chitinase system in insect molting fluid. Insect Biochemistry, 15, 1-7.

Fukamizo, T. & Kramer, K. J. (1985b). Mechanism of chitin hydrolysis by the binary chitinase system in insect molting fluid. Insect Biochemistry, 15, 141-145.

Gagou, M.E., Kapsetaki, M., Turberg, A., & Kafetzopoulos, D. (2002). Stage-specific expression of the chitin synthase *DmeChSA* and *DmeChSB* genes during the onset of *Drosophila* metamorphosis. Insect Biochemistry and Molecular Biology, 32, 141-146.

Gardner, K.H., & Blackwell, J. (1975). Refinement of the structure of beta-chitin. Biopolymers, 14, 1581-1595.

Giraud-Guille, M.M., & Bouligand, Y. (1986). Chitin-protein molecular organization in arthropods. In: Muzzarelli, R., Jeuniaux, C., Gooday, G.W. eds., Chitin in Nature and Technology. Plenum, New York, 29-35.

Haliscak, J. P., & Beeman, R. W. (1983). Status of malathion resistance in five genera of beetles infesting farm-stored corn, wheat, and oats in the United States. Journal of Economic Entomology, 76, 717-722.

Harris, M. T., & Fuhrman, J. A. (2002). Structure and expression of chitin synthase in the parasitic nematode *Dirofilaria immitis*. Molecular and Biochemical Parasitology, 122, 231-234.

Harris, M. T., Lai, K., Arnold, K., Martinex, H. F., Specht, C. A., Fuhrman, J. A. (2000). Chitin synthase in the filarial parasite, *Brugia malayi*. Molecular and Biochemical Parasitology, 111, 351-362.

Hinderlich, S., Berger, M., Schwarzkopf, M., Effertz, K., & Reutter, W. (2000). Molecular cloning and characterization of murine and human *N*-acetylglucosamine kinase. European Journal of Biochemistry, 267, 3301-3308.

Hogenkamp, D. G., Arakane, Y., Zimoch, L., Merzendorfer, H., Kramer, K. J., Beeman, R. W., Kanost, M. R., & Muthukrishnan, S. (2005). Chitin synthase genes in *Manduca sexta*: Characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development. Insect Biochemistry and Molecular Biology, 35, 529-540.

Howling, G.I., Dettmar, P.W., Goddard, P.A., Hampson, F.C., Dornish, M., & Wood, E.J. (2001). The effect of chitin and chitosan on the proliferation of human skin fibroblasts and keratinocytes *in vitro*. Biomaterials, 22, 2959-2966.

Ibrahim, G. H., Smartt, C. T., Kiley, L. M., Christensen, B. M. (2000). Cloning and characterization of a chitin synthase cDNA from the mosquito *Aedes aegypti*. Insect Biochemistry and Molecular Biology, 30, 1213-1222.

Jiang, H., Wang, Y., & Kanost, M. R. (1996). Primary structure of ribosomal proteins S3 and S7 from *Manduca sexta*. Insect Molecular Biology, 5, 31-38.

Kameda, T., Miyazawa, M. Ono, H., & Yoshida, M. (2005). Hydrogen bonding structure and stability of α-chitin studied by 13C solid-state NMR. Macromolecular Bioscience, 5, 103-106.

Kimura, S. (1976). The chitinase system in the cuticle of the silkworm *Bombyx mori*. Insect Biochemistry, 6, 479-482.

Koga, D., Jilka, J., & Kramer, K. J. (1983). Insect endochitinases: glycoproteins from moulting fluid, integument and pupal haemolymph of *Manduca sexta* L. Insect Biochemistry, 13, 295-305.

Koga, D., Mai, M. S., Dziadik-Turner, C., & Kramer, K. J. (1982). Kinetics and mechanism of exochitinase and *β-N*-acetylhexosaminidase from the tobacco hornworm, *Manduca sexta* L. (Lepidoptera: Sphingidae). Insect Biochemistry, 12, 493-499.

Kramer, K. J., & Aoki, H. (1987). Chitinolytic enzymes from pupae of the red flour beetle, *Tribolium castaneum*. Comparative Biochemistry and Physiology, 86B, 613-621.

Kramer, K. J., Corpuz, L., Choi, H. K., & Muthukrishnan, S. (1993). Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinases of *Manduca sexta*. Insect Biochemistry and Molecular Biology, 23, 691-701.

Kramer, K. J., Hopkins, T. L., & Schaefer, J. (1995). Applications of solids NMR to the analysis of insect sclerotized structures. Insect Biochemistry and Molecular Biology, 25, 1067-1080.

Kramer, K.J., Kanost, M.R., Hopkins, T.L., Jiang, H., Zhu, Y.C., Xu, R., Kerwin, J.L., & Turecek, F. (2001). Oxidative conjugation of catechols with proteins in insect skeletal systems. Tetrahedron, 57, 385-392.

Kramer, K. J., & Koga, D. (1986). Insect chitin: physical state, synthesis, degradation and metabolic regulation. Insect Biochemistry, 16, 851-877.

Kramer, K. J., & Muthukrishnan, S. (1997). Insect chitinases: Molecular biology and potential use as biopesticides. Insect Biochemistry and Molecular Biology, 27, 887-900.

Kramer, K. J., & Muthukrishnan, S. (2005). Chitin metabolism in insects: a revisit. In: Gilbert, L., I., Iatrou, K., & and Gill, S. eds., Comprehensive Molecular Insect Science. Vol. 4, Biochemistry and Molecular Biology, Chapter 3. Elsevier Press, Oxford, UK.

Kumar, S., Tamura, K., Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics, 5, 150-163.

Lehane, M.J. (1997). Peritrophic matrix structure and function. Annual Review of Entomology, 42, 525-550.

Léonard, R., Rendić, D., Rabouille, C., Wilson, I. B., Preat, T., & Altmann, F. (2006). The *Drosophila fused lobes* gene encodes an *N*-acetylglucosaminidase involved in *N*-glycan processing. The Journal of Biological Chemistry, 281, 4867-4875.

Locke, M. (1970). The molt/intermolt cycle in the epidermis and other tissues of an insect *Calpodes ethliu*s (lepidoptera, hesperiidae). Tissue & Cell, 2, 197-223.

Locke, M., & Huie, P. (1979). Apolysis and the turnover of plasma membrane plaques during cuticle formation in an insect. Tissue & Cell, 11, 277-291.

Locke, M., & Krishnan, N. (1973). The formation of the ecdysial droplets and the ecdysial membrane in an insect. Tissue & Cell, 5, 441-450.

Maier, T., Strater, N., Schuette, C. G., Klingenstein, R., Sandhoff, K., & Saenger, W. (2003). The X-ray crystal structure of human β-hexosaminidase B provides new insights into Sandhoff Disease. Journal of Molecular Biology, 328, 669-681.

Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S., & James, M. N. G. (2003). Crystal structure of human β-hexosaminidase B: Understanding the molecular basis of Sandhoff and Tay-Sachs Disease. Journal of Molecular Biology, 327, 1093-1109.

Martín-Udíroz, M., Madrid, M.P., Roncero, I.G. (2004). Role of chitin synthase genes in *Fusarium oxysporum*. Microbiology, 150, 3175-3187.

Merz, R. A., Horsch, M., Nyhlen, L. E., Rast, & D. M. (1999). Biochemistry of chitin synthase. In: Jollès, P. & Muzzarelli, R., A., A. eds., Chitin and Chitinases. Birkhäuser Verlag, Germany. 9-37.

Merzendorfer, H. (2006). Insect chitin synthases: a review. Journal of Comparative Physiology. Part B, 176, 1-15.

Moussian, B., Schwarz, H., Bartoszewski, S., Nüsslein-Volhard, C. (2005). Involvement of chitin in exoskeleton morphogenesis in *Drosophila melanogaster*. Journal of Morphology, 264, 117-130.

Munro, C. A. & Gow, N. A. (2001). Chitin synthesis in human pathogenic fungi. Medical Mycology, 39 (Suppl. 1), 41-53.

Muzzarelli, R. A. A. (1977). Chitin. Pergamon Press, New York, 5-16.

Muzzarelli, R. A. A. (1999). Native, industrial and fossil chitins. In: Jolles, P., Muzzarelli, R. A. A. (Eds), Chitin and Chitinases. Berkhauser Verlag, Basel, Switzerland, 1-3.

Nagahashi, S., Sudoh, M., Ono, N., Sawada, R., Yamaguchi, E., Uchida, Y., Mio, T., Takagi, M., Arisawa, M., & Yamada-Okabe, H. (1995). Characterization of chitin synthase 2 of *Saccharomyces cerevisiae*. Implication of two highly conserved domains as possible catalytic sites. Journal of Biological Chemistry, 270, 13961-13967.

Nagamatsu, Y., Yanagisawa, I., Kimoto, M., Okamoto, E., & Koga, D. (1995). Purification of a chitooligosaccharidolytic *beta-N*-acetylglucosaminidase from *Bombyx mori* larvae during metamorphosis and the nucleotide sequence of its cDNA. Bioscience, Biotechnology and Biochemistry, 59, 219-225.

Neufeld, E. F. (1989). Natural history and inherited disorders of a lysosomal enzyme, *beta* hexosaminidase. The Journal of Biological Chemistry, 264, 10927-10930.

Park, J.T. (2001). Identification of a dedicated recycling pathway for anhydro-*N*-acetylmuramic acid and *N*-acetylglucosamine derived from *Escherichia coli* cell wall murein. Journal of Bacteriology, 183, 3842-3847.

Peters, W. (1992). Peritrophic membranes. Zoophysiology series volume 30. New York: Springer-Verlag.

Richmond, T. (2000). Higher plant cellulose synthases. Genome Boil. 1, REVIEWS3001-3005.

Riddiford, L. M. (1994). Cellular and molecular actions of juvenile hormone I. General considerations and premetamorphic actions. Advances in Insect Physiology, 24, 213-274.

Riddiford, L.M., Hiruma, K., Zhou, X., & Nelson, C.A. (2003). Insights into the molecular basis of the hormonal control of molting and metamorphosis from *Manduca sexta* and *Drosophila melanogaster*. Insect Biochemistry and Molecular Biology, 33, 1327-1338.

Roncero, C. (2002). The genetic complexity of chitin synthesis in fungi. Current genetics, 41, 367-378.

Sahai, A. S., & Manocha, M. S. (1993). Chitinases of fungi and plants: their involvement in morphogenesis and host parasite interaction. FEMS Microbiology Reviews, 11, 317-338.

Saito, Y., Kumagai, H., Wada, M., & Kuga, S. (2002). Thermally reversible hydration of beta chitin. Biomacromolecules, 3, 407-410.

Sandhoff, K., Andreae, U., Jatzkewitz, H. (1968). Deficient hexosaminidase activity in an exceptional case of Tay-Sachs disease with additional storage of kidney globoside in visceral organs. Pathologia Europaea, 3, 278-285.

Saxena, I.M., Malcolm Brown, R. Jr., & Dandekar, T. (2001). Structure-function characterization of cellulose synthase: relationship to other glycosyltransferases. Phytochemistry 57, 1135-1148.

Shahidi, F., & Abuzaytoun, R. (2005). Chitin, chitosan, and co-products: chemistry, production, applications, and health effects. In: Taylor, S.L. ed. Advances in Food and Nutrition Research, Volume 49. Elsevier Academic Press, New York, 93-135.

Specht, C. A., Liu, Y., Robbins, P. W., Bulawa, C. E., Iartchouk, N., Winter, K. R., Riggle, P. J., Rhodes, J. C., Dodge, C. L., Culp, D. W., Borgio, P. T. (1996). The chsD and chsE genes of *Aspergillus nidulans* and their roles in chitin synthesis. Fungal Genetics and Biology, 20, 153-167.

Suzuki, T., Yano, K., Sugimoto, S., Kitajima, K., Lennarz, W. J., Inoue, S., Inoue, Y., & Emori, Y. (2002). Endo-*beta*-*N*-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. Proceedings of the National Academy of Sciences (USA), 99, 9691-9696.

Tellam, R.L., Vuocolo, T., Johnson, S.E., Jarmey, J., & Pearson, R.D. (2000). Insect chitin synthase: cDNA sequence, gene organization and expression. European Journal of Biochemistry, 267, 6025-6042.

Tellam, R.L., Wijffels, G., & Willadsen, P. (1999). Peritrophic matrix proteins. Insect Biochemistry and Molecular Biology, 29, 87-101.

Terra, W. R. (2001). The origin and functions of the insect peritrophic membrane and peritrophic gel. Archives of Insect Biochemistry and Physiology, 46, 47-61.

Tomoyasu, Y., & Denell, R. E. (2004). Larval RNAi in *Tribolium* (Coleoptera) for analyzing adult development. Development Genes and Evolution, 214, 575-578.

Tonning, A., Helms, S., Schwarz, H., Uv, A.E., & Moussian, B. (2005a). Hormonal regulation of *mummy* is needed for apical extracellular matrix formation and epithelial morphogenesis in *Drosophila*. Development, 133, 331-341.

Tonning, A., Hemphälä, J., Tång, E., Nannmark, U., Samakovlis, C., & Uv, A. (2005b). A transient luminal chitinous matrix is required to model epithelial tube diameter in the *Drosophila* trachea. Developmental Cell, 9, 423-430.

Valdivieso, M. H., Duran, A., Roncero, C. (1999). Chitin synthases in yeast and fungi. Experientia Supplementrum, 87, 55-69.

Van Dellen, K.L., Bulik, D.A., Specht, C.A., Robbins, P.W., & Samuelson, J.C. (2006). Heterologous expression of an *Entamoeba histolytica* chitin synthase in *Saccharomyces cerevisiae*. Eukaryotic Cell, 5, 203-206.

Veronico, P., Gray, L. J., Jones, J. T., Bazzicalupo, P., Arbucci, S., Cortese, M. R., Di Vito, M., De Giorgi, C. (2001). Nematode chitin synthases; gene structure, expression and function in *Caenorhabditis elegans* and the plant parasitic nematode *Meloidogyne artiellia*. Molecular Genetics and Genomics, 266, 28-34.

Wang, P., & Granados, R.R. (2001). Molecular structure of the peritrophic membrane (PM): identification of potential PM target sites for insect control. Archives of Insect Biochemistry and Physiology, 47, 110-118.

Vincent, J.F.V., & Wegst, U.G.K. (2004). Design and mechanical properties of insect cuticle. Arthropod Structure & Development, 33, 187-199.

Yalpani, R., Lehtinen, S., Wuolijoki, E., Ylitalo, P., & Lehtimaki, T. (2002). Antimicrobial activity of some chitosan derivatives. In: Brine, C.J., Sandford, P.A., Zikakis, J.P. eds., Advances in Chitin and Chitosan. Elsevier Applied Science, London, UK., 543-555.

Yeager, A.R., & Finney, N.S. (2003). The first direct evaluation of the two-active site mechanism for chitin synthase. Journal of Organic Chemistry, 69, 613-618.

Yeager, A.R., & Finney, N.S. (2004). Second-generation dimeric inhibitors of chitin synthase. Bioorganic & Medicinal Chemistry, 12, 6451-6460.

Zen, K. C., Choi, H. K., Krishnamachary, N., Muthukrishnan, S., & Kramer, K. J. (1996). Cloning, expression, and hormonal regulation of an insect *beta-N*-acetylglucosaminidase gene. Insect Biochemistry and Molecular Biology, 26, 435-444.

Zhang, W., Cao, P., Chen, S., Spence, A. M., Zhu, S., Staudacher, E., & Schachter, H. (2003). Synthesis of paucimannose N-glycans by Caenorhabditis elegans requires prior actions of UDP-*N*-acetyl-D-glucosamine:α-3-D-mannoside β1,2-*N*-acetylglucosaminyltransferase I, α3,6-mannosidase II and a specific membrane-bound β-*N*-acetylglucosaminidase. The Biochemical journal, 372, 53-64.

Zhu, Y. C., Muthukrishnan, S., Specht, C. A., Dittmer, N., Kanost, M. R., & Kramer, K. J., (2002). Sequence of a cDNA and expression of the gene encoding a putative epidermal chitin synthase of *Manduca sexta*. Insect Biochemistry and Molecular Biology, 32, 1497-1506.

Zhu, Q. (2005). Characterization of chitinase-like gene families and proteins from *Tribolium castaneum*, *Drosophila melanogaster* and *Anopheles gambiae*. Doctoral Dissertation, Kansas State University.

Zimoch, L., Hogenkamp, D., Kramer, K. J., Muthukrishnan, S., & Merzendorfer, H. (2005). Regulation of chitin synthesis in the larval midgut of *Manduca sexta*. Insect Biochemistry and Molecular Biology, 35, 515-527.

Zimoch, L. & Merzendorfer, H. (2002). Immunolocalization of chitin synthase in the tobacco hornworm. Cell and Tissue Research, 308, 287-297.

APPENDIX A: DEDUCED PROTEIN SEQUENCE OF TcHEX1

ANNISIRLDLIFHCSNFPKNSYLGVSSSSESTQIYKRKHLSFNKSVTRVSCDKRSVMTKL
VLYLLPLIISCNAIHPGPVVRATKGEVWPKPQVEEKTEQYYTVRPHGFNFKGPTNIGCPN
LLNDAFTRYWTIIATASSLERRGRLHEVGRKPKTKFWKADSNYLGDLTNLHVQLDDCASE
YVLPAFGDNENYTLSVTSEGASLTADTIWGVLRGLETFSQLIYLEQDTLIINATNVNDYP
RFSHRGLLLDTSRHFIPIYIILQTLDAMSYNKFNVFHWHITDDHSFPYKSRTYPELSDEG
AYHPVSKVYEQSDVSKVIEYARVRGIRVIPEFDTPGHTSSWGAAHPELLTTCYTNDKPDG
DLGPMDPTKNSTYDFLTKLFTEVVEVFPDSYFHIGGDEVDFTCWQHNPDIASFMKANNIS
TYEDLESYFIQHVVNLLDSLNSNYLVWEEVFVNGVTLPDSTLVHVWRDNGHETLNSVTKA
GKYGIFSSCWYLDHVSSGGDWEKFYECEALDFPGTEEQKKLVLGGEACMWSEAVNEYNVM
PRVWPRASAVAEKLWSAGNVNDTQAAKGRLEEHTCRMNNRGIAAQPPNGPG

APPENDIX B: DEDUCED PROTEIN SEQUENCE OF TcHEX2

WPKPQQQELNFQQPFNFPQGTDDALTRYWTIIATSITSKLEETPEANFWELDDNFLGYLE
TLTITLLGECPNENILPELHDNENYTLTVDSEGAFLESETIWGVLRGLETFSQLIYAEQD
TLMINTTKIVDFPRFPHRGFLLDTSRHFEPVRIILQMLDAMAYNKLNVFHWHITDDHSFP
YKSRTYHELSDKGAYHPVSGVYEQSDVMKIIEYARVRGIRVIPEFDTPGHTRSWGVAHPE
LLTSCFTDNVANGELGPMDPTKDTTYDFINNLFTEIVDVFPDSYFHIGGDEVEFDCWKSN
PDVSNFMKQNNFSTYEQLESYFIQHVVDILDNLSSKYLVWEEVFVNGVELPNSTVVHVWK
DNGLSTLNNVIKAGKYGLYSSCWYLSVLHSGSDWDAFYKCEPGLLLHTEEEKKLLLGGEA
CMWGEYVNEFSVIPRVWPRASAVAERLWSDENVVDISDAQIRLEEHACRMNKRGIAAQPP
NGPGMCF

APPENDIX C: DEDUCED PROTEIN SEQUENCE OF TcHEX3

KICSFYYTIMMFKLFFLLLIISFCSAFDFIFQPGPLVPASKGEIWPKPQHENKLDDGFFS
LLPTFFHFNPIGNICNTLTEALDRYRKLIIFNNRRIKEVYYKARSCYEGGDQNFLGYLTS
VEVELTGACNDEEYPSFEMKEECKFTRILFMQRISSDTIWGILRGLETFSQLIYLTDDYS
CHRIGTTSIHDYPRFAHRGLLLDTSRHYIPKEHILKLIETMSYNKLNVFHWHITDDYSFP
YVSKAFPQMSNKGAFHPTLMIYEQDFVSEVQEYARKRGIRVLAEFDTPGHTLSWGLGNPD
LLTDCHNVPQLKWGPINPIKNTTYDFIFKLFEEIKSVFKDEYTHLGGDEVDFSCWKSNPE
INQWMAEHQMEGDYVALQSHYIQKLINHVDSLGLNSIVWEEVFTNGVQLPKSTVVNVWIS
DDPKTTLKQVTEAGHPTIISSYWYLDILKTGGDWLKFYNADPQDFDGTDEQKRLVLGGEA
CMWSEVVDEYNLEPRVWPRASVAAERFWSPPDTPKSAQNLGELWTIASRLQEQTCRMNRR
GVAAQPPSGPSVCF

APPENDIX D: DEDUCED PROTEIN SEQUENCE OF TcENG

FFPHYQLPFPAVLQLPIMSTGNKTHFSPPVVSNLFVESDGPPAMKKLKLNQFECFPITSL
DNISEVIENPPQWVSRVAPLTERSPTVVQNIISDCHCDAKKFGPRTRLDKRFVPKTLVCH
DYKGGYLEDRFLSSSNIGNLYSFYNWQHIDIFVYFSHHLITIPPLCWINAAHQNGVKILG
TLITEFEPGKKICEKIFKDEDTMRIFATSLTQILKIFQFDGWLLNIENSLQDTEMLKKFV
AYLGVQTHTQNPDHLIIWYDSIIETGELKWQNELNPLNKFFFDNCDGIFLNYVWTEENIH
NSIEFAKHRTLDVYVGVDVFGRNTFGGGQFNCFKAAQMIRRHNLSMAIFAPGWTHETLPK
SENRQKFFEDFINRDSAFWNSLWPYLYTHPITRFFKTSFFTGVNQVFN

APPENDIX E: PROTEIN SEQUENCE OF HUMAN HEXB (1O7A)

PALWPLPLSVKMTPNLLHLAPENFYISHSPNSTAGPSCTLLEEAFRRYHGYIFGTQVQQLLVSI
TLQSECDAFPNISSDESYTLLVKEPVAVLKANRVWGALRGLETFSQLVYQDSYGTFTINESTII
DSPRFSHRGILIDTSRHYLPVKIILKTLDAMAFNKFNVLHWHIVDDQSFPYQSITFPELSNKGS
YSLSHVYTPNDVRMVIEYARLRGIRVLPEFDTPGHTLSWGKGQKDLLTPCYSKLDSFGPINPTL
NTTYSFLTTFFKEISEVFPDQFIHLGGDEVEFKCWESNPKIQDFMRQKGFGTDFKKLESFYIQK
VLDIIATINKGSIVWQEVFDDKAKLAPGTIVEVWKDSAYPEELSRVTASGFPVILSAPWYLDLI
SYGQDWRKYYKVEPLDFGGTQKQKQLFIGGEACLWGEYVDATNLTPRLWPRASAVGERLWSSKD
VRDMDDAYDRLTRHRCRMVERGIAAQPLYAGYCNH

APPENDIX F: PROTEIN SEQUENCE OF HUMAN HEXB (1NOU)

AKPGPALWPLPLSVKMTPNLLHLAPENFYISHSPNSTAGPSCTLLEEAFRRYHGYIFGFYKWHH
EPAEFQAKTQVQQLLVSITLQSECDAFPNISSDESYTLLVKEPVAVLKANRVWGALRGLETFSQ
LVYQDSYGTFTINESTIIDSPRFSHRGILIDTSRHYLPVKIILKTLDAMAFNKFNVLHWHIVDD
QSFPYQSITFPELSNKGSYSLSHVYTPNDVRMVIEYARLRGIRVLPEFDTPGHTLSWGKGQKDL
LTPCYSRQNKLDSFGPINPTLNTTYSFLTTFFKEISEVFPDQFIHLGGDEVEFKCWESNPKIQD
FMRQKGFGTDFKKLESFYIQKVLDIIATINKGSIVWQEVFDDKAKLAPGTIVEVWKDSAYPEEL
SRVTASGFPVILSAPWYLDLISYGQDWRKYYKVEPLDFGGTQKQKQLFIGGEACLWGEYVDATN
LTPRLWPRASAVGERLWSSKDVRDMDDAYDRLTRHRCRMVERGIAAQPLYAGYCNHENM