Characterization of two chitin synthase genes of the red flour beetle, *Tribolium castaneum*, and alternate exon usage in one of the genes during development


*Grain Marketing and Production Research Center, ARS-USDA, 1515 College Avenue, Manhattan, KS 66502, USA*

*Department of Biochemistry, 104 Willard Hall, Kansas State University, Manhattan, KS 66506-3702, USA*

*Jamie Whitten Delta States Research Center, ARS-USDA, Stoneville, MS 38776, USA*

*Department of Medicine, Section of Infectious Diseases, Boston University, Boston, MA 02118, USA*

Received 9 June 2003; received in revised form 6 November 2003; accepted 6 November 2003

**Abstract**

Two chitin synthase (CHS) genes of the red flour beetle, *Tribolium castaneum*, were sequenced and their transcription patterns during development examined. By screening a BAC library of genomic DNA from *T. castaneum* (Tc) with a DNA probe encoding the catalytic domain of a putative *Tribolium* CHS, several clones that contained CHS genes were identified. Two distinct PCR products were amplified from these BAC clones and confirmed to be highly similar to CHS genes from other insects, nematodes and fungi. The DNA sequences of these genes, TcCHS1 and TcCHS2, were determined by amplification of overlapping PCR fragments from two of the BAC DNAs and mapped to different linkage groups. Each ORF was identified and full-length cDNAs were also amplified, cloned and sequenced. TcCHS1 and TcCHS2 encode transmembrane proteins of 1558 and 1464 amino acids, respectively. The TcCHS1 gene was found to use alternate exons, each encoding 59 amino acids, a feature not found in the TcCHS2 gene. During development, *Tribolium* expressed TcCHS1 predominantly in the embryonic and pupal stages, whereas TcCHS2 was prevalent in the late larval and adult stages. The alternate exon 8a of TcCHS1 was utilized over a much broader range of development than exon 8b. We propose that the two isoforms of the TcCHS1 enzyme are used predominantly for the formation of chitin in embryonic and pupal cuticles, whereas TcCHS2 is utilized primarily for the synthesis of peritrophic membrane-associated chitin in the midgut.

© 2003 Elsevier Ltd. All rights reserved.

**Keywords:** Chitin synthase; Chitin; *Tribolium castaneum*; Insect; Development; *Manduca sexta*; Tobacco hornworm; Glycosyltransferase; RNA expression; Alternative exon splicing; Phylogeny; Nematode; Yeast; Malaria mosquito; Fruit fly; *Anopheles; Drosophila*; Cuticle; Epidermis; Transmembrane enzyme; Polysaccharide synthesis; Peritrophic membrane

1. **Introduction**

Many tons of chitin, a homopolymer of β-1,4-linked N-acetylglucosamine (GlcNAc), are recycled annually in the biosphere with fungi and arthropods being the principal producers (Muzzarelli, 1999). Chitin is synthesized from cytoplasmic pools of UDP-β-N-acetylglucosamine (UDP-GlcNAc) by chitin synthases (CHSs) (EC 2.4.1.16) located in the plasma membrane, which also ensure that the polymer is extruded outside of the cell into an extracellular matrix. In many species
of fungi, chitin is an essential component of the cell wall and septum, whereas for insects, it is an essential component of the exoskeleton and the peritrophic membrane (PM) that lines the midgut. CHSs have been identified in a variety of organisms including nematodes, fungi and insects. Amino acid sequence similarities have been the principal tools used for identifying CHSs, which form a subfamily within a larger group (family 2) of the glycosyltransferases (Coutinho and Henrissat, 1999).

Fungal CHSs have been studied extensively and various ones were found to function at different developmental stages such as cell division, sporulation and hyphal branching (Merz et al., 1999; Henar Valdivieso et al., 1999; Munro and Gow, 2001; Roncero, 2002). Consistent with their multifarious roles, fungal CHSs are encoded by a large family of genes and as many as eight different enzymes have been identified in a single species.

In contrast to fungi, insects and nematodes have fewer genes encoding CHSs (Zhu et al., 2002). The number of CHS genes in most insect species is likely to be two, based upon sequencing of the Drosophila and Anopheles genomes (Drosophila and Anopheles Genome Project databases at the National Center for Biotechnology Information). Genome sequencing of the nematode, Caenorhabditis elegans, also revealed two CHS genes (Veronico et al., 2001). Individual CHS genes from the blow fly (Lucilia cuprina), mosquito (Aedes aegypti), and tobacco hornworm (Manduca sexta) have been described recently based on cDNA sequencing, and preliminary evidence suggests the existence of a second gene in each of these species (Tellam et al., 2000; Zhu et al., 2002; our unpublished data). CHS genes have been reported for three other species of nematodes, Brugia malayi, Meloidogyne artiella, and Dirofilaria immitis (Harris et al., 2000; Veronico et al., 2001; Harris and Fuhrman, 2002), although it remains unclear whether any of these organisms have more than one CHS gene.

It has been suggested that insects have two CHSs, one responsible for synthesis of cuticular chitin and the other dedicated to the synthesis of chitin associated with the peritrophic membrane (Tellam et al., 2000; Zhu et al., 2002). However, to verify this suggestion, an extensive characterization of all CHS genes and their expression in different tissues and developmental stages in a single insect species must be completed. In this paper, we report the isolation and characterization of two distinct genes for CHS from the red flour beetle, Tribolium castaneum. We also demonstrate that even though there are only two CHS genes, Tribolium can generate different enzymes or isoforms as a result of alternate exon usage. This process appears to be developmentally regulated.

2. Materials and methods

2.1. Insect growth, developmental stages and BAC libraries

Beetles were reared at 30 °C in whole wheat flour fortified with 5% (v/v) Brewer’s yeast under standard conditions (Beeman and Stuart, 1990). The following life stages were used for analysis of CHS gene expression during development: embryos (0–3 days post-oviposition), early larvae (actively feeding, 6th–7th instar), last instar larvae (actively feeding, 7th–8th instar), prepupae (quiescent, post-feeding 7th–8th instar), pupae and adults. Instar numbering in Tribolium is indeterminant and cannot be known with certainty when mass culturing. The BAC library (a gift from Exelixis Pharmaceutical Co., South San Francisco, CA) was prepared in the plasmid vector pBACe3.6 using a partial EcoRI digest of genomic DNA from the highly inbred GA-2 strain. At the time of screening, the library consisted of approximately 8400 clones with an average insert size of about 129 kb and a genome coverage of ca. 5.4×.

2.2. Preparation of a DNA probe for TcCHS catalytic domain

To obtain a DNA probe for the identification of BAC clones containing TcCHS genes, we designed a pair of degenerate primers (forward: 5’-TTY-GARTAYGCNATHGGNCAYTGG-3’ and reverse: 5’-CCANCKRTCTCNCCYTGRTCRTAYTG-3’) corresponding to two highly conserved amino acid sequences in the catalytic domain of CHSs, FEYAIGHW and QYDQGDRW, respectively. These primers were used to amplify a 242 bp PCR product using Tribolium genomic DNA as the template. The PCR product was cloned into the pGEM-T vector (Promega). The cloned DNA was sequenced and shown to represent a CHS gene fragment by using the BLASTX search program of the GenBank protein database. The insert DNA from this clone was used subsequently for screening the Tc BAC library.

2.3. Screening of Tc BAC library with CHS DNA probe

The PCR fragment with sequence similarity to CHS genes obtained as described in Section 2.2 was labeled with 32P-dATP using the Promega primer-labeling kit and used as a probe for screening the Tribolium BAC library. A nylon membrane was double-loaded (diagonally or side by side) with a total of approximately 8400 colonies representing the Tc BAC library. The membrane was hybridized with the 32P-labeled probe in PerfectHyb Plus buffer (Sigma) at 62 °C for 18 h. The
membrane was washed at 55 °C for 1 h with three changes of 0.2x SSC/0.1% SDS and exposed to X-ray film to identify colonies that hybridized either strongly or weakly with the probe.

2.4. Determination of the DNA sequences of TcCHS genes

BAC plasmids containing the TcCHS genes were prepared by the alkaline-SDS method and used as templates for PCR amplification of different regions of the TcCHS genes. The sequences of degenerate primers and their locations in the TcCHS genes are shown in Table 1A and Fig. 1, respectively. For amplification of 5’- and 3’-end sequences, single primers, SF3, SR3, and WF3 (see Table 1B for sequences) were used. To obtain the 5’-end sequence of TcCHS2, inverse PCR was conducted using gene-specific primers, WF4 and WR4 (Table 1B). Approximately 2 μg of BAC DNA containing the TcCHS2 gene was digested with restriction enzyme PstI and the fragments were purified by phenol–chloroform extraction followed by ethanol precipitation. T4 DNA ligase was added to the fragments to allow the formation of circular DNA as a result of self-ligation. PCR reactions were conducted in a final volume of 50 μl containing 20 mM Tris–HCl, pH 8, 50 mM KCl, 2 mM MgCl2, 0.4 μM of the primers, 0.2 mM dNTPs, 2.5 units of Ex Taq polymerase and approximately 10 ng of the plasmid template. PCR reactions were conducted for 25 cycles with 1 min of denaturation at 94 °C, 1 min of annealing at 42–60 °C, and 3 min of extension at 72 °C. The amplified DNA fragments were subcloned into the pGEM-T vector (Promega) and sequenced using flanking vector primers T7 and SP6 and/or gene specific primers (Sequencing and Genotyping Facility, Kansas State University).

2.5. Cloning of cDNAs for CHS genes

To obtain the cDNAs corresponding to the entire protein coding regions of TcCHSIA, TcCHSIB and TcCHS2, mRNAs were purified from prepupae using the Oligotex Direct mRNA Kit (Qiagen) according to the manufacturer’s instructions. Reverse transcriptions were performed with SUPERSCRIPT II RNase H-Reverse Transcriptase (Invitrogen) using an oligo-(dT) primer. The primers used were 5’-ATGACATCCCGGG-GGGCTG-3’ and 5’-TCACATCTCCTGATTGCTGC-3’ for TcCHS1, and 5’-ATGCCCGCGCTACGG-3’ and 5’-TTATGCCTCCACGTCTGACC-3’ for TcCHS2. PCR reactions were conducted as follows: denaturation at 94 °C for 1 min, annealing at 60 °C for 1 min and polymerization at 72 °C for 4.5 min using Takara LA Taq polymerase for 30 cycles. The products of each reaction were subjected to electrophoresis on 0.8% agarose gel containing 1 μg/ml crystal violet (Invitrogen), excised and purified using the Freeze ’N Squeeze DNA Gel Extraction Spin Column (Bio-Rad).

Table 1
Primer sequences for the amplification of genomic DNA containing TcCHS1 and TcCHS2

<table>
<thead>
<tr>
<th>Name</th>
<th>Primer sequence (5’-3’)</th>
<th>Conserved AA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Degenerate primers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward primer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td>TTY GAR TAY GCN ATH GGN CAY TGG</td>
<td>FEYAIKHG</td>
</tr>
<tr>
<td>F2</td>
<td>TGY GCN ACN ATG TGG CAY G</td>
<td>CATMWHE</td>
</tr>
<tr>
<td>F3</td>
<td>GGN TGG TGG GAR AA</td>
<td>GW WEN</td>
</tr>
<tr>
<td>F4</td>
<td>CAR GAR ACN AAR GGN TGG GA</td>
<td>QETKGWE</td>
</tr>
<tr>
<td>F5</td>
<td>TGG GAY GTN TTY MGN GAY CCN CC</td>
<td>WD VFDDPP</td>
</tr>
<tr>
<td>Reverse primer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R1</td>
<td>CCA NCK RTC YTC NCC YTG RTC RTA YTG</td>
<td>QYDQGEDRW</td>
</tr>
<tr>
<td>R2</td>
<td>ACY TCN CKN GTN CCC CA</td>
<td>WGTRE</td>
</tr>
<tr>
<td>R3</td>
<td>AAN CKR TGR AAN ARC ATN GC</td>
<td>AMLHRF</td>
</tr>
<tr>
<td>R4</td>
<td>TTN GCN CCN TRN GTY TGC AT</td>
<td>MQTQ/YGAR</td>
</tr>
<tr>
<td>(B) Gene-specific primers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward primer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF1</td>
<td>CGG AAG CCA GGC ATT ACG TG</td>
<td>WF1</td>
</tr>
<tr>
<td>SF2</td>
<td>TGC TAC TGA TTT TGT ACT CC</td>
<td>WF2</td>
</tr>
<tr>
<td>SF3</td>
<td>TGG AAC CGA TCG GTG TAG TC</td>
<td>WF3</td>
</tr>
<tr>
<td>Reverse primer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR1</td>
<td>CAT GCT CGG TGG CCT TTT GC</td>
<td>WR1</td>
</tr>
<tr>
<td>SR2</td>
<td>CCA GTC GCA AAA TCG ACT TC</td>
<td>WR2</td>
</tr>
<tr>
<td>SR3</td>
<td>TTC TCC TTG ATT TTG CCC AG</td>
<td>WR4</td>
</tr>
</tbody>
</table>
The purified fragment was cloned into a pCR-XL-TOPO vector (Invitrogen). Furthermore, to identify the cDNAs of TcCHS1A and TcCHS1B, colony PCR was done using the common forward primer designed based on exon 6, 50-TGCTACTGATTTTGTACTCC-30, and either of the following exon 8a- or exon 8b-specific reverse primers: 50-TTGGGTGCTCTCGTCATAG-30 for exon 8a and 50-CGGACGTTTCCTCAATATAC-30 for exon 8b. Three cloned cDNA fragments corresponding to TcCHS1A, TcCHS1B, and TcCHS2 were fully sequenced using appropriate primers. Furthermore, to obtain the 50-end of cDNAs of TcCHS1 and TcCHS2, 50-RACE was performed using the 50-RACE system version 2.0 (Invitrogen) according to the manufacturer’s instructions. The following antisense gene-specific primers were used for the synthesis of the first strand cDNAs from total RNA of pupae and larvae: 50-TTCTCCTTGATTTTGCCCAG-30 (spanning positions 850 and 869) for TcCHS1 and 50-AACTCCTTTTTGCTCATAG-30 (spanning positions 823 and 842) for TcCHS2. Following the addition of a homopolymeric C-tail at the 30-end of the cDNAs, PCR was carried out using an adapter primer and the gene-specific forward primers, 50-TGACCATGAGGAGAGG-30 (spanning positions 4382 and 4398) for TcCHS1 and 50-TAGACTTTGATCTGTGACG-30 (spanning positions 3956 and 3975) for TcCHS2 and first strand cDNAs as the template.

### 2.6. Analysis of expression of TcCHS1 and TcCHS2 by RT-PCR during development

Total RNA was isolated from whole insects of different developmental stages and from male and female pupae and adults (see Section 2.1) using the RNaseasy Mini Kit (Qiagen). Two micrograms of total RNA were used as templates for first strand cDNA synthesis using an oligo-(dT) primer. This cDNA was used then as the template for amplification and detection of specific TcCHS sequences. The primers used were 50-TGCTACTGATTTTGTACTCC-30 (spanning positions 3131 and 3150 in exon 6) and 50-TGACCATGAGGAGAGG-30 (spanning positions 4382 and 4398) for TcCHS1 and 50-TAGACTTTGATCTGTGACG-30 (spanning positions 3956 and 3975) for TcCHS2 and first strand cDNAs as the template.
2.7. Analysis of TcCHS1 transcripts for alternate exon usage

The presence of TcCHS1 transcripts containing exon 8a or 8b in total RNA was analyzed using a PCR assay with primer combinations designed to discriminate between the two exons. cDNAs prepared from total RNA isolated from different stages of development were used as the template. A 20-nucleotide-long pri-
rRNA isolated from different stages of development between the two exons. cDNAs prepared from total RNA was analyzed using a PCR assay with primer combinations designed to discriminate 8a or 8b in total RNA was analyzed using a PCR assay with primer combinations designed to discriminate between the two exons. cDNAs prepared from total RNA isolated from different stages of development. A 20-nucleotide-long primer, 5'-TGCTACTGATTTTGTACTCC-3' (1F) in exon 5 of TcCHS1, was the forward primer. The reverse primer (3R) for exon 8a was 5'-TTGGGTGCTCTCGTCATAG-3' and the primer for exon 8b (4R) was 5'-CGGACGTTCCTCAATATACT-3'. PCR reactions were conducted using the following conditions: denaturation at 94 °C for 1 min, annealing at 54 °C for 1 min and polymerization at 72 °C for 1 min for 23 cycles. The products of each reaction were fractionated by electrophoresis in a 1% agarose gel and visualized by staining with ethidium bromide.

2.8. DNA and protein sequence analyses

The CHS genes of Drosophila and Anopheles were analyzed to determine probable exons and introns of each gene. Initially, a series of ‘‘TBLASTN’’ searches of the Drosophila and Anopheles Genome Project databases at NCBI were made by using as queries the protein sequences of M. sexta CHS1 (AAC38051), L. cuprina CHS1 (AA907912), A. aegypti CHS1 (AA934699) and the protein sequences for TcCHS1 and TcCHS2 as determined in this study. Exons were chosen that translated to protein sequences most similar to the queries and introns that had conventional 5' and 3' splice sites. Exons and introns were subsequently confirmed by recently submitted full-length cDNA sequences of DmCHS1 (NM_169052 and NM_079509), DmCHS2 (XM_079485) and AgCHS2 (XM_321336 and XM_321337); and a partial cDNA sequence of AgCHS1 (XM_321951). The presence of alternate exons corresponding to exon 8 of TcCHS1 in M. sexta CHS1 was confirmed by sequencing of a PCR-amplified genomic fragment (Hogenkamp et al., unpublished data).

Alignment of nucleotide sequences and deduced amino acids from cDNA clones was made using ClustalW software (PAM250). Protein sequences were analyzed for transmembrane helices using the TMHMM v.2.0 software available at http://www.cbs.dtu.dk/services/TMHMM-2.0/. Coiled-coil domains were identified using the Paircoil Program which can be accessed at http://theory.csc.mit.edu/~bab/paircoil/paircoil.html. Multiple sequence alignments to yield phylogenetic trees were generated with Megalign Software (DNAStar, Madison, WI) using ClustalW (PAM250).

2.9. Mapping of TcCHS genes to specific Tribolium linkage groups

CHS-positive BACs were genetically mapped by single-strand conformation polymorphism (SSCP) analysis. We identified SSCP dimorphisms between two highly inbred T. castaneum strains, GA-2 and ab2, using primer pairs specific for end-sequences from each BAC. The BACs were mapped onto a whole-genome recombination map at an average resolution of ca. 1.5 cM using a backcross family that consisted of 179 sib-
lings and using a marker set totaling more than 400 unique DNA sequences derived from BACs, cDNAs and other sources. Details of this mapping procedure will be published elsewhere (Beeman et al., unpublished data).

2.10. Southern blot analysis of Tribolium genomic DNA for CHS sequences

Tribolium genomic DNA (10 μg per sample) was digested with five different restriction enzymes, AseI, BamHI, ClaI, HindII, HindIII and XbaI and separated on 0.9% agarose gel. Nucleic acids were transferred onto Hybond N+ Nylon membrane (Amersham) under alkaline condition and hybridized with random primed 32P-labeled probes (Ready-To-Go DNA Labeling Beads, Amersham). The probe was the 506 bp PCR fragment spanning nucleotide positions 2233 through 2739 of TcCHS1 cDNA. The membrane was hybri-
dized to the probe at 50 °C and washed at 37 °C with 1× SSC/0.1% SDS and exposed to X-ray film for 3 d.

3. Results

3.1. Screening of Tc BAC library with the CHS probe

A 242 bp PCR fragment amplified from T. castaneum genomic DNA as outlined in Section 2.2 using two degenerate primers derived from two highly conserved regions in other insect chitin synthase genes was labeled with 32P-dATP and used to screen an ordered BAC library of Tc genomic DNA. Sixteen colonies were hybri-
dized to the probe, 12 strongly and four weakly. Plas-
mid DNA preparations were made from 12 of these BAC colonies. PCR amplification using these BAC DNAs as templates and one degenerate forward primer corresponding to FEYAIGHW and another reverse primer corresponding to another conserved block of residues, QYDQGEDRW, found in the catalytic domains of other insect CHSs (Tellam et al., 2000; Ibrahim et al., 2000; Zhu et al., 2002; Zimoch and Mer-
zendorfer, 2002; Gagou et al., 2002; Ostrowski et al., 2002), yielded amplification products with sizes of either 242 or 192 bp. The DNA from all eight colonies
with strong hybridization to the probe yielded the 242 bp product, whereas the four colonies with the weaker hybridization yielded the 192 bp PCR product. Sequencing of these fragments from four strongly hybridizing and four weakly hybridizing BAC clones indicated that the sequences of all of the members within the same size group were identical except in the degenerate primer regions. The DNA sequence of the larger PCR product was similar to the smaller PCR product (58.1% identity) with the additional 50 bp subsequently confirmed to be an intron. The encoded peptide sequences of both PCR fragments (see underlined sequences in Fig. 3) were highly similar to a conserved region found in insect and nematode CHSs. At this point, the gene that generated the 242 bp fragment was named TcCHS1 and that which amplified the 192 bp fragment was named TcCHS2.

3.2. Sequence and organization of TcCHS1 and TcCHS2 genes

One BAC clone containing the TcCHS1 gene and a second clone containing the TcCHS2 gene were chosen for the determination of the complete sequences of these genes. These two DNAs were used as templates in separate PCR reactions with different combinations of degenerate primers corresponding to conserved regions of insect/nematode CHSs. The designations of the primers, the sizes of the PCR products, and their relationship to the corresponding genes are indicated in Fig. 1. The sequences of the degenerate and gene-specific primers as well as the corresponding conserved amino acid sequences are shown in Table 1. In some cases, a single gene-specific CHS primer amplified the 5’- or 3’-end fragment presumably via a type of universal PCR strategy, one end of each such amplified fragment having been misprimed. In one case, an inverse PCR reaction was used to obtain the 5’-end sequence extending upstream of the protein-coding region (Fragment W-1). These strategies resulted in a series of overlapping PCR fragments and it was then possible to assemble the DNA sequences of the entire protein coding regions as well as both the 5’- and 3’-flanking sequences.

About 15.6 and 8.1 kb of genomic DNA sequence were determined for TcCHS1 and TcCHS2, respectively. By using BLASTX to align these sequences with protein sequences of insect CHSs in the NCBI database, it was possible to predict the ORFs and the exon–intron organization of the TcCHS1 and TcCHS2 genes (Fig. 2). For confirmation, we amplified the entire protein coding regions by RT-PCR using preupal poly(A)-RNA as template, and TcCHS1- and TcCHS2-specific forward and reverse primers. cDNAs of 4.7 kb and 4.4 kb were amplified corresponding to the ORFs of TcCHS1 and TcCHS2, respectively (data not shown). The TcCHS1 PCR product was actually a mixture of two sequences as a result of alternate exon splicing (see below). The DNA sequences of these PCR products and the 5'- and 3'-RACE PCR products confirmed the predicted exon–intron boundaries, as well as independently checking the accuracy of comparable sequences generated from the BAC clones. These three cDNA sequences have been deposited in the GenBank database with accession numbers AY291475, AY291476 and AY291477. The TcCHS1 and TcCHS2 genomic DNA sequences have been assigned GenBank accession numbers of AY295880 and AY295879, respectively.

The TcCHS1 gene has 10 exons and nine introns, whereas the TcCHS2 gene has eight exons and seven introns (Fig. 2). Approximately 1 kb of the promoter region of TcCHS1 and 1.1 kb of the promoter region of TcCHS2 have been sequenced. We identified putative TATA boxes upstream of exon 1 and polyadenylation sequences in the 3'-untranslated regions in both genes. The start of translation of TcCHS1 is in exon 2 (which is preceded by a long intron), whereas it is in exon 1 of TcCHS2. It is clear that the organizations of the two genes in Tribolium are different. Two introns are in identical positions in the two genes (introns 3 and 8 of TcCHS1 and introns 4 and 6 of TcCHS2, respectively), while others are at variable positions (Figs. 2 and 3). The introns range in size from 46 to 5623 bp. The most interesting difference between the two genes is the presence in tandem of two non-identical copies of exon 8 (denoted as 8a and 8b) in TcCHS1, whereas TcCHS2 has only one copy of the corresponding region as a part of exon 6 (Fig. 2).

Also shown in Figs. 2 and 3 is a comparison of the exon–intron organization of CHS genes from T. castaneum, together with those of D. melanogaster and A. gambiae. Both species have one gene related to TcCHS1 and one related to TcCHS2. Presented is a compilation from analysis of genomic sequences from their respective genome projects, using cDNAs available as separate sequence files submitted to GenBank and various BLAST queries to identify exons and introns. It is clear that the organization of insect chitin synthase genes has diverged among the three species compared, which included two dipterans and one coleopteran. Only one intron has remained in the same position in all six CHS genes (see Fig. 3). Like TcCHS1, both DmCHS1 and AgCHS2 have the start of translation in exon 2 and have alternate exons equivalent to exons 8a and 8b of TcCHS1. In all three of these genes, the positions of introns that flank the alternate exons are conserved and the exons are identical in length, encoding 59 amino acids. Tellam et al. (2000) previously reported the identification of the alternate exons (7a and 7b) in D. melanogaster CHS1. The Anopheles CHS2 gene, which has high sequence
similarity with DmCHS1 and TcCHS1, has the equivalent alternate exons 7a and 7b. The longest ORFs found in the cDNAs were 4674 bp in TcCHS1 (with either exon 8a or 8b present, see below) and 4392 bp in TcCHS2, which are capable of coding for proteins with 1558 and 1464 amino acids, respectively. The amino acid sequence identity between the two proteins, TcCHS1B and TcCHS2, is 47.5% and similarity is 71.0%. The identity and similarity between TcCHS1A and TcCHS2 are 47.8% and 71.1%, respectively. The sequences of these two TcCHS proteins are compared to CHSs from Drosophila and Anopheles in Fig. 3, which are the only insect species for which both CHS protein sequences are currently available. As with other insect CHSs including those of the sheep blow fly (Tellam et al., 2000), the TcCHSs have an N-terminal domain with several membrane-spanning regions, a central domain with high sequence identity to putative catalytic domains of CHSs of fungi, nematodes and other insects, and a C-terminal domain with multiple membrane spanning regions. TcCHS1 is predicted to have nine transmembrane segments in the N-terminal domain, whereas the TcCHS2 protein is expected to have only five such segments. Using the Paircoil Program (Lupas et al., 1991), a coiled-coil domain is predicted with high probability to occur in TcCHS1 (see Table 2) but not in the TcCHS2 protein.

3.3. Expression of TcCHS1 and TcCHS2 during development of Tribolium

To determine whether the two TcCHS genes are differentially expressed at various stages of development of Tribolium, we analyzed cDNA made from RNA isolated at different developmental stages including embryos, early larvae (penultimate larval instar), late larvae (last larval instar), prepupae, pupae and adults by PCR using primers specific for TcCHS1 and TcCHS2. PCR reactions were carried out with each cDNA preparation as the template and the two pairs of PCR primers for TcCHS1 and TcCHS2 in the same tube. Control experiments with a mixture of equal amounts of TcCHS1 and TcCHS2 cDNAs as templates indicated that the PCR amplification efficiencies were similar with the two sets of primers whose Tm’s, lengths and concentrations were nearly the same (bottom panel of Fig. 4B). As shown in Fig. 4A, the expected PCR product (1546 bp) corresponding to fully spliced TcCHS1 transcripts was detected in embryos, late larvae, prepupae and pupae but not in early instar larvae. Trace amounts of TcCHS1 transcripts were observed in early larvae and young adults but not in mature adults (Fig. 4C). No PCR products with the size expected for fully spliced transcripts of TcCHS2 (1345 bp) were seen at the embryonic stage, but they were detected in both early- and late larvae (Fig. 4A, lanes 2 and 3). TcCHS2 transcripts were not detected at the prepupal and pupal stages. However, in mature adults, only TcCHS2 transcripts were detected. Although the PCR data are only qualitative, they do suggest differences in the developmental pattern of expression of TcCHS1 and TcCHS2 genes. TcCHS1 is expressed at the pupal stage, whereas TcCHS2 is expressed at the late larval stage and in mature adults. Similar analyses with RNA preparations from male and female pupae, and

---

**Fig. 2.** Schematic diagram of the organization of the TcCHS1, TcCHS2, DmCHS1, DmCHS2, AgCHS1 and AgCHS2 genes. Boxes indicate exons. Lines indicate introns. The second of the two alternative exons (8b) of TcCHS1, DmCHS1(7b), and AgCHS2(7b) are indicated as closed boxes. About 15.6 kb of the TcCHS1 and 8.1 kb of TcCHS2 gDNA sequences were compared to their respective cDNA sequences to define the exons and introns. The exon-intron organization of the other four CHS genes was deduced partially from comparisons of available cDNA and genomic sequences but also from the sequence analysis described in Section 2.
Fig. 3. Alignment of deduced amino acid sequences of TcCHS1, TcCHS2, DmCHS1, DmCHS2, AgCHS1 and AgCHS2 using ClustalW software. Transmembrane regions predicted using TMHMM software (v. 2.0) are shaded. The positions in the protein sequences of TcCHS1 and TcCHS2 where coding regions are interrupted by introns are indicated by shaded arrow heads. Intron 1 of TcCHS1 lies in the 5' UTR region two nucleotides 5' of the translation start and is not indicated in this figure. The putative catalytic domains are boxed. Symbols below the aligned amino acid sequences indicate identical (\(\square\)), highly conserved (:), and conserved residues (.). The regions in TcCHS1 and TcCHS2 corresponding to the PCR probe made from two degenerate primers representing two highly conserved sequences in CHSs are underlined.
young adults and mature adults (Fig. 4C, top panel) indicated that there were no significant differences in expression of the two CHS genes between males and females.

3.4. Alternate exon usage in generation of TcCHS1 transcripts

The presence of two alternate forms of exon 8 in the TcCHS1 gene suggests the possibility of alternate exon usage in generating two different transcripts encoding two different TcCHS1 proteins. To investigate whether alternate exon usage occurs, we analyzed cDNA prepared from RNA isolated from Tribolium at different stages of development for the presence of exon 8a or 8b sequences of TcCHS1 using a forward primer in exon 5 and reverse primers specific for either exon 8a or 8b. A PCR fragment of the expected size (806 bp) for a mature transcript containing exon 8a was detected in all stages of development as described in Section 3.3. Transcripts were easily detected in prepupal and pupal cDNA (Fig. 4B, top panel, lanes 4 and 5), perhaps reflecting a high abundance at these stages. PCR products containing exon 8b sequences were prominent in pupal cDNA (lane 5), but they were barely detectable in embryonic, late larval and prepupal cDNAs (lanes 1, 3, and 4), and were undetectable in early larval and adult cDNAs (lanes 2 and 6). These data indicate that the utilization of exons 8a and 8b is variable throughout the different stages of Tribolium development. The developmental stages at which transcripts with exon 8b accumulate are apparently much more restricted than those in which exon 8a transcripts are expressed.

<table>
<thead>
<tr>
<th>Insect</th>
<th>Presence of alternate exon</th>
<th>Coiled-coil domain Probability</th>
<th>Position*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaCHS1</td>
<td>No</td>
<td>None</td>
<td>–</td>
</tr>
<tr>
<td>AgCHS1</td>
<td>No</td>
<td>None</td>
<td>–</td>
</tr>
<tr>
<td>AgCHS2</td>
<td>Yes</td>
<td>High I</td>
<td>I</td>
</tr>
<tr>
<td>DmCHS1</td>
<td>Yes</td>
<td>High I</td>
<td>I</td>
</tr>
<tr>
<td>DmCHS2</td>
<td>No</td>
<td>None</td>
<td>–</td>
</tr>
<tr>
<td>TcCHS1</td>
<td>Yes</td>
<td>High I</td>
<td>I</td>
</tr>
<tr>
<td>TcCHS2</td>
<td>No</td>
<td>Low II</td>
<td>II</td>
</tr>
<tr>
<td>MsCHS1</td>
<td>Yes</td>
<td>High I</td>
<td>I</td>
</tr>
<tr>
<td>LcCHS1</td>
<td>Yes</td>
<td>High I</td>
<td>I</td>
</tr>
</tbody>
</table>

* Position I is immediately after the 5-TMS region. Position II is 40–50 amino acid downstream from 5-TMS region.

b Prediction based on comparison with other insect CHSs.

3.5. Comparison of the sequences of exons 8a and 8b of TcCHS1

Exons 8a and 8b of TcCHS1 are both 177 bp-long and exhibit a nucleotide sequence identity of 63%.
The nucleotide sequence identities of exon 8a and 8b to the corresponding region in TcCHS2 are 58% and 55%, respectively. Both exons code for 59-amino acids and include a transmembrane segment of 20 amino acids in the middle of the sequence. The protein sequences encoded by exons 8a and 8b have an amino acid sequence identity of 70% and a similarity of 85% (Fig. 5). The N-terminal sequences of these exon-encoded peptides are postulated to be extracellular and the C-terminal segment is predicted to be exposed to the cytoplasm. The transmembrane sequence is more highly conserved than the two flanking sequences. Even though TcCHS2 has a stretch of amino acid sequence in exon 6 equivalent to exon 8a- or 8b-encoded segments including a transmembrane segment, it has a much lower level of amino acid sequence identity with these segments (only 54% and 44%, respectively).

3.6. Association between the presence of the coiled-coil domain and alternate exons

An alignment of the amino acid sequences corresponding to exons 8a and 8b of TcCHS1 and equivalent regions from several other insect CHSs is shown in Fig. 6. While there is substantial amino acid sequence identity among all of these sequences, the identity is greater within the two subgroups of sequences related to exon 8a or exon 8b of TcCHS1 (Fig. 6, top and bottom panels, respectively). Table 2 indicates that there is a correlation between the presence of the coiled-coil region (as predicted by the Paircoil Program) that immediately follows the catalytic domain and the presence of alternate exons. Only insect genes that code for CHSs with the coiled-coil region immediately following the five-transmembrane span (5-TMS) region have alternate exons corresponding to the exons 8a and 8b of TcCHS1 gene. The insect CHS genes, TcCHS2, DmCHS2, AaCHS1 and AgCHS1, which do not encode a CHS with the coiled-coil region immediately following the 5-TMS region, do not have the alternate exons. Included in this list of genes containing alternate exons is the LcCHS1 gene from L. cuprina whose genomic DNA sequence has not yet been published. However, the cDNA sequence of LcCHS1 (Tellam et al., 2000) encodes a CHS protein that has a high probability of having a coiled-coil domain. Based on this and the alignment of the alternate exon-encoded protein region to other insect CHSs, we predict that this gene also will have alternate exons. A similar prediction that the MsCHS1 gene has alternate exons was confirmed experimentally by PCR amplification and DNA sequencing of the region. The alternate exons from M. sexta CHS1 also encode 59-amino acid-long stretches with sequence similarities to correspond-
3.7. Mapping of TcCHS1 and TcCHS2 to Tribolium chromosomes

TcCHS1 and TcCHS2 were mapped to linkage groups 5 and 9, respectively, using SSCP analysis as described in Section 2.9 (at map positions 46.8/54 cM and 20.2/62.2 cM; Beeman et al., unpublished data). This result was unexpected because in the only two other insect species in which all of the CHS genes have been mapped, namely D. melanogaster and A. gambiae, the genes are linked. For D. melanogaster, they are positioned about 3 Mb apart on chromosome 3 and for A. gambiae, about 8 Mb apart on chromosome 2 (Gagou et al., 2002; results of TBLASTN search of A. gambiae genome database at NCBI).

3.8. Southern blot analysis of T. castaneum genomic DNA for detection of CHS genes

To support that there are only two CHS genes in the Tribolium genome, Southern blot analysis of genomic DNA was carried out using five different restriction enzymes and a radioactive probe amplified from the highly conserved catalytic domain TcCHS1 fragment as outlined in Section 2. In each digest, only two bands could be seen, one corresponding to TcCHS1 and the other corresponding to TcCHS2. The observed sizes, when compared with those that could be predicted from the TcCHS gene sequences, were in agreement (data not shown). No additional bands were detected under moderate stringency conditions. Hybridization of the probe to the other CHS gene (and to Drosophila CHS genes) was easily detected (Fig. 7) supporting our conclusion from the screening of the BAC library that there are only two genes for CHS in the Tribolium genome.

4. Discussion

We have identified several clones containing CHS genes of Tribolium from a screening of a BAC library.
Because the probe used for the screening was from the highly conserved region that encodes the catalytic domain of CHSs, we anticipated that it would hybridize under reduced stringency to different genes encoding CHS. Based on the strength of hybridization signals, length of PCR amplification products, and finally sequencing of the amplified DNAs from several BAC clones, we obtained evidence for the presence of only two CHS genes in Tribolium. The data from Southern blot analysis also indicate the presence of only two genes in Tribolium, which is consistent with the results obtained from studies of other insect species whose CHS genes have been identified. Search of the databases for the fully sequenced genomes of D. melanogaster, A. gambiae, and of the nematode C. elegans revealed the presence of two genes in each of these invertebrates. Even though only one CHS gene has been isolated from L. cuprina and M. sexta, it is likely that these insects also have two CHS genes (Tellam et al., 2000; Zhu et al., 2002; unpublished data from H. Merzendorfer, Osnabrueck University, and our laboratories).

The number and positions of introns in the two TcCHS genes are not conserved. Only two introns are in equivalent positions in the two TcCHS genes. Other exons are fused or split to form exons that differ in size between the two genes. The start of translation is in exon 2 of TcCHS1 and in exon 1 of TcCHS2. At the protein level, the C-terminal domain of the TcCHS2 protein is considerably shorter than that of TcCHS1. In Drosophila and Anopheles, the CHS1 and CHS2 genes are closely linked (approximately 3 and 8 Mb apart in their respective species), indicating that the two arose via a tandem duplication and divergence from a single progenitor gene. Our discovery that the two genes are not linked in Tribolium suggests more extensive genome rearrangement in beetles than in flies, at least for this region of the genome. In the nematode, C. elegans, the CHS genes occur on separate chromosomes (C. elegans genome database, NCBI), like in Tribolium.

The two CHS genes characterized in this study encode proteins that have amino acid sequences closely related to those from other insects and nematodes, and more distant from those of fungi (Fig. 8). All members of the CHS-A class for which genomic DNA sequences are available have two alternate forms of the exon corresponding to exon 8 of T. castaneum CHS1. The proteins encoded by the CHS-A class are slightly larger than the CHSs of the CHS-B class and are predicted to have a coiled-coil region immediately following the TMS region. No coiled-coil region was predicted for the CHS-B group of proteins. The coiled-coil regions are potential sites for protein–protein interactions and/or signals for vesicular trafficking (Melia et al., 2002). It is interesting that both alternate exons encode a highly conserved transmembrane domain, but have more variable sequences in the flanking regions that are predicted to be exposed to the external surface of the plasma membrane and to the cytosol. It is conceivable that these exposed domains interact with different regulatory molecules including those involved in vesicular trafficking. It is also worth noting that the corresponding region in TcCHS2 has fewer amino acid sequence identities with TcCHS1, even though it has a high similarity to the exon 8a- or 8b-encoded sequences (Fig. 5).

Alternate splicing has been shown to alter the localization of a human plasma membrane Ca\(^{2+}\) ATPase from a basolateral location to a predominantly apical location (Chicka and Strehler, 2003). In that case, a cytosolically exposed 45 amino acid-long region was shown to be responsible for this change in distribution of the Ca\(^{2+}\) ATPase.

The qualitative differences observed in the developmental patterns of expression of the two TcCHS genes are consistent with the enzymes encoded by these two genes having different physiological functions. In embryos, only TcCHSI transcripts are detectable suggesting that this enzyme is involved in the synthesis of chitin associated with embryonic cuticle and mouth-
The transcripts for TcCHS2 are not detectable in embryos or pupae but are expressed in the last larval stage and in adults when there is active PM synthesis, suggesting that TcCHS2 is associated with formation of PM-associated chitin. Because we have isolated RNA from whole larvae (and not from gut and epidermal tissues free of cross-contamination from other tissues), we are unable to ascertain whether TcCHS2 is exclusively associated with midgut chitin synthesis. Tellam et al. (2000) investigated tissue specificity of expression of LeCHS-1 (equivalent to TcCHS1) and concluded that this gene was expressed in the carcass (free of internal tissues) but not in cells of the midgut. Their in situ hybridization experiments also confirmed this tissue specificity of expression of LeCS-1. The data presented in this paper provide experimental support for the hypothesis that midgut chitin synthesis may be a major function of TcCHS2. Gagou et al. (2002) studied stage-specific expression of CHS genes in Drosophila using third instar larvae and prepupae, but they did not address the tissue specificity of expression of individual CHS genes. An increase in the level of TcCHS1 mRNA was observed during pupation followed by a decline to undetectable levels during the adult stage. Thus, TcCHS1 very likely plays an important role in the formation of the pupal cuticle. On the other hand, TcCHS2 transcript levels decline between the larval and pupal stages, and then reappear in adult insects. The developmental patterns of expression of TcCHS genes suggest that TcCHS2 is not essential during the embryonic stage and possibly during the pupal stage. The high levels of TcCHS2 mRNA in late larval stadia and in adults may be indicative of a role for the TcCHS2 protein in the production of the chitin-rich PM.

Developmental regulation of alternate exon usage appears to determine which form of the TcCHS1 transcripts accumulate during different growth stages. In the embryonic stage, all of the CHS transcripts detected are derived from TcCHS1 with either exon 8a or exon 8b. TcCHS1 transcripts are undetectable in early instar larvae but are present in the last instar larvae and prepupae, and these contain predominantly exon 8a. The greatest amounts of TcCHS1 transcripts (with either exon 8a or exon 8b) occur at the pupal stage when no TcCHS2 transcript was detectable, suggesting the requirement of either a different type or larger amounts of CHSs at this stage. The appearance of exon 8b-containing transcripts is confined predominantly to the pupal stage, which is just about the time ecldysteroid titers are expected to decline based on analogy with hormonal levels in the tobacco hornworm (Riddiford, 1994). The finding that transcripts with exon 8a are expressed over a wider range of developmental stages than transcripts with exon 8b suggests that alternate splicing may be under hormonal and/or developmental control. However, this possibility needs to be investigated further using tissues lacking endogenous sources of hormones and neurotransmitters (Kramer et al., 1993). DmCHS1 transcripts containing both exon 7a and 7b have been reported (NM_169052; Drosophila EST database: http://www.fruitfly.org/cgi-bin/annot/gene*kkv). We have found some evidence for the presence of TcCHS1 transcripts containing both exons 8a and 8b, but they are very rare relative to transcripts containing only exon 8a or exon 8b (Arakane et al., unpublished data).

We have analyzed the DNA sequence of the region immediately upstream of the putative transcription start site of TcCHS1 and TcCHS2 for the presence of sequences with consensus ecldysone-response elements, EcRE (Palli and Retnakaran, 1999), and consensus sequences for the binding of transcription factors, BR-C and E74a (Thummel, 1996). In the approximately 1 kb promoter regions of both TcCHS1 and TcCHS2, we have identified several putative EcRE and BR-C binding sequences that reasonably match the consensus sequences for these elements even though these matches were not perfect (data not shown).

Alternative splicing plays a major role in modulating gene function by expanding the diversity of expressed mRNA transcripts (Brett et al., 2000). While seeking to understand the repertoire of CHS genes and the proteins encoded by them in insects, we have obtained evidence for alternative splicing of pre-mRNA for CHS in Tribolium, Drosophila, Anopheles and Manduca also have CHS genes (CHS-A class, Fig. 8) that have alternate exons corresponding to Tribolium exon 8 and, therefore, the potential to form two CHS1 enzymes as a result of alternative splicing. The enzymes of the CHS-B class, on the other hand, are predicted to have only a single form. Alternate splicing to generate multiple isoforms of a serine protease inhibitor with different inhibitory specificities from a single gene with several alternate exons has been described in M. sexta (Jiang et al., 1996). Similarly, alternate splicing leads to different isoforms of tropomyosin in C. elegans and the Pdpl protein in D. melanogaster (Anyanful et al., 2001; Reddy et al., 2000). In this paper, we have provided experimental evidence that alternate exon usage does occur in T. castaneum for the purpose of chitin synthesis and that this process is developmentally regulated. In light of these findings, the contribution of alternative splicing to CHS diversity in other insect species and factors that control this process will need to be studied in the future. In addition, the significance of differences in biochemical properties and physiological functions of such alternate forms of CHSs in insects will be worthy of investigation.
6. Acknowledgments

The authors are grateful to Drs. Hans Merzendorfer, Juliet Fuhrman and Daizo Koga for critical comments. Support was provided in part by the NSF grant IBN-0316963. Yasuyuki Arakane was supported by an ARS-USDA postdoctoral fellowship. CAS acknowledges support from NIH grant RO1 A125780 to S.M. Levitz. Mention of a proprietary product does not constitute a recommendation or endorsement by the USDA. The USDA is an equal opportunity/affirmative action employer and all agency services are available without discrimination. This is contribution #03-391-J of the Kansas Agricultural Experiment Station.

References


