cDNA cloning and deduced amino acid sequences of three storage proteins in the common cutworm, *Spodoptera litura* (Lepidoptera: Noctuidae)

Yiping Zheng, Toyoshi Yoshiga and Sumio Tojo*

Laboratory of Applied Entomology, Department of Applied Biological Sciences, Saga University, Saga 840-8502, Japan
(Received 7 September 1999; Accepted 10 November 1999)

Abstract
Three storage proteins, hexamers of 80–90 kDa subunits, have been purified and characterized from the common cutworm, *Spodoptera litura*. We have cloned the cDNAs of these storage proteins by immunoscreening and rapid amplification of cDNA ends (RACE). The deduced amino acid sequence of SL-1, a basic protein containing 7.7% methionine showed 82.3 and 68.4% identities to the *Trichoplusia ni* basic juvenile hormone-suppressible storage protein (8.4% methionine) and to the *Bombyx mori* methionine-rich storage protein (11.8% methionine), respectively. The amino acid sequences of α and β subunits with ca 4% methionine of SL-2, another basic protein, were 84% identical to that of *T. ni* basic juvenile hormone-suppressible storage protein with moderate methionine content (5.3%), but showed only 45% identity to SL-1. SL-3, an arylphorin was very similar (54–63% identity) to known other arylphorins, but differed considerably (31–35% identity) from SL-1 and SL-2 in their amino acid alignments. The most parsimonious tree obtained by systematic analyses of the sequence alignments of 16 lepidopteran storage proteins demonstrated that three storage proteins of *S. litura* are clustered into three sister groups, although they share extensive similarities throughout the alignment.

Key words: Methionine-rich storage protein, arylphorin, cDNA nucleotide sequence, deduced amino acid sequence, *Spodoptera litura*

INTRODUCTION

In insects, storage proteins, hexamers of ca 80 kDa subunits, called hexamers, are synthesized in the fat body predominantly during the final larval stadium and partially or totally sequestered into the fat body before pupation. There are several different groups of storage proteins; in Lepidoptera, at least two representative groups of storage proteins are known, i.e., aromatic amino acid-rich arylphorins and methionine-rich storage proteins. Although arylphorins have been identified in many species of insects (Kanost et al., 1990; Telfer and Kunkel, 1991), methionine-rich storage proteins have been characterized in only a few species, such as *Hyalophora cecropia* (Tojo et al., 1978), *Bombyx mori* (Tojo et al., 1980), and *Manduca sexta* (Ryan et al., 1985). Phylogenetic relationships among these groups have been demonstrated by comparison of amino acid sequences (Beintema et al., 1994; Burmester and Scheller, 1996; Burmester et al., 1998) and also among arylphorins of Lepidoptera (Shimada et al., 1995).

In a previous study, we isolated three storage proteins SL-1, SL-2 and SL-3, from the haemolymph of the last stadium larvae of the common cutworm, *Spodoptera litura*. These proteins are hexamers of subunits which range in size from 70 to 80 kDa (Tojo and Yoshiga, 1994). SL-1 and SL-2 are basic proteins rich in methionine, but they react differently with antiserum against methionine-rich storage protein, SP-1 of *B. mori*. SL-3 is an arylphorin having 20% aromatic amino acids. It reacts with the *B. mori* SP-2 (arylphorin) antiserum and differs markedly from SL-1 and SL-2 in immunological characteristics (Tojo and Yoshiga, 1994). SL-1 and SL-2 first appear in the haemolymph of the final instars and are totally sequestered by the fat body shortly before pupa-

*To whom correspondence should be addressed at: E-mail: tojo@cc.saga-u.ac.jp*
tion (Yoshiga et al., 1998). Their synthesis is blocked by juvenile hormone (Tojo et al., 1985). Thus, these two proteins appear to be similar to two basic juvenile hormone-suppressible proteins of Trichoplusia ni, which are also rich in methionine (Jones et al., 1987, 1993).

In this paper we report the complete cDNA and deduced amino acid sequences for three storage proteins of S. litura. The relationships and phylogenies of these proteins are appraised by comparing the amino acid sequences with other storage proteins in Lepidoptera.

**MATERIALS AND METHODS**

**Experimental insects.** The common cutworm (S. litura) eggs were supplied by Sumitomo Chemical Co. Ltd. The larvae of S. litura were reared at 25°C under a 16 h light-8 h dark photoperiod on an artificial diet, as described by Okada (1977).

**Purification of storage proteins and preparation of antisera.** Three storage proteins, SL-1, SL-2 composed of α and β subunits and SL-3 were purified from the haemolymph of the sixth (final) stadium larvae and polyclonal antisera were prepared, as previously described (Tojo and Yoshiga, 1994). Purified proteins were run on SDS-PAGE (7.5% gel), transferred to a PVDF membrane and subjected to N-terminal sequencing (Matsudaira, 1987).

**Construction of a cDNA library and immunoscreening.** Messenger RNA was isolated directly from the fat bodies of the final stadium larvae (day 3) by using an mRNA purification kit (QuickPrep Micro, Pharmacia Biotech). cDNA was synthesized by reverse transcription of poly(A)+ RNA with an oligo(dT)12-18 primer (cDNA synthesis kit, Pharmacia Biotech) and a λ gt11 cDNA library was constructed using a commercially available system (NipponGene). To select positive phage clones, immunoscreening was used with polyclonal antibodies against SL-1, SL-2 and SL-3. After tertiary screening, phage DNA amplified by PCR was digested with NotI and the DNA was subcloned into pBluescript II SK(+) and KS(+) vectors.

**Rapid amplification of cDNA ends (RACE).** Messenger RNA was isolated from fat body of 2 day-old sixth stadium larvae by using MicroFast track (Invitrogen). Rapid amplifying of cDNA ends (RACE) was performed by using marathonone DNA synthesizing Kit according to the manufacturer’s instructions (Clontech). Two degenerated primers [5'-ACNGAYGAYAY-GAYGYYT-3' and 5'-GAYGAYTAYAYA-CNAAYAYTT-3'] (N stands for any, Y for C or T) were designed from the N-terminal sequences of SL-2 α subunit (TDDNDNL) and β subunit (DYNTNNF), respectively (Fig. 1, B, C). PCR conditions were 94°C for 1 min, 50°C for 30 s and 72°C for 1 min (29 cycles). To obtain the sequences of the 5' end for a subunit of SL-1 and α and β subunits of SL-2, gene specific primers were designed (SL-1, 5'-GCAGCTCTGCTCATGGAAAC-3'; SL-2, 5'-GCTTCTGAAATGACATCACGC-3'; Fig. 1, A, B, C), and touchdown PCR was performed of 94°C for 1 min, and for 26 cycles of 94°C for 30 s and 68°C for 3 min. PCR products were run on a 1.5% agarose gel, the expected DNA bands were excised from the gel, and the DNA was isolated by using QIA quick (Qiagen). DNA fragments were ligated into pGEM-T vector with a TA cloning kit (PROMEGA).

**Sequence and phylogenetic analysis.** DNA sequencing was done by using ABI PRISM DNA sequence kit (Applied Biosystems) and ABI 373A sequencers. Deletion clones were made by using Exo Mung deletion Kit (Stratagene). To confirm sequences of clones obtained by RACE, at least three clones were sequenced. DNA analysis was done with GENETYX-MAC/ATSQ (Software Development Co., Ltd.), and homology search was carried out with the BLAST (Pearson and Lipman, 1988) and FASTA (Altschul et al., 1994) programs. Highly scoring sequences were aligned using the ClustalX alignment program (NCBI Toolkit 1997), setting the gap opening at 3 and gap extension at 1 in the multiple alignment parameter option. This alignment was slightly modified manually and further examined. Distances between pairs of protein sequences were calculated from the ClustalX directly and corrected for multiple change according to Dayhoff’s (1979) PAM 001 matrix by using the PROTDIST of the PHYLIP package (Felsenstein, 1993). The distance trees were constructed by the neighbor-joining method.
(Saitou and Nei, 1987) using the NEIGHBOR program of PHYLIP. Bootstrap (Felsenstein, 1985) analysis with 100 replicates was conducted to assess the degree of support for each branch on the strict consensus tree.

The following lists the names of proteins utilized in these analyses with their abbreviations used in the table and figures, EMBL/GenBank database accession numbers and references: Galleria mellonella arylphorin [GmeAr; M73793 (Memmel et al., 1992)], M. sexta arylphorin a [MseaAr; P14296 (Willott et al., 1989)], M. sexta arylphorin β [MseaAr; P14296 (Willott et al., 1989)], M. sexta methionine-rich storage protein [MseSPM; L07690 (Wang et al., 1993)], H. cecropia arylphorin [HCeAr; AF032396 (unpublished)], H. cecropia methionine-rich storage protein [HceSP1M; AF032398 (unpublished)], H. cecropia methionine-moderate rich storage protein [HceSP2m; AF032399 (unpublished)], B. mori arylphorin [BmSP2Ar; P20613 (Fujii et al., 1989)], B. mori methionine-rich storage protein 1 [BmSP1M; P09179 (Sakurai et al., 1988)], Hyaehy whole hormone-suppressible storage protein 1 [TniBJHSP1M; L03280 (Jones et al., 1993)], T. ni basic juvenile hormone-suppressible storage protein 2 [TniBJHSP2mM; L03281 (Jones et al., 1993)].

RESULTS AND DISCUSSION

We obtained positive clones for SL-1 and SL-3 by immunoscreening of a cDNA library constructed from fat body of the final stadium larvae using antisera against SL-1, SL-2 and SL-3. While the clone for SL-3 (2.3 kbp) contained the whole open reading frame (ORF), the clone for SL-1 (2.3 kbp) did not contain the initiation codon. To obtain the 5' end of the SL-1 clone, we performed a 5' RACE using a gene specific primer. A clone containing the initiation codon was obtained. The deduced amino acid sequences of the clones for SL-1 and SL-3 matched the N-terminal sequences of the SL-1 (NVVKDTSNVFNIGKDNMNVD) and SL-3 (SAYPHHDLILLK) obtained by N-terminal sequencing (Fig. 1, A, D).

As we failed to obtain the clones for SL-2 subunits by immunoscreening, we tried 3' RACE using degenerated primers designed from the N-terminal sequences of two subunits of SL-2. About 2.2 kb clones for α and β subunits were obtained by the RACE experiment. Sequences encoding the 5' ends of each subunit were obtained by 5' RACE using a gene specific primer for SL-2 α subunit (5'-GGCTTTCTGAATGACCATCGCGC-3', see Fig. 1B). As the sequences encoding subunits of SL-2 were very similar to each other, clones encoding β subunits were also obtained with the same primer. The deduced amino acid sequences for two subunits were matched to the results of the protein sequencing from the N-terminus (α: TPLTDDNRLDVT; β: RRRTDDYNTNNFTMDIKQR).

The cDNA sequences encoding these three storage proteins thus obtained and the deduced amino acid sequences are shown in Figs. 1 and 2. The contents (mol%) of methionine in SL-1, SL-2 α-, β-subunits, and SL-3 were 7.7, 3.5, 3.9 and 1.9, respectively, while the contents of aromatic amino acids were 9.4, 13.1, 13.7 and 19.1. Thus, SL-1 could be characterized as being rich in methionine, SL-2 as moderately rich in methionine and SL-3 as arylphorin. One presumed recognition site for the N-glycosylation site was found in SL-3.

As we demonstrated in the previous study, only SL-1 showed a cross-reactivity to antisera against methionine-rich SP-1 (BmoSP1M) of B. mori under denatured condition, while both SL-1 and SL-2 reacted with the antisera on native molecules (Tojo and Yoshiga, 1994). Amino acid alignments deduced from nucleotide sequences were compared among them (Figs. 1 and 2). Identity of amino acid alignments between SL-1 and BmoSP1M was 68.2%, while the value between SL-2 (α- and β-subunits) and BmoSP1M was 44%. Henceforth, both subunits will be referred to as SL-2, if significant differences in description are not specified between the two subunits. As shown in Fig. 2, there were 20 alignments with identical sequences composed of more than 6 amino acid residues between SL-1 and BmoSP1M, among which 9 alignments were composed of more than 9 residues (maximum 17). Half of these long identical alignments were highly hydrophilic (data not shown), which may expose these
areas on the outer sides of the native molecules and contribute to their function as epitopes for the induction of the antibodies to be cross-reactive to both proteins, even at the peptide level.

On the other hand, only 2 identical alignments composed of a maximum of 7 residues were found between SL-2 and BmoSP1M (Fig. 3). Thus, the identical alignments in these molecules might be too short to function as the epitopes for induction of a common antiserum at the peptide level. It appears reasonable to assume that the epitopes bridging two subunit molecules of SL-2 functions to induce antibodies to be cross-reactive with both molecules, only in the native state.

SL-3 shows a serological similarity to BmoSP2Ar, but not to BmoSP1M (Tojo and Yoshiga, 1994). Coinciding with this fact, SL-3 showed rather high identity (58.4%) in amino acid sequence to BmoSP2Ar, but low identities (31–35%) to SL-1 and SL-2 (Figs. 1 and 3). Between SL-3 and BmoSP2Ar, there were 6 identical alignments composed of more than 9 ami-
no acid residues, among which 3 alignments were composed of more than 10 residues (maximum 16). These long identical sequences may have functioned as an epitope to induce an antibodies common to them. Among SL-1, SL-2 and BmoSP2Ar, only 2 identical alignments composed of more than 4 amino acid residues (maximum residues: 6) were detected (data not shown).

SL-1 and SL-2 are basic juvenile hor-
Fig. 2. Multiple alignments of the amino acid sequences from storage proteins, SL-1, SL-2 α and β subunits and SL-3 of *S. littoralis* (present study), BmoSP1M and BmoSP2Ar of *B. mori* (Sakurai et al., 1988; Fujii et al., 1989), and TniJHSP1M and TniJHSP2mM of *T. ni* (Jones et al., 1993). Shadows denote conserved positions.
mone-suppressible proteins (Tojo et al., 1985; Tojo and Yoshiga, 1994). In fact, values of isoelectric points for SL-1, SL-2 and SL-3 calculated from the deduced amino acid sequences (Fig. 1) were 9.3, 8.9–9.1 and 6.3, respectively. These values nearly corresponded to the values, 9.3, 8.6 and 6.3, estimated from the profiles of the proteins by isoelectrofocusing chromatography in a previous report (Tojo and Yoshiga, 1994). Thus, we can say that SL-1 and SL-2 are similar in their physiological and chemical characteristics to the two basic juvenile hormone-suppressible proteins (TniBJHSP1M and TniBJHSP2mM) of T. ni (Jones et al., 1987). Accordingly, alignments of amino acid sequences were also compared among storage proteins between two species (Fig. 3). SL-1 showed high identity to TniBJHSP1M (82.3%), but not to TniBJHSP2mM (43.9%), while SL-2 showed high identity to TniBJHSP2mM (84.5%), but not to TniBJHSP1M (45%) (Jones et al., 1993). The juvenile hormone-suppressible nature of SL-1 and SL-2, as in TniBJHSP1M and TniBJHSP2mM, will be published elsewhere (Zheng et al., in press).

The three storage proteins of S. litura appear to belong to three sister groups, although they share extensive similarities throughout the alignments, as judged from the comparison of their alignments mentioned above. Therefore, cladistic analysis was conducted with lepidopteran storage proteins registered in the Genebank, having identities over 50% in amino acid sequences to either of the S. litura proteins. The expected amino acid changes per site were also calculated (data not shown). Based on these data, a Neighbor-joining unrooted tree with the results of bootstrap
analyses on 100 resampled data sets was obtained (Fig. 4). Among 16 hexamersins, SL-1 (methionine: 7.7%) was grouped with TniJHSPlM (methionine: 8.4%; Jones et al., 1993) as the closest relative, followed by H. cunia HcuSPM (methionine: 6.0%, U60988) and H. cecropia HceSP1M (methionine: 8.5%, AF032399) and M. sexta MseSPM (methionine: 6.3%; Wang et al., 1993) and finally by BmoSP1M (methionine: 11.1%; Sakurai et al., 1988). Thus, the storage proteins included in this group are characterized by high contents of methionine (6.0 to 11.1%).

In another branch, SL-2 α and β-subunits (methionine: 3.5 and 3.9%, respectively) were closest to TniBJHS2mM (methionine: 5.3%; Jones et al., 1993), followed by H. cecropia HceSP2mM (methionine: 4.8%, AF032398). Thus, the second group of storage proteins is characterized as being moderately rich in methionine, considerably different from the first group in amino acid alignment, and monophyletic origin of the proteins in these two groups was supported by 100% of the bootstrapped data sets.

In the third group, SL-3 was clustered with arylphorin-type storage proteins of G. mellonella (GmeAr; Memmel et al., 1992), M. sexta (MseaAr and MseβAr; Willott et al., 1989), B. mori (BmoSP2Ar; Fujii et al., 1989), and H. cecropia (HceAr; AF032396). In the phylogenetic trees obtained from the arylphorin gene structures of silkworm species, high similarities have been demonstrated (Shimada et al., 1995).

The storage proteins of the second group were clustered together with those of the first group, or chained to the side of them in parsimonious trees obtained by the cluster analyses of the sequence alignments (Beintema et al., 1994; Burmester and Scheller, 1996). Burmester et al. (1998) proposed a tree essentially similar to ours, by adding the data on moderately methionine-rich storage protein of H. cecropia (HceSP1M). Further surveys on the moderately methionine-rich storage proteins in different insects are expected to confirm the robustness of this tree with three branches.

ACKNOWLEDGEMENTS

We would like to thank Dr. Yasuo Chinzei, Medical School of Mie University, for teaching us the analytical methods and to Dr. DeMar Taylor, University of Tsukuba, for refining the English of the manuscript.

REFERENCES


cDNA of Storage Proteins of *S. litura*


