Mol Gen Genet (1998) 257: 264–270

ORIGINAL PAPER

M. Žurovec · D. Kodrík · C. Yang · F. Sehnal
K. Scheller

The P25 component of *Galleria* silk

Received: 4 June 1997 / Accepted: 5 September 1997

Abstract The water-insoluble core of lepidopteran silk is composed of four major proteins, but only three genes have been identified. This study demonstrates that the 29- and 30-kDa components of *Galleria mellonella* silk are derived from a single gene designated *P25*. The gene is expressed exclusively in the posterior section of the silk glands as a 2-kb mRNA, which accumulates in the feeding larvae and declines at molting. The mRNA encodes a peptide of 24 864 Da that exhibits 51% identity with the putative product of the *P25* gene of *Bombyx*. The conservation of several amino acid stretches, including the relative positions of all 8 cysteines in the mature polypeptide, implies that the P25 proteins play similar, and apparently significant roles in silk formation in the two species. A *Galleria P25* cDNA yields a peptide of about 25 kDa when translated in vitro; the 29- and 30-kDa forms present in the silk are derived from this primary translation product by differential glycosylation.

Key words Fibroin · Glycosylation · DNA sequence · Evolution · Silkworm

Introduction

Several groups of arthropods possess the ability to secrete insoluble fibres known as silk. Larvae of many Lepidoptera produce several types of proteins that are assembled into the firm core and sticky coating of the silk thread (Sehnal and Akai 1990). In the waxmoth *Galleria mellonella* (superfamily Pyraloidea) and the domestic silkworm *Bombyx mori* (Bombycoidea), the silk fibre core, which is secreted in the posterior section of the silk glands, consists of heavy-chain fibroin and three other proteins that range in size from 25 to 30 kDa (Grzelak et al. 1988). In the SDS-PAGE profiles of silk extracts treated with 2-mercaptoethanol, the small components appear as 25-, 27- and 30-kDa bands in *Bombyx* and as 25-, 29- and 30-kDa bands in *Galleria*. The 25-kDa component of *Bombyx* was named light-chain fibroin; in native silk it is linked by disulfide bridges to the heavy-chain fibroin and this complex molecule was originally regarded as a single protein, the fibroin. Homology of the *Galleria* 25-kDa silk protein with *Bombyx* light-chain fibroin was established by comparing amino acid sequences of the two proteins (Žurovec et al. 1995).

Couble et al. (1983), however, described a cDNA that hybridized specifically with a 1.1-kb mRNA from the posterior silk gland section of *Bombyx*, and showed that it was translated into a 25-kDa peptide. The authors named the product silk protein P25 and assumed that it was identical with the 25-kDa silk fraction. Genomic clones of *P25* were then isolated (Couble et al. 1985) and sequencing confirmed that the putative translation product consists of 220 amino acids (Chevillard et al. 1986a). Chevillard et al. (1986b) recognized, however, that the deduced amino acid composition of this product did not match that established by direct amino acid analysis of the 25 kDa silk fraction. It was proposed that this fraction contained more proteins, possibly P25 and light-chain fibroin. This proved not to be the case: immunological studies revealed that the 25-kDa silk component was the product of *light-chain fibroin* gene, whereas the *P25* gene yielded the 27- and 30-kDa silk proteins (Tanaka et al. 1993). A similar relationship was disclosed by partial amino acid sequencing of the purified silk proteins in *Galleria* – the silk fraction of 25 kDa was identified as light-chain fibroin, whereas the 29- and 30-kDa proteins were both found to contain an N-ter-

Communicated by K. Illmensee

M. Žurovec · D. Kodrík · C. Yang · F. Sehnal (✉)
Institute of Entomology, Academy of Sciences, and Faculty of Biological Sciences, University of South Bohemia,
Branišovská 31, 370 05 České Budějovice, Czech Republic
Fax: +420-38/43625; e-mail: sehnal@entu.cas.cz

K. Scheller
Theodor-Boveri-Institute, University Würzburg, Am Hubland, 8700 Würzburg, Germany

minal sequence similar to that of *Bombyx* P25 peptides (Žurovec et al. 1995).

The occurrence of two P25 proteins in the gel profiles of the silk extracts contrasted with the identification of a single *P25* mRNA in *Bombyx* (Couble et al. 1983). In addition, both proteins were considerably larger than the putative translation product of the *P25* gene (Chevillard et al. 1986a). In search of an explanation for these discrepancies we have isolated a *P25* cDNA from *Galleria* and analyzed the corresponding mRNA and proteins. The results reported below show that a single translation product is converted into two proteins via glycosylation. A comparison of our data with those for *Bombyx* strongly indicates that a similar mechanism operates also in this species. The distribution of conserved amino acids in the P25 peptides of *Galleria* and *Bombyx* suggests that certain regions, including the positions of 8 out of 9 cysteine residues, are essential for the function of the mature P25 protein in the silk.

Materials and methods

Tissue preparation

Our present stock of the waxmoth *Galleria mellonella* L. (Lepidoptera, Pyralidae) was obtained from the Department of Animal Physiology of the Masaryk University (Brno, Czech Republic), where it had been grown for about 10 years as an offshoot of our original laboratory stock. Waxmoth cultures were maintained at 30° C and the larvae were fed on a semi-artificial diet (Sehnal 1966). Posterior parts of both silk glands were dissected from the water-anaesthetized penultimate and last-instar larvae, whose age was measured from the preceding ecdysis. In the case of fully grown last-instar larvae, the middle part of silk glands and the carcass remaining after the silk gland dissection were also analyzed. Northern analysis was further performed with RNA from the adults lacking the silk glands, which degenerate shortly after pupation. All dissected tissues were rapidly frozen in liquid nitrogen or on dry ice and stored at −80° C.

Isolation of cDNA clones and sequencing

A cDNA library in lambda gt10 was prepared from poly(A)$^+$-RNA isolated from posterior silk glands of day 3 last-instar larvae with the aid of a kit (Amersham). The library was screened with the degenerate oligoprobe 5′-GGI CCI GC(GATC) AA(CT) GTI GTI (AC)GI CC(GATC) CC(GATC) (AC)G(GATC) TT(GATC) GAT GAT GA-3′, which was derived from the N-terminal sequence GPANNVVRPPRLDDD of *Galleria* 29- and 30-kDa silk fibre proteins (Žurovec et al. 1995). The probe (10 pmol) was end-labelled with 50 μCi of [γ-^{32}P]ATP; free nucleotides were removed by chromatography on a Sephadex G-25 column. Hybridization was performed at 45° C in 5 × SSPE, 5 × Denhardt's solution, 0.5% SDS, and washing at room temperature twice in 2 × SSPE, 0.1% SDS and twice in 1 × SSPE, 0.1% SDS. About 20 000 plaques were screened and three positive phages were identified. Their inserts were subcloned in pBluescript SK(+/−) (Stratagene) and sequenced with Sequenase Version 2.0 (Amersham). Radiolabelled [γ-^{32}P]ATP, [α-^{32}P]dATP and [α-^{35}S]dATP were purchased from Amersham. The sequence was analyzed with the program DNA Star and with the aid of EMBL network services.

RNA and genomic DNA analyses

The largest cDNA insert was radiolabelled with [α-^{32}P]dATP using the Random Prime Labeling kit (Amersham) and employed as a probe in nucleic acid analyses. For the Northern analysis, total RNA was fractionated by electrophoresis and blotted onto Hybond-N membranes (Amersham) as described previously (Žurovec et al. 1992). Hybridization was carried out under high-stringency conditions at 65° C in 6 × SSC, 0.5% SDS; washing was done at the same temperature, twice in 2 × SSC, 0.2% SDS and twice in 0.2 × SSC, 0.1% SDS. In one series of experiments, Northern blots of the RNA from posterior silk gland section were hybridized separately with the 5′ (328 bp) and 3′ (322 bp) fragments of the cDNA.

The content of *P25* mRNA was established using a dot-blot assay. Total RNA was extracted from the silk glands with the RNA Clean system (Angewandte Gentechnologie Systeme). Different amounts of RNA (2 μg, 0.5 μg, and 0.125 μg) were applied with a BioDot applicator (Bio-Rad) onto Hybond-N membrane that was subjected to the hybridization procedure described above for the Northern analysis. Relative amounts of hybridized RNA were assessed by scanning the developed X-ray film with a CAMAG TLC Scanner II. The maximal content found in slowly mobile prepupae, i.e. larvae that are finishing cocoon spinning, was taken as 100%, and values established at other times were expressed in per cent relative to this number.

Radiolabelled cDNA was also used to probe Southern blots of the genomic DNA according to the protocol specified by Žurovec et al. (1992).

In vitro transcription/translation

P25 cDNA in pBluescript SK(−) was transcribed and translated in vitro using the T_NT Coupled Reticulocyte Lysate system (Promega) with T3 RNA polymerase. A sample of 0.5 μg of cDNA was added to 25 μl of reaction mixture containing 8 μCi of [^{35}S]cysteine as label, and the mixture was incubated at 30° C for 90 min. Transcription/translation of *Galleria light-chain fibroin* cDNA was run in parallel for comparison. The [^{35}S]cysteine was chosen as radioactive marker because P25 contains 9 (Fig. 1), and light-chain fibroin 3 (Žurovec et al. 1995) cysteine residues. The translation products were subjected to SDS-PAGE and visualized by fluorography. The molecular weights of the products were assessed using standard molecular markers (Serva) and pre-stained markers (Sigma).

Protein analyses

Silk proteins were solubilized by soaking chopped cocoons for 48 h in a solution containing 10 mM TRIS-HCl (pH 7.0), 2% SDS, 8 M urea and 5% 2-mercaptoethanol. Insoluble heavy-chain fibroin was removed by centrifugation (5 min at 60 × *g*) and the dissolved proteins were analyzed by SDS-PAGE. Proteins electroblotted onto a nitrocellulose membrane were probed for the presence of sugars with the lectins concanavalin A, lentil agglutinin, and peanut agglutinin (all from Lectinola, Prague), which were conjugated to horseradish peroxidase (Grubhoffer et al. 1990). Conjugates bound to the blots were revealed with 3,3-diaminobenzidine.

A portion of the dissolved cocoon proteins was subjected to chemical deglycosylation with the GlycoFree Chemical kit (Oxford Glyco Systems); the protein solution was dialyzed against 0.1% trifluoroacetic acid, then thoroughly lyophilized and treated with trifluoromethansulphonic acid, which cleaves both the N- and O-linked oligosaccharides. Deglycosylated proteins were precipitated with 0.5% ammonium bicarbonate and separated from dissolved reagents by centrifugation. Samples of the original extract of soluble proteins and of the deglycosylated aliquot were dissolved in the solubilizing buffer and subjected to electrophoresis in 20% polyacrylamide gels.

Results

Isolation of *P25* cDNA

Three cDNA clones ranging in size from 0.5 to 1 kb were detected with the oligonucleotide probe, which was designed based on the N-terminal sequences of the 29- and 30-kDa silk proteins. The sequencing of these cDNAs revealed that all were derived from the 5′ end of the same mRNA species and that none of them included a poly(A) tail. Further screening of the cDNA library with one of these clones led to isolation of three additional clones that were also truncated in the 3′ region. We surmised that the 3′ region of the target mRNA contained multiple oligo(A) sequences and that reverse transcription during the cDNA preparation was initiated from those preceding the poly(A) tail. Hence, a full-length cDNA had probably not been produced and further screening of the library was point less. Most of the cDNA clones we isolated contained the entire coding region and this was sufficient for our study.

In the longest cDNA clone, a translation initiation codon was identified in position 18 from the 5′ end and the coding sequence continued to the termination codon TAA at position 672 (Fig. 1). The non-coding tail comprised 370 bp but no polyadenylation signal was included. Alignment of the coding sequence with the homologous part of the *Bombyx* gene *P25* (Chevillard et al. 1986a) disclosed 55.3% nucleotide sequence identity, confirming that the identified cDNA corresponded to the *Galleria P25* gene.

The putative translation product

The open reading frame of the *Galleria P25* cDNA encodes a peptide of 119 amino acids. The amino acid sequence GPANNVVRPPRLDDD, which was established by analysing the N-termini of *Galleria* 29- and 30-kDa silk proteins (Žurovec et al. 1995), was found to begin in amino acid position 17 of the putative product. This proved that we had isolated cDNA encoding the 29- and 30-kDa silk proteins, and indicated that signal peptide cleavage occurred between alanine (position 16) and glycine (position 17) residues of the primary translation product. The N-terminal sequence of the putative mature peptide differed from the stretch of 15 amino acids identified by peptide sequencing, in having a cys-

Fig. 1 Nucleotide sequence of *Galleria P25* cDNA (numbered from the most distal 5′ nucleotide of the identified cDNA) and sequence of the deduced peptide, which is aligned with that of *Bombyx* P25 (amino acid residues homologous to *Galleria* P25 are *underlined*). Cysteine residues are printed in *italics* and the putative N-glycosylation sites are *double underlined*. The *arrow* indicates the signal peptide cleavage site and the beginning of N-terminal sequences of the mature 29- and 30-kDa silk proteins

```
Galleria cDNA   CTAACTGCGTACCAAACATGTTGAAGTTTATCATATTTGCTCTCACAGTGGCTCTGTGTGAA   GCTGGGCCAGCCAACAAT    80
Galleria protein                MetLeuLysPheIleIlePheAlaLeuThrValAlaLeuCysGlu   AlaGlyProAlaAsnAsn    21
Bombyx protein                  MetLeuAlaArgCysLeuAlaValAlaAlaValAlaValLeuAlaSerAlaGlyProProSerPro    22
                                                                                ⇑

GTGGTGAGACCATGCAGACTTGATGACCTGAAATGTATCCGTGACAACATATCCGCTAACTCGAACTGCAACGCCAACGTTCGGGGTTCGATTCCCTCG   179
ValValArgProCysArgLeuAspAspLeuLysCysIleArgAspAsnIleSerAlaAsnSerAsnCysAsnAlaAsnValArgGlySerIleProSer    54
IleTyrArgProCysTyrLeuAspAspTyrLysCysIleSerAspHisLeuAlaAlaAsnSerLysCysIleProGly   ArgGlyGlnIleProSer    54

GAATATGTCATCCCCAGATTCAATTTCGAGACGCCATTTTTCAACGCTTCGTACATAGACAACAACTTGATCATCAGAAACAACGATGCTTGCCGGGTC   278
GluTyrValIleProArgPheAsnPheGluThrProPhePheAsnAlaSerTyrIleAspAsnAsnLeuIleIleArgAsnAsnAspAlaCysArgVal    87
GlnTyrGluIleProValPheGlnPheGluIleProTyrPheAsnAlaThrTyrValAspHisAsnLeuIleThrArgAsnHisAspGlnCysArgVal    87

TCGGAATTTTTCTTCAATGTAAAAGCTGACACATCTGTCCTTGCTGTCGACTGCCCCAACTTGGACTTGGAATCCGACCGCACCCTGATCCAGCACGCA   377
SerGluPhePhePheAsnValLysAlaAspThrSerValLeuAlaValAspCysProAsnLeuAspLeuGluSerAspArgThrLeuIleGlnHisAla   120
SerGluPheTyrAspAsnValArgThrLeuLysThrValLeuThrValAspCysProTrpLeuAsnPheGluSerAsnArgThrLeuAlaGlnHisMet   120

TCATTACAGGAGGAGACCACGTACAACTATCACATCCGAGGAATCTACCCACTGATCCGTCTAACGACCAATCTCCTCAACGCAGACCGACTAAATCTT   476
SerLeuGlnGluGluThrThrTyrAsnTyrHisIleArgGlyIleTyrProLeuIleArgLeuThrThrAsnLeuLeuAsnAlaAspArgLeuAsnLeu   153
SerPheLysGluAspValValLeuSerPheTyrIleAsnGlySerTyrProLeuIleArgLeuThrThrValPheAspLysGlyAsnAsnPheAspLeu   153

TGTAACGCATTCACATATGCCGACGTCACTGCT   CTGCCGATATTCAAGATCGATCCTAAAGACAGACCAACTGCCAACTTCCTGTCCcGCGACCTG   572
CysAsnAlaPheThrTyrAlaAspValThrAla   LeuProIlePheLysIleAspProLysAspArgProThrAlaAsnPheLeuSerArgAspLeu   185
CysSerAlaPheThrPheAlaAspLeuAlaGlyGlyLeuProIlePheHisIleAsnProAsnAspGlnArgThrAlaGlnTrpLeuSerLysAspLeu   186

TCCTTGCTCAACATCTACGAGAGAGAGACGTTCGCGTACCGGCCTCCCCAGCTCATACGCCAATTTGTTAACTCACTTATTTGTGATTTTGGCTGCCAA   671
SerLeuLeuAsnIleTyrGluArgGluThrPheAlaTyrArgProProGluLeuIleArgGlnPheValAsnSerLeuIleCysAspPheGlyCysGln   218
ThrLeuLeuHisIleTyrGluArgGluHisIlePheGlyLysArgAsnTrpLeuAlaArgSerPheIleSerArgThrLeuCysAspPheGlyCysHisHis220

TAAGATTACCTTTAAAAGTCCGTTCGGACAAATTTTTTTTATTTGTGTTTTTGTTTTTGTCCTTTTTCTCTCTGCAGGCAAATTTGTGTAGGCTATAAA   770

CTAAATATTTAAAAAATGAAAGATATGATTGTTTGTTTGTAACGGATAAATTAAAAAATACTCAACCGGTTTTAAAAATTCTCTCGCATATAAAAAGTT   869

ACATTATAAGCAAATAACTTTGGGCTGTATTTTGAAACAAAAAGATATCTATCTATGCGAAGCTAGTGCAGTTCTCTAGCTCTATATACAGACAAGATC   968

AAAATCAGACTGAGAAAATTAATCTATGCAACCATTTATATGCATCAATACTTTAAAAGGAATTTATATTCAT    1041
```

teine in place of proline in the second position, and a leucine in place of aspartate in the last position. This difference was not surprising; we expected ambiguities in the N-terminal sequencing which was performed with small amounts of insufficiently purified proteins.

Utilization of the reading frame opened by the Met codon at nucleotide position 18, and the homology of isolated cDNAs with *Bombyx P25* gene, were further confirmed by aligning the deduced peptide sequence with that of the primary translation product of the *Bombyx P25* gene (Fig. 1). We found that *Galleria* and *Bombyx* P25s are 73% similar, and 51% identical in amino acid sequence. All eight cysteine residues and several stretches of amino acids of the mature peptides are conserved. The alignment showed that *Galleria* P25 differs from that of *Bombyx* by two internal deletions (one of them in the signal region) and one C-terminal deletions, and by one internal amino acid insertion.

Northern and dot blot analyses

The following experiments were done to establish the tissue specificity of the expression of *Galleria P25*, and to ascertain that the 29- and 30-kDa silk proteins are not encoded by different *P25* mRNAs. Northern analysis of total RNA from the selected tissues demonstrated that the *P25* cDNA hybridized exclusively with a single 2-kb mRNA from the posterior section of silk glands (Fig. 2, lanes P, M, C, I). Identical results were obtained when the total RNA from posterior silk gland was probed with the 5′ and 3′ fragments of the cDNA (data not shown). The blot analysis thus proved that all isolated cDNAs were truncated reverse transcripts of the same mRNA. This *P25* mRNA was detected in the posterior silk gland of all developmental stages examined (Fig. 2, lanes 1–4).

Since the Northern data indicated that the amount of *P25* mRNA in the silk glands fluctuated, the relative content of this mRNA in total RNA was quantified by dot blot assays in course of the penultimate and final larval instars. We found that the *P25* mRNA content was low at ecdysis to the penultimate larval instar, peaked in the feeding larvae of this instar, and decreased again at the molt to the last larval instar (Fig. 3). It rose rapidly after ecdysis when the larvae initiated feeding, and reached a maximum in slowly mobile prepupae at the end of cocoon spinning. The following tenfold drop of *P25* mRNA coincided with the start of silk gland degeneration.

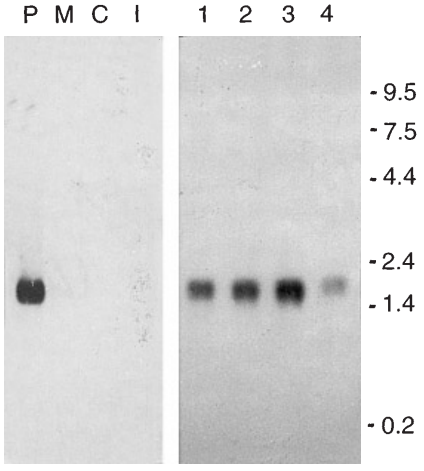

Fig. 2 Northern analysis of *P25* expression. Demonstration of *P25* mRNA in the posterior silk gland extract (lane P); no signal is found in identical amounts of total RNA extracted from the middle silk gland (lane M), larval body carcass (lane C) or whole adults (lane I). Detection of *P25* mRNA in RNA extracts from the posterior silk glands of larvae dissected in mid-penultimate instar (lane 1) and on days 1 (resumption of feeding), 5 (termination of feeding) and 6.5 (cocoon spinning) of the last instar (lanes 2, 3, and 4, respectively). Positions of molecular weight markers are shown on the right

Analysis of the P25 silk proteins

The presence of considerable amounts of *P25* mRNA in the silk glands of spinning larvae indicated that representation of the corresponding proteins in the silk would be high. Indeed, SDS-PAGE analysis of the solubilized cocoons confirmed that the P25 proteins of 29- and 30-kDa belonged to the dominant silk components (Fig. 4, lane 2). In contrast to the previously analyzed *Galleria* silk, which contained a single protein in the 30 kDa region (Kodrík 1992), the larvae of our present stock produced two proteins whose size was close to 30 kDa. The remaining major bands seen in lane 2 of Fig. 4 represented the 25-kDa light chain fibroin (Žurovec et al. 1995) and a 17.5-kDa protein called seroin (Žurovec et al. 1996). Heavy-chain fibroin was not included in the present analysis.

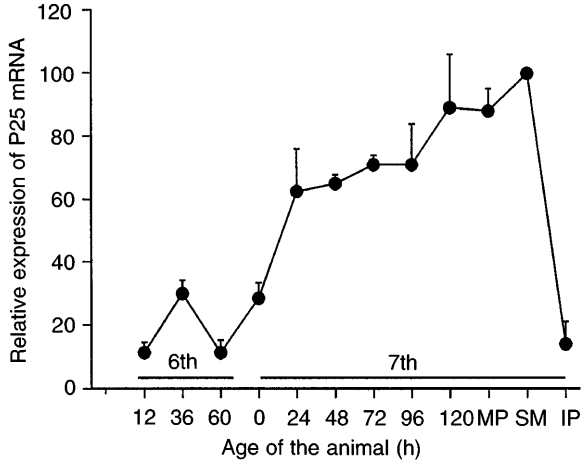

Fig. 3 Changes in the relative amounts of *P25* mRNA in the course of the penultimate and last larval instars as established by dot blot assay (average values ± S.D. of 4–5 measurements)

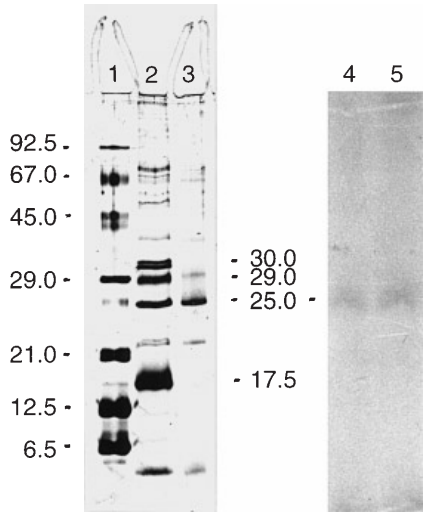

Fig. 4 Electrophoretic analysis of *P25* and light-chain fibroin. SDS-PAGE of peptide size markers (lane 1); cocoon extract with major bands at 30, 29, 25, and 17.5 kDa (lane 2); cocoon extract after chemical deglycosylation; note the absence of the 30-kDa band, the faint 29-kDa band, and the strong 25-kDa band (lane 3); products of in vitro translation of *P25* cDNA (lane 4), and *light-chain fibroin* cDNA (lane 5). Proteins in lanes 1–3 were stained with Coomassie Briliant Blue, peptides synthesized in vitro were radiolabelled with [^{35}S]cysteine and visualized by fluorography

While the silk contained P25 proteins of 29 and 30 kDa, the deduced molecular weight of the secreted translation product (lacking the signal peptide) of the *Galleria P25* mRNA was only 23 098 Da (Fig. 1). In an attempt to clarify this discrepancy, we decided to analyze the in vitro translation product of *P25* cDNA and to examine if the P25 proteins present in the cocoon were post-translationally modified. We found that *P25* cDNA generated in vitro a single peptide of about 25 kDa (Fig. 4, lane 4), consistent with the estimated size of 24 863 Da. Our assessment of the size of the in vitro product was corroborated by a comparison with the product of the *light-chain fibroin* cDNA (Fig. 4, lane 5), which encodes a peptide of 25 442 Da (Žurovec et al. 1995). As expected, the mobilities of peptides obtained with *P25* cDNA and *light-chain fibroin* cDNA were similar (Fig. 5).

Since the sequence of P25 contained several potential glycosylation sites (see Discussion), the size increase of the secreted products could be caused by attached sugar moieties. This proved to be the case, for chemical deglycosylation of the crude mixture of solubilized cocoon proteins led to a complete disappearance of the 30-kDa bands and to a pronounced reduction in the intensity of the 29-kDa band (Fig. 4, lane 3). The remaining 25-kDa band probably contained both light-chain fibroin and the deglycosylated 29- and 30-kDa proteins; the relatively low staining intensity of the 25-kDa band in lane 3 was obviously due to considerable protein loss associated with the deglycosylation procedure.

The nature of the glycosylation adducts in the 29- and 30-kDa proteins was investigated using lectin binding to Western blots of solubilized cocoon proteins. Several types of glycosylation were identified in the sericin fractions above 45 kDa (Fig. 5.) The 30-kDa protein reacted strongly with concanavalin A (Fig. 5, lane 1), whereas the 29-kDa protein was not clearly recognized by any of lectins tested. Chemical deglycosylation caused the loss of all lectin reactions (Fig. 5, lanes 2, 4, and 6). Similar results were obtained when the 29- to 30-kDa fraction was electroeluted from the gel, treated with bacterial glycosidase F, which specifically cleaves N-glycosides, and analyzed by PAGE (data not shown). Lack of glycosylation in light chain fibroin was confirmed in the similarly eluted 25 kDa fraction of the cocoon extract (unpublished data).

Discussion

Using an oligonucleotide derived from a partial sequence of two *G. mellonella* silk proteins, we identified in a silk gland-specific cDNA library a homolog of the *P25* gene known from *B. mori* (Fig. 1). This present finding of *Galleria P25* and our earlier description of the *Galleria light-chain fibroin* gene (Žurovec et al. 1995) contrast with the report of Tamura and Kubota (1988), who found no small silk proteins associated with heavy-chain fibroin in several saturniid silkmoths. Saturniids belong to the same superfamily, Bombycoidea, as *Bombyx*, whereas *Galleria* represents a rather distant superfamily Pyraloidea (Nielsen and Common 1991). From this limited information it seems that small silk components, such as P25 proteins, evolved before Pyraloidea and

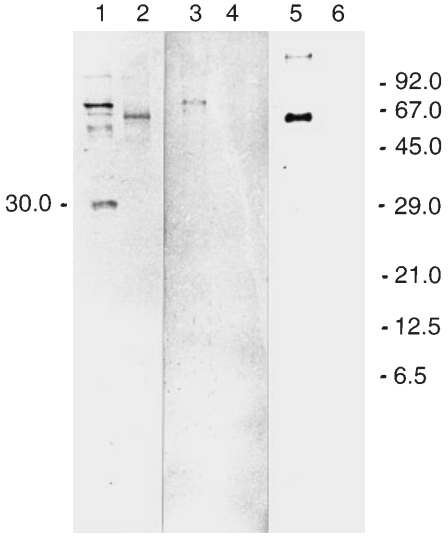

Fig. 5 Detection of sugar moieties in silk proteins fractionated by SDS-PAGE. Cocoon extracts were electroblotted onto a nitrocellulose membrane and treated with lectins conjugated to horseradish peroxidase. Silk proteins were reacted before (odd lane numbers) and after chemical deglycosylation (even lane numbers), with concanavalin A (lanes 1 and 2), lentil agglutinin (lanes 3 and 4), or peanut agglutinin (lanes 5 and 6). The position of the 30-kDa protein is indicated on the left, positions of protein markers on the right

Bombycoidea separated in phylogeny and were secondarily lost in the saturniids.

The longest of our *P25* cDNA clones included a 5′ non-translated region of 17 nucleotides. From a comparison with *Bombyx P25*, in which the ATG initiation codon lies in position +23 (Couble et al. 1985), we assume that the identified non-coding 5′ end of *Galleria* cDNA is nearly complete. The comparison with *Bombyx P25* further proves that the coding region of 671 nucleotides, detected in our cDNAs, encodes the whole P25 peptide. In the 3′ non-coding region, however, we identified only a stretch of 370 nucleotides, whereas the size of *Galleria P25* mRNA (2 kb) suggests that the 3′ tail is about 1.3 kb long. The 3′ non-coding region of *Galleria P25* mRNA appears to be much longer than the 474-nucleotide 3′ tail of *Bombyx P25* mRNA (Chevillard et al. 1986a).

Developmental changes in the content of *P25* mRNA in the silk glands of *Galleria* resemble those reported for *Bombyx* (Couble et al. 1983), in that the content declines at the molt to the last larval instar (Fig. 3). The rise that occurs after ecdysis is rather gradual in *Bombyx*, whereas in *Galleria* it is steep in the first 24 h and gradual afterwards, reaching a maximum in larvae that are about to finish cocoon spinning. A similar difference between *Bombyx* and *Galleria* was established in the developmental profile of *light-chain fibroin* mRNA (Kimura et al. 1985; Žurovec et al. 1995). The rapid post-ecdysial rise of mRNAs encoding silk proteins apparently enables *Galleria* larvae to produce large volumes of silk nearly continuously (Jindra and Sehnal 1989), while the spinning of *Bombyx* larvae is largely restricted to cocoon construction.

Putative translation products of *Galleria P25* and *Bombyx P25* cDNAs are composed of 218 and 220 amino acids, respectively. Representations of individual amino acids are similar, large differences being found only in tryptophan (1.5% in *Bombyx*, 0% in *Galleria*) and histidine (4.4% and 0.9%, respectively). Positions of about half of all amino acids are conserved, including those of all 8 cysteine residues in the mature peptide. Signal peptide cleavage occurs in *Galleria* between alanine and glycine (positions 16 and 17) in the AlaGlyPro sequence, which is conserved also in *Bombyx*. Since cleavage between alanine and glycine preceding a proline residue is common (von Heijne 1986), this site is probably used also in *Bombyx*. Interestingly, the signal peptide sequence exhibits relatively low conservation; only 5 out of 16 (*Galleria*) or 17 (*Bombyx*) amino acid residues are shared by the two species. This contrasts with the conservation of 12 out of 16 amino acid residues in the signal peptide of light-chain fibroins (Žurovec et al. 1995).

High conservation of a number of amino acid residues in secreted P25 indicates that their positions are important for the protein function, but no details are known. For *Bombyx* P25 it was suggested that cysteines are not used for covalent linkage to heavy-chain fibroin (as is the case of light-chain fibroin), but probably allow P25 to attain a conformation necessary for its interaction with other fibroin components (Tanaka et al. 1993). The 53.7% identity between *Galleria* P25 and *Bombyx* P25 indicates that P25 proteins are important for silk formation in the two species. It is interesting to note that light-chain fibroin, whose significance and role in silk production were proved in *Bombyx* (Takei et al. 1987), exhibits only 40.2% sequence identity between *Galleria* and *Bombyx*.

Sinohara et al. (1971) demonstrated that *Bombyx* fibroin (heavy-chain was not separated from the light-chain fibroin and the P25 proteins) is associated with sugars composed of glucosamine and mannose, which bind to Asn in the SerAsnThr motif. In *Galleria*, this motif is present in the heavy-chain fibroin (our unpublished data) but not in the light-chain fibroin (Žurovec et al. 1995) or the P25 peptide (Fig. 1). However, the latter contains two N-glycosylation motifs of the type Asn-X-Ser/Thr (X is any amino acid except proline); the reaction with concanavalin A (Fig. 5) indicated that mannose-rich N-glycans are actually bound to the asparagine residue in one or both these motifs in case of the 30-kDa derivative of P25. No sugar moiety was detected with lectin assays in the 29-kDa derivative, but changes in protein mobility after deglycosylation (lanes 2 and 3 in Fig. 4) disclosed that both the 29- and the 30-kDa silk proteins contain sugars. We conclude that the primary translation product of *Galleria P25* is converted into two silk proteins by differential glycosylation, the nature of which has been elucidated only in part.

The deduced size of *Bombyx* P25 is about the same as in *Galleria* and the corresponding proteins appear in the SDS-PAGE of silk extracts treated with 2-mercaptoethanol as 27- and 30-kDa bands. Tanaka et al. (1993) assumed that P25 exhibits unusual electrophoretic behavior but the situation in *Galleria* indicates that the occurrence of two final products and the discrepancy between the deduced and the apparent molecular weights in *Bombyx* P25 proteins is due to glycosylation. *Bombyx* P25 contains three N-glycosylation sites, one of them in an identical position to one of the motifs in the *Galleria* P25 (Fig. 1).

Acknowledgements We wish to thank Dr. L. Grubhoffer, Institute of Parasitology AV ČR, for the supply of lectin-horseradish peroxidase conjugates and for advising us on the methods of glycoprotein analysis. We further acknowledge the gift of *Galleria mellonella* larvae from Mgr. Jana Benešová of the Masaryk University, Brno. Our study was supported by grant 204/96/1100 of the Grant Agency of the Czech Republic and by Volkswagenstiftung grant I68 514.

References

Chevillard M, Couble P, Prudhomme JC (1986a) Complete nucleotide sequence of the gene encoding the *Bombyx mori* silk protein P25 and predicted amino acid sequence of the protein. Nucleic Acid Res 14:6341–6342

Chevillard M, Deleage G, Couble P (1986b) Amino acid sequence and putative conformational characteristics of the 25 kD silk protein of *Bombyx mori*. Sericologia 26:435–449

Couble P, Moine A, Garel A, Prudhomme JC (1983) Developmental variations of a nonfibroin mRNA of *Bombyx mori* silk gland, encoding for a low-molecular-weight silk protein. Dev Biol 97:398–407

Couble P, Chevillard M, Moine A, Ravel-Chapuis P, Prudhomme JC (1985) Structural organization of the P_{25} gene of *Bombyx mori* and comparative analysis of its 5′ flanking DNA with that of the fibroin gene. Nucleic Acid Res 13:1801–1814

Grubhoffer L, Guirakhoo F, Heinz FX, Kunz C (1990) Interaction of tick-borne encephalitis virus protein E with labelled lectins. In: Kocourek J, Freed DL (eds) Lectins: biology, biochemistry, clinical biochemistry, vol 7. Sigma Chemical Company, St. Louis, Mo., pp 313–319

Grzelak K, Couble P, Garel A, Kludkiewicz B, Alrouz H (1988) Low molecular weight silk proteins in *Galleria mellonella*. Insect Biochem 18:223–228

Jindra M, Sehnal F (1989) Larval growth, food consumption, and utilization of dietery protein and energy in *Galleria mellonella*. J Insect Physiol 35:719–724

Kimura K, Oyama F, Ueda H, Mizuno S, Shimura K (!985) Molecular cloning of the fibroin light chain complementary DNA and its use in the study of the expression of the light chain gene in the posterior silk gland of *Bombyx mori*. Experientia 41:1167–1171

Kodrik D (1992) Small protein components of the cocoons in *Galleria mellonella* (Lepidoptera, Pyralidae) and *Bombyx mori* (Lepidoptera, Bombycidae). Acta Ent Bohemoslov 89:269–273

Nielsen ES, Common IFB (1991) Lepidoptera (moths and butterflies). In: Naumann JD (ed) The Insects of Australia (2nd edn) Melbourne University Press, Carlton (Victoria, Australia) pp 817–915

Sehnal F (1966) Kritisches Studium der Bionomie und Biometrik der in verschiedenen Lebensbedingungen gezüchteten Wachsmotte, *Galleria mellonella* L. (Lepidoptera). Z Wiss Zool 174:53–82

Sehnal F, Akai H (1990) Insect silk glands: their types, development, and function, and effects of environmental factors and morphogenetic hormones on them. Int J Insect Morphol Embryol 19:79–132

Sinohara H, Asan Y, Fukui A (1971) Glycopeptides from silk fibroin of *Bombyx mori*. Biochem Biophys Acta 237:273–279

Takei F, Kikuchi Y, Kikuchi A, Mizuno S, Shimura K (1987) Further evidence for importance of the subunit combination of silk fibroin in its efficient secretion from the posterior silk gland cells. J Cell Biol 105:175–180

Tamura T, Kubota T (1988) A determination of molecular weight of fibroin polypeptides in the saturniid silkworms, *Antheraea yamamai*, *Antheraea pernyi* and *Philosamia ricini* by SDS-PAGE. In: Akai H (ed) Wild Silkmoth '88. National Institute of Sericulture and Entomological Science, Tsukuba, Japan, pp 67–72

Tanaka K, Mori K, Mizuno S (1993) Immunological identification of the major disulfide-linked light component of silk fibroin. J Biochem 114:1–4

von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683–4690

Žurovec M, Sehnal F, Scheller K, Kumaran AK (1992) Silk gland specific cDNAs from *Galleria mellonella* L. Insect Biochem Mol Biol 22:55–67

Žurovec M, Vašková M, Kodrík D, Sehnal F, Kumaran AK (1995) Light-chain fibroin of *Galleria mellonella*. Gen Mol Biol 247:1–6

Žurovec M, Yang C, Sehnal F, Scheller K (1996) A new component of the silk fibre produced in the middle and posterior silk glands. Proceedings XX International Congress on Entomology, p 425