ELSEVIER

Gene 209 (1998) 157–165

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Characterization of the *P25* silk gene and associated insertion elements in *Galleria mellonella*

Changsong Yang [a,1], Xiuyin Teng [a], Michal Zurovec [a], Klaus Scheller [b], Frantisek Sehnal [a,*]

[a] *Institute of Entomology, Academy of Sciences, and Faculty of Biological Sciences, University of South Bohemia, 370 05 Ceske Budejovice, Czech Republic*

[b] *Department of Zoology, University of Wurzburg, 8700 Wurzburg, Germany*

Received 2 September 1997; accepted 16 December 1997; Received by B. Dujon

Abstract

Insect silk genes attract attention by their precise territorial and developmental regulations and extremely high expression rates. Our present investigations demonstrated that the *P25* silk gene of *Galleria mellonella* is down-regulated by ecdysteroid hormones. The gene was identified within 5217 nucleotides (nt) of two genomic clones. In contrast to other silk genes, *Galleria P25* lacks the canonical TATA box. Transcription is initiated wihin a region of three nucleotides that lie at the end of a capsite initiator sequence ACAGT and about 90 nt downstream from a CAAT box. A stretch of 32 nt with a core sequence CTTTT was detected in the 5′ region of *Galleria P25* as well as in the presumptive regulatory regions of all other silk genes that are expressed in the posterior silk gland. However, consensus sequences reported for the regulatory regions of *Bombyx* silk genes are not obvious in *Galleria P25.* The coding sequence of this gene includes 654 nt, is interrupted by 4 introns, and ends in position +3369; a potential polyadenylation signal starts at +4382. The gene contains 3 copies of a short interspersed nuclear element (SINE), which are located in the upstream region (−833 to −579) and in the first (+542 to +840) and second (+2259 to +2556) introns. The repeat, which was named *Gm1*, occurs in some other *Galleria* genes and exhibits homology to *Bm1* SINE of the silkworm and to a similar element of a spider. Another insertion of at least 150 nt and with loosely defined borders is present in the 3′ untranslated region (*UTR*) of *Galleria P25*. It includes a box (+3453 to +3552) of 99 nt that is tentatively called *Lep1* because it was disclosed also in some other Lepidoptera. *Lep1* seems to represent the core region of insertion elements that occur in the genomes of lepidopteran insects in various species specific and region specific modifications. © 1998 Elsevier Science B.V.

Keywords: DNA repeats; Ecdysone; Fibroin; Genomic evolution; SINE

1. Introduction

The salivary glands of some insect larvae produce proteinaceous threads known as silk. In the larvae of lepidopteran insects, the posterior section of such silk glands secretes a core of the silk fibre, whereas the middle section provides the core with a sticky coating made of several sericin proteins. For two distantly related lepidopteran species, the waxmoth, *Galleria mellonella*, and the silkworm, *Bombyx mori*, it has been shown that the core is composed of three types of proteins known as heavy-chain fibroin, light-chain fibroin, and P25 (Tanaka et al., 1993; Zurovec et al., 1995). Proper assemblage of these components provides the strength, resistance, and contractibility of the fibre.

The expression of silk genes is restricted to appropriate sections of the silk glands and fluctuates during development in dependence on nutrition supply and in response to hormones. DNA sequences indispensable for this spatial and temporal specificity of expression were studied in *Bombyx* silk genes. Most work was done with the *Fib-H* gene, which encodes heavy chain fibroin, and with the *P25* gene. These two genes share a number of short, presumably regulatory motifs in their 5′ flanking regions (Couble et al., 1985). Y. Suzuki and co-workers

* Corresponding author. Tel: +42 38 40822; Fax: +42 38 43625; e-mail: sehnal@entu.cas.cz

[1] Present address: Biology Department, University of Pennsylvania, 415 S. University Ave., Philadelphia, PA 19104-6018, USA.

Abbreviations: *Fib-H*, gene for heavy-chain fibroin; *Fib-L*, gene for light chain fibroin; 20E, 20-hydroxyecdysone; kb, kilobase pairs; *Lhp76*, gene for larval hemolymph protein 76; *Lhp82*, gene for larval hemolymph protein 82; nt, nucleotide; P25, P25 silk protein; SINE, short interspersed nuclear element; tsp, transcription start point; *UTR*, untranslated region.

0378-1119/98/$19.00 © 1998 Elsevier Science B.V. All rights reserved.
PII S0378-1119(98)00029-8

pioneered investigations on *Fib-H* expression in a cell-free expression system. Over the years, they have identified a minimal promoter at −37 to +10, and ascertained that an upstream region within −200 nt and an intronic element laying at +156 to +454 enhance transcription (Takiya et al., 1990). Couble and Prudhomme's group introduced modified *P25* gene of *Bombyx* into *Drosophila* and identified the region −437 to +76 as sufficient to drive the tissue specific expression [reviewed by Prudhomme and Couble (1995)]. Several putative *cis*-elements were identified in *Fib-H* and *P25* by gel shift assays, and several nuclear proteins binding to certain DNA motifs in the silk genes were isolated [reviewed by Nony et al. (1995) and Takiya et al. (1997)]. Experiments with the transfer of *Bombyx P25* gene into *Drosophila* showed that *cis*-elements of the silk genes are properly recognized by the transcription factors present in the salivary glands of fly larvae (Bello and Couble, 1990).

In our previous work, we demonstrated that expressions of *Galleria* silk genes exhibited identical territorial specificities (Zurovec et al., 1992, 1995) and similar developmental fluctuations (Yang et al., 1995, 1996) as the silk genes of *Bombyx*. In experiments described in this paper, we verified that expression of the newly identified *P25* gene of *Galleria* is regulated by ecdysteroid hormones. This result and other data on *P25* expression (Zurovec et al., 1998) led us to believe that a comparison of the regulatory regions of *Galleria P25* with the known silk genes of *Bombyx* may disclose conserved regulatory DNA motifs. This perspective prompted us to elucidate the full sequence of *Galleria P25* gene. We report here that this gene has a similar structure to the *Bombyx P25*, but homologies in presumed regulatory regions are not obvious. Outcomes of our work include descriptions of a short interspersed nuclear element (SINE), which seems to be shared by spiders and insects, and of an insertion element that is shared by various moths.

2. Materials and methods

2.1. Effect of 20-hydroxyecdysone on P25 expression

All experiments were performed with the wax moth, *Galleria mellonella* L., reared on a semiartificial diet (Sehnal, 1966). Fully grown *Galleria* larvae were anaesthetized in water and ligated with a thread across mesothorax; body region anterior to the ligation was cut off. The removal of head and prothorax deprived the insects of ecdysteroids and other hormones and caused a cessation of development. The abdomens, which survived for several weeks, were taken for experiments 3 days after ligation. Experimental specimens were each injected with 1.6 or 2.4 μg of the natural ecdysteroid hormone, 20-hydroxyecdysone, which was dissolved in 2 or 3 μl of 8% ethanol; controls were injected with the solvent alone. Silk glands were dissected 24 h after the injection. RNA extractions and detection of the *P25* mRNA with a cDNA probe (Zurovec et al., 1998) were done as decribed previously (Yang et al., 1996).

2.2. Preparation and screening of a genomic library

High-molecular-weight DNA was prepared from newly ecdysed last instar *Galleria* larvae that had a low fat content and nearly empty guts. DNA was partially digested with *Sau*3A; fragments of 15–20 kb were isolated by centrifugation in a sucrose gradient and ligated to *Xho*I half-site arms of LambdaGEM™-12 vector (Promega). Phages assembled with the Packagene system (Promega) were introduced into the LE392 strain of *Escherichia coli*. Transformed cells were screened with the 5′ (328 bp) and 3′ (322 bp) fragments of *Galleria* P25 cDNA (Zurovec et al., 1998). The probes were digoxigenin- or ^{32}P- labelled with the Random Primed DNA Labelling Kit (USB). Two genomic clones, designated GGE1 and GGE2, were found by screening 4.0×10^5 plaques with the 5′ and 3′ cDNA fragments, respectively.

2.3. Sequencing, primer extension, and Southern analysis

Standard methods of DNA analysis were used (Sambrook et al., 1989). The genomic clones were mapped with restrictases and subcloned; three sets of nested subclones were generated with unidirectional deletion method (Erase-a-Base System, Promega) and sequenced as depicted in Fig. 2. Single- or double-stranded templates were sequenced by the dideoxy chain termination reaction using [α-^{35}S]dATP, T3 or T7 primers, and the Sequenase Version 2.0 DNA Sequencing Kit (Amersham). Specifically designed primers were used to confirm the junctions of particular subclones. The sequence was usually read from two overlapping subclones of one DNA strand and, in cases of doubt, also from subclones of the opposite DNA strand.

For Southern analysis, aliquots of 5 μg genomic DNA from *Galleria* larvae were digested with *Xba*I, *Sal*I, and *Pst*I, respectively, and electrophoresed on 0.8% agarose gel. The blots on nylon membrane were probed with the ^{32}P-radiolabeled 5′ fragment (328 bp) of *Galleria* P25 cDNA. Relative positions of *Hind*III-digested λ phage DNA fragments were used as size markers.

A 22-nt synthetic oligonucleotide complementary to the region +68 to +89 of the gene (Fig. 3) was employed for primer extension analysis. Using T4 polynucleotide kinase, the primer was 5′-labeled with [γ-^{32}P]ATP, and a dose of 5×10^4 cpm was hybridized

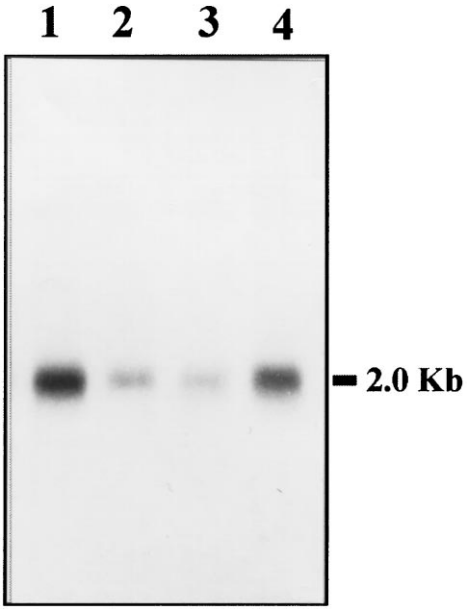

Fig. 1. Effect of 20-hydroxyecdysone (20E) on *Galleria P25* expression. The source of endogenous ecdysteroids was removed by ligating off the head and thorax of fully grown larvae. Relative content of *P25* mRNA in the glands was established 3 days after ligation when the insects were injected (lane 1) and 24 h after the injection of 1.6 μg 20E (lane 2), 2.4 μg 20E (lane 3), and the solvent alone (lane 4), respectively.

to 50 μg total RNA (prepared from the posterior section of silk glands of day 3 last instar *Galleria* larvae) for 16 h at 32°C. The extension was carried out with M-MLV reverse transcriptase (Amersham) for 2 h at 42°C. The products were analyzed on a sequencing gel (8% polyacrylamide, 7 M urea) with the aid of G, A, T, and C ladders of appropriate genomic subclone that was sequenced with the same primer.

3. Results and discussion

3.1. *Down-regulation of* P25 *by 20-hydroxyecdysone*

In a previous study, we found that the silk gland content of *P25* mRNA decreased at the last larval molt, increased during the intermolt period, and dropped before the pupal ecdysis of *Galleria* (Zurovec et al., 1998). This pattern indicated that the surges of ecdysteroids, which precede and evoke ecdyses (Sehnal et al., 1981), are causally associated with the declines of *P25* mRNA.

To prove that high ecdysteroid concentrations cause a *P25* turn-off, we prepared isolated abdomens that are devoid of these hormones. The content of *P25* mRNA in the silk glands of such abdomens was gradually reduced, obviously as a consequence of insufficient nutrient supply, but 3 days after ligation it was still relatively high (Fig. 1, lane 1). Injections of 20E in physiological doses caused a dramatic decrease in mRNA (Fig. 1, lanes 2 and 3), whereas injection of the solvent had no appreciable effect on the gradual decline of the mRNA content (Fig. 1, lane 4). These results were consistent with our previous observations on the ecdysteroid effects on other silk genes of *Galleria* (Yang et al., 1995, 1996). Since ecdysteroids are known to terminate expression of *Bombyx* silk genes (Sehnal and Akai, 1990), we

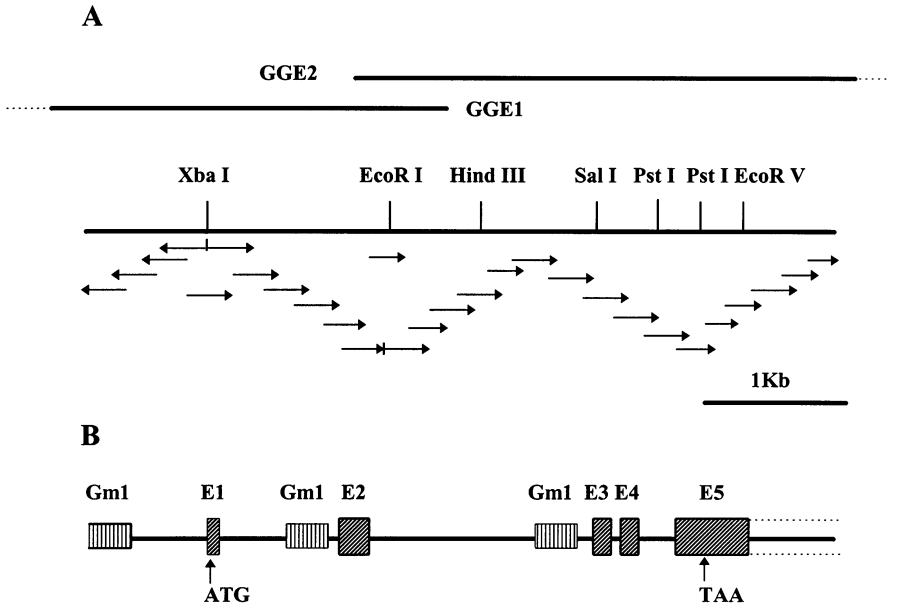

Fig. 2. Structure and sequencing strategy of *Galleria P25* gene. (**A**) Partial restriction map of two genomic clones (GGE1 and GGE2); horizontal arrows indicate the direction and extent of sequencing. (**B**) Schematic chart of gene organization; exons (E1–E5) are indicated by hatched boxes, the *Gm1* elements by boxes with vertical lines.

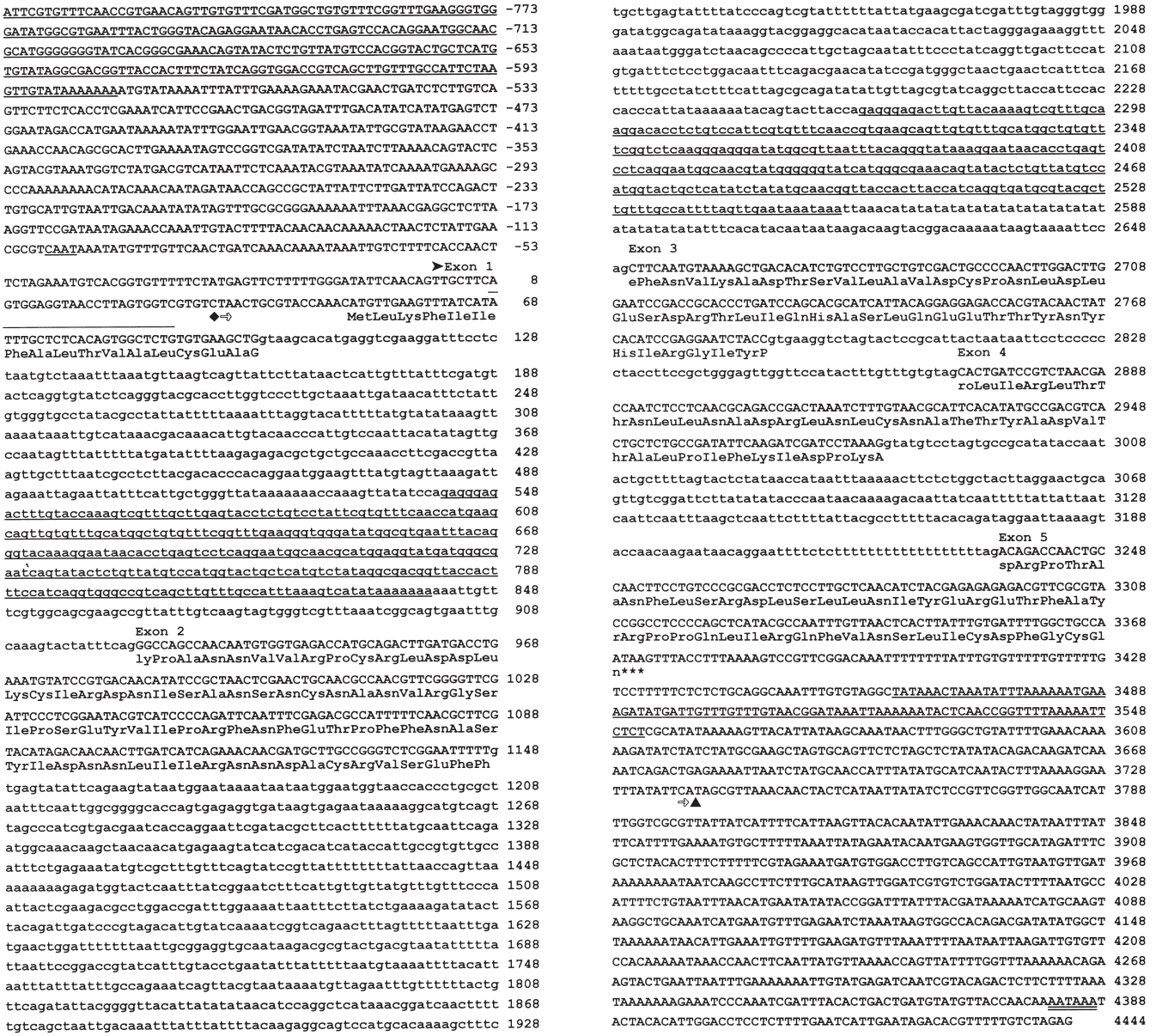

Fig. 3. Nucleotide sequence of *Galleria P25* gene with the 5′-flanking region and the deduced amino acid sequence. Nucleotides are numbered (thick arrow = +1) from the central position of three nucleotides that were identified as alternative transcription starts (Fig. 4). The CAAT box is underlined and the potential polyadenylation signal double-underlined, the introns are shown in lower-case letters, and the stop codon is marked with asterisks. The repetitive elements *Gm1* in 5′ *UTR* and the 1st and 2nd introns, and the insertion element *Lep1* in 3′ *UTR* are underlined. The complementary oligonucleotide primer used for the primer extension experiment is overlined. The beginning of sequenced cDNA is earmarked with a diamond, and the end with a triangle. The gene sequence has EMBL Accession Number AF009677.

hoped that a comparison of *Galleria P25* gene with the *Bombyx* silk genes would reveal *cis*-motifs involved in this regulation.

3.2. *Isolation of* GalleriaP25 *gene*

Several genomic clones were found to hybridize with our *P25* cDNA probes. On the basis of restriction maping and Southern blot analysis, we chose to sequence appropriate parts of two overlapping clones GGE1 and GGE2, which included the entire *P25* gene. The restriction map and sequencing strategy of the clones are shown in Fig. 2. A comparison of the established sequence of 5217 nt (Fig. 3) with the *P25* cDNA sequence (Zurovec et al., 1998) confirmed that these sequences corresponded to the same gene. The previously reported cDNA sequence began in position +34 of our gene numbering, included all coding exons, and ended at +3739. The cDNA and the genomic sequences differed only in 3 nt; two at silent codon positions (C in place of T at +1043 and G in place of C at +3258), and one (T instead of A at +3374) in the 3′ *untranslated*

region (*UTR*). These discrepancies probably reflected an actual difference between the two *Galleria* populations that were used in our research.

Distribution of sequences homologous to the cDNA revealed that the genomic clone GGE1 contained several kilobases (kb) long 5′ flanking region and extended into the second intron. Clone GGE2 ranged from the start of the second exon far beyond the known 3′ *UTR*. The overlap included the second exon and about half of the following intron.

Genomic Southern blot analysis indicated that *Galleria P25* is a single copy gene (Fig. 4), like all the other silk genes that have been analyzed (Lizardi and Brown, 1975; Okamoto et al., 1982; Couble et al., 1985; Kikuchi et al., 1992; Zurovec et al., 1992).

3.3. The structure of the P25 gene

Comparison of the genomic sequence with the *P25* cDNA (Zurovec et al., 1998) disclosed that the *P25* gene was interrupted by four introns. The donor and acceptor sequences of all exon–intron junctions conformed to the GT/AG rule (Breathnach et al., 1978). Exon 1 (101 nt) encoded the 5′-*UTR* and the signal peptide; exon 2 (223 nt) the N-terminus of the secreted protein; exon 3 (136 nt) and exon 4 (110 nt) the internal part of the protein; and exon 5 encoded protein C-terminus and contained 3′ *UTR*.

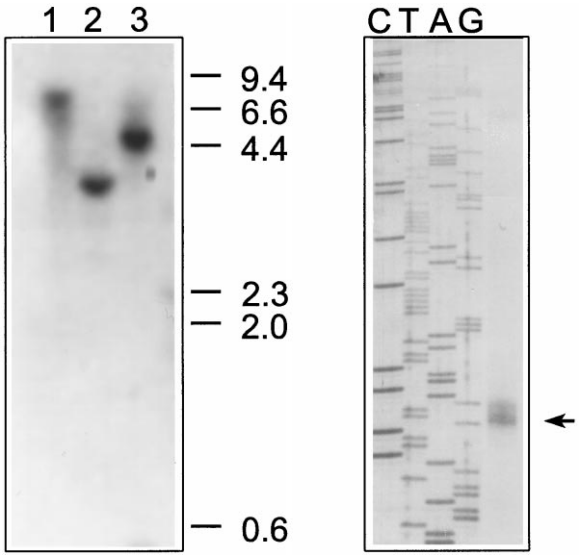

Fig. 4. Southern blot analysis of the genomic DNA (left) and primer extension mapping (right). The Southern analysis was performed with 5 μg DNA digested with *Xba*I (lane 1), *Sal*I (lane 2), and *Pst*I (lane 3), respectively, and hybridized to radiolabeled 5′ fragment (328 bp) of *P25* cDNA. Fragments of *Hind*III-digested λ phage DNA were used to construct the molecular size ladder (on the right). Oligonucleotide complementary to nt 68–89 of the gene (Fig. 3) was used for the extension mapping. The product was analyzed in a sequencing gel along with the G, A, T, C ladders of a genomic subclone that was sequenced with the same primer. The result reveals three possible transcription starts at G, T (used for sequence numbering in Fig. 3), and T (arrow).

The primer extension analysis, which was performed on total RNA from the posterior section of silk glands, demonstrated that transcription may be initiated at one of three adjacent nucleotides (Fig. 4). The T residue 50 nt upstream from the ATG codon seems to be used as major transcription start point (tsp); therefore, our sequence numbering begins at this nt. However, transcription may also begin at the preceding G or the following T residues. Such a variation of tsp position is not uncommon in genes that lack the TATA box (Reynolds et al., 1984). The pentamer ACAGT, which begins in position −3 to −5 to the transcription starts (G, T, and T, respectively, cf. Fig. 3), was determined as a consensus capsite for many eukaryotic genes (Bucher, 1990). In *Drosophila*, the pentamer T(A,G)CAGT was described as a transcription initiator or a facultative promoter element, which is located within −10 and +10 of tsp and whose function depends on the array of other promoter *cis*-elements present in the gene (Cherbas and Cherbas, 1993).

We did not establish the precise length of exon 5, but a putative polyadenylation signal AATAAA (Proudfoot, 1991) was found in position +4382. We assume that the poly(A) tail begins, as in other silk genes, 20–40 nt after this signal, and the sequence shown in Fig. 3 covers the whole gene. Hence, the putative 3′ *UTR* includes about 800 nt, and the entire transcript nearly 1800 nt; the deviation from the assessed 2-kb size of P25 mRNA (Zurovec et al., 1998) can be accounted for by the length of a poly(A) tail.

The exon–intron boundary domains and the length of exons of *Galleria P25* gene are almost identical as described for the *Bombyx P25* (Chevillard et al., 1986). The *Galleria* gene, however, contains much longer noncoding sequences than the *Bombyx* counterpart. Extended lengths of the first and second introns of *Galleria P25* are due to insertions of a SINE motif, which is also present in the upstream region (see Section 3.5); a different insertion element extends the length of the 3′-*UTR* (see Section 3.6). *P25* genes of *Galleria* and *Bombyx* also seem to differ in their regulatory regions.

3.4. Features of the 5′-flanking region

No canonical TATA box sequence was found around position −30 of the 5′-flanking region (Fig. 3). Two typical TATA recognition sequences TATAAAA, however, were evident at −588 to −584 and −575 to −571 positions from the major tsp. A similar situation is known from some other genes, for example a gene encoding human thrombin receptor (Schmidt et al., 1996). Since the relevance of a TATA box at such a distance from the tsp has not been proven, we regard

the *Galleria P25* gene as TATA-less. This is common in the 'housekeeping' genes (Azizkhan et al., 1993)—about one-half of known *Drosophila* promoters are TATA-less (Arkhipova, 1995)—but the absence of a TATA box in a silk gene is exceptional. The 5′ flanking region of *Galleria P25* contains, however, a CAAT box (Fig. 3) that is located in the usual position around 80 nt upstream from the tsp (Breathnach and Chambon, 1981).

Using sequence alignment, we did not detect in *Galleria P25* any of the *cis*-elements responsible for the spatial and temporal control of *Bombyx P25* and *Fib-H* expression (Nony et al., 1995; Takiya et al., 1997). However, juxtaposing 5′ flanking region of the *Galleria P25* gene to the sequences of other silk genes revealed in position −76 to −45 a putative element of 32 nt (Fig. 5). The element was over 65% identical with a sequence present in the 5′-flanking region of *Galleria Fib-H* (unpublished data). The core motif GTCTTTT and identical spacing of a few other nucleotides were also revealed in the upstream regions of *Bombyx* genes *P25* (Couble et al., 1985) and *light chain fibroin* (*Fib-L*) (Kikuchi et al., 1992), as well as in the intron of *Fib-H* (Hui et al., 1990b). Since all these genes are specifically expressed in the posterior section of silk glands, the motif may play some role in the territorial regulation of gene activity.

A search for transcription factor recognition elements using the MatInspector (Quandt et al., 1995) revealed a number of putative binding sites for homeoproteins (data not shown). Binding of several homeoproteins to regulatory silk gene sequences was demonstrated in *Bombyx* (Hui et al., 1990a; Mach et al., 1995, etc.), but the established *cis*-elements differed from the putative homeoprotein-binding sites that we found in *Galleria P25* in both their sequences and positions.

3.5. Short interspersed nuclear element

The genomes of higher eukaryotes contain multiple copies of short interspersed nuclear elements (SINEs) (Davidson et al., 1973). A repetitive SINE-type element of about 300 nt, with a characteristic stretch of adenines at its 3′ end, was found in the upstream region and in the first and second introns of *Galleria P25* (Figs. 2 and 3). The copy in the upstream region obviously extended beyond the DNA sequence shown in Fig. 3. We named this new SINE-type element *Gm1* for *Galleria mellonella* insertion element No. 1.

```
GP25(-74~-43)        AAATAAATTGTCTTTTCACCAACTTCTAGAAA
Gfib-H(-346~-316)    ACTCAATGTTGCTTTTCACAAAGTACTA-AAA
P25(-940~-909)       ATTCCTTTAGTCTTTTATTTATCGACTAGCGA
Fib-L(-51~-23)       ATTTTAATTGTCTTTTT---ATATATAAATAG
Fib-H(468~499)       ATTGCTATTGCCTTTTTTTCGCAAATTATAAT
Fib-L(-113~-83)      AATGCAAAAGTCTTTTGAACGTTAGATGCTGT
```

Fig. 5. A 32-nt sequence motif identified in the 5′-flanking regions of *Galleria* genes *P25* (GP25) and *Fib-H* (GFib-H), in the 5′-flanking regions of *Bombyx P25* and (twice) *Fib-L* genes, and in the intron of *Bombyx Fib-H*. Multiple alignment was performed using Clustal method in the DNASTAR program. Numbers in parentheses indicate nt positions.

Analyses of the data retrived from the GEN/EMBL database revealed that *Gm1* occurred also in other genes (Fig. 6). Copies inserted into two *Galleria* genes for larval hemolymph proteins—in the upstream region of *Lhp76* (Memmel et al., 1992) and in the intron 6 of *Lhp82* (Memmel et al., 1994)—proved to be identical with the *Gm1* copies in the *P25* gene. Truncated versions of *Gm1* were found in the upstream region and in the intron of *Galleria Fib-H* gene and in the second intron of a *Galleria* sericin gene (unpublished data, not included in Fig. 6).

Gm1 is highly homologous with a repetitive DNA element of the spider *Nephila clavipes* (GenBank Accession No. U54613) as well as with the *Bm1* repetitive elements of *Bombyx* (Adams et al., 1986; Wilson et al., 1988). Interestingly, *Bm1* is more similar to the repetitive element of the evolutionary remote spider than to *Gm1* (cf. Fig. 6). Although SINE-like elements were rarely found in arthropods (Robertson and Lampe, 1995), our data suggest that a family of *Gm1*-like elements, which must have evolved before the ancestors of spiders diversified from those of insects, is widespread at least in some insect species.

3.6. Insertion element shared by lepidopteran insects

A second sequence regarded as an insertion was identified in a single copy in the 3′ *UTR* of *Galleria P25*. The borders of the element were not precisely defined, but alignment with similar sequences that were detected in the DNA of several lepidopteran species revealed a highly conserved box between nt +3453 and +3552. About 20 lepidopteran DNA sequences containing this box or a substantial part of it were found in the GEN/EMBL Database. To emphasize that homologous sequences were identified exclusively in Lepidoptera, we called the box *Lep1*.

Ten *Lep1* boxes were disclosed in the known DNA sequences of *Bombyx mori*; six of them are compared in Fig. 7. Nearly identical boxes, which are regarded here as *Lep1* consensus, were found in the intron of *storage protein 2* gene (Fujii et al., 1989) and in the spacer separating genes *bombyxin B-9* and *C-1* (Kondo et al., 1996). A full match of these sequences begins in position 10 of the box and continues for 48 nt beyond the box; subsequent 88 positions are also nearly identical. In total, a high homology of these two insertions extends for about 250 nt. A high match (over 90%) with the box consensus and with about 40 nt flanking the box down-

```
UP      ----------------------------------------------TCGTGTTTCAACCGTGAA-CAGTTGTGTTTCGATGGCTGTGTTTCGGTTTGAAG
INTRON1 GGGAGACTTTGTACCAAAGTCGTTTGCTTGAGTACCTCTGTCCTATTCGTGTTTCAACCATGAAGCAGTTGTGTTTGCATGGCTGTGTTTCGGTTTGAAG
INTRON2 GGGAGACTTGTTACAAAAGTCGTTTGCAAGGACACCTCTGTCC-ATTCGTGTTTCAACCGTGAAGCAGTTGTGTTTGCATGGCTGTGTTTCGGTCTCAAG
Lhp76   GGGAGACTTTGTACAAAATTCGATTACTTGGATACCTCTGTCCCATTCGTGTTTCAACCGTAAGGCAGTTGTGTTTGCATGGCTGTGTTTCGGTTTGAAG
Lhp82   GGGAGACTTTGTACAAAAGTCGATTACTTGGGTACCTCTGTCCCATTCGTGTTTCAACCGTAAAGCAGTTGTGTTTGCATGGCTGTGTTTCGGTTTGAAG
Bm1     GGGAGGACCTCTTGTGAGTCCGCACGGGTAGGTACCACCGCCCTGCC--TATTTCTGCCGTGAAGCAGTAATGC------------GTTTCGGTTTGAAG
Nc1     ----------------------------------------------CC--TATTTCTGCCGTGAAGCAGTAATGC------------GTTTCGGTTTAAAG

UP      GGTGGGATATGGCGTGAATTTACTGGGTACAGAGGAATAACACCTGAGTCCACAGGAATGGCAACGCATGGGGGGTATCACGGGCGAAACAGTATACTCT
INTRON1 GGTGGGATATGGCGTGAATTTACAGGGTACAAAGGAATAACACCTGAGTCCTCAGGAATGGCAACGCATGGAG-GTATGATGGGCGAATCAGTATACTCT
INTRON2 GGAGGGATATGGCGTTAATTTACAGGGTATAAAGGAATAACACCTGAGTCCTCAGGAATGGCAACGTATGGGGGGTATCATGGGCGAAACAGTATACTCT
Lhp76   GGTGGGATATGGTGTGAATTTACTGGGTACAGAGGAATACCACCTGAATCGTCAGGAATGGCAACACATGGGGTGTATCATGGGCGAAATAGTATACTCT
Lhp82   GGTGGGATATGGCGTGAATTTACAGGGTACAGAGGAATACCACCTGAATCCTCAGGAATGGCAACGTATGGGGGGCATCATGGACGAAACAGTATACTCT
Bm1     GGTGGGGCA-GCCGT------------TGTA-----ACTATACTG-AGACCTTAGAACTTATATCTCAAGGTGGGT------GGCG------CAT-----
Nc1     GGTGGGGCA-GCCGT------------TGTA-----ACTACACTTCAAACCTTAGAGCTTATATCTCAAGGTGGGT------GGCG------CAT-----

UP      GTTATGTCCACGGTACTGCTCATGTGTATAGGCGACGGTTACCACTTTCTATCAGGTGGACCGTCAGCTTGTTTGCCATTCTAAGTTGTATAAAAAAA
INTRON1 GTTATGTCCATGGTACTGCTCATGTCTATAGGCGACGGTTACCACTTTCCATCAGGTGGGCCGTCAGCTTGTTTGCCATTTAAAGTCATATAAAAAAA
INTRON2 GTTATGTCCATGGTACTGCTCATATCTATATGCAACGGTTACCACTTACCATCAGGTGATGCGTACGCTTGTTTGCCATTTTA-GTTGAATAAATAAA
Lhp76   GTTATGTCCATGGTACTGCTCATGTGTAAAGGCGACGGTTATTACTTTCCATCAGGTGGGTCGTCAGCTTGTTTGCCATTCTAAGTTGTATAAAAAAA
Lhp82   CTTATGTCCATGATACTGCTCTTGTCTATAGGCGACGGTTATTACTTTCCATCAGGTGGGCCGTCAGCTTGTTTGCCATTCTAAGTAGAATAAAAAAA
Bm1     -TTACGTT-GTAG--------ATGTCTATGGGCTCCGGTAACCACTTAACTTCAGGTGGGCTGTGAGCTCGTTTATCCATCTAAGCAATAAAAAAAAA
Nc1     -TAACGTT-GTAG--------ATGTGCGTGGGCTCCAGTAACCACTTAACACCAGGTGGGCTGTGAGCTCGTCCACCCATCTAAGCTATAAAAAAAAA
```

Fig. 6. Multiple alignment of *Gm1* SINE elements in the upstream region (UP), first intron (Intron 1) and second intron (Intron 2) of *P25*, in the upstream region of *Lhp76*, and in intron 6 of *Lhp82 Galleria* genes. Sequences of the repeats *Bm1* from the genome of *Bombyx mori* and *Nc1* from the spider *Nephila clavipes* are also juxtaposed.

```
Bm SP2    TT-CATGATTGACTTCCACGGTAAAGGAATAAC--ATCTATACTAATATTATAAAGAGGAAAGATTTGTTTGTTTGTTTGTTTCGAATAG
Bm Space  TATTA--ATTGAAA---ACATGAAATTTATAAT-TATCTATACTAATATTATAAAGAGGAAAGATTTGTTTGTTTGTTTCGAATAG
Bm RAPD   CTGTAAATCATT-TGAAGGGTCGTATATAGAATATATCTATACTAATATTATAAAGAGGAAAGATTTGTTTGTTTGTTTGTATTGAATAA
Bm Fib-L  TCTTATTATATCACATCCAATTGCTTATCTA-TACTTCTATACTAATATTATAAAGAGGAAAGATTTGTTTGTTTGTTTGTATTAAATAT
Bm Sorb   AAAAATCAT-CTATACTAATCTACTAATAAAATACTTCTATACTAATATTATAAAGAGGAAAGCTTTGTTTGTTTGTTTGTATTGGATAG
Gm P25    TCCTTTTTCTCTCTGCAGGCAAATTTGTGTAG-GCTATAAACTAAATATTTAAAAAATGAAAGATATGATTGTTTGTTTGTAACGGATAA
Bm P25    AACTACCA-AACATTGTGTACCCTTTA-ATGATA-TTCTAATCTATATATATAAAAATGAATTGCT-GTTCGTTAGTCTCGCT--AAAA-
Hc Cecr                                ... ATA-TAGGGTAATACATTTAAAAAAAAAATTG-TGTGTCTGTTTGTTCCGGC-TAAT--
Hc Atac                                  ...TA-AAAATTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTTTGTTCCGGG-TAAT--
Hc Cut12                               ... CTGTCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTTTGTTCCGGG-TAAT--

Bm SP2    GCTCCGAAACTACTGGACCGATTT-GAAAAA-TTCTTTTTCCATTAGAAGCAACATTGTCCCTGATGAACATAGGCTACATTTTTTAATTT
Bm Space  GCTCCGAAACTACTGGACCGATTT-GAAAAA-TTCTTTTTCCATTAGAAGCGACATTGTCCCTGATGAACATAGGCTACTTTTTTTATTTA
Bm RAPD   GCTCCGAAACTACTGAACCGATTT-GAAAAA-TTCCTTCACTGTTTGGAAGTACATCATTCGCGAGTGACTTAGGCTATAATCTTT-TTTG
Bm Fib-L  TCTCCGAAACTGCTGAACCGATTT-GAAAAA-TTCTTTCACTGTTTGGA-GTACACTATTCCCGAGTGACATAGGCTATAATATTT-TTTG
Bm Sorb   GTTCCGAAACTGCTGAACCGATTT-GAAAAA-TTCTTACACTGTTTGGAAGTACACTATTCCTGAGTGACATAGGGGTGGTGCGATCAATG
Gm P25    ATTAAAAAA-TACTCAACCGGTTT--TAAAAATTCTCTCGCATATAAAAAGTACATTATAAGCAAATAACTTTGGGCTGTATTTTGAAACA
Bm P25    -CTCGAGAACGGCCGGACCGATTT-GGCTAA-TT-TTGGTCTTGAATTATTTGTGGAAGTCCAGAGAAGATTTAGAAGGTTTAAATAAATA
Hc Cecr   -CTTCTGAACGGCTGGACCGATATTGATGGAATTTTTTTAGGCAGGTAGTGAAATTGATTTGCC--TAACTTAGGCTACTTTTATTGCTTG
Hc Atac   -CTCCGGAACG-CTGGACGGATTTTGACGGGACTTTTTTTTGGCAGAGAGTGAAGTTACT--GCGGGTAACATAGGGTACTTTTTATTTTCG
Hc Cut12  -CTCCGGAACGGCTGGACGGATTTTGACGGGACTTTTTTTTGGCAGATAGTGAAGTTACT--GCGGGTAACTTAGGGTACTTTTTATTTTCG

Bm SP2    TTTTTTTT-------TTTTTTGT-TTCATGTGTGTTTTAATGTTTCCGAAGCGAAG-CGAGGGCGGGTCGCTAGTCGTGTAATAAAAATCAA
Bm Space  TTTTTTTATTTTTGATTTTTGGT-TTCATGTGTGTTTTAATGTTTCCGAAGCGAAGACGAGGGCGGGTCGCTAGTAATTATATAATATATGA
Bm RAPD   AAAAAAAAATAGTGCTCTTTACTAAAACTCCAG----TAATGTAACCCAAGGTGTAAAAAAATTACATAAAATATTCATTACACCGGGTGCC
Bm Fib-L  AAAAAATT-AGGGA-TCCTTACTAAAACTCCAA----TAATGTAACCAAAAGTGTAAAAAAATTACCTAAAATACTCTTTACATCGCGTGCC
Bm Sorb   GCCTTATACGGGG--CGCCC-GTATGGGGCCAGTCCTGACCGTGGGGGTACC-GAAGCTGCTGCAAC--GCCGCAACGCACCATCGGGGTCA
Gm P25    CAAAAAGATATCTATCTATGCGAAGCTAGTGCAGTTCTCTAGCTCTATATACAGACAAGATCAAAATCAGACTGAGAAAATTAATCTATGCA
Bm P25    TGAAAATGCTCGGAATTAAATAAAAATAACAATTTTGTTTTTTCTTTGATGTGTTCCCCTCGGACGGATTCCTTTAGTCTTTTATTTATCGA
Hc Cecr   GCAATAATTAAGGTTCCGTGTAAAAAA
Hc Atac   AAAATAATTAGGGTTCCGCGTAAAAAA
Hc Cut12  AAAATAATTAGGGTTCCGCGTAAAAAA
```

Fig. 7. Multiple alignment of the *Lep1* box and adjacent sequences detected in *Galleria P25* 3′ *UTR* (*Gm* P25) with *Bombyx* sequences present in the *storage protein 2* gene (*Bm* SP2), a spacer in *bombyxin* locus (*Bm* Space), randomly amplified DNA (*Bm* RAPD), *Fib-L* gene (*Bm* Fib-L), *sorbitol dehydrogenase* gene (*Bm* Sorb), and *P25* gene (*Bm* P25); and with *Hyalophora cecropin* gene (*Hc* Cecr), attacin locus (*Hc* Attac), and *cuticular protein 12* gene (*Hc* Cut12).

stream was detected in the first intron of *Fib-L* (Kikuchi et al., 1992), in a randomly amplified DNA sequence (Abe et al., 1995), in the fourth intron of the *sorbitol dehydrogenase* gene (Niimi et al., 1996), and in the third intron of the *xanthin dehydrogenase* gene (GenBank Accession No. AB005911; Komoto et al., 1997). The insertions in the *Fib-L* intron and in the randomly amplified sequence are nearly identical for additional 100 nt; matching of these insertions spans about 250 nt (Fig. 7), whereas insertions in the intron of *sorbitol dehydrogenase* (Fig. 7) and in the intron of *xanthin dehydrogenase* (data not shown) are rather dissimilar

beyond about 10 nt upstream and 50 nt downstream of the box.

Only about 50% homology within the box consensus and occasional stretches of matching nucleotides downstream of the box (Fig. 7) were found in the 5′ *UTR* of *Bombyx p25* (Chevillard et al., 1986), as well as in the presently identified *Lep1* insertion in the 3′ *UTR* of *Galleria P25* (Fig. 3).

A somewhat different group of elements lacking the first third of the *Lep1* box was disclosed in the moth *Hyalophora cecropia*. All detected sequences included only two copies of the TGTT motif that was present four times in the *Lep1* consensus (Fig. 7). In the sequence between the *attacin* pseudogene and the functional gene for basic attacin (Sun et al., 1991), and in the 5′ *UTR* of the gene for cuticle protein 12 (Binger and Willis, 1994), the TGTT motif is preceded by seven or eight TGTC repeats (Fig. 7). By contrast, unique sequences terminated with just one TGTC motif precede the *Lep1* homology in the introns of the genes for cecropin (Gudmundsson et al., 1991) (Fig. 7) and for cuticle protein 66 (Lampe and Willis, 1994) (not shown). All *Hyalophora* sequences with the *Lep1*-like box are over 80% identical also in about 80 nt downstream of the box (Fig. 7).

The TGTC repeats were recognized in the *Hyalophora attacin* locus by Sun et al. (1991) as part of a 1264-nt sequence ACATCTA... $(TGTC)_7$... TAGACAT (underlined parts are inverted repeats). The authors suggested that the direct repeat ACAT was the target site of an insertion. This is consistent with our assumption that the *Lep1* box is a small internal part of a larger inserted element. It seems that a mobile element common to Lepidoptera was modified during evolution to diverse forms. For a majority of the elements that we compared, *Lep1* was the most conserved part. However, a stretch of cca 140 nt, which is shared by the *Hyalophora* elements, does not include the first third of *Lep1*.

The alignments in Fig. 7 show that *Lep1*-type insertions detected in *Bombyx* could be classified into several categories that differ in the degree of box homology and particularly in the diversification of the regions flanking the box. A single *Lep1*-type insertion presently identified in *Galleria* represents a related type, whereas the insertions known from *Hyalophora* are of two other kinds. It seems that insertions are residues of an ancient mobile element. Some of their modifications, such as the lack of the first third of *Lep1* box, are taxon-specific, whereas others are found side by side within a single genome.

4. Conclusions

(1) The expression of *Galleria* silk gene *P25* is down-regulated by molt-inducing concentrations of ecdysteroid hormones.

(2) The gene, which was isolated as two overlapping genomic clones, is present in one copy per genome, includes about 4440 nt and is interrupted by four introns.

(3) No canonical TATA box sequence is located at appropriate position of the gene. The 5′-flanking region contains a 32-nt sequence that is partly conserved in all genes that are expressed in posterior silk gland section.

(4) DNA repeat *Gm1* of 300 nt, which is homologous to SINE-like element *Bm1* of the silkworm and to a spider repetitive element, was disclosed in the upstream region, intron 1 and intron 2 of the *P25* gene, and in several other *Galleria* genes.

(5) An apparent insertion detected in the 3′ *UTR* of *Galleria P25* gene contains a conserved *Lep1* box of cca 100 nt that is widespread in lepidopteran insects. It represents the core region of a DNA element that occurs in Lepidoptera in species-specific modifications.

Acknowledgement

Our study was supported by grant 204/96/1100 of the Grant Agency of the Czech Republic, and by Volkswagenstiftung grant I68 514.

References

Abe, H., Shimada, T., Yokoyama, T., Oshiki, T., Kobayashi, M., 1995. Identification of random amplified polymorphic DNA on the W chromosome of the Chinese 137 strain of the silkworm, *Bombyx mori*. J. Sericult. Sci. Jpn. 64, 19–22.

Adams, D.S., Eickbush, T.H., Herrera, R.J., Lizardi, P.M., 1986. A highly reiterated family of transcribed oligo(A)-terminated, interspersed DNA elements in the genome of *Bombyx mori*. J. Mol. Biol. 187, 465–478.

Arkhipova, I.R., 1995. Promoter elements in *Drosophila melanogaster* revealed by sequence analysis. Genetics 139, 1359–1369.

Azizkhan, J.C., Jensen, D.E., Pierce, A.J., Wade, M., 1993. Transcription from TATA-less promoters: dihydrofolate reductase as a model. Crit. Rev. Eukaryot. Gene Expr. 3, 229–254.

Bello, B., Couble, P., 1990. Specific expression of a silk-encoding gene of *Bombyx* in the anterior salivary gland of *Drosophila*. Nature 346, 480–482.

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., Chambon, P., 1978. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon–intron boundaries. Proc. Natl. Acad. Sci. USA 75, 4853–4857.

Breathnach, R., Chambon, P., 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349–383.

Binger, L.C., Willis, J.H., 1994. Identification of the cDNA, gene and promoter for a major protein from flexible cuticles of the giant silkmoth *Hyalophora cecropia*. Insect Biochem. Mol. Biol. 24, 989–1000.

Bucher, P., 1990. Weight matrix descriptions of four eukaryotic RNA

polymerase II elements derived from 502 unrelated promoter sequences. J. Mol. Biol. 212, 563–578.
Cherbas, L., Cherbas, P., 1993. The arthropod initiator: the capsite consensus play an important role in transcription. Insect Biochem. Mol. Biol. 23, 81–90.
Chevillard, M., Couble, P., Prudhomme, J.C., 1986. Complete nucleotide sequence of the gene encoding the *Bombyx mori* silk protein P25 and predicted amino acid sequence of the protein. Nucleic Acids Res. 14, 6341–6342.
Couble, P., Chevillard, M., Moine, A., Ravel-Chapuis, P., Prudhomme, J.C., 1985. Structural organization of the *P25* gene of *Bombyx mori* and comparative analysis of its 5′ flanking DNA with that of the fibroin gene. Nucleic Acids Res. 13, 1801–1814.
Davidson, E.H., Hough, B.R., Amenson, C.S., Britten, R.J., 1973. General interspersion of repetive with non-repetive sequence elements in the DNA of *Xenopus laevis*. Eur. J. Mol. Biol. 77, 1–23.
Fujii, T., Sakurai, H., Izumi, S., Tomino, S., 1989. Structure of the gene for the arylphorin-type storage protein Sp 2 of *Bombyx mori*. J. Biol. Chem. 264, 11020–11025.
Gudmundsson, G.H., Lidholm, D.-A., Asling, B., Gan, R., Boman, H.G., 1991. The cecropin locus: Cloning and expression of a gene cluster encoding three antibacterial peptides in *Hyalophora cecropia*. J. Biol. Chem. 266, 11510–11517.
Hui, C.-C., Suzuki, Y., Kikuchi, Y., Mizuno, S., 1990a. Homeodomain binding sites in the 5′ flanking region of the *Bombyx mori* silk fibroin light-chain gene. J. Mol. Biol. 213, 395–398.
Hui, C.-C., Matsuno, K., Suzuki, Y., 1990b. Fibroin gene promoter contains a cluster of homeodomain binding sites that interact with three silk gland factors. J. Mol. Biol. 213, 651–670.
Kikuchi, Y., Mori, K., Suzuki, S., Yamaguchi, K., Mizuno, S., 1992. Structure of the *Bombyx mori* fibroin light-chain-encoding gene: upstream sequence elements common to the light and heavy chain. Gene 110, 151–158.
Komoto, N., Yukuhiro, K., Tamura, T., 1997. Gene structure and expression pattern of tandemly duplicated xanthine dehydrogenase genes of the silkworm (*Bombyx mori*). GenBank Accession No. AB005911.
Kondo, H., Ino, M., Suzuki, A., Ishizaki, H., Iwami, M., 1996. Multiple gene copies for bombyxin, an insulin-related peptide of the silkmoth *Bombyx mori*: structural signs for gene rearrangement and suplication responsible for generation of multiple molecular forms of bombyxin. J. Mol. Biol. 259, 926–937.
Lampe, D.J., Willis, J.H., 1994. Characterization of a cDNA and gene encoding a cuticular protein from rigid cuticles of the giant silkmoth, *Hyalophora cecropia*. Insect Biochem. Mol. Biol. 24, 419–435.
Lizardi, P.M., Brown, D.D., 1975. The length of the fibroin gene in the *Bombyx mori* genome. Cell 4, 207–215.
Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T., Suzuki, Y., 1995. Silk gland factor-I involved in the regulation of *Bombyx Sericin-1* gene contains fork head motif. J. Biol. Chem. 270, 9340–9346.
Memmel, N.A., Trewitt, P.M., Grzelak, K., Rajaratnam, V.S., Kumaran, A.K., 1994. Nucleotide sequence, structure and developmental regulation of LHP82, a juvenile hormone-suppressible hexamerin gene from the waxmoth, *Galleria mellonella*. Insect Biochem. Physiol. Mol. Biol. 24, 133–144.
Memmel, N.A., Trewitt, P.M., Silhacek, D.L., Kumaran, A.K., 1992. Nucleotide sequence and structure of the arylphorin gene from the waxmoth, *Galleria mellonella*. Insect Biochem. Physiol. Mol. Biol. 22, 133–144.
Niimi, T., Yamashita, O., Yaginuma, T., 1996. Structure of the *Bombyx* sorbitol dehydrogenase gene: A possible alternative use of the promotor. Insect Mol. Biol. 5, 269–280.
Nony, P., Prudhomme, J-C., Couble, P., 1995. Regulation of the P25 gene transcription in the silk gland of *Bombyx*. Biol. Cell 84, 43–52.
Okamoto, H., Ishikawa, E., Suzuki, Y., 1982. Structural analysis of sericin genes. Homologies with fibroin gene in the 5′ flanking nucleotide sequences. J. Biol. Chem. 257, 15192–15199.
Prudhomme, J.C., Couble, P., 1995. Regulation of the transcription of the *P25* gene in the silk glands of *Bombyx mori*. In: Ohnishi, E., Sonobe, H., Takahashi, S.Y. (Eds.), Recent Advances in Insect Biochemistry and Molecular Biology. University of Nagoya Press, Nagoya, pp. 369–382.
Proudfoot, N., 1991. Poly (A) signals. Cell 64, 671–674.
Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T., 1995. MatInd and MatInspector—New fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23, 4878–4884.
Reynolds, G.A., Basu, S.K., Osborne, T.F., Chin, D.J., Gil, G., Brown, M.S., Goldstein, J.L., Luskey, K.L., 1984. HMG CoA reductase: A negatively regulated gene with unusual promoter and 5′ untranslated region. Cell 38, 275–285.
Robertson, H.M., Lampe, D.J., 1995. Distribution of transposable elements in arthropods. Annu. Rev. Entomol. 40, 333–357.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schmidt, V.A., Vitale, E., Bahou, W.F., 1996. Genomic cloning and characterization of the human thrombin receptor gene. J. Biol. Chem. 271, 9307–9312.
Sehnal, F., 1966. Kritisches Studium der Bionomie und Biometrik der in verschiedenen Lebensbedingungen gezuchteten 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.
Sehnal, F., Maroy, P., Mala, J., 1981. Regulation and significance of ecdysteroid titre fluctuations in lepidopterous larvae and pupae. J. Insect Physiol. 27, 535–544.
Takiya, S., Hui, C.-C., Suzuki, Y., 1990. A contribution of the core-promoter and its surrounding regions to the preferential transcription of the fibroin gene in posterior silk gland extracts. EMBO J. 9, 489–496.
Takiya, S., Kokubo, H., Suzuki, Y., 1997. Transcriptional regulatory elements in the upstream and intron of the fibroin gene bind three specific factors POU-M1, Bm Fkh and FMBP-1. Biochem. J. 321, 645–653.
Tanaka, K., Mori, K., Mizuno, S., 1993. Immunological identification of the major disulfide-linked light component of silk fibroin. J. Biochem. 114, 1–4.
Sun, S.-C., Lindstrom, I., Lee, J.-Y., Faye, I., 1991. Structure and expression of the attacin genes in *Hyalophora cecropia*. Eur. J. Biochem. 196, 247–254.
Wilson, E., Conliffe, D.P., Sprague, K., 1988. Transcriptional properties of *BmX*, a moderately repetitive silkworm gene that is an RNA polymerase III template. Mol. Cell. Biol. 8, 624–631.
Yang, C., Krishnan, M., Zurovec, M., Sehnal, F., Scheller, K., 1995. Correlations between ecdysteroid titre and silk gene expression in the last instar larvae of *Galleria mellonella* (Lepidoptera: Pyralidae). Eur. J. Entomol. 92, 229–234.
Yang, C., Sehnal, F., Scheller, K., 1996. Juvenile hormone restores larval pattern of sericin gene transcripts. Arch. Insect Biochem. Physiol. 22, 353–362.
Zurovec, M., Sehnal, F., Scheller, K., Kumaran, A.K., 1992. Silk gland specific cDNAs from *Galleria mellonella* L. Insect Biochem. Molec. Biol. 22, 55–67.
Zurovec, M., Vaskova, M., Kodrık, D., Sehnal, F., Kumaran, A.K., 1995. Light-chain fibroin of *Galleria mellonella*. Gen. Mol. Biol. 247, 1–6.
Zurovec, M., Kodrık, D., Yang, C., Sehnal, F., Scheller, K., 1998. The P25 component of *Galleria* silk. Gen. Mol. Biol. 257, 264–270.