Eur. J. Biochem. **268**, 2064–2073 (2001) © FEBS 2001

Insect silk contains both a Kunitz-type and a unique Kazal-type proteinase inhibitor

Xavier Nirmala, Dalibor Kodrík, Michal Žurovec and František Sehnal

Entomological Institute, Academy of Sciences, and the Faculty of Biological Sciences, University of South Bohemia, Branišovská, České Budějovice, Czech Republic

Insect silk is made up of structural fibrous (fibroins) and sticky (sericins) proteins, and contains a few small peptides of hitherto unknown functions. We demonstrate that two of these peptides inhibit bacterial and fungal proteinases (subtilisin, proteinase K and pronase). These 'silk proteinase inhibitors' 1 and 2 (SPI 1 and 2) are produced in the middle section of the silk-secreting glands prior to cocoon spinning and their production is controlled at transcription level. The full length cDNA of pre-SPI 1 contains 443 nucleotides and encodes a peptide of 76 amino-acid residues, of which 20 make up a signal sequence. The mature SPI 1 (6056.7 Da, 56 residues) is a typical thermostable Kunitz-type proteinase inhibitor with Arg in P1 position. The cDNA of pre-SPI 2 consists of 260 nucleotides and yields a putative secretory peptide of 58 amino-acid residues. The functional SPI 2 (3993 Da, 36 residues) is a single-domain Kazal-type proteinase inhibitor with unique structural features: free segment of the N-terminus is reduced to a single amino-acid residue, lack of CysI and CysV precludes formation of the A-ring and provides increased flexibility to the C-ring, and absence of several residues around the normal position of CysV shortens and changes the α helix segment of the protein. The structure reveals that the length and arrangement of the B-ring, including exposure of the P1 residue, and the position of the C-terminus relative to the B-loop, are essential for the activity of the Kazal-type inhibitors.

Keywords: antimicrobial defense; *Galleria mellonella*; Kazal domain; proteinase inhibitors; silk.

Typical insect silk, which is product of a pair of labial silk glands of caterpillars, consists of water insoluble core and sticky coating [1]. The core represents actual silk fibre, while the coating provides the glue allowing formation of the cocoon, in which the caterpillar pupates. Most of our knowledge of silk composition comes from studies on the major commercial silk producer, domestic silkworm *Bombyx mori*, and on a model insect, the waxmoth *Galleria mellonella*. The silk core, which is derived from the posterior section of silk glands, includes large heavy chain fibroin that is disulfide-linked to light chain fibroin and is also associated with a protein called P25 [2]. The coating is provided by several glycoproteins, which are derived from two sericin genes expressed in the middle gland section [3,4]. In addition to these major silk proteins, several low molecular weight components were detected in the silk of both *Bombyx mori* and *Galleria mellonella* [5,6]. One of them, designated seroin, proved to be product of a discrete gene that is expressed specifically both in the middle and in the posterior sections of silk glands [7]. Seroin is hardly involved in silk fibre construction and coating. It has been suggested that it may play a role in protecting silk against microbial degradation.

In the present paper we report on the identification of two small silk components that belong to serine proteinase inhibitors, a class of proteins that were originally discovered in mammalian blood serum and are widespread in both plants and animals. They were grouped into 10 families [8], of which the Kunitz- and Kazal-type inhibitors are represented by the bovine pancreatic trypsin inhibitor and the ovomucoids, respectively [9]. Inhibitors of both these families mostly consist of several repeated domains, each made of 50–60 amino acids with six conserved cysteines, which form three intradomain disulfide bridges [10]. The size of the resulting rings and conservation of a few amino-acid residues characterize the family. No consensus sequence can be recognized in the reactive site that is exposed on the domain surface and mediates interaction with cognate proteinases; the specificity to certain kinds of proteinases is determined by P1 amino-acid residue at the reactive site [11].

A number of inhibitors belonging to the Kunitz family were found in the insects and proposed to have diverse functions. For example, those occurring in the hemolymph were reported to block activation of the prophenoloxidase and thereby regulate insect defense against pathogens and injury [12,13]. A Kunitz-type trypsin inhibitor detected in the silk of *Bombyx mori* was proposed to prevent inappropriate degradation of silk proteins during their secretion [14]. Kazal-type inhibitors were also identified in the hemolymph as blockers of the chymotrypsin but their real function is unknown [15,16]. A thrombin inhibitor consisting of two

Correspondence to F. Sehnal, Entomological Institute, Academy of Sciences, and the Faculty of Biological Sciences, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic. Fax: + 420 38 5300354, Tel.: + 420 38 5300350, E-mail: sehnal@entu.cas.cz

Abbreviations: LMCI, *Locusta migratoria* chymotrypsin inhibitor; SGPI, *Schistocerca gregaria* proteinase inhibitor; SPI 1, SPI 2, silk proteinase inhibitors 1 and 2, respectively; UTR, untranslated region.

(Received 24 November 2000, revised 2 February 2001, accepted 9 February 2001)

Kazal domains was identified in the salivary glands of the blood-sucking bug *Rhodnius prolixus* as an agent preventing blood-clotting during feeding [17].

In this paper we demonstrate that the silk of *Galleria mellonella* contains a typical Kunitz-type inhibitor, which we named SPI 1, and a strucurally unique Kazal-type inhibitor called SPI 2. Production of two different proteinase inhibitors and the persistence of their activity in the cocoons indicate that they have important functions. Based on their effect on the bacterial and fungal proteinases, and on the developmental profile of their production, we assume that they protect silk proteins from microbial destruction.

EXPERIMENTAL PROCEDURES

Animals

Larvae of the waxmoth, *Galleria mellonella* L. (Lepidoptera: Pyralidae), were reared on a semiartificial diet at 30 °C [18]. Water-anaesthetized final instar larvae or newly ecdysed pupae were used for tissue collection. Fat body, integument, and the middle and posterior parts of silk glands were dissected from 15 to 20 individuals, immediately frozen in liquid nitrogen, and stored at −80 °C. Freshly spun or aged cocoons were taken for protein analysis.

Extraction and purification of proteinase inhibitors

Six to seven clean cocoons (75 mg) were shredded into tiny pieces and immersed for 48 h in either (a) 1 mL 10 mM tris buffer, pH 7.0, containing 8 M urea, 2% SDS, and 5% 2-mercaptoethanol (added just before use) or (b) 1.2 mL 0.11% trifluoroacetic acid. The insoluble parts were removed, the extract (a) was desalted on C18 Sep-Pak Plus cartridge (Waters), and 100–500 μL aliquots of the extracts were taken for HPLC analysis on the Merck-Hitachi D-6000 chromatography system at a flow rate 1 $mL \cdot min^{-1}$ and UV monitoring at 215 nm.

The desalted extract (a) was fractionated on the Purospher STAR PR-18 encapped column (Merck), 55 × 4 mm, particles 3 μm, with a gradient 0–2 min 0% B, 2–32 min 0–100% B (A = 0.11% trifluoroacetic acid in water, B = 60% acetonitrile in 0.1% trifluoroacetic acid). Further purification of the fraction exhibiting proteinase inhibitory activity was performed on the same column with the gradient 0–2 min 40% B, 2–6 min 40–55% B, 6–20 min 55–70% B.

The trifluoroacetic acid extract (b) was purified by RP-HPLC on the LiChrospher WP 300 RP-18 column (Merck), 250 × 4 mm, particles 5 μm, with a gradient 0–2 min 10% B, 2–20 min 10–30% B and 20–40 min 30–40% B. The active fraction was re-chromatographed on the Purospher STAR RP-18 encapped column (see above), with a gradient 0–2 min 0% B, 2–20 min 0–30% B and 20–40 min 30–40% B.

Determination of the proteinase inhibitory activity

Trypsin was used as the target enzyme of proteinase inhibitors and its inhibition was measured with azocoll assay [19] modified for microplate ELISA photometric readers [20]. HPLC fractions were evaporated to dryness and the residues dissolved and serially diluted in 50 mM tris/HCl buffer, pH 7.5. Azocoll was washed (75 mg per 10 mL) in 50 mM tris/HCl buffer, pH 7.5, for 2 h, centrifuged, and resuspended in fresh buffer (15 $mg \cdot mL^{-1}$). Analyzed protein samples were pipetted into a 96-well microplate (100 μL per well). After addition of 100 μL bovine trypsin solution (2.5 $\mu g \cdot mL^{-1}$) and 100 μL homogenous azocoll suspension to each well, the plate was incubated under shaking at 37 °C for 2 h. Azocoll was allowed to settle for 15 min, and 100 μL aliquots of clear supernatants of the reaction mixtures were transferred onto another plate for optical density measurements at 490 nm (Spectra Max340 PC reader, Molecular Devices). Appropriate negative controls without trypsin and positive controls without the inhibitor were assayed simultaneously, and the difference in absorbance between the positive and the negative control was taken as 100% proteinase activity. A decrease of this activity indicated that the respective fraction contained a proteinase inhibitor. The potency of inhibition was eventually expressed in D_m values (peptide amount causing a median inhibition) [21]. Activities of purified proteinase inhibitors were also assayed with other proteinases as described in Results.

Protein quantification and analysis

Protein concentration in the analysed samples was determined with the bicinchoninic acid photometric method [22] on the UV 1601 Shimadzu spectrophotometer, using a commercial kit (Pierce Chemical Company) and bovine serum albumin (BSA) as standard. Purified proteins were desiccated in SpeedVac (Jouan, St Herblain, France). Mass spectrometry was performed on Biflex MALDI-TOF spectrometer (Bruber) using the α-cyno-3-hydroxy-cinnamic acid matrix. N-Terminal sequencing was performed in the commercial facility of the Medical College of Wisconsin (Milwaukee, WI, USA).

Detection of cDNAs encoding proteinase inhibitors

Total RNA was isolated from the whole silk glands of last instar *Galleria* larvae as described previously [23] and stored at −80 °C until required. Synthesis of the first and second strand cDNA was performed with the SMART III PCR cDNA synthesis kit (Clontech) on the thermal cycler GeneAmp PCR System 2400 (Perkin Elmer). The resulting enriched full length ds cDNA, flanked by the 5′ and 3′ Smart III adaptors, was diluted 100-fold and used as template for PCR. Alternatively, the first strand cDNA was synthesized using an adapter-poly (T) primer designed in our lab (5′-TGAGCAAGTTCAGCCTGGTTATTTTTTTT TTTTTTTTT-3′). To amplify the 3′ region of the desired cDNA, a degenerate forward primer was combined with the 3′ SMART III adaptor or with our adaptor. To obtain the 5′ region, Smart III 5′ oligonucleotide primer was used with a reverse primer complementary to a sequence in the identified 3′ region. In all cases, about 50 ng of the first strand or ds cDNA was amplified in a 50-μL reaction mixture containing 1 × PCR buffer (Top-Bio), 200 μM dNTP, 2 mM $MgCl_2$, 2.5 U *Taq* polymerase (Top-Bio), and 20 pmol of each primer. Temperature cycling was adjusted to the oligonucleotide composition of the primers but in general it consisted of 2 min initial denaturation at 94 °C, 30 cycles each including 30 s at 94 °C, 30 s at 50–60 °C,

and 1 min at 72 °C, and the final extension at 72 °C for 7 min.

Nucleotide sequencing and computational analysis of sequence data

PCR products were gel extracted with an extraction kit (Qiagen), ligated into the pGEM T easy vector (Promega), and sequenced using M13 reverse and forward primers in an ABI prism sequencer (Perkin Elmer model 310). BLAST & BEAUTY [24] software was used to screen the SWISS PROT database in the homology search. Alignments of nucleotides and amino-acid residues were carried out using CLUSTAL as implemented by the MEGALIGN program of the LASERGENE package (DNASTAR). Protein conformation was modelled with the Swiss Pdb Viewer [25].

Northern Blotting

Aliquots of 5 μg total RNA from various tissues were separated on 1% agarose gel containing formaldehyde and Mops, and capillary blotted to Hybond-N^+ nylon membrane (Amersham). PCR products containing the 5′ end and the entire open reading frame were labelled with [^{32}P]dATP using the multi prime labelling kit (Amersham) and taken as hybridization probes. Hybridization was performed at 65 °C in 1% BSA, 7% SDS, 1 mM EDTA and 0.25 M Na_2PO_4 for 16 h. Membranes were washed twice in 1% BSA, 5% SDS, 1 mM EDTA and 25 mM Na_2PO_4, at 65 °C and twice again in 1%SDS, 25 mM Na_2PO_4 and 1 mM EDTA at 65 °C.

RESULTS

Purification of silk proteinase inhibitors

RP-HPLC analysis of the desalted urea extract of the silk proteins (see Experimental procedures) yielded a number of UV-absorbing fractions that were tested for the inhibition of trypsin activity. The results obtained with fractions containing most of the eluted material are shown in top of Fig. 1. No activity was found in the remaining fractions (data not shown). Fraction 5 was most active and reduced azocoll hydrolysis by trypsin to 10.7% of the control value. Further RP-HPLC separation of this fraction provided highly pure material (Fig. 1, bottom), which was designated silk proteinase inhibitor 1 (SPI 1). N-terminal amino-acid sequencing of SPI 1 disclosed a sequence (D)DDICSLPLKTG. It could not be determined if it began with two or three Asp residues.

The silk proteins extracted with 0.1% trifluoroacetic acid and fractionated on a LiChrospher column were also subjected to the azocoll assay. Most of the inhibitory activity was confined to fraction no. 4 and to a small extent to fraction no. 5 (Fig. 2). All fractions that are unmarked in Fig. 2 were inactive. Further HPLC purification of the active material on the Purospher column (Fig. 2, bottom) delivered a peptide of molecular mass 3993.2 ± 1.0 Da. A stretch of 20 amino-acid residues VCTTQXDPVCGKDGKTYYNL was identified with its N-terminal sequencing. The peptide was named silk proteinase inhibitor 2 (SPI 2).

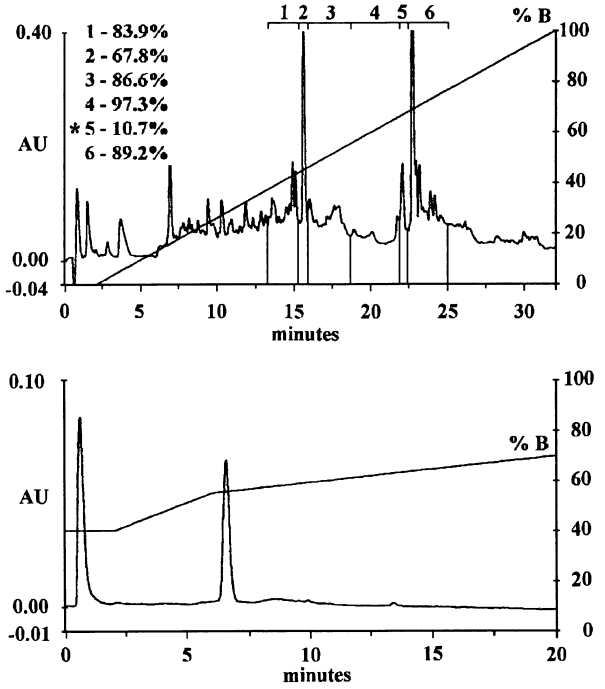

Fig. 1. Top panel: RP-HPLC elution profile (Purospher column) of the silk proteins extracted from the cocoons with 8 M urea, 2% SDS, and 5% 2-mercaptoethanol in 10 mM tris buffer, pH 7.0 (200 μL extract equivalent to 15 mg of raw silk) and trypsin inhibition with fractions 1–6. The active and further fraction is marked with *. Bottom panel: Purified fraction from 300 μL of the raw extract. For details see Experimental procedures.

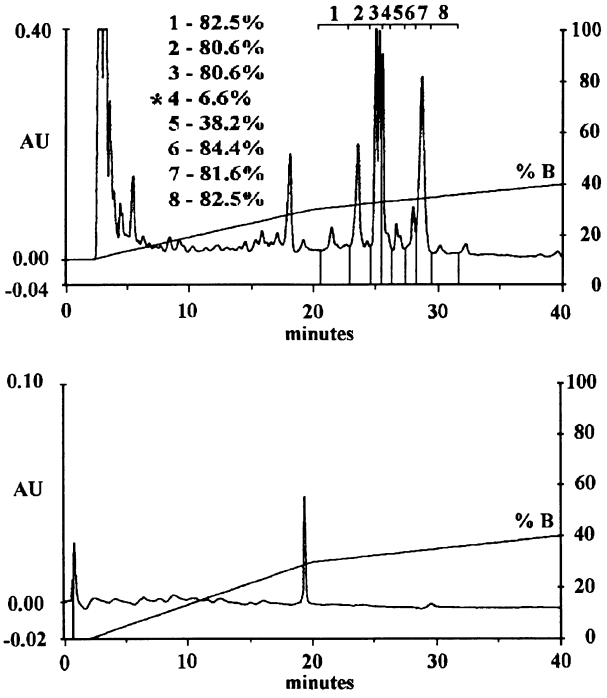

Fig. 2. Top panel: RP-HPLC elution profile of the 0.11% trifluoroacetic acid cocoon extract (200 μL extract, equivalent to 12.5 mg of raw silk, LiChrospher column) and trypsin inhibition with fractions 1–8. The active and further purified fraction is marked with *. Bottom panel: purification of the active fraction (50 μL raw extract equivalent, Purospher column). For details see Experimental procedures.

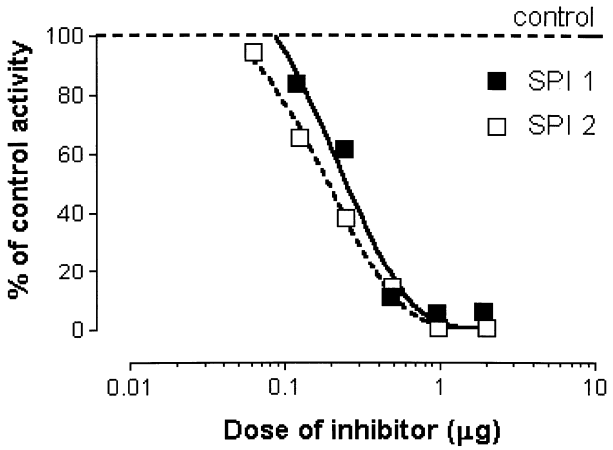

Fig. 3. Inhibitory effects of increasing doses of SPI 1 and SPI 2 on the trypsin activity. SPI 1 and SPI 2 were quantified with the bicinchoninic acid method and bovine serum albumin as standard, trypsin activity was measured with the azocoll assay.

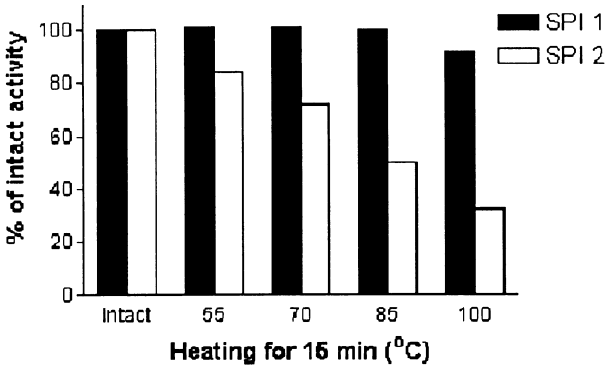

Fig. 4. Trypsin inhibition activities of SPI 1 (1.91 μg) and SPI 2 (1.99 μg) after a 15-min exposure to indicated temperatures. Measured with the azocoll assay and expressed in percentage of the control activity (= 100%).

Quantification and properties of SPI 1 and SPI 2

The amounts of SPI 1 and SPI 2 in the cocoons were quantified with the bicinchoninic acid method against the BSA standard. The results revealed that the doses used in the initial tests (see Figs 1 and 2) corresponded to 0.48 μg SPI 1 and 1.99 μg SPI 2. The recalculated contents per one cocoon (12.5 mg silk) were 7.63 μg SPI 1 and 7.96 μg SPI 2. Dose response curves were constructed using these assessments of the SPIs amounts in BSA equivalents (Fig. 3). Doses required to produce a median effect (D_m) were computed to be 0.251 μg for SPI 1 and 0.201 μg for SPI 2.

The established trypsin-inhibitory activities of SPI 1 and SPI 2 were about three to nine times lower that those of the standard, commercially available trypsin inhibitors (Table 1). The difference was approximate because the preparations of SPI 1 and SPI 2 were quantified only by comparison with BSA, and the assays were run under conditions optimized for the commercial inhibitors. We nevertheless suspected that SPIs may be tuned to other types of proteinases than trypsin. This was confirmed in azocoll assay performed with several proteinases from bacteria, fungi, and mammals (Table 1). Bacterial subtilisin and fungal proteinase K proved to be most sensitive, with SPI 2 being more active than SPI 1.

The cocoons stored at room temperature for 6 months did not show any appreciable loss of the proteinase inhibitory activity, and the activity of purified SPIs was not altered during two weeks of laboratory handling. SPI 1 and SPI 2, however, differed in thermostability (Fig. 4). SPI 1 fully resisted 15 min heating to 55 °C and its inhibitory potention was reduced only by 8.3% after heating to 100 °C. By contrast, the trypsin-inhibitory activity of SPI 2 declined already at 55 °C and a 15-min exposure to 100 °C reduced the activity by 67.9%.

Isolation of cDNAs encoding SPI 1 and SPI 2

Total RNA extracted from the silk glands was used as starting material to identify cDNAs of the proteinase inhibitors. First strand cDNA was used as template for PCR performed with combinations of the degenerate oligonucleotide primers and the adaptor primers. The degenerate primers were designed from the N-terminal amino-acids sequences of the silk proteinase inhibitors: Primer 5′-GAYGAYATHTGYTCNYTNCC-3′ was derived from the DDICSLP amino-acid sequence of the SPI 1 N-terminus, and primer 5′-TGYGGNAARGAYGGNAARAC-3′ from the CGKDGKT sequence of the SPI 2 N-terminus. Products of 350 and 150 nucleotides, which were amplified with the SPI 1 and SPI 2 degenerate primers, respectively, were cloned and sequenced (Fig. 5). Subsequently, the 5′ ends of the respective cDNA species were obtained from

Table 1. Activities of SPI 1 and SPI 2 in comparison with the standard trypsin inhibitors (first column) and the inhibition of various proteases with SPI 1 and SPI 2 (2nd and 3rd columns, respectively). D_m, dose causing a 50% inhibition (median effect), TI, trypsin inhibitor, NA, not active.

Trypsin inhibition		SPI 1 inhibition of proteases		SPI 2 inhibition of proteases	
Inhibitor	D_m	Protease	D_m	Protease	D_m
Aprotinin	0.050	Pronase	NA	Pronase	0.624
Hen egg white TI	0.071	Proteinase K	0.093	Proteinase K	0.039
Soybean TI	0.029	α-chymotrypsin	0.294	α-chymotrypsin	NA
SPI 1	0.251	Subtilisin	0.150	Subtilisin	0.014
SPI 2	0.201	Trypsin	0.251	Trypsin	0.201

SPI 1

```
GGACATTTCCAATCGACACCAATCGACTCCAATCGATTCGCTCCACCACACGTGTGTAAAAATGTATAAATTCATCGCTCTTATTGTTATTTTATTTATC 100
                                                          MetTyrLysPheIleAlaLeuIleValIleLeuPheIle  13
                        ---------------------->
ATGTACAATATTTGTAACGGAGACGACGACATTTGCTCCTTGCCTTTAAAGACAGGACCGTGTAGAGCGGCATTTCAAAGATACGGCTACGTCGAAGGCA 200
MetTyrAsnIleCysAsnGlyAspAspAspIleCysSerLeuProLeuLysThrGlyProCysArgAlaAlaPheGlnArgTyrGlyTyrValGluGlyL  47
                      *  *  *  *  *  *  *  *  *  *  *
                                                                                <-----------------
AAGGCTGCGTACTATTCACGTACGGCGGCTGTCAAGGGAACGCGAACAATTTCGAGACATTAGAAGCGTGCAAAAACGCGTGCGAGAACTAGTGTTACAT 300
ysGlyCysValLeuPheThrTyrGlyGlyCysGlnGlyAsnAlaAsnAsnPheGluThrLeuGluAlaCysLysAsnAlaCysGluAsn .           76

------
AAACCTGTCATTACTTTAAACGATTTTTATTTATTTATTAAAAAAAAATAAGTGTAAATAAATTATTTATATACAAAAATAAGATTGAAAATTGTTCAAA 400

TAGTTGTTACATTAATAAATGAAAAGTGTTGTTGAAAAAAAAA                                                          443
```

SPI 2

```
AGTGATATCAAGCTGTAATCACCATGAAGTACTACTTATATGCAATTATATTTCTGATGGCGATATTTATATGGAATGAAGCGACCGCAGTTTGCACCAC 100
                       MetLysTyrTyrLeuTyrAlaIleIlePheLeuMetAlaIlePheIleTrpAsnGluAlaThrAlaValCysThrTh  26
                                                                                      *  *  *  *
                 ---------------------->                                  <--------------------
CGAGTGGGACCCTGTTTGTGGGAAAGATGGCAAGACGTACAGCAATCTTTGCTGGTTAAACGAGGCTGGCGTCGGGTTGGATCATGAAGGTGAATGCTGA 200
rGluTrpAspProValCysGlyLysAspGlyLysThrTyrSerAsnLeuCysTrpLeuAsnGluAlaGlyValGlyLeuAspHisGluGlyGluCys .   58
  *  X  *  *  *  *  *  *  *  *  *  *  *  *  Y  *  *  *

TTTACATCAAAATATGTTTGAGTACACAAATAAACTGAAATGATTTTAAAGTAAAAAAAA                                         260
```

Fig. 5. Complete cDNAs and the deduced amino-acid sequences of pre-SPI 1 and pre-SPI 2. Arrows indicate regions of the degenerate forward and the specific reverse primers used for cDNA identifications by PCR. The initiation and termination codons are printed in bold and the putative polyadenylation signals are underlined. Amino-acid residues identified with N-terminal sequencing of the SPI 1 and SPI 2 peptides, respectively, are marked with asterisks (X denotes that the respective residue could not be identified with certainty, and Y signifies that amino-acid sequencing indicated Tyr where cDNA analysis disclosed Ser). Nucleotide sequences of SPI 1 and SPI 2 were registered by the GenBank under the accession numbers AF292098and AF292099, respectively.

the enriched ds cDNA by PCR with specific reverse primers (Fig. 5), which were combined with the 5′ SMART III primer. Sequences of the products of 300 nucleotides (SPI 1) and 200 nucleotides (SPI 2) showed expected overlaps with the previously obtained 3′ terminal sequences of SPI 1 and SPI 2, respectively. The cloning of both SPI 1 and SPI 2 cDNAs was repeated to verify the nucleotide sequences presented in Fig. 5.

Molecular characterization of SPI 1

The complete cDNA sequence of *pre-SPI 1* was 443 nucleotides long, including an ORF encompassing 231 nucleotides (Fig. 5). The 5′ end contained a 61 nucleotides long UTR; the 3′ UTR was 151 nucleotides long, included a typical polyadenylation signal [26] and was terminated with a poly A tail. The open reading frame of *pre-SPI 1*, initiated by the first ATG codon, encoded a preprotein of 76 amino-acid residues. Signal peptide cleavage, predicted to occur between Gly20 and Asp21 within a typical motif for the recognition of signal peptidases [27], yielded a signal peptide of 20 amino acids. Putative mature peptide SPI 1 began with three Asp residues, confirming the preferred interpretation of N-terminal amino-acid sequencing of the purified SPI 1. Amino-acid residues in the following nine positions were exactly as established with the amino-acid sequencing.

The deduced mature SPI 1 contained 56 amino acids, of which six were Cys residues in positions characteristic for the Kunitz-type proteinase inhibitors. An alignment of SPI 1 with other Kunitz inhibitors (Fig. 6) disclosed relatively high homology with the *Manduca sexta* B chymotrypsin

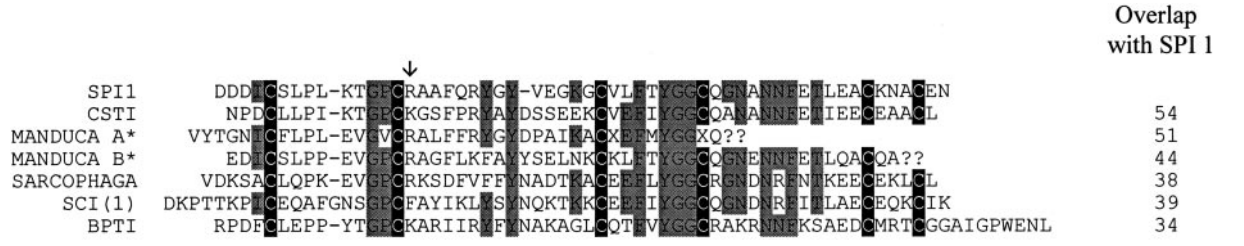

Fig. 6. Amino-acid sequence alignment of SPI 1 and percentage overlap with the shown domains of selected Kunitz-type proteinase inhibitors: CSTI, *Bombyx mori* cocoon specific trypsin inhibitor [14]; Manduca A, Manduca B, and Sarcophaga, inhibitors from the hemolymph of *Manduca sexta* and *Sarcophaga bullata*, respectively [28,29]; SCI (1), first domain of the chymotrypsin inhibitor from *Bombyx mori* hemolymph [30]; BPTI, bovine pancreatic trypsin inhibitor [8]. Conserved Cys residues are shaded black and other conserved amino-acid residues are shaded gray; putative P1 residue is marked with ↓; ?? indicates that the C-terminal sequence of the respective domain is not known (in this case, homology was established between the known sequence and the corresponding region of SPI 1).

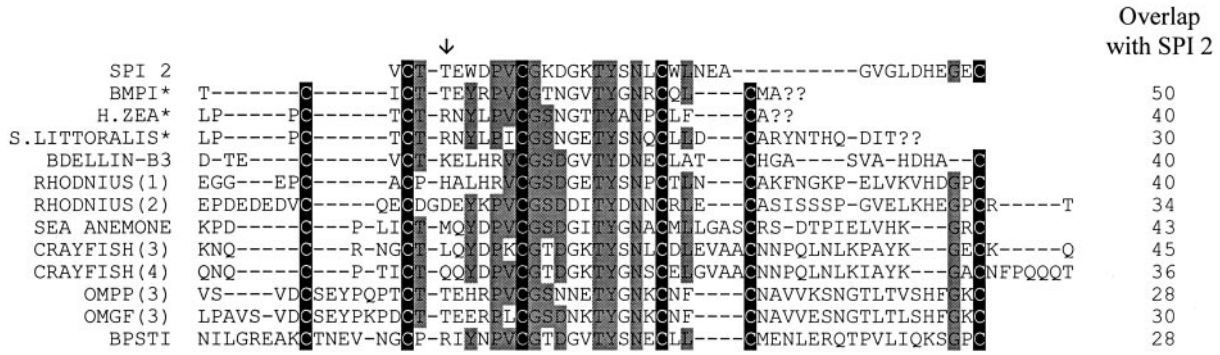

Fig. 7. Amino-acid sequence alignment of SPI 2 and percentage overlap with selected domains (in parentheses; not shown for the single-domain inhibitors) of the Kazal-type inhibitors from the insects *Bombyx mori* (BMPI) [32], *Heliothis zea* [15], *Spodoptera littoralis* [16], and *Rhodnius prolixus* [33], the leech *Hirudo medicinalis* (bdellin B3) [34], the sea anemone *Anemonia sulcata* [35], the crayfish *Pacifastacus leniusculus* [36], the bovine pancreas (BPSTI) [37], and the egg white of gray francolin and peacock phesant (OMGF, OMPP) [9]. Conserved Cys residues are shaded black and other conserved amino-acid residues are shaded gray; putative P1 residue is marked with ↓; ?? indicates that the C-terminal sequence of the respective domain is not known (in this case, homology was established between the known sequence and the corresponding region of SPI 2). See text for details.

inhibitor, which was isolated from the hemolymph [28], and with the cocoon-associated trypsin inhibitor of *Bombyx mori* [14]. The alignment specified the position of putative P1 residue in the reactive site of SPI 1 and identified it as Arg. Calculated mass of SPI 1 was 6056.7 Da.

Molecular characterization of SPI 2

The complete cDNA of *pre-SPI 2* (Fig. 5) was 260 nucleotides long and included 23 nucleotides of 5′ UTR, an open reading frame of 177 nucleotides, 60 nucleotides of a 3′ UTR with a polyadenylation signal and a poly A tail. The deduced translation product, starting from the first ATG codon, comprised 58 amino acids (Fig. 5). As the N-terminus of the sequenced SPI 2 peptide began with Val23, and the amino-acid residues around positions 22–23 of the deduced translation product complied with the rules for signal peptide cleavage, it was evident that the preprotein contained a 22 amino-acid long signal sequence followed by secreted peptide of 36 amino acids. Eighteen out of the first 20 amino acids of the deduced mature peptide were consistent with the result of the N-terminal sequencing of the cocoon-purified SPI 2. The residue in position 6, which could not be determined with the amino-acid sequencing, was deduced to be Trp, and the double Tyr in positions 17 and 18 proved to be Tyr followed by Ser (a strong Tyr signal in amino-acid sequencing apparently masked the adjacent Ser). The molecular mass of the deduced secretory protein

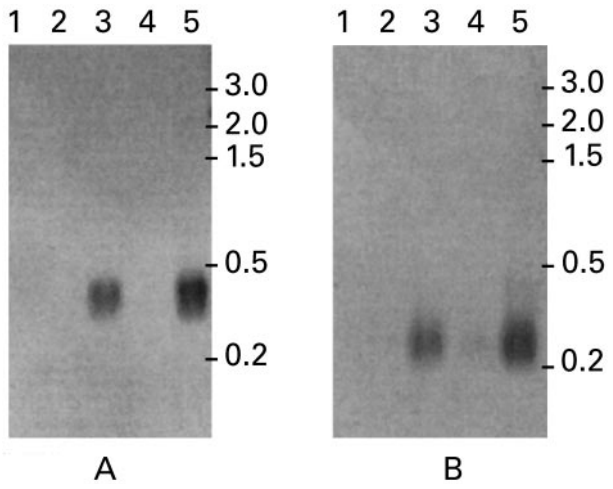

Fig. 8. Tissue specific expression of SPI 1 and SPI 2. Northern blots of 5 μg total RNA extracted from the integument (lane 1), fat body (lane 2), middle silk gland (lane 3) and posterior silk gland (lane 4) or 10 μg of total RNA from the middle silk gland (lane 5), were hybridized either to the 306 nucleotides fragment of SPI 1 (panel a) or to the 200 nucleotides fragment of SPI 2 (panel b). Positions of an RNA ladder are given on the right.

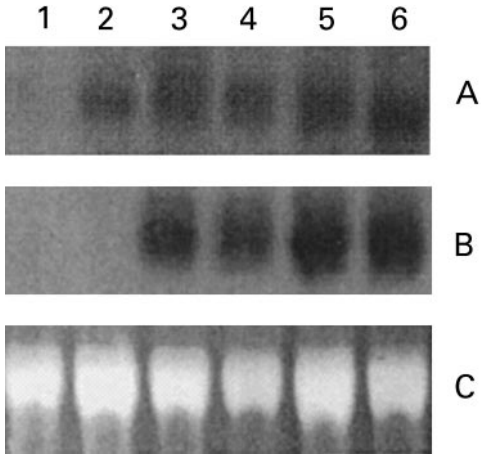

Fig. 9. Developmental profile of SPI 1 and SPI 2 mRNA during the penultimate and last larval instars. Total RNA (5 μg), which was extracted from the middle silk gland of the penultimate instar larvae (1), and from larvae at the start (2), and end of feeding (3), wandering (4), spinning (5), and postspinning (6) periods of the last instar, was blotted and probed either with the 306 nucleotides fragment of SPI 1 (panel a) or the 200 nucleotides fragment of SPI 2 (panel b). The lower panel shows rRNA stained in the nylon membrane with methylene blue.

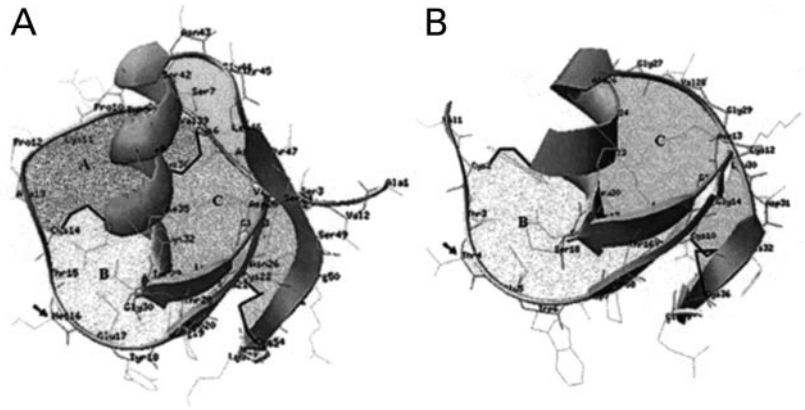

Fig. 10. Molecular conformations of (a) the ovomucoid third domain of the mikado pheasant; representing a classical member of the Kazal family [42], and (b) hypothetical conformation of SPI 2. Disulfide bridges are indicated by black lines, the P1 residue by an arrow, and the A, B, and C rings are stippled.

was calculated to be 3990 Da, which is in good agreement with the 3993 Da mass measured with MALDI-TOF.

An alignment of the deduced amino-acid sequence with the proteinase inhibitors revealed that SPI 2 corresponded to a single domain of the Kazal-type inhibitors. However, a typical domain contains 6 Cys residues, while the SPI 2 was truncated and lacked a Cys at the N-terminus and another Cys in the C-terminal region (Fig. 7). The existing four Cys residues in SPI 2 aligned with the 2nd, 3rd, 4th, and 6th cysteines in typical Kazal-type inhibitors. The overall homology of SPI 2 with the domains of known Kazal inhibitors did not exceed 45%. The Thr in position 4 of the mature peptide was identified as putative P1 residue.

Spatial and developmental expression of SPI 1 and SPI 2

The 5′ end with the entire open reading frame of *pre-SPI 1* (306 nucleotides) and *pre-SPI 2* (200 nucleotides) cDNAs were used as probes to detect the respective mRNAs in the Northern blots. Total RNA from the fat body, integument, and the middle and posterior sections of the silk gland were examined. We found that both inhibitors were expressed exclusively in the middle silk gland section, where a single transcript of approximately 450 nucleotides was detected with the *pre-SPI 1* probe, and another single transcript of cca 300 nucleotides with the *pre-SPI 2* probe (Fig. 8). To study the developmental expression profiles, silk glands were dissected on day 2 of the penultimate instar (total instar length 3.5 days) and on days 1 (early feeding), 4, 5 (end of feeding), 6 (wandering period) and 7 (termination of cocoon spinning) of the last larval instar. Figure 9 shows that the mRNAs of both proteinase inhibitors were absent in the penultimate instar larvae. Their expression began from early last larval instar and continued to the postspinning stage when most of the silk had been spun out as cocoon.

DISCUSSION

The thermostable Kunitz proteinase inhibitor

SPI 1 is a typical, thermostable proteinase inhibitor of the Kunitz-type, with Arg in the P1 position (Fig. 6). Proteinase inhibitors with P1Arg were reported to be specific towards trypsin and trypsin-like enzymes [10], but our results show that SPI 1 is more effective on the bacterial and fungal proteinases. The other insect inhibitors with P1 Arg, which were isolated from the hemolymph of the hornworm *Manduca sexta* [28] and the fleshfly *Sarcophaga bullata* [29], were found to inhibit trypsin, chymotrypsin and plasmin, but their activities on microbial proteinases have not been tested. This is true also for the cocoon specific trypsin inhibitor (CSTI) of *Bombyx mori* that is highly homologous to SPI 1 except for Lys in the P1 position [14]. There are obviously not enough data to appreciate the significance of P1 residue for the function of insect proteinase inhibitors. Information available on Lepidoptera shows that various amino acids can occur in this position. For example, a chymotrypsin inhibitor from the hemolymph of *B. mori* contains P1 Phe in all of its three domains [30,31], whereas the extracellular matrix protein lacunin of *Manduca sexta*, which includes a stretch of 11 Kunitz domains, contains in P1 positions Arg, Leu, Ser, Ala, Thr, Gln, and Gly, respectively [38]. Biological activities of the lacunin domains are not known but the authors assumed that their diversity confers on lacunin the ability to inhibit a multitude of proteinases.

Unique structure of the Kazal-type proteinase inhibitor

SPI 2 was identified as the shortest Kazal-type proteinase inhibitor in animals. Unlike most Kazal inhibitors, where each functional domain consists of 50–60 amino-acid residues, SPI 2 is a single domain inhibitor of only 36 residues (Fig. 5). Bdellin B-3 from the leech *Hirudo medicinalis* [34], which is apparently the second shortest Kazal domain, contains 40 residues (Fig. 7). SPI 2 is comparable in size only to structurally unrelated proteinase inhibitors LMCI I and LMCI II (36 and 35 residues, respectively) from the migratory locust, *Locusta migratoria* [39,40] and SGPIs 1–5 (35–40 residues) from the desert locust, *Schistocerca gregaria* [41].

Each domain of typical Kazal-type inhibitors includes 6 cysteines (Fig. 7) that form disulfide bridges dictating the molecular conformation of the domain [8]. The conformation includes 3 loops A, B and C that are stabilized and closed into rings by disulfide bridges between CysI–CysV, CysII–CysIV, and CysIII–CysVI (Fig. 10). The CysI–CysV bridge closing the A-ring, and the CysII–CysIV bridge separating the A- and B-rings, link the N-terminus and the reactive site region, respectively, to the central α helix, while the CysIII–CysVI bridge closes the C-ring by linking C-terminus to the central β sheet (Fig. 10a). The size of

A-loop is variable, as the number of amino-acid residues between CysI and CysII ranges from 18 in avian ovomucoids to just one in the insect inhibitors (Fig. 7). The length of the C-loop varies only between 31 and 34 residues, despite that CysV and thereby the position of the disulfide anchoring point of the C-loop is shifted by one turn of the α helix towards the C-terminus in the sea anemone [35] and apparently also in the crayfish inhibitors (cf. Figure 7). Only bdellin-B3 from the leech [34] contains a C-loop of just 27 residues and the loop is properly anchored by the CysIII–CysVI bridge.

Hypothetical conformation of SPI 2 was generated with the aid of the Swiss Pdb Viewer [25]. As shown in Fig. 10b, the absence of CysI and CysV in SPI 2 precludes the closure of ring A and provides increased flexibility to the C-ring (Fig. 10). SPI 2 is also exceptional by having only a single amino-acid residue preceding CysII at the N-terminus. These unusual features of SPI 2 demonstrate that the presence of A-loop is not essential for the activity. The absence of CysV and adjacent residues not only shortens the C-ring to 27 residues as in bdellin-B3 (Fig. 7), but it also alters conformation of the α-helical segment of the molecule (Fig. 10). By contrast, the length and conformation of the B-ring, including exposure of the reactive site with P1 residue, and the position of the C-terminus relative to the B-loop, are conserved in SPI 2 (cf. structures in Fig. 10). These features are probably crucial for the inhibitory activity. A different length of the B-loop was found only in the second domain of the thrombin inhibitor from *Rhodnius prolixus* that includes additional Gly between CysII and the P1 residue, but it is not known if this domain is functional [33].

The specificity of each proteinase inhibitor depends chiefly on the P1 residue at the reactive site [11]. It has been proposed that inhibitors with P1 Lys and Arg tend to inhibit trypsin and trypsin-like enzymes, those with Pro, Tyr, Phe, Leu and Met inhibit chymotrypsin and chymotrypsin-like enzymes, and those with Ala and Ser inhibit elastase-like enzymes [8]. The role of Thr in P1 position, as it occurs in SPI 2, has not been specified. Considering the high inhibitory activity of SPI 2 against proteinase K and subtilisin, it is possible that P1 Thr confers specificity against microbial proteinases. This is consistent with P1 Thr in the new family of fungal proteinase inhibitors (F-FPI) isolated from the integument and hemolymph of *Bombyx mori* [43]. Thr in P1 also occurs in some Kazal-type inhibitors, such as the third ovomucoid domain of some birds [9], and in the partially sequenced inhibitor isolated from the cocoons of *Bombyx mori* [32]; however, target proteinases of these inhibitors have not been identified. It seems that Thr in P1 position is common in the proteinase inhibitors present in secretory products.

Significance of proteinase inhibitors in the silk

The silk spun by caterpillars is the secretion from the pair of silk glands [1]. Fibroins, which make up the water insoluble core of the silk fibre, are produced in the posterior gland section, while sericins, which provide sticky coating of the fibre, are derived from the middle gland section. In addition to these large structural proteins, both the middle and the posterior gland sections secrete seroin, a protein proposed to have antibacterial properties [7]. The resistance of silk to the destruction by microbes, fungi, mites, and insects is well known [44]. It seems likely that the proteinase inhibitors SPI 1 and SPI 2 are involved in the antimicrobial defense, as indicated by their inhibition of bacterial subtilisin and fungal proteinase K (Table 1). Simultaneous occurrence of two or more proteinase inhibitors with similar activities is common in various tissues and secretions and has been interpretted as functional backup [45]. However, differences in the activities of SPI 1 and SPI 2, notably the inhibition of the α-chymotrypsin only by SPI 1 and of the pronase just by SPI 2, imply that these two inhibitors are functionally complementary.

The caterpillars of *Galleria mellonella*, which live in the bee nests, spin considerable amounts of silk through most of their development to produce tubes protecting them against the attack by bees [46]. No expression of SPIs, however, was detected prior to the last larval instar. This suggests that SPIs lend special protection to the cocoon, which is spun at the end of the last larval instar. In contrast to the larval silk tubes, which are renewed continuously, the cocoon must endure and provide shelter to prepupa and pupa for a period ranging from one week to many months (depending on the season). High stability of SPIs secreted by the silk glands apparently evolved to match the need for long protection. We assume that homologs of SPIs are present in the cocoons of all insects but data comparable to our investigations are available only for the silkworm, *Bombyx mori*. The cocoons of this insect contain a typical Kunitz-type trypsin inhibitor [14], exhibiting 54% homology to SPI 1, and a Kazal-type inhibitor that might be related to SPI 2 but whose sequence is incompletely known [32]. Activities of these inhibitors on the bacterial and fungal proteinases remain to be established.

ACKNOWLEDGEMENTS

We thank Dr Petr Kopáček of the Institute of Parasitology, Czech Academy of Sciences, for advice and guidance in course of this work, and Dr Vladimír Havlíček of the Institute of Microbiology, for MALDI-TOF analysis of SPI 2. The research was supported by grant ME 204 of program Kontakt and grant 204/00/0019 provided by the Grant Agency of the Czech Republic.

REFERENCES

1. 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 Morph. Embryol.* **19**, 79–132.
2. Tanaka, K., Mori, K. & Mizuno, S. (1993) Immunological identification of the major disulfide-linked light component of silk fibroin. *J. Biochem.* **114**, 1–4.
3. Michaille, J.J., Couble, P., Prudhomme, J.-C. & Garel, A. (1986) A single gene produces multiple sericin messenger RNAs in the silkgland of *Bombyx mori*. *Biochimie* **68**, 1165–1173.
4. Couble, P., Michaille, J.J., Couble, M.L. & Proudhomme, J.-C. (1987) Developmental switches of sericin mRNA splicing in individual cells of *Bombyx mori* silkgland. *Dev. Biol.* **124**, 431–440.
5. Sasaki, T. & Noda, H. (1973) Studies on silk fibroin of *Bombyx mori* directly extracted from the silk gland. I. Molecular weight determination in guanidine hydrochloride or urea solutions. *Biochim. Biophys. Acta* **310**, 91–103.

6. Grzelak, K., Couble, P., Garel, A., Kludkiewicz, B. & Alrouz, H. (1988) Low molecular weight silk proteins in *Galleria mellonella*. *Insect Biochem.* **18**, 223–228.
7. Žurovec, M., Yang, C., Kodrík, D. & Sehnal, F. (1998) Identification of a novel type of silk protein and regulation of its expression. *J. Biol. Chem.* **273**, 15423–15428.
8. Laskowski, M. & Kato, I. (1980) Protein inhibitors of proteinases. *Ann. Rev. Biochem.* **49**, 593–626.
9. Laskowski, M., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M., Kohr, W., Park, S., Parks, K., Schatzley, B., Schoenberger, O., Tashiro, M., Vichot, G., Whatley, H., Wieczorek, A. & Wieczorek, M. (1987) Ovomucoid third domains from 100 avian species: isolation, sequences, and hypervariability of enzyme-inhibitor contact residues. *Biochemistry.* **26**, 202–221.
10. Bode, W. & Huber, R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204**, 433–451.
11. Kellenberger, C., Boudier, C., Bermudez, I., Bieth, J.G., Luu, B. & Hietter, H. (1995) Serine protease inhibition by insect peptides containing a cysteine knot and a triple-stranded β-sheet. *J. Biol. Chem.* **270**, 25514–25519.
12. Saul, S.J. & Sugumaran, M. (1986) Protease inhibitor controls prophenoloxidase activation in *Manduca sexta*. *FEBS Lett.* **208**, 113–116.
13. Aso, Y., Yamashita, T., Meno, K. & Murakami, M. (1994) Inhibition of prophenoloxidase-activating enzyme from *Bombyx mori* by endogenous chymotrypsin inhibitors. *Biochem. Mol. Biol. International* **33**, 751–758.
14. Kurioka, A., Yamazaki, M. & Hirano, H. (1999) Primary structure and possible functions of a trypsin inhibitor of *Bombyx mori*. *Eur. J. Biochem.* **259**, 120–126.
15. Polanowski, A., Wilusz, T., Blum, M.S. & Travis, J. (1994) Two serine proteinase inhibitors from the larval hemolymph of *Heliothis zea*. *Acta. Biochim. Polon.* **41**, 178–180.
16. Boigegrain, R., Pugniere, M., Paroutaud, P., Castro, B. & Brehelin, M. (2000) Low molecular weight serine protease inhibitors from insects are proteins with highly conserved sequences. *Insect Biochem. Mol. Biol.* **30**, 145–152.
17. Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H.G., Hoffken, H.W., Reuschenbach, P., Otte, M. & Dodt, J. (1993) A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. *J. Biol. Chem.* **268**, 16216–16222.
18. Sehnal, F. (1966) Kritisches Studium der Bionomie und Biometrik der in verschiedenen Lebensbedingungen gezüchteten Wachsmotte, *Galleria mellonella*. *Z. Wiss. Zool.* **174**, 53–82.
19. Chavira, R.J., Burnett, T.J. & Hagemann, J.H. (1984) Assaying proteinases with azocoll. *Anal. Biochem.* **136**, 446–450.
20. Wedde, M., Weise, C.H., Kopáèek, P., Franke, P. & Vilcinskas, A. (1998) Purification and characterization of an inducible metalloprotease inhibitor from the hemolymph of greater wax moth larvae, *Galleria mellonella*. *Eur. J. Biochem.* **255**, 535–543.
21. Chou, J. & Chou, T.-Ch (1987) Dose-effect analysis with microcomputers. Biosoft, Cambridge, UK.
22. Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J. & Klenk, D.C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76–85.
23. Yang, C., Sehnal, F. & Scheller, K. (1996) Juvenile hormone restores larval pattern of sericin gene transcripts. *Arch. Insect Biochem. Physiol.* **33**, 353–362.
24. Worley, K.C., Culpepper, P., Wiese, B.A. & Smith, R.F. (1998) BEAUTY-X: enhanced BLAST searches for DNA queries. *Bioinformatics* **14**, 890–891.
25. Guex, N. & Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.
26. Proudfoot, N. (1991) Poly (A) signals. *Cell* **64**, 671–674.
27. Von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**, 17–21.
28. Ramesh, N., Sugumaran, M. & Mole, J.E. (1988) Purification and characterization of two trypsin inhibitors from the hemolymph of *Manduca sexta* larvae. *J. Biol. Chem.* **263**, 11523–11527.
29. Sugumaran, M., Saul, S.J. & Ramesh, N. (1985) Endogenous protease inhibitors prevent undesired activation of prophenolase in insect hemolymph. *Biochem. Biophys. Res. Comm.* **132**, 1124–1129.
30. Sasaki, T. (1984) Amino acid sequence of a novel Kunitz-type chymotrypsin inhibitor from hemolymph of silkworm larvae, *Bombyx mori*. *FEBS Lett.* **168**, 227–230.
31. Sasaki, T. (1988) Amino acid sequences of two basic chymotrypsin inhibitors from silkworm larval hemolymph. *Biol. Chem. Hoppe Seyler* **369**, 1235–1241.
32. Kurioka, A. & Hirano, H. (1995) Partial amino acid sequence of a 11kDa cocoon shell Kazal-type protein (in Japanese). *J. Seric. Sci. Jpn.* **64**, 392–394.
33. Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kroger, B., Hoffken, W. & Bode, W. (1995) Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. *EMBO J.* **14**, 5149–5157.
34. Fritz, H. & Krejci, K. (1976) Trypsin-plasmin inhibitors (bdellins) from leeches. In *Methods in Enzymology* (L. Lorand, ed.), 45, pp. 797–806. Academic Press, New York, USA.
35. Tschesche, H., Kolkenbrock, H. & Bode, W. (1987) The covalent structure of the elastase inhibitor from *Anemonia sulcata* – a 'non-classical' Kazal-type protein. *Biol. Chem. Hoppe Seyler* **368**, 1297–1304.
36. Johansson, W., Keyser, P. & Soderhall, K. (1994) Purification and cDNA cloning of a four-domain Kazal proteinase inhibitor from crayfish blood cells. *Eur. J. Biochem.* **223**, 389–394.
37. Greene, L.J. & Bartelt, D.C. (1969) The structure of the bovine pancreatic secretory trypsin inhibitor – Kazal's inhibitor. II. The order of the tryptic peptides. *J. Biol. Chem.* **244**, 2646–2657.
38. Nardi, J.B., Martos, R., Walden, K.K.O., Lampe, D.J. & Robertson, H.M. (1999) Expression of lacunin, a large multi-domain extracellular matrix protein, accompanies morphogenesis of epithelial monolayers in *Manduca sexta*. *Insect Biochem. Mol. Biol.* **29**, 883–897.
39. Brehelin, M., Boigegrain, R., Drif, L. & Coletti-Previero, M. (1991) Purification of a protease inhibitor which controls prophenoloxidase activation in hemolymph of *Locusta migratoria* (Insecta). *Biochem. Biophys. Res. Commun.* **179**, 841–846.
40. Boigegrain, R., Mattras, H., Brehelin, M., Paroutaud, P. & Coletti-Previero, M. (1992) Insect immunity: two proteinase inhibitors from hemolymph of. *Locusta Migratoria. Biochem. Biophys. Res. Commun.* **189**, 790–793.
41. Hamadaoui, A., Wataleb, S., Deveresse, B., Chiou, S.-J., Vanden Broeck, J., Van Beeumen, J., Loof, A. & Schoofs, L. (1998) Purification and characterization of a group of five novel peptide serine protease inhibitors from ovaries of the desert locust, *Schistocerca gregaria*. *FEBS Lett.* **422**, 74–78.
42. Laskowski, M., Apostol, I., Ardelt, W., Cook, J., Giletto, A., Kelly, C.A., Lu, W., Park, S.J., Quasim, M.A., Whatley, H.E., Wieczorek, A. & Wynn, R. (1990) Amino acid sequences of ovomucoid third domain from 25 additional species of birds. *J. Protein Chem.* **9**, 715–725.
43. Pham, T., Hayashi, K., Takano, R., Itoh, M., Eguchi, M., Shibata, H., Tanaka, T. & Hara, S. (1996) A new family of serine protease inhibitors (Bombyx family) as established from the unique

topological relation between the positions of disulphide bridges and reactive site. *J. Biochem.* **119**, 428–434.
44. Akai, H. (1997) Anti-bacteria function of natural silk materials. *Int. J. Wild Silkmoth Silk* **3**, 79–81.
45. Kanost, M. & Jiang, H. (1996) Proteinase inhibitors in invertebrate immunity. In *New Directions in Invertebrate Immunology* (Soderhall, K., Iwanaga, S. & Vanta, G., eds), pp. 155–173. SOS Publications, Fairhaven, NY, USA.
46. Jindra, M. & Sehnal, F. (1989) Larval growth, food consumption, and utilization of dietary protein and energy in *Galleria mellonella*. *J. Insect Physiol.* **35**, 719–724.