Genes encoding two cystatins in the flesh fly *Sarcophaga crassipalpis* and their distinct expression patterns in relation to pupal diapause

Shin G. Goto and David L. Denlinger

Citation	Gene. 292(1-2); 121–127		
Issue Date	2002-06-12		
Туре	Journal Article		
Textversion	n author		
	\bigcirc 2002. All rights reserved. This manuscript version is made available under the		
	CC-BY-NC-ND 4.0 License. <u>https://creativecommons.org/licenses/by-nc-nd/4.0/</u> .		
	This is the accepted manuscript version. The article has been published in final form		
Rights	at <u>https://doi.org/10.1016/S0378-1119(02)00652-2</u> .		
	Please cite only the published version.		
	引用の際には出版社版をご確認ご利用ください。		
DOI	10.1016/S0378-1119(02)00652-2		

Self-Archiving by Author(s) Placed on: Osaka City University

~ 1 ~

Genes encoding two cystatins in the flesh fly *Sarcophaga crassipalpis* and their distinct expression patterns in relation to pupal diapause

Shin G. Goto, David L. Denlinger

Department of Entomology, Ohio State University, 1735 Neil Avenue, Columbus, OH 43210, USA

Abstract

Two genes encoding cystatins, cysteine proteinase inhibitors, were cloned from the flesh fly *Sarcophaga crassipalpis*. One, *sarcocystatin A* (abbreviated *Scys-A* and SCYS-A for the gene and protein, respectively) was previously cloned from *Sarcophaga peregrina*, but the other is novel. Here the novel gene is named *sarcocystatin B* (abbreviated *Scys-B* and SCYS-B for the gene and protein, respectively). Tissue distribution of the two sarcocystatins differs: the fat body is the major site of *Scys-A* expression, while *Scys-B* is present in both the fat body and gut. Both *Scys-A* and *-B* are developmentally regulated, but their expression patterns also differ. *Scys-A* transcripts are present in both diapause- and non-diapause-destined third instar wandering larvae, and are then downregulated throughout pupal diapause. By contrast, *Scys-B* in early diapause suggests a possible role for this proteinase inhibitor in halting development.

Keywords: Cysteine proteinase inhibitor (cystatin); Diapause; Early diapause gene; Gene expression; Tissue specificity; Lipopolysaccharide injection

1. Introduction

The cystatins are low molecular mass cysteine proteinase inhibitors known from a range of arthropods and vertebrates (Rawlings and Barrett, 1990; Turk and Bode, 1991; Brown and Dziegielewska, 1997). In the flesh fly *Sarcophaga peregrina*, a cystatin purified from the hemolymph is referred to as sarcocystatin A (Suzuki and Natori, 1985). Sarcocystatin A inhibits activity of cysteine proteases (papain and ficin) and serine protease (cathepsin B) (Suzuki and Natori, 1985; Kurata et al., 1990). In *S. peregrina*, one of the cysteine proteinases, cathepsin L, acts as a regulator of metamorphosis, and blocking the action of cathepsin L can inhibit differentiation of imaginal discs (Homma and Natori, 1994), and one of the serine proteases, cathepsin B, participates in dissociation of the fat body during metamorphosis (Kurata et al., 1990). The degradation of larval tissue proceeds simultaneously with the development of new adult tissue, thus precise regulation of these proteinases is likely to be needed during metamorphosis to protect developing adult tissues while scavenging larval tissues. Natori and coworkers suggest that sarcocystatin A is an important regulator of metamorphosis (Suzuki and Natori, 1985, 1986; Kurata et al., 1990); Saito et al., 1989).

Our interest in the cystatins was prompted by the discovery that one of the cystatin genes found in the flesh fly, *Sarcophaga crassipalpis*, is highly expressed early in pupal diapause. Diapause is a form of developmental arrest analogous to hibernation in mammals (Denlinger, 2000). At the molecular level, diapause is characterized by both a shutdown in the expression of many genes as well as the upregulation of others (Denlinger, 2002). Different categories of upregulated genes have been noted: expression of some upregulated genes persists throughout diapause, others are upregulated only in an early or late stage of diapause. One of the cystatin genes we report on here (*sarcocystatin B*) is the first identified, diapause-upregulated gene with an expression pattern restricted to the very early phase of diapause. High expression at this time is of special interest because such genes have the potential to play a critical role in initiating the arrest in development. The other, a homolog of the sarcocystatin A (Saito et al., 1989), is highly expressed at metamorphosis but is expressed at low levels during pupal diapause. The two cystatin genes we report in this study exhibit distinct expression patterns in association with development and have different patterns of tissue distribution.

2. Materials and methods

2.1. Flies

The colony of *S. crassipalpis* Macquart was maintained as described (Denlinger, 1972). Flies were reared under diapauseaverting (15 h light-9 h dark, 25°C) or diapause-inducing (12 h light-12 h dark, 25°C for adults and 20°C for larvae and pupae) conditions. Pupal diapause was terminated by directly applying 5 µl of hexane to the heads of pupae that had been in diapause for 25 days (Denlinger et al., 1980).

2.2. RNA isolation

RNA was isolated by TRIzol (Gibco BRL) according to supplier's instructions.

2.3. Cloning

The fragment of the *sarcocystatin B* gene (*Scys-B*) from *S. crassipalpis* was co-amplified in a rapid amplification of DNA complementary to RNA (cDNA) ends (RACE) reaction for the *period* gene, a reaction that will be published elsewhere.

The fragment of the *sarcocystatin A* gene (*Scys-A*) was amplified by reverse transcription polymerase chain reaction (RT-PCR). cDNA was synthesized with RNA derived from pupae using M-MLV reverse transcriptase and oligodeoxyribonucleotide(dT) (Gibco BRL) according to the supplier's instructions. The PCR reaction used 1 μ l of cDNA sample and 1.25 U of Platinum Taq DNA polymerase (Gibco BRL) and a final concentration of 1x PCR buffer as formulated by Gibco BRL, 0.2 μ M of SCA-F1 and -R1 primers (Table 1), 0.1 mM of deoxyribonucleotide triphosphate and 1.5 mM of MgCl₂ in a total volume of 50 μ l.Amplification was achieved with a preheat for 2 min at 94°C and 35 cycles of 30 s at 94°C, 30 s at 56°C and 1 min at 72°C.

The fragment of *ribosomal protein L32* gene (RpL32) in *S. crassipalpis* was also amplified by RT-PCR as described for *sarcocystatin A*, except that the primers for PCR were RP49L-F1 and -R1 primers (Table 1) (Warman et al., 2000) and the annealing temperature for PCR was 58°C.

The fragments were subcloned into pCR2.1 vectors that were used to transform competent cells (TA cloning kit, Invitrogen) according to supplier's instructions. The clones were sequenced on an ABI3700 DNA analyzer with BigDye terminator cycle sequencing chemistry.

2.4. Northern hybridization

RNAs were electrophoresed on denaturing gels and transferred to nylon membrane according to Sambrook et al. (1989).

Sarcocystatin A, B and RpL32 fragments were amplified by PCR from the plasmid. The reaction was the same as for the cloning of sarcocystatin A, except that the template was the plasmid containing the sarcocystatin A, B or RpL32 fragment as an insert, and the primers were M13-F and -R (Table 1). Amplification was achieved with a preheat at 94°C for 2 min and 25 cycles of 15 s at 94°C, 10 s at 55°C and 30 s at 72°C.

PCR products were applied to NucleoSpin Extraction kit (Clontech) and used as templates for probes. The probe labeling, hybridization and detection were performed using a DIG-High Prime DNA Labeling and Detection kit I (Roche Molecular Biochemicals).

2.5. Lipopolysaccharide (LPS) injection

One µl of LPS (Sigma) solution (0.1 µg/ml in water) was injected into the head of pupae that had been in diapause for 25 days.

3. Results and discussion

3.1. Ribosomal protein L32

As a control for Northern hybridization, a fragment of the *RpL32* was cloned from *S. crassipalpis* (DDBJ/GenBank/ EMBL accession number AB074534). The nucleotide and putative amino acid sequences were quite similar to *RpL32* from other insects (Fig. 1).

3.2. Sequences of sarcocystatins

One of the cDNA clones (DDBJ/GenBank/EMBL accessionnumberAB074535, 3rd row in Figs. 2A, B) derived from the RACE product library had a sequence showing similarity to the gene encoding the cysteine proteinase inhibitor, *sarcocystatin A* (abbreviated as *Scys-A* and SCYS-A for the gene and protein), from *S. peregrina* (DDBJ/GenBank/EMBL accession number J02847, 1st row in Figs. 2A, B). SMART analysis (Schultz et al., 2000) detected the cystatin-like domain in the putative amino acid sequence (position 2–80 in the 3rd row, Fig. 2B). Amino acid sequences of QxVxG (x can be one of several amino acids) and PW that are highly conserved in various cystatins (Brown and Dziegielewska, 1997) were also observed in the putative amino acids (3rd row in Fig. 2B). These two regions appear, from X-ray crystallographic analysis of the chicken cystatin, to form a 'wedge' that can interact with the active site of papain (Bode et al., 1988). In addition, the putative amino acid sequence contains two cysteine residues at the C-terminus (positions 90and110in the 3rd row, Fig. 2B). These residues were predicted to form the B-type disulfide loop in *S. peregrina* SCYS-A (Brown and Dziegielewska, 1997).

However, the similarity between this gene in *S. crassipalpis* and *Scys-A* in *S. peregrina* was considerably lower than that between other orthologous genes in these species (Figs. 2A, B and Table 2). For example, the *S. crassipalpis* anterior fat body protein gene (Goto, S.G. and Denlinger, D.L., unpublished data) had identities of 96.1% for the nucleotides and 98.1% for the amino acids with the *S. peregrina* gene (DDBJ/GenBank/EMBL accession number AB036903) (Nakajima and Natori,2000), and for the *mitochondrial cytochrome c oxydase subunit I* gene, the identity was 91.6 and 99.6% for the nucleotides and amino acids, respectively (DDBJ/GenBank/EMBL accession numbers are AF259510 and AF259509, respectively) (Wells et al., 2001). This indicated that the gene cloned from *S. crassipalpis* is not a homolog of *Scys-A* in *S. peregrina*, but is a novel gene.

To further address this issue, *Scys-A* in *S. crassipalpis* was amplified with RT-PCR. The nucleotide and putative amino acid sequences of the obtained fragment (2nd row in Figs. 2A, B) revealed high similarity with *S. peregrina ScysA* (Fig. 2 and Table 2), indicating that this is a homolog of *Scys-A*. Here we name the novel gene *sarcocystatin B* (abbreviated as *Scys-B* and SCYS-B for the gene and protein). *S. crassipalpis* SCYS-A also contains the conserved amino acid sequences QxVxG and PW (2nd row in Fig. 2B). Like *S. peregrina, S. crassipalpis* SCYS-A has two cysteine residues at the N-terminus, separated by only two amino acids (positions 22 and 25 in the 2nd row, Fig. 2B). A C-type disulfide loop is thought to be formed by this sequence (Brown and Dziegielewska, 1997).

Delbridge and Kelly (1990) reported a gene, *cystatin-like* (*cys*), in *Drosophila melanogaster* (DDBJ/GenBank/EMBL accession number X55178). In addition, the Berkeley *Drosophila* Genome project recently revealed that *D. melanogaster* has three other unidentified genes (CG15369, CG12163 and CG8066: DDBJ accession numbers are AE003466, AE003603 and AE003705, respectively) that are similar to the *cys* gene previously reported in *D. melanogaster* (Berkeley Drosophila Genome project, personal communication). Together, these four genes formed a cystatin family in *D. melanogaster*. However, *Scys-A* and *-B* in *Sarcophaga* show low similarity to these *Drosophila* genes (Table 2), thus it is not yet possible to determine which genes in *Drosophila* are orthologs of the *Sarcophaga* genes.

3.3. Tissue specificity

The low similarity between SCYS-A and -B in *S. crassipalpis* (Fig. 2 and Table 2) suggests the possibility of functional differences. To approach this question, tissue distribution of their transcripts was investigated in nondiapausing pupae.

In *S. peregrina*, SCYS-A protein is present in the fat body (Suzuki and Natori, 1986). The present study confirmed this profile at the messenger RNA (mRNA) level for *Scys-A* in *S. crassipalpis* (Fig. 3A). The transcripts are approximately 0.7 kb in length. *Scys-A* mRNA was also present in the gut and epithelium. In contrast, expression of *Scys-B* transcripts was high in both the fat body and gut, but was low in the epithelium. Two bands, approximately 0.8 and 0.7 kb in length, were observed for *Scys-B*, but the signal for the 0.7 kb band was weaker and sometimes undetectable. Brain expressed quite low levels of both the *Scys-A* and *-B* transcripts. These differences in tissue specificity suggest that SCYS-A and -B proteins play distinct roles in each tissue.

3.4. LPS injection

A cysteine protease inhibitor from the horseshoe crab *Tachypleus tridentatus*, limulus (L)-cystatin, has antimicrobial activity against Gram-negative bacteria (Agarwala et al., 1996). In addition, cysteine proteinase inhibitors from the silkmoth, *Bombyx mori*, act as defense proteins against invading pathogens and parasites, many of which use cysteine proteinases to enter their hosts. For example, *B. mori* nuclear polyhedrosis virus has a gene encoding a cysteine proteinase that is similar to lysosomal cysteine proteinases (cathepsins B, H, L, and S), and it is postulated that the proteinase plays an important role in the degradation of host tissues (Yamamoto et al., 1999).

Northern hybridization, however, revealed that *Scys-A* and *-B* expression in *S. crassipalpis* was not altered by injection of LPS, at least not within 6 h after injection (data not shown). LPS is commonly used as an immune stimulant in other insects, and our protocol is effective in inducing sarcotoxin II, an antimicrobial peptide gene in *S. crassipalpis* (Rinehart, J.P., et al., personal

communication). In addition, Kobayashi et al. (1993) reported that expression of *Scys-A* in NIH-Sape-4 cells (embryonic cell line from *S. peregrina*) was not altered by the presence of β -glucan, another inducer of antimicrobial peptides. These results suggest that SCYS-A and -B are not involved in defense responses to microbial invasion.

3.5. Developmental regulation of sarcocystatins

In non-diapausing individuals of *S. peregrina*, it was reported that SCYS-A protein and Scys-A transcripts are developmentally regulated: expression increased markedly at puparium formation, reached a maximum in early pupae, and then gradually decreased (Suzuki and Natori, 1986; Saito et al., 1989). The present study with *S. crassipalpis* also shows a similar pattern of expression for *Scys-A* transcripts in non-diapausing pupae: expression was already high in the wandering stage of the third (final) larval instar and was upregulated further in the pupae. By contrast, expression of *Scys-A* was low in wandering larvae destined for pupal diapause and was downregulated throughout diapause (Fig. 3B). Expression of *Scys-B* was nearly undetectable in non-diapausing pupae in non-diapausing pupae (Fig. 3B). Expression was likewise nearly undetectable in wandering larvae programmed for diapause, but interestingly, expression of *Scys-B* was highly upregulated early in pupal diapause (Fig. 3B). Expression gradually decreased by mid and late diapause. Neither *Scys-A* nor *-B* showed increased expression when diapause was terminated, at least not within 24 h (data not shown).

Other diapause upregulated genes in *S. crassipalpis* include genes encoding two heat-shock proteins, *heat shock protein* 70 (*Hsp70*) (Rinehart et al., 2000) and *Hsp23* (Yocum et al., 1998), both of which are upregulated throughout diapause. A gene encoding *ultraspiracle (usp)*, a receptor protein for the insect growth hormone, 20-hydroxyecdysone, is upregulated late in diapause (Rinehart et al., 2001), presumably as a preparative step for diapause termination, but *Scys-B* is the first identified gene that is expressed strongly only in early diapause. Early diapause genes may be of interest for several reasons. Such genes are possibly involved in bringing development to a halt. And, their eventual decline in expression likely marks a transition to a new phase of diapause. For example, midway through diapause, the pupae of *S. crassipalpis* switch from lipid to non-lipid resources for energy utilization (Adedokun and Denlinger, 1985).

In *S. peregrina*, one of the cysteine proteinases, cathepsin L, acts as a regulator of metamorphosis, and blocking the action of cathepsin L can inhibit differentiation of imaginal discs (Homma and Natori, 1994). One of the serine proteases, cathepsin B, participates in dissociation of the fat body during metamorphosis (Kurata et al., 1990), and blocking this action would serve to maintain the integrity of the pupal fat body. High expression of a gene encoding a proteinase inhibitor in early diapause is thus likely to assure the shutdown in development that characterizes the diapause state. It is also clear that the roles for SCYS-A and -B must be a bit different. *Scys-A* is downregulated before diapause begins, while *Scys-B* is highly upregulated in early diapause. SCYS-B thus emerges as the potentially important homolog contributing to the dramatic halt in protein synthesis noted in early diapause (Joplin and Denlinger, 1989).

4. Conclusion

1. Two genes encoding cystatins were cloned from *S. crassipalpis*. One was previously cloned from *S. peregrina*, but the other is novel. Here the novel gene is named *sarcocystatin B*.

2. *Sarcocystatin A* transcripts are mainly present in the fat body, while *sarcocystatin B* transcripts are highly expressed in both the fat body and gut. Low expression of *sarcocystatin A* and *B* are noted in the brain.

3. LPS injection did not alter their expression patterns.

4. Both *sarcocystatins A* and *B* are developmentally regulated but their expression patterns differ. *Sarcocystatin A* transcripts are present in both diapause- and non-diapause-destined wandering larvae, but *sarcocystatin B* transcripts are not. *Sarcocystatin A* is downregulated throughout diapause, but *sarcocystatin B* is upregulated in early diapause. The upregulation of *sarcocystatin B* in early diapause suggests a possible role in halting development.

Acknowledgements

The continuous support from members of the Insect Physiology Laboratory in the Department of Entomology, Ohio State University is gratefully acknowledged. This study was supported in part by USDA-NRI grant 98-35302-6659.

References

- Adedokun, T.A., Denlinger, D.L., 1985. Metabolic reserves associated with pupal diapause in the flesh fly, Sarcophaga crassipalpis. J. Insect Physiol. 31, 229–233.
- Agarwala, K.L., Kawabata, S.-I., Hirata, M., Miyagi, M., Tsunasawa, S., Iwanaga, S., 1996. A cysteine protease inhibitor stored in the large granules of horseshoe crab hemocytes: purification, characterization, cDNA cloning and tissue localization. J. Biochem. 119, 85–94.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
- Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., Turk, V., 1988. The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. EMBO J. 7, 2593–2599.
- Brown, W.M., Dziegielewska, K.M., 1997. Friends and relations of the cystatin superfamily new members and their evolution. Protein Sci. 6, 5-12.
- Delbridge, M.L., Kelly, L.E., 1990. Sequence analysis, and chromosomal localization of a gene encoding a cystatin-like protein from Drosophila melanogaster. FEBS Lett. 274, 141–145.
- Denlinger, D.L., 1972. Induction and termination of pupal diapause in Sarcophaga (Diptera: Sarcophagidae). Biol. Bull. Woods Hole 142, 11-24.
- Denlinger, D.L., 2000. Molecular regulation of insect diapause. In: Storey, K.B., Storey, J. (Eds.). Environmental Stressors and Gene Responses, Elsevier Science, Amsterdam, pp. 259–275.
- Denlinger, D.L., 2002. Regulation of diapause. Annu. Rev. Entomol. 47, 93-122.
- Denlinger, D.L., Campbell, J.J., Bradfield, J.Y., 1980. Stimulatory effect of organic solvents on initiating development in diapausing pupae of the flesh fly, Sarcophaga crassipalpis, and the tobacco hornworm Manduca sexta. Physiol. Entomol. 5, 7–15.
- Homma, K.-I., Natori, S., 1994. Purification, characterization, and cDNA cloning of procathepsin L from the culture medium of NIH-Sape-4, an embryonic cell line of Sarcophaga peregrina (flesh fly), and its involvement in the differentiation of imaginal discs. J. Biol. Chem. 269, 15258–15264.
- Joplin, K.H., Denlinger, D.L., 1989. Cycles of protein synthesis during pupal diapause on the flesh fly, Sarcophaga crassipalpis. Arch. Insect Biochem. Physiol. 12, 111–122.
- Kobayashi, A., Matsui, M., Kubo, T., Natori, S., 1993. Purification and characterization of a 59-kilodalton protein that specifically binds to NF-kB-binding motifs of the defense protein genes of Sarcophaga peregrina (the flesh fly). Mol. Cell. Biol. 13, 4049–4056.
- Kurata, S., Saito, H., Natori, S., 1990. Participation of hemocyte proteinase in dissociation of the fat body on pupation of Sarcophaga peregrina (flesh fly). Insect Biochem. 20, 461–465.
- Nakajima, Y., Natori, S., 2000. Identification and characterization of an anterior fat body protein in an insect. J. Biochem. 127, 901–908.
- Rawlings, N.D., Barrett, A.J., 1990. Evolution of proteins of the cystatin superfamily. J. Mol. Evol. 30, 60-71.
- Rinehart, J.P., Yocum, G.D., Denlinger, D.L., 2000. Developmental upregulation of inducible hsp70 transcripts, but not the cognate form, during pupal diapause in the flesh fly, Sarcophaga crassipalpis. Insect Biochem. Mol. Biol. 30, 515–521.
- Rinehart, J.P., Cikra-Ireland, R.A., Flannagan, R.D., Denlinger, D.L., 2001. Expression of ecdysone receptor is unaffected by pupal diapause in the flesh fly, Sarcophaga crassipalpis, while its dimerization partner, USP, is downregulated. J. Insect Physiol. 47, 915–921.
- Saito, H., Suzuki, T., Ueno, K., Kubo, T., Natori, S., 1989. Molecular cloning of cDNA for sarcocystatin A and analysis of the expression of the sarcocystatin A gene during development of Sarcophaga peregrina. Biochemistry 28, 1749–1755.
- Sambrook, J., Fritsch, E.G., Maniatis, T., 1989. Molecular Cloning: a Laboratory Manual, 2nd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., Bork, P., 2000. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. 28, 231–234.
- Suzuki, T., Natori, S., 1985. Purification and characterization of an inhibitor of the cysteine protease from the hemolymph of Sarcophaga peregrina larvae. J. Biol. Chem. 260, 5115–5120.
- Suzuki, T., Natori, S., 1986. Changes in the amount of sarcocystatin A, a new cycteine proteinase inhibitor, during the development of adult Sarcophaga peregrina. Insect Biochem. 16, 589–595.
- Turk, V., Bode, W., 1991. The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett. 285, 213-219.
- Warman, G.R., Newcomb, R.D., Lewis, R.D., Evans, C.W., 2000. Analysis of the circadian clock gene period in the sheep blow fly Lucilia cuprina. Genet. Res. Camb. 75, 257–267.
- Wells, J.D., Pape, T., Sperling, F.A., 2001. DNA-based identification and molecular systematics of forensically important Sarcophagidae (Diptera). J. Forens. Sci. 46, 1098–1102.
- Yamamoto, Y., Watabe, S., Kageyama, T., Takahashi, S., 1999. Purification and characterization of Bombyx cysteine proteinase specific inhibitors from the hemolymph of Bombyx mori. Arch. Insect Biochem. Physiol. 41, 119–129.
- Yocum, G.D., Joplin, K.H., Denlinger, D.L., 1998. Upregulation of a 23 kDa small heat shock protein transcript during pupal diapause in the flesh fly Sarcophaga crassipalpis. Insect Biochem. Mol. Biol. 28, 677–682.

Table 1 Sequences of primers

Oligonucleotides	Sequence $(5^0 q 3^0)$	
SCA-F1	TGGCCAGTCGTTATTTGAATC	
SCA-R1	GTTCACCGGGACAATTGAAG	
RP49L-F1	CACCAGTCGGATCGNTATGCC	
RP49L-F2	AGCTGCTTGGCNCGNTC	
M13-F	AAACGACGGCCAG	
M13-R	CAGGAAACAGCTATGAC	

Table 2

Percentage of similarities (identity/positive) and e-values (BlastP analysis, Altschul et al., 1997) of pair-wise alignment for sarcocystatins A and B amino acid sequences when compared to other known members of the cystatin family in *D. melanogaster*^a

	Accession #	Comparison		
		S.p. Scys-A	S.c. Scys-A	S.c. Scys-B
S.p. Scys-A	J02847			
S.c. Scys-A	AB074535	97.2/99.1 (2e-54)		
S.c. Scys-B	AB074536	57.6/82.6 (2e-24)	56.0/81.3 (2e-16)	
D.m. CG15369	AE003446	49.6/82.6 (2e-27)	46.2/78.3 (2e-19)	52.2/79.3 (2e-21)
D.m. CG12163	AE003603	35.2/67.2 (4e-12)	34.0/66.0 (2e-8)	40.2/72.8 (2e-13)
D.m. cys	X55178	38.3/76.7 (3e-17)	42.5/77.4 (2e-15)	37.1/76.4 (5e-10)
D.m. CG8066	AE003705	34.2/70.2 (5e-11)	34.6/65.4 (5e-10)	37.1/68.5 (3e-7)

a e-Values are shown in parentheses. S.p.: S. peregrina; S.c.: S. crassipalpis; and D.m.: D. melanogaster.

S. D. B.	crassipalpis subobscura mori	AAGTTGTCGCACAAATGGAGAAAACCTAAAGGTATTGACAACAGAGTACGTCGTCGTTTC TGTTTGCAGCACAAATGGCGCAAGCCTAAAGGTATTGACAACAGAGTACGTCGTCGCTCC AAACTTAAGAGGAATTGGCGTAAACCTAGAGGTATTGACAACAGAGTCCGCAGGCGGTTC * * ** *** * ** *** ****	60 1044 85
		AAGGGTCAATACTTGATGCCCAACATTGGTTATGGTTCCAACAAGCGTACCCGCCATATG AAGGGCCAGTACCTGATGCCCAACATCGGTTATGGCTCCAACAAGCGTACCCGTCATATG AAGGGTCAATACTTGATGCCCAACATTGGTTACGGTTCCAACAAGAAGACCCGTCATATG ***** ** *** ************************	120 1104 145
		CTCCCAATTGGCTTCAAGAAGTTTGTTGTTGTTCATAATGTTAAGGAACTTGAAGTTTTGATG CTGCCCACTGGCTTCAAGAAATTCCTGGTGCACAATGTGCGAGAGCTGGAGGTTCTGCTC CTCCCAAATGGATTCCGTAAGGTCCTAGTTCACAATGTTAAAGAGCTGGAAATCTTGATG ** ** * *** *** *** ** * * * * ** ******	180 1164 205
		ATGCAAAATCGTGTCTATTGCGGTGAAATTGCCCATGGTGACTCCTCAAAGAAACGTAAG ATGCAGAACCGCATCTACTGCGGTGAGATCGCCCACGCCGTCTCCTCAAAGAAGCGCAAG ATGCAAAACAGGAAG ***** ** *	240 1224 220
S. D.	crassipalpis subobscura	GAAATCGTT GAAATCGTC	249 1233
B))		
S. D. B.	crassipalpis subobscura mori	KLSHKWRKPKGIDNRVRRRFKGQYLMPNIGYGSNKRTRHMLPIGFKKFVVHNVKELEVLM KLSHKWRKPKGIDNRVRRFKGQYLMPNIGYGSNKRTRHMLPTGFKKFLVHNVRELEVLL KLKRNWRKPRGIDNRVRRFKGQYLMPNIGYGSNKKTRHMLPNGFRKVLVHNVKELEILM	60 88 68

ibobscura	KLSHKWRKPKGIDNRVRRRFKGQYLMPNIGYGSNKRTRHMLPTGFKKFLVHNVRELEVLL	88
ori	KLKRNWRKPRGIDNRVRRRFKGQYLMPNIGYGSNKKTRHMLPNGFRKVLVHNVKELEILM	68
	** • • • • * * * * * * * * * * * * * *	
	MQNRVYCGEIAHGDSSKKRKEIV	83
	MQNRIYCGEIAHAVSSKKRKEIV	111
	MQNRK	73
	* * * *	

Fig. 1. Nucleotide (A); and putative amino acid (B) sequences of *RpL32* in *S. crassipalpis*, *Drosophila subobscura* and *Bombyx mori* (DDBJ/GenBank/EMBL accession numbers AB074534, AJ228908 and AB048205, respectively). '*': same nucleotide or amino acid residue; ':': strong amino acid positive; and '.': weaker amino acid positive.

Δ)	S.p. Scys-A	AAAATACTGGCCAGTCGTTATTTGAATCTTCATTGATTAACAATATCTTCATCGGACATT	60
~)	S.c. Scys-A S.c. Scvs-B	TTCATTAGTGAACAACAGCTTCGAACATT	29
		ATGAAATACGTTTTGATTTTGTGCGTTATCACTTTAGCTACTGTGGCTTATGCCCAGCCA	. 120
			89
			100
		CAGTGCGTCGGTTGTCCCCAGTGAAGTGCAGGGCGATAAACTTAAGCAATCGGAAGAAACT	149
		CGTTGATATTGAGAAAGCGGCAAAAGAT ** * ** ** ** ** ** ** *	28
		TTAAACAAGTCACTATCGAAGTTAGCAGCTGGTGATGGACCCACTTACAAACTGGTTAAA	240
		TTAAATAAGTCGCTATCGAAGTTAGCAGCTGGTGATGGGCCCCACTTACAAACTGGTTAAA	209
		TTGACTGCTTCTTTGGAGAAATTAGCTGCTGGTGATGGACCAAATTTTAAACTGGGTAAA ** * * * * * **** ***** ***********	88
		ATCAACTCAGCCACCACACAAGTTGTTTCTGGCAGTAAGGACGTAATAAATGCTGATTTA	300
		ATCAACTCAGCCACCACCAAGTTGTTTCTGGCAGTAAAGACGTAATAAATGCTGATTTA ATACACTCAGCTGAGACACAAGTGGCGGGGCACCTTTTAAGCTAAAATGCTGGCGGGTTCTC	269 148
		** ******* ******* ** * ** * * * *** ****	140
		AAAGATGAAAACGATAAGACTAAGACCTGCGACATAACCATTTGGTCTCAACCTTGGTTG	360
		AAAGATGAAAAACGATAAGATTAAGACCTGCGATATAACCATTTGGTCTCAACCTTGGTTG ATTGATGCTGATGGCAAGACCAAAGAATGCAAAGTTACCATTTGGAGTCGCCCTTGGTTG	329 208
	· .	* **** * * **** ** *** * * ************	200
		GAAAATGCCATCGAAGTTACCTTCAATTCTCCCGCTGAACCGAAAGTCCTTAAAGAAGAAC	420
		GAAAATGGCATTGAAGTTAC	349
		CCCAATGGCGTTGAAGTAACCTTTGAGTGTGAAGGCGAAGACAAAATTGTTAGGAAACAT ****** * ***** **	268
	S.p. Scys-A	AGTGCCTAAAATCAATATAAAACAGTCGATTGCATTGGCTGACTGTTAATGCTCACTGAT	480
	S.c. Scys-B	AGTGCCTAAGATGATCAAGAACATTTCTTATGTATT <u>ATTTA</u> CTTGATCCACAAAC	323
		ATTCTACGTTCGAATACATAATTTTATAAATCAAATTGAATATTTTGTTTG	540
		A <u>ATTTA</u> TATGTGTATGTATTATA <u>ATTTA</u> AATGGATGCGTCATATCTATGTGATAAGCCTTT	383
		ATTTGCATAAAAAAGTGA <u>AATAAA</u> AAAAGAAAAAC	575
		CTTTGTGTGTGAAGCTTAATATGTAATCATTTCTTATTCTATTTATATGCA <u>AATAAA</u> C	443
	S.c. Scys-B	AGTAAATATTTTATGTGTCCGACTATATAAACAATC	479
D)	S D SCVS-A		60
D)	S.c. SCYS-A	MKIVIIICVIIIAIVAIAGPQCVGCFSEVAGDAIAQSEEIIAASDSALAAGDGFIIAIVA MKYVLILCVIILATVAYARPQCVGCPSEVQGDKLKQSEEILNKSLSKLAAGDGPIYKLVK	60
	S.C. SCYS-B	VDIEKAAKDLTASLEKLAAGDGPNFKLGK	29
		.::::: * ** ***************************	
		INSATT <mark>QVVSG</mark> SKDVINADLKDENDKTKTCDITIWSQ <mark>PW</mark> LENGIEVTFNCPGEPKVVKKH	120
		INSATT OVVSC SKDVINADLKDENDKIKTCDITIWSQPWLENGIEV IHSAETOVVACTILTKIVADLIDADGKTKECKVTIWSRDM.DNGVEVTFECECEDKIVPKH	106 89
		*:** ***:*: * *** * :.* * *.:**********	0.5
	S.p. SCYS-A	SA	122
	S.c. SCYS-B	SA	91

Fig. 2. Nucleotide (A); and putative amino acid (B) sequences of *Scys-A* and *-B* in *S. peregrina* (*S.p.*) and *S. crassipalpis* (*S.c.*). '*': same nucleotide or amino acid residue; ':': strong amino acid positive; and '.' weaker amino acid positive. Bars indicate the alignment gaps. Polyadenylation signal and RNA instability motif were indicated by underline and double-underline, respectively. Conserved motifs of QxVxG and PW were indicated by black boxes. Sequences are available from DDBJ/GenBank/EMBL accession numbers, J02847, AB074535 and AB074536 for *S.p. sarcocystatin A*, *S.c. sarcocystatin A* and *S.c. sarcocystatin B*, respectively.



Fig. 3. Northern hybridization for tissue (A); and developmental stage (B) with *Scys-A*, -*B* or *RpL32* probe. ND: non-diapause; D: diapause; w1: wandering larvae; and P: pupae.