Biochimie 92 (2010) 1749–1759

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Review

Protease inhibitors and proteolytic signalling cascades in insects

David Gubb [a,*], Arantza Sanz-Parra [a], Laura Barcena [a], Laurent Troxler [b], Ane Fullaondo [a,**]

[a] Department of Functional Genomics, CICbioGUNE Institute, Bizkaia Technology Park, Ed. 801A, 48160 Derio, Spain
[b] IBMC, UPR 9022 du CNRS, 15 rue Rene Descartes, F67084 Strasbourg Cedex, France

ARTICLE INFO

Article history:
Received 16 July 2010
Accepted 3 September 2010
Available online 17 September 2010

Keywords:
Insect
Protease inhibitors
Serpin
Kazal
Signalling

ABSTRACT

Proteolytic signalling cascades control a wide range of physiological responses. In order to respond rapidly, protease activity must be maintained at a basal level: the component zymogens must be sequentially activated and actively degraded. At the same time, signalling cascades must respond precisely: high target specificity is required. The insects have a wide range of trapping- and tight-binding protease inhibitors, which can regulate the activity of individual proteases. In addition, the interactions between component proteases of a signalling cascade can be modified by serine protease homologues. The suicide-inhibition mechanism of serpin family inhibitors gives rapid turnover of both protease and inhibitor, but target specificity is inherently broad. Similarly, the TEP/macroglobulins have extremely broad target specificity, which suits them for roles as hormone transport proteins and sensors of pathogenic virulence factors. The tight-binding inhibitors, on the other hand, have a lock-and-key mechanism capable of high target specificity. In addition, proteins containing multiple tight-binding inhibitory domains may act as scaffolds for the assembly of signalling complexes. Proteolytic cascades regulated by combinations of different types of inhibitor could combine the rapidity of suicide-inhibitors with the specificity lock-and-key inhibitors. This would allow precise control of physiological responses and may turn out to be a general rule.

1. The different families of protease inhibitors and their functions

Protease inhibitors are found in all eukaryotes and play critical roles in regulating many biological processes. Their main function is to limit the activity of proteases, both those with broad proteolytic actions (such as digestive enzymes and pathogen-encoded virulence factors) and those with specific target substrates (such as the component proteases of signalling cascades). In mammals, many physiological responses are activated by extra-cellular signalling cascades, such as blood-clotting, angiogenesis, wound-healing and the inflammatory and complement responses [1–3]. Similar signalling cascades occur in insects, but few have been studied in detail. In general, the component proteases of extracellular signalling cascades are secreted as inactive zymogens, with each successive zymogen being activated by the previous protease in the cascade. Individual zymogens may be autocatalytic, particularly the first protease in a pathway, but amplification cascades can only respond to physiological challenges if they retain a low-level basal activity. These constraints imply that the activity of proteases involved in signalling cascades should be tightly regulated at the post-translational level. With continued activation of a pathway, changes may occur at the transcriptional level in some, or all, of its components. Such changes, however, are secondary, and understanding the post-translational regulation of protease activity remains a major challenge.

There are two major types of protease inhibitor, trapping inhibitors and tight-binding inhibitors. The first type form irreversible complexes with their target proteases, while the second type form strong, but reversible, interactions [4]. The are only two large families of trapping inhibitors, the Serpins (**Ser**ine **p**rotease **in**hibitors, MEROPS inhibitor family I4, http://merops.sanger.ac.uk) and the macroglobins (MEROPS inhibitor family I39); with the single example of a baculovirus trapping inhibitor with a novel inhibitory mechanism [5]. These inhibitors are large molecular weight proteins in which cleavage of a peptide bond triggers a conformational change. In the case of the serpin family, an irreversible, covalently linked, acyl–enzyme complex is formed with the target protease (see below). In the macroglobin-family inhibitors, the catalytic site of the target protease is generally sequestered within the re-folded inhibitor molecule and remains intact [6], although there are a few examples of covalent macroglobulin/protease complexes [7]. Both

* Corresponding author. Tel.: +34 944061309; fax: +34 944061301.
** Corresponding author. Sequencing and Genotyping Service, CICbioGUNE Institute, Bizkaia Technology Park, Ed. 502, 48160 Derio, Spain. Tel.: +34 946572524; fax: +34 946568732.
E-mail address: dgubb@cicbiogune.es (D. Gubb).

0300-9084/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2010.09.004

families of trapping inhibitors contain a single inhibitory domain, but the macroglobulin proteins include several other domains conserved within the family. In contrast, the tight-binding inhibitors are short folded domains that form high-affinity lock-and-key interactions with their target proteases. Most of the 48 families of inhibitors recognised by [4], and listed in the MEROPS database [8], fall into this group. Many tight-binding inhibitors are simple proteins of less than 100 amino acids. In contrast to the trapping inhibitors, tight-binding inhibitory domains are often present as component modules within compound inhibitors, or complex inhibitors that include one, or more, inhibitory domains together with other conserved protein domains.

The mechanisms of action of the short protein tight-binding inhibitors have been investigated for many years [9]. More recently, the serpins and macroglobulins have been intensively studied because of their importance in human disease mechanisms. In humans, many rapid physiological responses are regulated by serpins [1]. The coagulation, inflammatory and complement pathways are controlled by antithrombin, α_1-antitrypsin and C1-inhibitor, respectively [2], while plasminogen activator inhibitor-1 (PAI-1) modulates angiogenesis, affecting both wound-healing and tumour growth [3]. It is notable that the α_2-macroglobulin trapping inhibitor also regulates blood coagulation, via thrombin, and fibrinolysis, via plasmin. It is likely that physiological responses may be regulated by the combination of tight-binding and trapping inhibitors, to greater extent than has been recognised. We propose that it is the combination of interactions between various classes of inhibitors with active serine proteases and inactive serine protease homologues that allows rapid and precise regulation of the activity of zymogen/protease signalling cascades.

2. Inhibitory mechanisms and constraints

2.1. Trapping inhibitors, serpins

Serpins interact with their target proteases via a unique "suicide-cleavage" mechanism, which results in the formation of an inactive, covalently linked, serpin/protease complex [10]. The serpin fold consists of three β-sheets (A–C) with 8 or 9 α-helical linkers and an exposed reactive centre loop (RCL) of about 20 amino acids, which acts as bait for the target protease. The complete fold is between 350 and 450 amino acids in length. In the native state, serpins are in a metastable, stressed (S) conformation. Following RCL cleavage, the serpin adopts a relaxed conformation (R), which distorts the active site of the protease and traps it in a covalent serpin/protease complex. During this process, the protease is translocated through 70 Å, from the "top" to the "bottom" pole of the serpin, and the RCL inserts as an extra strand within β-sheetA. The protease is denatured by crushing against the bottom of the serpin core and the denatured serpin/protease complex is targeted for degradation [11], Fig. 1.

The major function of the serpin core is to deliver the potential energy store held in the metastable conformation, but the constraints that the S→R transition places on the serpin fold leave little potential for modifications of the core structure to contribute to protease target specificity. In practice, this specificity resides largely in the primary amino-acid sequence of the RCL, which presents an idealized bait to the target protease. The protease cuts between the P1 and P1′ residues of the RCL, and it is the identity of the P1 residue, which is the major determinant of substrate specificity (with a lesser contribution from the P1′ and other adjacent residues). This is quite different from the mechanism of tight-binding inhibitors, such as those of the Kazal and Kunitz families, in which a precise lock-and-key fit between inhibitor and protease can give very high specificity (see below).

One particular constraint on the serpin structure is that the S→R transition requires a flexible "hinge region" of small amino acids to allow insertion of the cut RCL loop within β-sheetA [12]. This hinge region extends from P17 to P9 (17 to 9 residues N-terminal to the P1/P1′ protease cleavage site), with a consensus E[EKR]G[TS].$(AGS)_4$ sequence, based mainly on mammalian sequences [12]. There are several groups of protein which conserve the β-sheet plus α-helical linker serpin fold, but which lack the consensus hinge sequence, such as ovalbumin, Hsp47 and Maspin

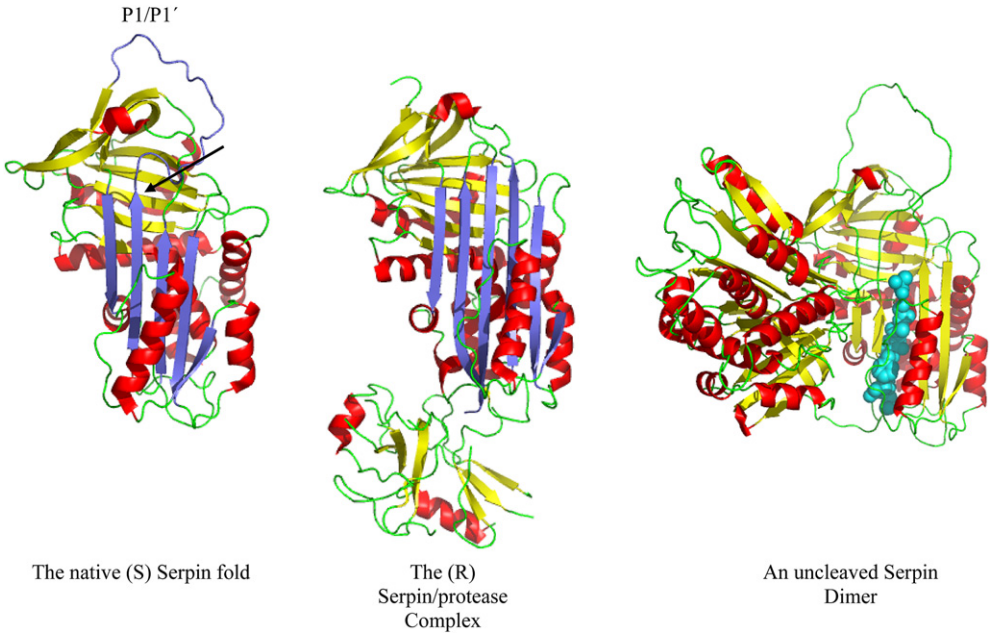

Fig. 1. The native (S) serpin structure carries an exposed RCL, which is cleaved between the P1 and P1′ sites. Inhibitory serpins have a flexible hinge region, consisting of amino acids with short side-chains (arrow). Following cleavage, the protease is translocated and crushed against the bottom of the (R) serpin as a denatured covalent complex as the RCL inserts between β-sheetA (purple). In the absence of cleavage, the RCL of one serpin (blue) can insert between the strands of β-sheetA (yellow) of another serpin to form polymers. This model shows the classical single-strand insertion mechanism, rather than the novel domain-swapping mechanism proposed by Yamasaki et al. [68].

[12]. These proteins have varied roles as storage proteins, molecular chaperones, hormone transporters and tumour suppressors [13]. In many cases it is unclear what role the serpin fold plays in the activity of these non-inhibitory molecules.

2.2. Trapping inhibitors, macroglobulins

The macroglobulin-family inhibitors inactivate a wide range of proteases and can also act as hormone transporters. Most function as tetramers consisting of disulphide-linked sub-units, although monomeric and dimeric α-macroglobulins are also known [7]. As with serpins, these inhibitors carry a bait region that is susceptible to protease cleavage. The resultant conformational change enfolds the protease within the inhibitor and exposes sites that are recognised by high-affinity macroglobulin receptors, leading to rapid uptake and degradation of the complex. The bait region tends to be longer than the RCL of serpins and contain multiple cleavage sites; as a result of which α-macroglobulins can form complexes with a range of proteases from all classes [14].

In the native conformation, human α_2-macroglobulin is stabilised by a thioester link between Cys949 and Gln952. This link becomes much more susceptible to nucleophilic attack after the conformational change which follows cleavage within the bait region [15,16]. A striking feature of α_2-macroglobulin is the wide range of serine, cysteine, aspartic and metallo-proteases that it is able to inhibit. Under physiological conditions, α_2-macroglobulin inhibits blood coagulation, via thrombin and fibrinolysis, via plasmin and kallikrein [17]. Both these physiological responses are also regulated by serpins. Notably, while α_2-macroglobulin and the plasma serpins are both secreted from the liver, their degradation routes are distinct. The α_2-macroglobulin/protease complexes are cleared from circulation primarily via macrophages [7], while serpin/protease complexes are taken up again and degraded in the liver.

2.3. Tight-binding inhibitors

The mechanism of action of tight-binding inhibitors has been extensively studied for many years, [4,9,18–20]. Tight-binding inhibitors tend to be shorter than 100 amino acids. As with trapping inhibitors, these peptides carry a bait region, but the inhibitor forms a precise fit with the catalytic site of their target protease and acts as a particularly stable substrate/protease complex [19]. The structural constraints on the core fold are relatively loose and a wide variety of protein folds can support the exposed RCL loop of the tight-binding inhibitors. A common feature of these core folds is that they are cross-linked, generally with cysteine bridges. As a consequence, cleavage at the P1/P1′ site within their bait region opens a ring structure, but does not release separate peptide fragments. The cut ends of the inhibitor tend to remain in place, so that the hydrolysis reaction is reversible. Most of the tight-binding inhibitors function via the “standard” mechanism of Laskowski and Qasim [21]. A diagnostic feature of this mechanism is that the inhibitor/protease complex forms an equilibrium mixture of cleaved and un-cleaved forms. The protease molecule can be released intact, while the released inhibitor can be in the native, or the cleaved form. Some Kunitz inhibitors, however, have such strong target affinities that inhibitor/protease binding becomes essentially irreversible [22]. Although the mechanism is completely different to that of the formation of serpin “suicide-binding complexes”, the physiological consequences are similar.

The Kunitz family inhibitors were first described as simple peptides present at high concentrations in beans, where they block the activity of the digestive enzymes of insects and other herbivores [18]. These inhibitors typically consist of short proteins of about 20 kDa, but have also been found as individual domains within larger proteins. Similarly, the first Kazal family inhibitors were identified as short, single domain proteins [9], but larger proteins containing from 2 to 15 Kazal domains have since been described [4]. The tendency to form compound protease inhibitors (CI) with multiple inhibitory domains is common in other tight-binding inhibitor families [4]. A few examples are known of heterotypic compound inhibitors (hCI), containing inhibitory domains from more than one family, but this is rare. In addition, single or multiple inhibitory domains can be identified within complex proteins containing heterogeneous non-inhibitory domains. We shall refer to this sub-group of compound inhibitors as complex protease inhibitors (CxI). The biological function of most CxIs is obscure as is the role of the inhibitory domains in their activity.

3. Insect protease inhibitors

The protease inhibitors of insects have received much less attention than the human trapping inhibitors, or the plant and mammalian tight-binding inhibitors. Despite this there is an increasing body of experimental work and comparative genomic data. In general, the insect protease inhibitors appear to regulate a similar range of responses as their mammalian counterparts. In addition to inhibitors of digestive enzymes and pathogen-encoded virulence factors, protease inhibitors in *Drosophila melanogaster* are involved in morphogenetic pathways, spermatogenesis, organic anion transport, cellular necrosis, apoptosis, long-term memory, longevity and many aspects of the innate immune response (Tables 1 and 2). These results are probably general for most insects, but there are also specialized functions for protease inhibitors incorporated into silk fibres and the blood-sucking insects require to inhibitors to block activation of their host’s blood-clotting cascades. In the rest of this review, we will summarize the information from genetic analysis and functional genomic studies in *Drosophila*, with additional work in other insects.

The *D. melanogaster* genome contains orthologues of 12 families: the serpin and macroglobulin, trapping inhibitors (MEROPS I4 and 39) as well as the Kazal, Kunitz and 8 other tight-binding inhibitor families (MEROPS I1, 2, 8, 17, 21, 25, 29, 32, 35 and 51), Tables 1 and 2. In addition, there is a single I3 inhibitory domain within a complex Kunitz inhibitor, Table 2. The list for *Aedes aegypti* is similar, with one tight-binding family not present in *D. melanogaster* (I43), but missing I25 and I29. *Culex pipiens* has 3 families not found in *D. melanogaster* (I19, 31 and 43), but lacks I25 and 29; while *Anopheles gambiae* has 2 families not present in *D. melanogaster* (I19 and 77), but lacks I25. *Pediculus humanus* has I32, 35 and 63 families, while lacking I8, 21 and 29. Several of these groups contain complex proteins with Kazal-like domains, but the I19 class (in *Culex* and *Anopheles*) is notable in that it has only been identified in insects, and the I31 family comprises tight-binding inhibitors of cysteine peptidases (C1). The major classes of serpin, Kazal and Kunitz inhibitors are similarly represented in all these insects, but there is a notable expansion of the macroglobulin class from 5 members in *D. melanogaster* and *A. aegypti* to 10 in *C. pipiens* and 39 in *A. gambiae*, presumably relating to the macrophage-based cellular immune response. In the tick, *Ixodes scapularis*, the Kazal and macroglobulin families are moderately increased in comparison with *D. melanogaster*, but there has been a massive increase in the number of inhibitory serpins [23].

3.1. The Drosophila trapping inhibitors

3.1.1. Serpins

The *D. melanogaster* genome contains 29 serpin genes [24] and 211 chymotrypsin-fold serine proteases [25], compared with 37

Table 1
The *Drosophila* Trapping inhibitors.

MEROPS Family I4 Serpins	FlyBase Identifier	AAs	P1/P1'	Biological function/comment	mRNA Regulation	Reference
Spn27A	CG11331	447	K/F	Toll/PO		[29,30,40]
Spn28B	CG6717	378	K/K			
Spn28Da	CG31902	384	L/S			
Spn28Db	CG33121	346	–			
Spn28Dc	CG7219	536	S/G	Melanisation	+ I.C	[31,40,77]
Spn28F	CG8137	375	Y/S			
Spn31A	CG4804	382	–			
Spn38F	CG9334	372	K/S	Immunity		
Spn42Da	CG9453	379	A/S R/AT/S V/A	Multiple, seetext	+ I.C	[40]
Spn42Db	CG9454	388	K/G			
Spn42Dc	CG9455	403	M/M			
Spn42Dd	CG9456	372	R/A	Immunity		[77]
Spn42De	CG9460	404	E/S			
Spn43Aa	CG12172	390	M/S	Morphogenesis?	+ eye	[81]
Spn43Ab	CG1865	393	–			
Spn43Ac	CG1857	476	L/S	Toll, immunity		[40,77]
Spn43Ad	CG1859	407	–			
Spn47C	CG7722	382	–			
Spn53F	CG10956	379	–			
Spn55B	CG10913	374	R/M			
Spn75F	CG32203	356	–			
Spn76A	CG3801	388	–	Immunity		[77]
Spn77Ba	CG6680	450	K/A	Melanisation		[32]
Spn77Bb	CG6663	362	–			
Spn77Bc	CG6289	416	–	Seminal fluid		
Spn85F	CG12807	640	?	SerXXpin		
Spn88Ea	CG18525	427	S/A	PO, wing		[40,77]
Spn88Eb	CG6687	426	S/S			[40,77]
Spn100A	CG1342	649	?	SerXXpin		

I39

Macroglobulins		AAs	Biological Function	mRNA	Reference
Tep1	CG18096	1354		+ I.C	[39]
Tep2	CG7052	1388	Phagocytosis *E. coli*	+ I.C	[30,38,39,77]
Tep3	CG7068	1469	Phagocytosis *S. aureus*	+ I.C	[38,39]
Tep4	CG10364	1496		+ I.C	[30,77]
Mcr	CG7586	1760	Phagocytosis *C.albicans*	+ I.C	[38,39]

and 176 in humans, and 64 and 227 in mice. Thirteen of the fly serpin genes, however, encode serpin folds that lack the consensus hinge region sequence typical of inhibitory serpins [24]. The *D. melanogaster Spn42Da* gene (also known as *Spn4*, *CG9453*), however, is exceptional in having alternatively-spliced exons, encoding RCL sequences and N-terminal localisation signals. This complex organisation allows 11 inhibitory variants to be generated that share the same serpin core, but have 4 different RCL peptides localised to different cellular compartments. On this basis, *D. melanogaster* has 28 active serpin inhibitors, 15 of which are putative extra-cellular proteins, Table 1.

The *D. melanogaster* serpins have lengths within the normal range of 350–450 amino acids, with the exceptions of Spn28Dc, Spn43Ac (necrotic), Spn85F and Spn100A. Given the normal invariance of the serpin fold, these exceptions are surprising. In the cases of Spn28Dc and necrotic, the increased lengths correspond to long N-terminal extensions, of about 93 and 77 amino acids, respectively. Such extensions are unusual within the serpin family, although a few cases of short extensions are known. Human heparin cofactor II, for example, has a short N-terminal extension, which interacts with the anion-binding exosite of thrombin. The serpin shows reduced substrate specificity for thrombin after N-terminal cleavage [26]. With necrotic, the N-terminal is cleaved on immune challenge, which increases substrate specificity for porcine pancreatic elastase *in vitro*. Spn28Dc also has an immune-related function, but whether its N-terminal can be cleaved is unknown. With Spn100A the increased length results from the insertion of an extra segment of 294 amino acids within the serpin structure to produce a novel serpin-related fold [27], while Spn85F has an extra inserted segment of about 245, in a similar position. The insertion of large peptide segments within a conserved protein fold is an extremely rare occurrence [28]. Presumably, Spn85F and Spn100A proteins derive from a single event, followed by gene duplication. The two genes are well diverged from each other in *D. melanogaster* and are present within all the *Drosophilid* species [24,27]. On the other hand, we find no orthologues of *Spn85F* or *Spn100A* in the mosquito genome assemblies, or *Bombyx*. The sequence in the putative hinge regions of both *Spn85F* and *Spn100A* has diverged from the *Drosophilid* inhibitory serpin consensus of (E [EKR]G[TSG][ET]AGS[YAGS][AGS][VAGS][TS]), but in both cases the region remains predominantly composed of short side-chain amino acids [27]. On this basis, Spn85F and Spn100A remain potential protease inhibitors.

Several of the *Drosophila* serpins have immune-related functions: Spn43Ac (Necrotic) [29] activates the Toll-related immune response in the adult [30], while Spn27A [31], Spn28Dc [32], Spn77Ba [33] and Spn88Ea [34] have roles in the phenol–oxidase (PO) cascade, or tracheal melanisation. In addition, Spn27A regulates the Toll-mediated morphogenesis of the dorso-ventral axis during embryogenesis [35,36] and Spn88Ea affects wing expansion after hatching from the pupal case [34]. Spn42Da (Spn4) inhibits furin, in the secretory pathway [37] as well as subtilase (S8) [38], chymotrypsin (S1) and papain-like cysteine proteases (C1) [38], see below.

3.1.2. Macroglobulins

The *Drosophila* genome encodes five macroglobulin orthologues, the thioester-containing proteins (TEP1–TEP4) and macroglobulin complement related (Mcr), Table 1. The TEPs bind

Table 2
The *Drosophila* Tight-binding Inhibitors.

MEROPS Family	Gene name	FlyBase identifier	AAs	Domains	Biological function/comment	mRNA regulation	Reference
I1							
Kazal		*CG31704*	68	S		+ I.C	[28]
		CG14933	77	S			
		CG42486	77	S			
		CG17278	80	S	Wnt receptor signalling	+ I.C	[30,78]
	Kaz1-B	*CG1220*	97	S			
	Sfp33A3	*CG42474*	99	S	Seminal fluid protein		
	Kaz1-A	*CG33790*	103	S			
		CG7695	145	S			
	E(spl)m1	*CG8342*	165	S			
		CG7906	354	C (2x)	Ecdysone induced		[30]
		CG7924	321	C (2x)	Ecdysone Induced		
		CG12716	418	S			
		CG7173	564	C			
		CG31758	605	Cx	Cyclic nucleotide		
	magu	*CG2264*	613	Cx	Longevity		[46]
		CG13830	629	Cx			
		CG32354	662	C (5x)			
	Oatp58a	*CG30277*	684	Cx	Org. anion transport		
	Oatp26F	*CG11332*	692	Cx	Org. anion transport		
	Oatp58b	*CG3382*	722	Cx	Org. anion transport		
	Anon-f	*CG1077*	730	Cx			
	Oatp33Ea	*CG5427*	745	Cx	Organic anion transport		
	Follistatin	*CG12956*	767	Cx + Ig	Dpp signalling		[47]
	Oatp58Dc	*CG3380*	789	Cx	Organic anion transport		
	Oatp30B	*CG3811*	791	Cx	Organic anion transport		
	Oatp33Eb	*CG6417*	814	Cx	Organic anion transport		
	Oatp74D	*CG7571*	819	Cx	Organic anion transport		
	Reck	*CG5392*	971	Cx + oat	Oncogene orthologue		
I2							
Kunitz		*CG42465*	73	S			
	Acp24A4	*CG31779*	78	S	Seminal fluid protein		
		CG16704	79	S			
		CG42466	80	S			
		CG16712	82	S		+ I.C	[41]
		CG16713	82	S		+ I.C	[40,41]
		CG42467	82	S			
		CG2816	84	S			
		CG42464	87	S			
		CG3513	88	S			
		CG42537	90	S	Moulting fluid	+ I.C	[77,78]
		CG42717	96	S			
		CG42716	97	S			
		CG15418	97	S			
	Sfp24Bc	*CG42602*	100	S	Seminal fluid protein		
		CG31778	102	S			
		CG31777	109	S			
	Sfp24Ba	*CG42463*	111	S	Seminal fluid protein		
		CG14298	111	S			
		CG13748	113	S			
		CG6784	116	S			
		CG17380	119	S			
		CG10031	129	S		+ I. C	[77]
		CG3604	132	S		+ I. C	[39]
		CG31515	148	S			
	Fatspondin	*CG6953*	763	Cx	Cytoskeleton	- I. C	[77]
		CG17739	837	Cx			
		CG30203	904	Cx			
		CG5639	1511	hC I2 + I3			
		CG31609	1589	Cx			
	axo	*CG18296*	2148	Cx	Neuronal		[79]
	Ppn	*CG33103*	2174	Cx	Extra-cellular matrix		
I8							
	Acp62F	*CG1262*	115	S	Male fertility		
		CG5267	178	C			
	Hml	*CG7002*	11768	Cx	Hemolectin		
I17							
	Kal-1	*CG6173*	525	Cx			
I21							
	7B2	*CG1168*	271	S	Neuroendocrine		
I25							
Cystatin		*CG8066*	104	S	Cysteine Protease Inhibitor		
		CG15369	122	S	Cysteine Protease Inhibitor		

(*continued on next page*)

Table 2 (*continued*)

MEROPS Family	Gene name	FlyBase identifier	AAs	Domains	Biological function/comment	mRNA regulation	Reference
		CG31313	124	S			
		CG8050	126	S			
I29							
	cer	*CG10460*	79	S	Long-term memory		
		CG1075	274	Cx + prt	Inhibitory + proteolytic		
		CG11459	339	Cx + prt	Inhibitory + proteolytic		
	Cp1	*CG6692*	341	Cx	Autophagic death		
		CG6347	352	C			
		CG6357	439	C?			
		CG12163	614	Cx + prt	Inhibitory + proteolytic		
I32							
	deterin	*12265*	153	S	Anti-apoptosis		
	Th	*12284*	438	C?	Anti-apopt., develop.		
	Iap2	*8293*	498	C?	Anti-apopt., immunity		
	Bruce	*6303*	4876	Cx			
I35							
	timp	*CG6281*	210	Cx	Extra-cellular matrix		
	143	*CG30188*	1140	C			
I51							
		CG7054	179	S			
		CG17919	202	S			
		CG30060	202	S			
	a5	*CG5430*	210	S	Phosphoethanolamine binding		
		CG17917	211	S	Phosphoethanolamine binding		
		CG18594	–				
		CG6180	257	S	Phosphoethanolamine binding		
	Mrpl3A	*CG15871*	416	C	Mitochondrial ribosome		

Each family of inhibitors is ordered by increasing length of the putative amino-acid coding sequence (AAs), so that short peptide inhibitors (S) tend to be separated from compound inhibitors (C), with multiple inhibitory domains, and Cx proteins, with a mix of inhibitory and other conserved domains. Transcripts that are up-reguulated following immune challenge are designated (+I.C), while those that are down-regulated are (−I.C). Data extracted from MEROPS and FlyBase, and checked against the NCBI conserved domain server (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

specifically to fungal and bacterial pathogens and stimulate phagocytosis by macrophages: Mcr binds *Candida albicans*, Tep2 promotes phagocytosis of *E. coli* and Tep3 that of *S. aureus* [39]. Consistent with this immune function, the levels of TEP transcripts are up-regulated following immune challenge [40]; in particular, TEP2 levels in the fat-body (FlyBase, Affymetrix Dros2 array data).

As with Spn42Dd, the *Tep2* transcription unit contains alternatively-spliced exons, allowing 5 separate bait regions to be generated from a single transcription unit. Given that macroglobulin bait regions tend to have multiple cleavage sites anyway [14], the range of pathogen virulence-factor proteases targeted by the five Tep2 protein isoforms is potentially very large. It may be that one of the 5 alternative bait regions targets several *E. coli* virulence factors and that sets of virulence factors from other pathogens are targeted by the other bait regions.

3.2. The Drosophila tight-binding inhibitors

The functions of the tight-binding inhibitors remain largely unexplored, but those that have been characterised appear to affect a similar range of responses as the serpin inhibitors.

3.2.1. Kazal

As in other organisms, the majority of *Drosophila* Kazal domains are found within 16 CxIs proteins, together with other conserved domains. In addition, there are three CIs, with two to five Kazal domains, and ten short peptide, SIs, Table 2. Two of the SI transcripts (CG31704 and CG17278) are up-regulated after bacterial and fungal infection [41,42] and CG17278 has additional functions in morphogenesis of the wing veins and the adhesion of dorsal and ventral wing surfaces [43]. The CIs, *CG7906* and *CG7924*, are expressed at the onset of metamorphosis, under control of the ecdysone hormone [44]; while the CxI, *magu* regulates longevity [45], *Follistatin* regulates Dpp and activin signalling [46] and *Reck* is orthologous to a mammalian oncogene.

Kaz1 is another example of a transcription unit that encodes multiple inhibitory functions [47]. In principle, the *Kaz1* transcription unit is di-cistronic, encoding two SIs, Kaz1-A (CG33790) and Kaz1-B (CG1220). However, the Kaz1-B gene contains alternatively-spliced exons, which give rise to six peptide isoforms targeted to different putative target proteases and cellular compartments.

3.2.2. Kunitz

The majority of *Drosophila* Kunitz domains are found in simple peptides, but one hCI combines a Kunitz and an I3 inhibitory domain and there are 6 CxIs, including *Ppn* (with 12 Kazal domains with 3 Ig domains and a WAP domain), which forms a structural component of the extra-cellular matrix. Two of the SI peptide genes are regulated at the transcriptional level following immune challenge, Table 2.

3.2.3. The I8, I17, I21, I25, I29, I32, I35 and I51 families

The remaining families of tight-binding inhibitors contain relatively few members.

I29 contains two genes with characterised phenotypes: a CxI affecting autophagic cell death and an SI affecting long-term memory. In addition three of the I29 family are CxIs containing both inhibitory and proteolytic domains, Table 2.

I32 contains three members that correspond to well-characterised anti-apoptotic genes, with additional roles in immunity and development. Despite this related set of functions, the domain structure is heterologous. The I32 family consists of two CxIs, one with an ubiquitin conjugating enzyme domain (CG6303) and the other with a RING finger domain (CG12284), a CI with three I32 domains (CG8293) and a short peptide SI (CG12265).

4. Tight-binding inhibitors in other insects and ticks

Two Kunitz inhibitors have been characterised from the haemolymph of *Manduca sexta* larvae that inhibit trypsin and plasmin

[48]. The SP1 (Kunitz) and SP2 (Kazal) inhibitors are incorporated into the silk fibres of the waxmoth [49] and a similar silk protein inhibitor is known from *Bombyx mori* [50]. SP1 and SP2 are broad range inhibitors that protect the cocoon silk proteins from degradation by bacterial and fungal proteases (subtilisin, protease K and pronase) [49]. Most insects do not produce silk, but the protein-rich salivary gland secretion, which glues the pupal case to the substrate might be expected to incorporate similar inhibitors. In addition, tight-binding inhibitors might well be incorporated into insect cuticle, where they could provide protection against pathogen penetration of the exoskeleton.

The saliva of the tick *I. scapularis* contains a tissue factor pathway inhibitor (TFPI), ixolaris, which binds both human factor X and factor Xa and specifically inhibits the FVIIa/TF complex (coagulation factor VIIa/membrane-bound tissue factor complex) [51]. The human TFPI is a plasma protease inhibitor which carries three tandem Kunitz domains, the first of which inhibits factor VIIa and the second factor Xa; while the third Kunitz domain not required for these activities [52]. The Tick ixolaris inhibitor contains of the two critical Kunitz domains, which bind to the FV11a/TF complex and FXa. The proposed mechanism of action is that the second ixolaris Kunitz domain binds FXa, after which the first Kunitz domain docks with the active site of the FV11a/TF complex and stabilises the formation of a tightly-bound quaternary complex [51]. Both the human and tick TFPI proteins demonstrate the potential for tandem tight-binding domains to form a scaffold for the assembly of protein complexes. An additional tick salivary CxI from *Dermacentor variabilis*, carries five Kunitz domains and reduces the transmission of *Rickettsia montanensis* from the saliva into the cells of its host [53]. It is unclear whether the potential chelating activity of the multiple Kunitz domains is critical for this activity, but reducing the transmission of bacterial infections to its host would be of benefit to the tick parasite.

5. Serpins in other insects

A comparative genomic study of the *B. mori* genome assembly identified 34 serpins, 11 of which are up-regulated after immune challenge [54]. Three of these genes encode secreted proteins between 450 and 519 amino acids in length, but none are long enough to accommodate the inserted segments of *Drosophila Spn85F* or *Spn100A* [54]. A similar study on *Apis mellifera* identified only 5 serpins, but there is a serpin-related fold of 612 amino acids (GB15070) that is orthologous to *Spn85F*, and a protein of 1976 amino acids (GB10078) that is orthologous to an additional serpin-related gene in *Drosophila* (CG14470), that has not been previously identified. Thirty-one serpins have been identified in *Tribolium castaneum*, including a cluster of 16 duplicated genes within a 50 kb segment [54]. A biochemical study of *Tenebrio molitor* larval haemolymph has identified three serpins which form specific complexes with Toll cascade protease [55]. Surprisingly, it appears that each successive step in the zymogen cascade is regulated by a different serpin.

5.1. Alternatively-spliced serpin transcripts and redundant genetic functions

The generation of multiple inhibitory activities from a single gene was first described in *B. mori* [56] and the genetic mechanism identified for the *Manduca sexta Spn1* gene, which has 12 alternative RCL exons [57]. Since then, similar complex serpin transcription units have been described in *Anopheles* [58] and *Caenorhabditis elegans* [59]. Similarly, alternatively-spliced exons for the bait regions of TEP2 and Kaz1-B generate a range of target specificities attached to the same core fold (see above).

Within the 12 sequenced *Drosophilid* genomes, the *Spn42Da* orthologues maintain their complex structure, but the number of alternative RCL-encoding exons is variable and the target proteases they inhibit differ, by the criterion that their putative P1/P1′ sites are not conserved [27,60]. This variation in the number of RCL exons reflects the mechanism by which new serpin/target protease interactions have arisen.

Given the crucial role of many serpin-regulated physiological responses, the inhibition of a novel target protease would generally require a duplication/divergence mechanism, leaving the original target protease regulated by one of the original serpin duplicates. This process seems to have driven the massive duplication and divergence of the vertebrate Clade A and B serpins [61–64]. There is little evidence of a similar process within the *Drosophilid* group in general, where the RCL sequences of inhibitory serpin orthologues are strongly conserved [27]. The alternatively-spliced RCL exons of the *Spn42Dd* orthologues represent a mechanism that allows flexibility on an evolutionary time-scale, within a compact genome.

If most serpin genes in the *Drosophilids* represent non-redundant genetic functions, then alteration in the target protease specificity of individual serpins would leave critical signalling pathways unregulated. In this case, lack of function (LOF) serpin mutations should be lethal, or give a visible mutant phenotype. Until recently, no mutant phenotypes had been associated with serpin genes in *Drosophila*, but in 1999 a LOF phenotype was described for the *necrotic* serpin (*nec*, *Spn43Ac*) and LOF phenotypes have since been associated with *Spn27A*, *Spn28Dc*, *Spn77Ba*, *Spn88Ea* mutants. To survey for additional serpin LOF phenotypes, we used RNAi knockdown strains from the Vienna and Tokyo stock centres. Such gene knockdowns may be incomplete, or give off-target effects, but using this approach we estimate that 12 of the *Drosophila* serpins give lethal or visible mutant phenotypes (Table 3), while 17 give no visible phenotype. A significant proportion of this later group may be seminal fluid components, which would not show a visible phenotype in adult flies. Ten serpins are expressed at high levels in the fat-body, as would be expected from haemolymph components; while eleven serpins are expressed at high levels in the male accessory glands as putative components of seminal fluid (FlyAtlas, http://130.209.54.32/atlas/atlas.cgi). By comparison, most of the variants associated with human serpinopathies, represent gain of function (GOF) mutations and LOF (deficiency) phenotypes have been identified for only 6 of the 37 serpins [65].

5.2. Serpin inhibitors in the innate immune response

In mammals, the immune response consists of an immediate “innate” response (mediated via antimicrobial peptides) and a delayed “acquired” response (mediated via antibodies). Insects lack an antibody-mediated response, but multiple cellular and humoral responses are activated on immune challenge. These responses include phagocytosis, coagulation, melanisation, activation of NF-κB transcription factors, synthesis of antimicrobial peptides, production of reactive oxygen species and the regulation of iron metabolism [66]. In *Drosophila*, the response to fungal and Gram-positive bacterial infections is mediated by the Toll receptor, while the phenol–oxidase cascade is associated with melanisation and the macrophage-mediated killing of bacteria and wasp parasitoids. The response to Gram-negative infection is mediated through a separate receptor activated by a phosphatase/kinase regulated cascade [67].

The innate immune response is strongly conserved and that of *Drosophila* has proved to be an extremely powerful model of the human response. In particular, the mammalian Toll-like receptors (TLRs) were identified by homology to *Drosophila* Toll and the

Table 3
RNAi knockdown phenotypes and tissue expression levels of *Drosophila* Serpin transcripts.

MEROPS I4 Serpins	FlyBase identifier	RNAi			RNAi strain	Transcript expression: fat-body	Male accessory gland
		18°	25°	29°			
Spn27A	CG11331	+	L	L	4804R-1	**330 ± 51**	
Spn28B	CG6717		NA		–	12 ± 2	**4190 ± 202**
Spn28Da	CG31902	+	+	+	109323	2 ± 2	**2871 ± 75**
Spn28Db	CG33121	+	+	+	12375	0 ± 0	**1078 ± 85**
Spn28Dc	CG7219	+	+	L	12377	**185 ± 50**	
Spn28F	CG8137	+	+	+	8137R-3	7 ± 5	**9520 ± 327**
Spn31A	CG4804	+	L	L	4804R-1	16 ± 4	
Spn38F	CG9334	L	L	L	9334R-3	2 ± 1	**10476 ± 296**
Spn42Da	CG9453	L	L	L	47262	**997 ± 100**	
Spn42Db	CG9454	+	+	+	24033	**13 ± 3**	
Spn42Dc	CG9455	+	+	+	13263	14 ± 4	
Spn42Dd	CG9456	+	+	+	37955	**556 ± 23**	
Spn42De	CG9460	+	+	+	9460R-3	10 ± 2	
Spn43Aa	CG12172	+	L	L	127172R-1	2 ± 0	
Spn43Ab	CG1865	+	+	+	45315	**4223 ± 413**	
Spn43Ac	CG1857	V	v	V	f::nec-3	**1884 ± 66**	
Spn43Ad	CG1859	+	+	+	20113	**87 ± 12**	
Spn47C	CG7722	+	+	+	25534	10 ± 3	
Spn53F	CG10956	+	+	+	37978	4 ± 2	**1305 ± 77**
Spn55B	CG10913	L	L	L	10913R-1	51 ± 22	
Spn75F	CG32203	+	+	+	34444	8 ± 3	**8625 ± 193**
Spn76A	CG3801	+	+	+	51599	5 ± 3	**11566 ± 348**
Spn77Ba	CG6680	L	L	L	43280	**510 ± 40**	
Spn77Bb	CG6663	+	+	+	15954	8 ± 3	**3375 ± 381**
Spn77Bc	CG6289	L	L	L	6289R-2	6 ± 5	**1699 ± 344**
Spn85F	CG12807	+	+	+	12807R-1	8 ± 3	
Spn88Ea	CG18525	+	+	+	29048	85 ± 13	
Spn88Eb	CG6687	+	V	V	6687R-1	**2228 ± 200**	
Spn100A	CG1342	+	+	V	1342R-1	2 ± 1	**20 ± 4**

The putative inhibitory serpins are marked in **bold**. RNAi knockdown used $RNAi^{UAS}$ strains from the Japanese NIG-Fly or Vienna stock collections, as identified, except for the $f.nec^{UAS}$ strain (FlyBase). These strains were crossed to the *Gal4-Act5C* driver strain and cultured at 18, 25 and 29°. (The *Gal4-UAS* system is inherently temperature sensitive, so that knockdown is more complete at higher temperatures.) + designates apparently wild-type $RNAi^{UAS}$ *Gal4-Act5C* fly., L designates lethality of $RNAi^{UAS}$ *Gal4-Act5C*, V designates a visible phenotype. Transcript expression data is Affymetrix Dros2 array data from FlyAtlas. Transcripts with high relative expression in the fat-body are marked in **bold** and are putative secreted components of the haemolymph. Transcript levels for those serpins that are predominantly expressed in the male accessory gland are shown in the right-hand column. These serpins are putative seminal fluid components.

intracellular components of the Toll-signalling pathway are conserved between flies and mammals. As in the fly, the mammalian "acute-phase" inflammatory response activates macrophages and production of broad range antimicrobial peptides, which allow host survival until specific antibodies can be generated. The insects, however, are completely dependent on their innate response and the activation of their Toll receptor is quite different to that of mammalian TLRs. In humans, TLRs bind pathogen associated molecular patterns (PAMPs) directly, while in *Drosophila* the activated Spatzle (Spz) ligand binds to the Toll receptor. Spz itself is activated by an extra-cellular, proteolytic cascade that is regulated by the necrotic serpin [30]. Necrotic is most similar to the human acute phase serpin α_1-antitrypsin, with a broad inhibitory spectrum that includes elastases [24]. It has recently become clear that this extra-cellular Toll-signalling pathway has three side-branches, all of which converge on Spatzle Processing Enzyme. One side-branch, the "danger signalling" pathway, is sensitive to pathogen-encoded virulence-factor proteases and is regulated by necrotic. The other two side-branches are activated by Gram-positive and fungal cell wall PAMP receptors [67] independently of necrotic. In addition, Spn27A, Spn28Dc, Spn77Ba and Spn88Ea regulate the immune-related PO cascade at different points (see above).

The increase in the number of chymotrypsin-fold proteases in the *Drosophila* compared to humans is the opposite of what is normally found for orthologous gene families between these two species and it may be that many of these extra proteases have immune-related functions. In the absence of an antibody-mediated response, a flexible sensing mechanism for the innate immune response is even more critical. A complex set of regulated protease cascades converging on a single Toll receptor would be difficult for a particular pathogen to evade and would allow rapid divergence and evolution of the upstream elements.

5.3. Serpin polymers, serpinopathies and conformational disease

One consequence of the unique suicide-cleavage inhibition mechanism of serpins is that they show an inherent tendency to form polymers [1]. Instead of the RCL loop of a serpin inserting within its own β-sheetA, the uncut RCL of one serpin can insert within β-sheetA of another, in the absence of enzyme cleavage, Fig. 1. A novel domain-swapping model for serpin polymerisation has recently been proposed, in which both the RCL and strand 5 of β-sheetA are exchanged between serpins [68]. This extremely elegant model is based on the crystal structure of antithrombin dimers, but it remains uncertain whether disease-associated sepin polymers are formed by this domain-swapping mechanism or the more classical single-strand insertion model. Whatever the details of the molecular mechanism, most wild-type serpins polymerise only slowly, so that the native monomer remains the prevalent form under physiological conditions. The serpin fold is, however, susceptible to mutations that alter conformational stability and facilitate polymer formation. Such polymerogenic variants are associated with many serpin-related human diseases. In particular, variants of α_1-antitrypsin, α_1-antichymotrypsin, antithrombin, C1-inhibitor and neuroserpin underlie particular cirrhosis, emphysema, thrombosis and angio-oedema pathologies and an early-onset inclusion-body dementia [1]. A common factor is that many

serpins are expressed at high levels in health and are rapidly degraded by their target proteases. Once polymerised, however, normal serpin clearance ceases. The secreting cells fill with excess inactive polymer and there is a serpin deficiency in extra-cellular fluids. In the case of Z-variant α_1-antitrypsin retained polymers damage the liver, leading to cirrhosis; while the lack of serpin activity in the lungs leads to emphysema. This serpinopathy is the major cause of neonatal liver failure in Caucasian populations. In surviving adults, the lack of circulating serpin leads to reduced inhibition of elastase in the lungs, so that the localised inflammatory response spreads and causes damage to host tissue.

Strikingly, the same amino-acid substitutions associated with human serpinopathies have been identified in mutant alleles of *necrotic* [69,70]. In particular, two *nec* mutations (*nec*^{9} and *nec*^{20}) carry the homologous Glu^{342}→Lys substitution as that responsible for human Z-variant α_1-antitrypsin disease. The fat-body of *Drosophila* is not impaired by accumulation of necrotic polymers, but lack of necrotic activity in the haemolymph leads to a cellular necrosis phenotype in epithelia and activation of the Toll pathway [30,71].

The conformational stability of the necrotic serpin is also affected by the *nec*^{1} mutation, which has a 2 amino-acid deletion in α-helix A. This deletion apparently destabilises β-sheetA of the serpin. The resulting *nec*^{1} protein is almost, but not completely, inactive *in vivo*. As with Z-allele antitrypsin monomers, the nec^1 and nec^9 mutant serpins lack the S→R transition (diagnostic of active inhibitors), when run on transverse urea gradient gels [69]. In addition, the *nec*^{7} and *nec*^{22} mutations carry a Gly→Ser sustitution homologous to an antithrombin variant that primarily gives dimers [69,70]. When the Ser→F substitution associated with the extreme Siiyama-variant antitrypsin disease is engineered into the nec serpin, the resultant inhibitor is inactive in transgenic flies, but produces a dominant, temperature-sensitive inactivation of the endogenous wild-type nec serpin.

These results establish that homologous amino-acid substitutions within human and fly serpins cause analogous defects in serpin function [69]. The fly is a particularly good model for investigating the affects of temperature on serpin stability and polymer formation under physiological conditions.

6. Protein clearance from insect haemolymph

The balance between the activities of extra-cellular proteases and their inhibitors is influenced by both their synthesis and degradation rates. The major site of synthesis of insect haemolymph proteins is the fat-body, but little work has been done on the protein degradation mechanism. In mammals, many plasma serpins are both secreted from the liver and taken up again and degraded in the liver by endocytosis. This is the route of degradation of both serpin/protease complexes and serpin polymers, which are recognised by trafficking receptors of the LDLR family [73,74]. Two recent publications have shown that haemolymph proteins are taken up in athrocytes [75,76]. These cells resemble those of the human reticulo-endothelial system, in that they form small clusters of endocytotically active cells, but athrocytes also share some features of the mammalian kidney glomurula cells [77]. The necrotic serpin is taken up in garland and pericardial athrocytes by the LpR1 trafficking receptor and targeted for lysosomal degradation [75,76]. Similarly, the peritracheal athrocytes of *M. sexta* take up and degrade, both a *Manduca* storage protein and a mammalian protein, IgG. Strikingly, mutations in *LpR1* modify the Toll-mediated immune response, indicating that the scavenging of serpin/proteinase complexes might be a critical step in the regulation of proteolytic cascades [75,76].

7. Protease inhibitors and regulatory cascades in insects

Our understanding of the regulation of proteases and the activity of proteolytic signalling cascades in the insects is incomplete. Part of the problem is that there are large numbers of both proteases and inhibitors and the small size of insects makes direct biochemical approaches difficult. Despite this there have been significant biochemical studies on some of the larger insect species [78,79]. Comparative genomic approaches in *Drosophila* and mosquitoes have identified a large number of inhibitors and proteases, but while the full complement of serpins has probably been identified in *D. melanogaster*, the numbers of orthologues identified in the other 11 *Drosophilid* genome assemblies is reduced to about 71% of the expected number [27] and the number of serpins identified in the mosquito species is lower still. Part of the problem is that although the serpin-fold is strongly conserved, the constraints on the primary sequence are relatively loose [12]. As a result, identifying orthologues between distantly-related species is difficult. A large proportion of serpin-folds are non-inhibitory, but the conserved hinge-region motif can be used to identify putative inhibitors. With the tight-binding inhibitory domains, the sequence of known core folds can be identified by homology, but any novel core folds will be missed. Furthermore, core-fold homology, in itself, does not distinguish between active inhibitors and putative transport proteins [4].

Similarly, with the target proteases themselves, sequence alignment shows that 27% (57/211) of the *Drosophila* chymotrypsin folds lack one, or more, of the amino acids that form the catalytic triad and are therefore not active proteases [25,80]. Such serine protease homologues (SPH) could form reversible complexes with protease inhibitors and sequester them from interactions with active proteases. Alternatively, many of the *Drosophila* chymotrypsin folds are associated with CLIP domains, which are characteristic of insect immune-related proteases and may be implicated in protease/zymogen interactions. Thus, the haemolymph serine protease homologues, SPH-1 and SPH-2, of *Manduca* bind to a C-type lectin and are necessary for activation of the PO cascade serine proteases, PAP-1 and PAP-2, by bacterial lipopolysaccaride [72,81] and SPH-1 of *Tenebrio* is necessary for production of melanin on the surface of bacterial pathogens [82]. All these SPHs carry CLIP domains and may be acting as scaffolds for the assembly of components of signalling cascades. The *Drosophila* SPH masquerade is strongly up-regulated on immune challenge [41] as well as having roles in morphogenesis of neurones and muscles [83].

Proteolytic cascades control a similar range of physiological responses in insects as in mammals and these responses are regulated by a variety of trapping and tight-binding inhibitors. The serpins play a critical role in many signalling cascades, but their target-specificity is inherently broad. Similarly, the TEP/macroglobulins have extremely broad target-specificity, which suits them for roles as hormone transport proteins and to form complexes with a wide range of pathogenic virulence factors, which then activate macrophages. The tight-binding inhibitors, on the other hand, are generally competitive inhibitors of their target proteases, with much higher specificity than the serpins or macroglobulins. The presence of multiple tight-binding domains, within compound and complex inhibitor proteins, allows the potential for these proteins to act as scaffolds for the assembly of signalling complexes. A number of proteolytic cascades are known to regulated by both trapping- and tight-binding inhibitors and this may be a general rule. Such dual regulation could combine the rapidity of response of the suicide-inhibitors with the specificity of lock-and-key inhibitors. This would allow precise control of proteolytic cascades under physiological conditions.

Acknowledgements

This work was supported by a Spanish Ministry of Education and Science grant and a Basque Government Department of Education, Universities and Research grant to D. Gubb. Additional support was provided from The Department of Industry, Tourism and Trade of the Government of the Autonomous Community of the Basque Country (Etortek Research Programs) and from the Innovation Technology Department of Bizkaia County. L. Troxler was supported by the French CNRS. Fig. 1 was the kind gift of Dr. T. Dafforn.

References

[1] D.A. Lomas, D. Belorgey, M. Mallya, E. Miranda, K.J. Kinghorn, L.K. Sharp, R.L. Phillips, R. Page, A.S. Robertson, D.C. Crowther, Molecular mousetraps and the serpinopathies, Biochem. Soc. Trans. 33 (2005) 321–330.
[2] D.A. Lomas, New insights into the structural basis of alpha 1-antitrypsin deficiency, Qjm 89 (1996) 807–812.
[3] K.M. Providence, P.J. Higgins, PAI-1 expression is required for epithelial cell migration in two distinct phases of in vitro wound repair, J. Cell Physiol. 200 (2004) 297–308.
[4] N.D. Rawlings, D.P. Tolle, A.J. Barrett, Evolutionary families of peptidase inhibitors, Biochem. J. 378 (2004) 705–716.
[5] G. Xu, R.L. Rich, C. Steegborn, T. Min, Y. Huang, D.G. Myszka, H. Wu, Mutational analyses of the p35-caspase interaction. A bowstring kinetic model of caspase inhibition by p35, J. Biol. Chem. 278 (2003) 5455–5461.
[6] A.J. Barrett, Alpha 2-macroglobulin, Methods Enzymol. 80 (Pt C) (1981) 737–754.
[7] L. Sottrup-Jensen, Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation, J. Biol. Chem. 264 (1989) 11539–11542.
[8] N.D. Rawlings, F.R. Morton, C.Y. Kok, J. Kong, A.J. Barrett, MEROPS: the peptidase database, Nucleic Acids Res. 36 (2008) D320–D325.
[9] L.A. Kazal, D.S. Spicer, R.A. Brahinsky, Isolation of a crystalline trypsin inhibitor–anticoagulant protein from pancreas, J. Am. Chem. Soc. 70 (1948) 3034–3040.
[10] P.G. Gettins, Serpin structure, mechanism, and function, Chem. Rev. 102 (2002) 4751–4804.
[11] J.A. Huntington, R.J. Read, R.W. Carrell, Structure of a serpin–protease complex shows inhibition by deformation, Nature 407 (2000) 923–926.
[12] J.A. Irving, R.N. Pike, A.M. Lesk, J.C. Whisstock, Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function, Genome Res. 10 (2000) 1845–1864.
[13] G.A. Silverman, P.I. Bird, R.W. Carrell, F.C. Church, P.B. Coughlin, P.G. Gettins, J.A. Irving, D.A. Lomas, C.J. Luke, R.W. Moyer, P.A. Pemberton, E. Remold-O'Donnell, G.S. Salvesen, J. Travis, J.C. Whisstock, The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature, J. Biol. Chem. 276 (2001) 33293–33296.
[14] L. Sottrup-Jensen, O. Sand, L. Kristensen, G.H. Fey, The alpha-macroglobulin bait region. Sequence diversity and localization of cleavage sites for proteinases in five mammalian alpha-macroglobulins, J. Biol. Chem. 264 (1989) 15781–15789.
[15] P.G. Gettins, Thiol ester cleavage-dependent conformational change in human alpha 2-macroglobulin. Influence of attacking nucleophile and of Cys949 modification, Biochemistry 34 (1995) 12233–12240.
[16] P.G. Gettins, B. Crews, A.H. Beth, K. Hideg, Bait region-thiol ester mapping in human alpha 2-macroglobulin, FEBS Lett. 367 (1995) 137–140.
[17] P.C. Harpel, M.F. Lewin, A.P. Kaplan, Distribution of plasma kallikrein between C-1 inactivator and alpha 2-macroglobulin in plasma utilizing a new assay for alpha 2-macroglobulin–kallikrein complexes, J. Biol. Chem. 260 (1985) 4257–4263.
[18] M. Kunitz, Crystallization of a trypsin inhibitor from soybean, Science 101 (1945) 668–669.
[19] J. Sri Ram, L. Terminiello, M. Bier, F.F. Nord, On the mechanism of enzyme action.LVII. Interaction between trypsin and ovomucoid, Arch. Biochem. Biophys. 52 (1954) 451–463.
[20] R.M. Roberts, N. Mathialagan, J.Y. Duffy, G.W. Smith, Regulation and regulatory role of proteinase inhibitors, Crit. Rev. Eukaryot. Gene Expr. 5 (1995) 385–436.
[21] M. Laskowski, M.A. Qasim, What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? Biochim. Biophys. Acta 1477 (2000) 324–337.
[22] R.J. Beynon, J.S. Bond, Proteolytic Enzymes. IRL at Oxford University press, Oxford, 1998.
[23] A. Mulenga, R. Khumthong, K.C. Chalaire, Ixodes scapularis tick serine proteinase inhibitor (serpin) gene family; annotation and transcriptional analysis, BMC Genomics 10 (2009) 217.
[24] J.M. Reichhart, Tip of another iceberg: *Drosophila serpins*, Trends Cell Biol. 15 (2005) 659–665.
[25] J. Ross, H. Jiang, M.R. Kanost, Y. Wang, Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships, Gene 304 (2003) 117–131.
[26] T.P. Baglin, R.W. Carrell, F.C. Church, C.T. Esmon, J.A. Huntington, Crystal structures of native and thrombin-complexed heparin cofactor II reveal a multistep allosteric mechanism, Proc. Natl. Acad. Sci. U S A 99 (2002) 11079–11084.
[27] M. Garrett, A. Fullaondo, L. Troxler, G. Micklem, D. Gubb, Identification and analysis of serpin-family genes by homology and synteny across the 12 sequenced Drosophilid genomes, BMC Genomics 10 (2009) 489.
[28] V. Alva, K.K. Koretke, M. Coles, A.N. Lupas, Cradle-loop barrels and the concept of metafolds in protein classification by natural descent, Curr. Opin. Struct. Biol. 18 (2008) 358–365.
[29] E. De Gregorio, S.J. Han, W.J. Lee, M.J. Baek, T. Osaki, S. Kawabata, B.L. Lee, S. Iwanaga, B. Lemaitre, P.T. Brey, An immune-responsive serpin regulates the melanization cascade in *Drosophila*, Dev. Cell 3 (2002) 581–592.
[30] E.A. Levashina, E. Langley, C. Green, D. Gubb, M. Ashburner, J.A. Hoffmann, J.M. Reichhart, Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*, Science 285 (1999) 1917–1919.
[31] P. Ligoxygakis, N. Pelte, C. Ji, V. Leclerc, B. Duvic, M. Belvin, H. Jiang, J.A. Hoffmann, J.M. Reichhart, A serpin mutant links Toll activation to melanization in the host defence of *Drosophila*, EMBO J. 21 (2002) 6330–6337.
[32] C. Scherfer, H. Tang, Z. Kambris, N. Lhocine, C. Hashimoto, B. Lemaitre, *Drosophila* serpin-28D regulates hemolymph phenoloxidase activity and adult pigmentation, Dev. Biol. (2008).
[33] H. Tang, Z. Kambris, B. Lemaitre, C. Hashimoto, A serpin that regulates immune melanization in the respiratory system of *Drosophila*, Dev. Cell 15 (2008) 617–626.
[34] S.T. Ahmad, S.T. Sweeney, J.A. Lee, N.T. Sweeney, F.B. Gao, Genetic screen identifies serpin5 as a regulator of the toll pathway and CHMP2B toxicity associated with frontotemporal dementia, Proc. Natl. Acad. Sci. U S A 106 (2009) 12168–12173.
[35] P. Ligoxygakis, S. Roth, J.M. Reichhart, A serpin regulates dorsal-ventral axis formation in the *Drosophila* embryo, Curr. Biol. 13 (2003) 2097–2102.
[36] C. Hashimoto, D.R. Kim, L.A. Weiss, J.W. Miller, D. Morisato, Spatial regulation of developmental signaling by a serpin, Dev. Cell 5 (2003) 945–950.
[37] M.J. Richer, C.A. Keays, J. Waterhouse, J. Minhas, C. Hashimoto, F. Jean, The Spn4 gene of *Drosophila* encodes a potent furin-directed secretory pathway serpin, Proc. Natl. Acad. Sci. U S A 101 (2004) 10560–10565.
[38] M. Bruning, M. Lummer, C. Bentele, M.M. Smolenaars, K.W. Rodenburg, H. Ragg, The Spn4 gene from *Drosophila melanogaster* is a multipurpose defence tool directed against proteases from three different peptidase families, Biochem. J. 401 (2007) 325–331.
[39] S.L. Stroschein-Stevenson, E. Foley, P.H. O'Farrell, A.D. Johnson, Identification of Drosophila gene products required for phagocytosis of *Candida albicans*, PLoS Biol. 4 (2006) e4.
[40] M. Lagueux, E. Perrodou, E.A. Levashina, M. Capovilla, J.A. Hoffmann, Constitutive expression of a complement-like protein in toll and JAK gain-of-function mutants of *Drosophila*, Proc. Natl. Acad. Sci. U S A 97 (2000) 11427–11432.
[41] E. De Gregorio, P.T. Spellman, G.M. Rubin, B. Lemaitre, Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays, Proc. Natl. Acad. Sci. U S A 98 (2001) 12590–12595.
[42] N. Buchon, N.A. Broderick, M. Poidevin, S. Pradervand, B. Lemaitre, Drosophila intestinal response to bacterial infection: activation of host defense and stem cell proliferation, Cell Host Microbe 5 (2009) 200–211.
[43] T.L. Jacobsen, D. Cain, L. Paul, S. Justiniano, A. Alli, J.S. Mullins, C.P. Wang, J.P. Butchar, A. Simcox, Functional analysis of genes differentially expressed in the *Drosophila* wing disc: role of transcripts enriched in the wing region, Genetics 174 (2006) 1973–1982.
[44] R.B. Beckstead, G. Lam, C.S. Thummel, Specific transcriptional responses to juvenile hormone and ecdysone in *Drosophila*, Insect Biochem. Mol. Biol. 37 (2007) 570–578.
[45] Y. Li, J. Tower, Adult-specific over-expression of the *Drosophila* genes magu and hebe increases life span and modulates late-age female fecundity, Mol. Genet. Genomics 281 (2009) 147–162.
[46] J. Pentek, L. Parker, A. Wu, K. Arora, Follistatin preferentially antagonizes activin rather than BMP signaling in *Drosophila*, Genesis 47 (2009) 261–273.
[47] T. Niimi, H. Yokoyama, A. Goto, K. Beck, Y. Kitagawa, A *Drosophila* gene encoding multiple splice variants of Kazal-type serine protease inhibitor-like proteins with potential destinations of mitochondria, cytosol and the secretory pathway, Eur. J. Biochem. 266 (1999) 282–292.
[48] N. Ramesh, M. Sugumaran, J.E. Mole, Purification and characterization of two trypsin inhibitors from the hemolymph of *Manduca sexta* larvae, J. Biol. Chem. 263 (1988) 11523–11527.
[49] X. Nirmala, D. Kodrik, M. Zurovec, F. Sehnal, Insect silk contains both a Kunitz-type and a unique Kazal-type proteinase inhibitor, Eur. J. Biochem. 268 (2001) 2064–2073.
[50] A. Kurioka, M. Yamazaki, H. Hirano, Primary structure and possible functions of a trypsin inhibitor of *Bombyx mori*, Eur. J. Biochem. 259 (1999) 120–126.
[51] I.M. Francischetti, J.G. Valenzuela, J.F. Andersen, T.N. Mather, J.M. Ribeiro, Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, *Ixodes scapularis*: identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex, Blood 99 (2002) 3602–3612.
[52] T. Hamamoto, M. Yamamoto, O. Nordfang, J.G. Petersen, D.C. Foster, W. Kisiel, Inhibitory properties of full-length and truncated recombinant tissue factor pathway inhibitor (TFPI). Evidence that the third Kunitz-type domain of TFPI

is not essential for the inhibition of factor VIIa-tissue factor complexes on cell surfaces, J. Biol. Chem. 268 (1993) 8704–8710.
[53] S.M. Ceraul, S.M. Dreher-Lesnick, A. Mulenga, M.S. Rahman, A.F. Azad, Functional characterization and novel rickettsiostatic effects of a Kunitz-type serine protease inhibitor from the tick *Dermacentor variabilis*, Infect. Immun. 76 (2008) 5429–5435.
[54] Z. Zou, Z. Picheng, H. Weng, K. Mita, H. Jiang, A comparative analysis of serpin genes in the silkworm genome, Genomics 93 (2009) 367–375.
[55] R. Jiang, E.H. Kim, J.H. Gong, H.M. Kwon, C.H. Kim, K.H. Ryu, J.W. Park, K. Kurokawa, J. Zhang, D. Gubb, B.L. Lee, Three pairs of protease-serpin complexes cooperatively regulate the insect innate immune responses, J. Biol. Chem. 284 (2009) 35652–35658.
[56] T. Sasaki, Patchwork-structure serpins from silkworm (*Bombyx mori*) larval hemolymph, Eur. J. Biochem. 202 (1991) 255–261.
[57] H. Jiang, M.R. Kanost, Characterization and functional analysis of 12 naturally occurring reactive site variants of serpin-1 from *Manduca sexta*, J. Biol. Chem. 272 (1997) 1082–1087.
[58] A. Danielli, F.C. Kafatos, T.G. Loukeris, Cloning and characterization of four *Anopheles gambiae* serpin isoforms, differentially induced in the midgut by *Plasmodium berghei* invasion, J. Biol. Chem. 278 (2003) 4184–4193.
[59] O. Kruger, J. Ladewig, K. Koster, H. Ragg, Widespread occurrence of serpin genes with multiple reactive centre-containing exon cassettes in insects and nematodes, Gene 293 (2002) 97–105.
[60] S. Borner, H. Ragg, Functional diversification of a protease inhibitor gene in the genus *Drosophila* and its molecular basis, Gene 415 (2008) 23–31.
[61] D.J. Askew, Y.S. Askew, Y. Kato, R.F. Turner, K. Dewar, J. Lehoczky, G.A. Silverman, Comparative genomic analysis of the clade B serpin cluster at human chromosome 18q21: amplification within the mouse squamous cell carcinoma antigen gene locus, Genomics 84 (2004) 176–184.
[62] D.J. Askew, P. Coughlin, P.I. Bird, Mouse serpins and transgenic studies. in: G.A. Silverman, D.A. Lomas (Eds.), Molecular and Cellular Aspects of the Serpinopathies and Disorders in Serpin Activity (2007) Singapore.
[63] D. Kaiserman, P.I. Bird, Analysis of vertebrate genomes suggests a new model for clade B serpin evolution, BMC Genomics 6 (2005) 167.
[64] J.D. Inglis, R.E. Hill, The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain, EMBO J. 10 (1991) 255–261.
[65] R.H. Law, Q. Zhang, S. McGowan, A.M. Buckle, G.A. Silverman, W. Wong, C.J. Rosado, C.G. Langendorf, R.N. Pike, P.I. Bird, J.C. Whisstock, An overview of the serpin superfamily, Genome Biol. 7 (2006) 216.
[66] J.A. Hoffmann, J.M. Reichhart, Drosophila innate immunity: an evolutionary perspective, Nat. Immunol. 3 (2002) 121–126.
[67] M. Gottar, V. Gobert, A.A. Matskevich, J.M. Reichhart, C. Wang, T.M. Butt, M. Belvin, J.A. Hoffmann, D. Ferrandon, Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors, Cell 127 (2006) 1425–1437.
[68] M. Yamasaki, W. Li, D.J. Johnson, J.A. Huntington, Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization, Nature 455 (2008) 1255–1258.
[69] C. Green, G. Brown, T.R. Dafforn, J.M. Reichhart, T. Morley, D.A. Lomas, D. Gubb, Drosophila necrotic mutations mirror disease-associated variants of human serpins, Development 130 (2003) 1473–1478.
[70] R. Carrell, J. Corral, What can *Drosophila* tell us about serpins, thrombosis and dementia? Bioessays 26 (2004) 1–5.
[71] C. Green, E. Levashina, C. McKimmie, T. Dafforn, J.M. Reichhart, D. Gubb, The necrotic gene in *Drosophila* corresponds to one of a cluster of three serpin transcripts mapping at 43A1.2, Genetics 156 (2000) 1117–1127.
[72] X.Q. Yu, H. Jiang, Y. Wang, M.R. Kanost, Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, Manduca sexta, Insect Biochem. Mol. Biol. 33 (2003) 197–208.
[73] M.Z. Kounnas, F.C. Church, W.S. Argraves, D.K. Strickland, Cellular internalization and degradation of antithrombin III–thrombin, heparin cofactor II-thrombin, and alpha 1-antitrypsin–trypsin complexes is mediated by the low density lipoprotein receptor-related protein, J. Biol. Chem. 271 (1996) 6523–6529.
[74] A. Kasza, H.H. Petersen, C.W. Heegaard, K. Oka, A. Christensen, A. Dubin, L. Chan, P.A. Andreasen, Specificity of serine proteinase/serpin complex binding to very-low-density lipoprotein receptor and alpha2-macroglobulin receptor/low-density-lipoprotein-receptor-related protein, Eur. J. Biochem. 248 (1997) 270–281.
[75] S.F. Soukup, J. Culi, D. Gubb, Uptake of the necrotic serpin in *Drosophila melanogaster* via the lipophorin receptor-1, PLoS Genet. 5 (2009) e1000532.
[76] C. Owa, F. Aoki, M. Nagata, Protein degradation in silkworm peritracheal athrocytes and its physiological role in metamorphosis, J. Insect Physiol. (2010).
[77] H. Weavers, S. Prieto-Sanchez, F. Grawe, A. Garcia-Lopez, R. Artero, M. Wilsch-Brauninger, M. Ruiz-Gomez, H. Skaer, B. Denholm, The insect nephrocyte is a podocyte-like cell with a filtration slit diaphragm, Nature 457 (2009) 322–326.
[78] N. Jiang, S. Thangamani, C.F. Chor, S.Y. Wang, I. Winarsih, R.J. Du, J. Sivaraman, B. Ho, J.L. Ding, A novel serine protease inhibitor acts as an immunomodulatory switch while maintaining homeostasis, J. Innate Immun. 1 (2009) 465–479.
[79] H. Jiang, A.B. Mulnix, M.R. Kanost, Expression and characterization of recombinant *Manduca sexta* serpin-1B and site-directed mutants that change its inhibitory selectivity, Insect Biochem. Mol. Biol. 25 (1995) 1093–1100.
[80] N.D. Rawlings, F.R. Morton, The MEROPS batch BLAST: a tool to detect peptidases and their non-peptidase homologues in a genome, Biochimie 90 (2008) 243–259.
[81] H. Jiang, Y. Wang, X.Q. Yu, M.R. Kanost, Prophenoloxidase-activating proteinase-2 from hemolymph of *Manduca sexta*. A bacteria-inducible serine proteinase containing two clip domains, J. Biol. Chem. 278 (2003) 3552–3561.
[82] H. Kan, C.H. Kim, H.M. Kwon, J.W. Park, K.B. Roh, H. Lee, B.J. Park, R. Zhang, J. Zhang, K. Soderhall, N.C. Ha, B.L. Lee, Molecular control of phenoloxidase-induced melanin synthesis in an insect, J. Biol. Chem. 283 (2008) 25316–25323.
[83] B. Murugasu-Oei, V. Rodrigues, X. Yang, W. Chia, Masquerade: a novel secreted serine protease-like molecule is required for somatic muscle attachment in the *Drosophila* embryo, Genes Dev. 9 (1995) 139–154.