Trypsin and Chymotrypsin Inhibitors in Insects and Gut Leeches

Liliane Schoofs¹, Elke Clynen¹ and Michel Salzet²

1. Laboratory of Developmental Physiology and Molecular Biology, K.U.Leuven, Naamsestraat 59, B-3000 Leuven, Belgium

2. Laboratoire de Neuroimmunologie des Annélides, ESA CNRS 8017, SN3, IFR 17 INSERM, Université des Sciences et Technologies de Lille, 59655 Villeneuve d’Ascq Cédex, France

Abstract: The constant increase of life expectancy is associated with major ageing of developed populations. This indicates that the new century will have one of most epidemic progressions of cardiovascular, cancer and inflammatory diseases. The high challenge for medical research is to compress such morbidity. In these conditions, invertebrates have demonstrated to be truly useful models in drug discovery for such ageing diseases. The last decade, drug discovery in leeches has opened the gate of new molecules to treat emphysema, coagulation, inflammation, dermitis and cancer. Also other invertebrates such as insects, which evolved from the annelids, harvest potential interesting molecules, such as serine protease inhibitors that can be exploited by the medical industry.

In all metazoan species, proteases play a prominent role in a wide array of physiological processes such as food digestion, blood clotting, embryogenesis, tissue reorganization (e.g. wound healing, regeneration, melting, metamorphosis etc.), defense mechanisms and immune responses. Many of these processes are proteolytic cascades, which, once set in action, lead very rapidly and irreversibly to a specific cellular response. Activation and inactivation of protease cascades have to be closely controlled at different regulatory levels being protease gene transcription, mRNA translation, zymogen activation, substrate specificity, enzyme kinetics and by means of enzyme-inhibitors. Most animal species synthesize a variety of protease inhibitors with different specificities, whose function is to prevent unwanted proteolysis. It follows that, and evidence for this is accumulating,-proteases are involved in various disease states. For instance, the destruction of the extracellular matrix of articular cartilage and bone in arthritic joints is thought to be mediated by excessive proteolytic activity [1]. In emphysema, gingivitis, tumour invasion and inflammatory infections, it is suggested that tissue destruction is caused by proteases [1]. Among the enzymes involved in extracellular matrix degradation, a few serine proteases (elastase, collagenase, cathepsin G) are able to solubilize fibrous proteins such as elastin and collagen [2,3]. Given the specific recognition by proteases of defined amino acid sequences, it may be possible to inhibit these enzymes when they are involved in pathological processes. Potent inhibitors have the potential to be developed as new therapeutic agents. In vertebrates, serine protease inhibitors have been studied for many years and they are known to be involved in phagocytosis, coagulation, complement activation, fibrinolysis, blood pressure regulation, etc. In the last decade, it became obvious that in invertebrates, serine proteases and their inhibitors are also involved in parallel physiological processes (e.g. blood clotting cascade in Limulus [4] and the innate immune response [5]). Moreover, some of the protease inhibitors isolated from invertebrate sources are quite specific towards individual mammalian serine proteases. This also offers huge opportunities for medicine. Thus, the development of non-toxic protease inhibitors extracted from invertebrates for in vivo application may be quite important [1]. In the future, it is likely that numerous specific protease inhibitors will be tested clinically for the treatment of human disease such like emphysema, inflammation, dermitis and cancer.

SERINE PROTEASE INHIBITORS

The amino acid sequences of the currently identified serine protease inhibitors differ significantly, with the number of constituent amino acids ranging from 29 to about 400. Despite this, only two fundamentally different inhibiting mechanisms seem to exist. Most inhibitors bind with their cognate enzyme(s) according to a common, substrate-like standard mechanism. They are all relatively small (from 29 to 190 amino acids) and share an exposed, rigid binding loop with a very characteristic ‘canonical’ conformation [6]. These are the Kunitz-type inhibitors. The serpins on the other hand constitute a family of large (glyco)proteins (typically about 400 amino acid residues in length) which function as suicide substrate inhibitors. Upon binding, they are cleaved by the target protease within a reactive center loop region of about 20 amino acids near the C-terminus. The amino acid N-terminal to the scissile bond (the P1 position) appears to be important for determining the specificity of serpins for particular proteases. Further N-terminally within the reactive center loop region are the residues of the hinge region that are highly conserved among...
inhibitory serpins [7]. Serpins are involved in diverse biological processes [8]. Many serpins found in human plasma regulate proteolytic reactions important in blood coagulation, fibrinolysis, the immune response, and inflammation. A serpin called maspin identified in mammalian breast tissue has been implicated in tumor suppression [9].

**SERINE PROTEASE INHIBITORS IN INSECTS**

Insect haemolymph, like vertebrate serum, contains several serine protease inhibitors [10]. Most of the insect serine protease inhibitors examined hitherto were identified or partially characterized from haemolymph extracts and can be grouped into two families based on their amino acid sequence and their protease inhibition characteristics: low molecular mass proteins (below 10 kDa) related to the Kunitz-type inhibitors [6] and proteins of about 45 kDa which belong to the serpin superfamily [7]. The genome of the fruit fly, *Drosophila melanogaster*, contains many serpin genes [11]. The first serpin from this organism to be reported was the Acp76A protein, a component of the male accessory gland that is transferred to the female during mating [12]. Recently, six new *Drosophila* serpins were identified that according to their sequence are likely to function as serine protease inhibitors involved in dorsoventral patterning [13]. These serine proteases are perhaps acting sequentially in a cascade like the mammalian blood clotting proteases. The authors provide the first biochemical evidence that at least one of them is a potent inhibitor of trypsin-like proteases *in vitro*. One *Drosophila* serpin displays sequence similarities with human neuroserpin [14], whereas others appear to be most similar in sequence to members of the ov-serpin sub-family, which includes inhibitory serpins such as human plasminogen activator inhibitor-2 [15].

Although details on the function of the insect serine protease inhibitors are as yet not available, evidence is accumulating that they play a role in insect anti-microbial defence mechanisms, digestion, metamorphosis and development [16]. Recently, two new families of low molecular weight serine protease inhibitors have been discovered. The first, designated as the *Bombyx* family was discovered in silkworms [17]. The second is designated as the locust serine protease inhibitor peptide family (Fig. 1). So far, the family consists of low molecular weight peptides of around 4 kDa with three inhibitors in *Locusta migratoria* designated as LMC1 1 and 2 (or *Locusta migratoria* chymotrypsin inhibitor 1 and 2), and HI [18-22] and five in *Schistocerca gregaria* designated as SGPIs or *Schistocerca gregaria* protease inhibitors 1-5 [16]. The locust peptide inhibitors display sequence similarities with nine cysteine-rich domains (PLDs) in the light chain of pacifastin, a heterodimeric serine protease inhibitor isolated from the haemolymph of the crayfish *Pacifastacus leniusculus* [23]. The PLD’s and the 4 kDa locust inhibitors share a conserved pattern of six cysteine residues (Cys-Xaa-12-Cys-Asn-Xaa-Cys-Xaa-3-Gly-Xaa3-4-Cys-Thr-Xaa3-Cys), which form three disulfide bridges (Cys1-Cys4, Cys2-Cys6, Cys3-Cys5) (Fig. 1). Below, we review the latest data on the characterisation of these low molecular weight serine protease inhibitors.

**Structure Analysis**

The studies of Mer et al. [24-27] contributed significantly to the elucidation of the specific structural characteristics of the locust low molecular weight inhibitors. The core region adopts a compact, globular fold, which consists of three strands (β1, β2 and β3) arranged in an anti-parallel β-sheet, that demarcates a cavity and an amino-terminal segment, orientated almost perpendicular to the β-sheet. Inside the cavity, hydrophobic residues are clustered with an aromatic ring in the center of the hydrophobic core. The protease binding loop, located between two cysteine residues in the carboxy-terminal segment exhibits an

**Fig. (1).** Sequence alignment of the locust 4 kDa serine protease inhibitors. The conserved (cysteine) residues are indicated. A consensus sequence displaying the disulfide bridges is below. The P1 residue is indicated by an arrow. Fucosylated Thr-residues are in italic.
extended conformation and is anchored to the core region by two cysteine bridges [22,25]. These structural characteristics are commonly encountered in small ‘canonical’ inhibitors. They all have a conserved (‘canonical’) and extended binding loop, which permits an anti-parallel β-strand interaction with the protease active site. This interaction is stabilized by intramolecular bonds (i.e. disulfide bonds) between the binding loop and the core of the inhibitor. The locust peptides are stabilized by three disulfide bridges forming a ‘cysteine knot’ [9, 10]. They adopt a tertiary fold [14] hitherto unobserved in the large group of small “canonical” proteinase inhibitors (Kunitz-type). Therefore, these inhibitors constitute a new subfamily within the large group of small ‘canonical’ inhibitors [16, 25]

Four of the eight locust inhibitors (LMCI-2, SGPI 2,4,5) are post-translationally modified by a deoxyhexose moiety on the threonine residue on the fifth position after the first cysteine residue [16, 19]. As the non-fucosylated peptides retain their full activity [22,28], it remains obscure why some inhibitors of this newly identified family contain this sugar while others lack it. Inspired by other conformational studies on glycosylated proteins, Mer et al. proposed a stabilizing effect of the fucose moiety [27]

**Specificity of the Locust Inhibitors**

The locust enzyme inhibitors display specificity differences towards locust endogenous enzymes [16] but also towards mammalian serine proteases as shown in table 1. These differences can be readily attributed to the amino acid sequence within the active site. Indeed, a difference of one or two amino acid residues around the reactive sites often results in considerable alteration of the inhibitory specificity [9]. The P1 residue is mainly responsible for the inhibitor’s specificity for a particular protease. Five of the 4 kDa inhibitors (LMCI-2, SGPI 2,3,4 and 5) have a Leu as P1 residue and are potent or moderately potent inhibitors of α-chymotrypsin [16,20,22]. The other inhibitors (HI, LMCI-1 and SGPI-1) with an Arg at the P1 position are potent trypsin inhibitors and display no or very weak chymotrypsin inhibiting activity [16,22]. Structure activity studies by Kellenberger et al [22] provided strong evidence for the close relation between the P1 residue and the inhibitor’s specificity for a particular enzyme. They demonstrated that the substitution of an Arg-P1 residue for a Leu residue resulted in an enormous decrease in the Ki value (by a factor 2000) for chymotrypsin, converting both trypsin inhibitors into potent chymotrypsin inhibitors. Vice versa, the replacement of the P1-Leu by Arg in SGPI-2 converted this potent chymotrypsin inhibitor (Ki = 6.2 nM) into a moderate (Ki = 51 nM) trypsin inhibitor, whereas the activity towards chymotrypsin (Ki = 5.5 nM) decreased a 1000fold.

Structure-activity studies with analogues of SGPI-2 demonstrated that a Met residue instead of a Lys at the P’1 position increases the affinity of the inhibitor for trypsin [28]. The rather weak trypsin inhibition (Ki = 210 nM) of the naturally occurring SGPI-1, in contrast to SGPI-5, can be attributed to the favourable P’1 interaction (positively charged Lys) of SGPI-1 (CTR-KGC) with the S’1 residue (Lys) of trypsin.

**Precursor Genes**

In *L. migratoria*, LMCI 1 and 2 are derived from a single precursor polypeptide [29]. Similarly, SGPI-1 and SGPI-2 are derived from a single polypeptide precursor in *S. gregaria* [30], whereas SGPI-3, 4 and 5 are encoded by three additional precursors ([30], Simonet et al., unpublished data), which contain putative additional protease inhibitors displaying the same cysteine pattern. The transcripts are present in several locust tissues, but not in the midgut. Important changes in transcript levels occur during development. Vanden Broeck et al also provided evidence indicating that the expression of the inhibitors is hormonally regulated [30].

**Tissue Distribution of the Locust Inhibitors**

LMCI-1, LMCI-2 and HI were isolated from *Locusta* haemolymph and brain tissue [19,20,22]. The five *Schistocerca* inhibitors were isolated from ovarian tissue [16]. However, the spatial and temporal distribution data as analysed by HPLC indicated the presence of each inhibitor in gonads, haemolymph and fat body of adults and in larval

---

**Table 1. Comparison of Active Site Sequences of the Insect 4 kDa Serine Protease Inhibitors**

<table>
<thead>
<tr>
<th></th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1’</th>
<th>P2’</th>
<th>P3’</th>
<th>Inhibitors of</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGPI-1</td>
<td>A</td>
<td>C</td>
<td>T</td>
<td>R</td>
<td>K</td>
<td>G</td>
<td>C</td>
<td>Trypsin</td>
</tr>
<tr>
<td>SGPI-2</td>
<td>A</td>
<td>C</td>
<td>T</td>
<td>L</td>
<td>K</td>
<td>A</td>
<td>C</td>
<td>Chymotrypsin, Elastase</td>
</tr>
<tr>
<td>SGPI-3</td>
<td>I</td>
<td>C</td>
<td>T</td>
<td>L</td>
<td>K</td>
<td>Y</td>
<td>C</td>
<td>Chymotrypsin</td>
</tr>
<tr>
<td>SGPI-4</td>
<td>V</td>
<td>C</td>
<td>T</td>
<td>L</td>
<td>K</td>
<td>A</td>
<td>C</td>
<td>Chymotrypsin</td>
</tr>
<tr>
<td>SGPI-5</td>
<td>S</td>
<td>C</td>
<td>T</td>
<td>L</td>
<td>M</td>
<td>A</td>
<td>C</td>
<td>Chymotrypsin</td>
</tr>
<tr>
<td>HI</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>R</td>
<td>K</td>
<td>A</td>
<td>C</td>
<td>Trypsin/Chymotrypsin (weak)</td>
</tr>
<tr>
<td>LMCI-1</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>R</td>
<td>K</td>
<td>G</td>
<td>C</td>
<td>Trypsin</td>
</tr>
<tr>
<td>LMCI-2</td>
<td>A</td>
<td>C</td>
<td>T</td>
<td>L</td>
<td>K</td>
<td>A</td>
<td>C</td>
<td>Chymotrypsin, Elastase</td>
</tr>
</tbody>
</table>
hemolymph and fat body [16]. The midgut was devoid of this type of 4 kDa inhibitors. Our recent data indicate that the locust midgut contains its own chymotrypsin/elastase inhibitors, which according to elution characteristics, differ from the 4 kDa locust enzyme inhibitors (Vercammen et al., unpublished data).

LMCI-1 has been localized in neurosecretory cells of the brain (pars intercerebralis) as shown by immunocytochemistry [31]. Peptidomic analysis (liquid chromatography on line with mass spectrometry) unequivocally proved that LMCI-1 and 2 as well as SGPI 1 and 2 are synthesized in the brain, more specifically in the pars intercerebralis-corpus cardiacum (CC) axis (Fig. 2), which is the equivalent of the mammalian hypothalamus-pituitary system [32]. These data suggest that the inhibitors might be involved in the regulation of neuropeptide precursor processing by serine proteases (convertases). Mass spectrometric analysis also indicated the release of both LMCIs from the CC in vitro. Therefore, it is likely that LMCI-1 and LMCI-2 are released into the haemolymph via a neuro-endocrine pathway.

Besides the typical 4 kDa inhibitors, another 14 kDa serine protease inhibitor was isolated from Schistocerca ovaries by a combination of trypsin-affinity chromatography and reverse-phase high performance liquid chromatography [33]. The N-terminal sequence (Y) XAEXDELA (A) EEY (Y)Q(Q)X(I)(L)M (X being a Cys, an irregular or modified amino acid) revealed no similarities with any other protease inhibitors isolated from invertebrate or vertebrate source. The 14 kDa inhibitor was found to be heat-stable. The purified molecule, which could be extracted in water but not in acidic methanol, potently inhibited bovine trypsin and chymotrypsin, but not pancreatic elastase. The cDNA for this 14kDa inhibitor has not been cloned as yet. Once these types of insect serine protease genes can be introduced in specific expression systems, large amounts of these molecules can be produced.

**Biological Function**

The 4 kDa inhibitors are assumed to play a role in the innate immune defense system [18,20]. They are potent inhibitors of the proteolytic cascade activating prophenoloxidase (proPO), which is present as a proenzyme in the hemolymph. Upon activation, it leads to the formation of antifungal quinones and the local production of melanine around invading parasites.

Of particular interest is that LMCI-1 and LMCI-2 have also been shown to act as inhibitors of high voltage-activated calcium channels on mammalian cells [31, 34].

As the locust inhibitors are expressed in several tissues (haemolymph, brain, fat body and gonads) we may assume that they probably are involved in many physiological processes where proteolytic cascades have to be fine-tuned.

**SERINE PROTEASE INHIBITORS IN LEECHES**

In hematophageous leeches, studies of serine protease inhibitors and their substrates have been extensively performed. Two groups of serine protease inhibitors can be distinguished. The first is related to specific inhibitors that interfere in the activation of the blood clotting system and
Table 2. Comparison of Active Site Sequences of Leech Protease Inhibitors

<table>
<thead>
<tr>
<th></th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1'</th>
<th>P2'</th>
<th>P3'</th>
<th>Inhibitors of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antistasin I</td>
<td>V</td>
<td>R</td>
<td>C</td>
<td>R</td>
<td>I</td>
<td>H</td>
<td>C</td>
<td>Factor Xa, Trypsin</td>
</tr>
<tr>
<td>Antistasin II</td>
<td>I</td>
<td>N</td>
<td>C</td>
<td>R</td>
<td>K</td>
<td>T</td>
<td>C</td>
<td>Factor Xa, Trypsin</td>
</tr>
<tr>
<td>Ghilanten</td>
<td>V</td>
<td>R</td>
<td>C</td>
<td>R</td>
<td>V</td>
<td>Y</td>
<td>C</td>
<td>Factor Xa, Trypsin</td>
</tr>
<tr>
<td>Guamerin I</td>
<td>I</td>
<td>R</td>
<td>C</td>
<td>M</td>
<td>I</td>
<td>F</td>
<td>C</td>
<td>Elastase</td>
</tr>
<tr>
<td>Guamerin II</td>
<td>I</td>
<td>R</td>
<td>C</td>
<td>M</td>
<td>I</td>
<td>F</td>
<td>C</td>
<td>Chymotrypsin</td>
</tr>
<tr>
<td>Hirustasin</td>
<td>V</td>
<td>H</td>
<td>C</td>
<td>R</td>
<td>I</td>
<td>R</td>
<td>C</td>
<td>Capthepsin G, Trypsin</td>
</tr>
<tr>
<td>Tessulin</td>
<td>C</td>
<td>L</td>
<td>C</td>
<td>K</td>
<td>E</td>
<td>P</td>
<td>C</td>
<td>Chymotrypsin, Trypsin</td>
</tr>
<tr>
<td>Therin</td>
<td>Y</td>
<td>L</td>
<td>C</td>
<td>K</td>
<td>M</td>
<td>A</td>
<td>C</td>
<td>Trypsin</td>
</tr>
<tr>
<td>Therostasin</td>
<td>A</td>
<td>Q</td>
<td>C</td>
<td>R</td>
<td>I</td>
<td>Y</td>
<td>C</td>
<td>Factor Xa</td>
</tr>
</tbody>
</table>

the second group related to inhibitors that work on the extracellular matrix [35].

In gut leeches (Fig. 3), Cytin [36] is the first chymotrypsin inhibitor having two chains isolated from leeches. The B chain exhibits 16% sequence identity with eglin from *H. medicinalis* [37]. No sequence homology was found with gelin from *H. manillensis* [38] or with the other chymotrypsin inhibitors isolated from jawed Hirudinae (Fig. 4) *i.e.* bdellins [39]. Cytin, as does eglin [37], possesses sequence identity with the substilisin/chymotrypsin inhibitor family isolated from barley seeds (CI, [40]). Alignment of N-terminal sequences (first 22 residues) of the B chain of Cytin with CI-2a revealed 73% residue identities (Table 3) and 52% to 67% residue identities with CI-1 (a, b, c) and CI-2b, respectively. At the level of the A chain, no sequence homology was found with other chymotrypsin inhibitors. Thus, this structural feature- 2 chains linked by a disulfide bridge- is novel among chymotrypsin inhibitors. Additionally, this structure appears to be essential for the inhibitory activity of Cytin. In fact, after reduction and S-pyridylethylation, the two chains became separated and the peptide lost its activity [36]. Furthermore, this unique structure makes Cytin different even from the potato chymotrypsin inhibitors (PI-1, 41), of which eglin and CI-2 are members [42], as well as from other known serine proteases inhibitors. Moreover, the oxidation of Cytin provoked a loss of activity of around 85%, reflecting the fact that the Met residue in the chain A is important for the molecule’s activity.

Chymotrypsin inhibitors isolated from jawed leeches (*H. medicinalis* and *H. manillensis*), *i.e.*, gelin and eglin [37, 38] display no sequence homology with Cytin. This result is in contrast to thrombin inhibitors from these animals where there is a 70% structural homology [35]. Therefore, we can conclude that eglin and gelin did not originate from the same ancestral gene as Cytin. Besides this chymotrypsin

---

Fig. (3). Photographs of *Theromyzon tessulatum* gut.
inhibitor, a specific small peptide trypsin inhibitor (therin, [42]) as well as trypsin-chymotrypsin inhibitors ([43], tessulin) were isolated from the same species. Therin exhibits ca. 30% of sequence identity and spacing of the 8 cysteines residues with molecules of the antistasin-type inhibitors family [44]. The sequence similarity is much higher with the first domain of antistasin (33%) [44]. Indeed, if therin is aligned with these potentially homologous inhibitors on the basis of its cysteines, the positions of the cysteines and the putative P1 active site match best with all the antistasin-type protease inhibitors. The P1 residue of the reactive site of the inhibitor generally determines its specificity [46]. In antistasin, only the N-terminal domain is inhibitory and the P1 residue has been determined as Arg34 [47]. In the therin sequence, the P1 residue could be Lys28. Therin inhibits bovine trypsin with high affinity and specificity (Ki value of 45 ± 12 pM). This value is much higher than those obtained with other potent trypsin-chymotrypsin inhibitors found in leeches (isolated from respectively *H. medicinalis* (hirustasin; bellastasin [48, 49]), *Hirudinaria nipponia* (guamerin and piguamerin [50, 51]) and the non blood sucker leech *whitmania edentula* (guamerin II, [52]) (Table 2).

Tessulin [43] is a 9-kDa peptide of 81 amino acids. It possesses 16 cysteines and displays 16% sequence similarity with antistasin-type inhibitors (Fig. 5). It inhibits trypsin (Ki 1 pM) and chymotrypsin (Ki 150 pM) and exhibits no activity with thrombin, factor Xa, cathepsin G and elastase. Tessulin, like Cytin [36] and Therin [41] are the only trypsin-chymotrypsin inhibitors isolated from leeches that do not inhibit elastase or cathepsin G. Furthermore, Tessulin, in conjunction with other serine-protease inhibitors isolated from *Theromyzon i.e.* Therin, Cytin, Therostasin and Theromin [42, 43, 45-47], significantly diminishes the level of human granulocyte and monocyte activation induced by lipopolysaccharides (10 µg). The combined level of inhibition is higher than that of aprotinin, another serine-protease inhibitor used biomedically. Sequence comparisons were carried out for Tessulin with the five different protease inhibitors isolated from the leech *T. tessulatum*: Therin, Therostasin, Cytin, antitrypsin A, B and Tessulin [40]. This revealed that, for three of the five peptides, being Therostasin, Theromin and Tessulin, display a high degree of sequence similarities (>80%), except for the amino acid residues surrounding the putative active site. Therefore, they probably constitute a new protease inhibitor family (Fig. 6).

**Table 3. Sequence Comparison of Cytin with Other Chymotrypsin Inhibitors Isolated from Barley Seeds (Cl-) or from Potatoes (PI-). Amino Acid Residues Underlined are Identical in Cytin**

<table>
<thead>
<tr>
<th>Molecules</th>
<th>Sequence</th>
<th>Residue Identities (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytin</td>
<td>LKCEWPELVTRGEEAKETIER</td>
<td></td>
</tr>
<tr>
<td>CI-2a</td>
<td>LKTEWPELVGSMVEAKKVER</td>
<td>73</td>
</tr>
<tr>
<td>CI-1c</td>
<td>KTEDEVVQGMSAEKARTIER</td>
<td>52</td>
</tr>
<tr>
<td>CI-1a</td>
<td>KTEDEVVQGMSAEKARTIER</td>
<td>62</td>
</tr>
<tr>
<td>CI-1b</td>
<td>KTEDEVVQGMSAEKARTIER</td>
<td>67</td>
</tr>
<tr>
<td>CI-2b</td>
<td>KTEDEVVQGMSAEKARTIER</td>
<td>67</td>
</tr>
<tr>
<td>PI-1a</td>
<td>KERWPELGPARKAMQII</td>
<td>47</td>
</tr>
</tbody>
</table>

**Fig. (4). Photographs of Hirudo medicinalis jaws.**
Fig. (5). SDS-PAGE analysis of tessulin.

Fig. (6). Sequence alignment between Theromin, Therostasin and Tessulin obtained using Multalin software: http://www.toulouse.inra.fr/centre/serWWW.htm
Fig. (7). SDS-PAGE analysis of anti-trypsin A (1) and anti-trypsin B (2).

Apart from this family of small peptides (< 10 kDa), we also demonstrated the presence of two specific trypsin-chymotrypsin inhibitors with a molecular mass ranging between 14 and 15 kDa, from head parts of the rhynchobdellid leech *Theromyzon tessulatum* (Fig. 7, [46]). Two proteins, anti-trypsin-chymotrypsin A : ATCA (14636.6 ± 131 Da) and anti-trypsin-chymotrypsin B : ATCB (14368 ± 95 Da) were purified by size exclusion and anion-exchange chromatography followed by reversed-phase HPLC. Based on amino-acid composition, N-terminal sequence determination (MELCELQGSRSNPQPSNM), matrix assisted laser desorption-time of flight measurement (MALDI-TOF), trypsin mapping comparison, inhibition constant determination (K_i), and influence on amidolytic activity of different serine proteases, we demonstrate that ATCA and ATCB are novel and highly potent serine-protease inhibitors of trypsin and chymotrypsin (ATCA : 350 fM towards trypsin and chymotrypsin; ATCB : 400 and 75 fM towards trypsin and chymotrypsin, respectively). We further surmise that ATCA and ATCB are linked, in that ATCB would lead to the formation of ATCA after loss of few amino acid residues.

Taken together, these data reflect that invertebrates contain several highly specific serine protease inhibitors that could be used as lead compounds for the development of powerful drugs.

ACKNOWLEDGEMENTS

This work was supported in part by the Centre National de la Recherche Scientifique, and the MNERT. L. Schoofs acknowledges funding by the FWO and the Research Council of the K.U.Leuven (G.0187.00, GOA/2000/04).

REFERENCES

Trypsin and Chymotrypsin Inhibitors


RELATED ARTICLES RECENTLY PUBLISHED IN CURRENT PHARMACEUTICAL DESIGN