Molecular diversity of anticoagulants from haematophagous animals

Cho Yeow Koh1; R. Manjunatha Kini1,2

1Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore; 2Department of Biochemistry, Virginia Commonwealth University Medical Center, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia, USA

Summary
To obtain blood meals, haematophagous animals are armed with potent pharmacological molecules to overcome several of their hosts’ response systems. Among them, a large number of exogenous anticoagulants which target the haemostatic system have been identified and characterised. Studies on these anticoagulants have expanded our knowledge on the blood coagulation system, and provided a valuable source of antithrombotic therapeutic agents. Advances in genomic, transcriptomic, structural and proteomic tools greatly accelerated the discovery and analysis of the exogenous anticoagulants in recent years. The molecular diversity observed in these molecules is huge and is constantly expanding. In this review, we will provide an overview on the structure, function and mechanism of the exogenous anticoagulants from haematophagous animals and rationalise their molecular diversity.

Keywords
Anticoagulants, haematophagous animals, structure-function relationships, exogenous factors, molecular diversity

Haemostasis and anticoagulants
The circulation of blood is essential for our survival. Vascular injury triggers formation of a platelet plug, thrombin and fibrin clot through blood coagulation to arrest the bleeding (1–3). In this tightly regulated system, any imbalance could lead to haemorrhagic disorders or thrombosis. Thrombosis in particular causes high morbidity and mortality. Globally, with changing food habits and lifestyles, atherosclerosis and thromboembolic disorders are becoming increasingly common (4–6). Antithrombotic drugs, including anticoagulants and antiplatelets, are used to prevent and treat thrombosis. Anticoagulants are effective for initial and long-term management of both arterial and venous thrombosis (7) but antiplatelets are less efficacious in the prevention of venous thromboembolism (8, 9).

Heparin (10) and vitamin K antagonists (such as warfarin) (11) are the cornerstones of anticoagulation therapy. Unfortunately, both classes of drugs have well-documented limitations, which drives the continual and intense efforts to develop new, efficacious and safe anticoagulants, especially those targeting specific coagulation factors (6, 12). Some of the leaders in the market are direct thrombin inhibitors, such as hirudin (13), bivalirudin (14), argatroban (15) and dabigatran (16) as well as the factor (F)Xa inhibitor rivaroxaban (17). Despite the emergence of these new anticoagulants, the complicated nature and clinical settings of thrombosis (e.g. arterial vs. venous thrombosis, acute vs. long-term management, thrombosis in pregnant, nursing, renal-impaired or cancer patients) continue to call for more new and safe anticoagulants with different pharmacological and pharmacokinetic properties. Thus, the search for new lead compounds for the development of anticoagulants is still very relevant (9).

Haematophagous animals
Extensive research is focused on isolating and characterising highly specific anticoagulants from blood-feeding (haematophagous) animals. The success of recombinant hirudin (initially isolated from the medicinal leech), and to a greater extent, bivalirudin (designed based on hirudin), demonstrates the utility of these natural products in drug design. Haematophagous animals consist mainly of arthropods in the orders of Ixodida (Ixodidae – hard ticks; Argasidae – soft ticks), Diptera (Culicidae – mosquitoes; Ceratopogonidae – biting midges; Tabanidae – horseflies; Glossinidae – tsetse flies; Simuliidae – blackflies; Phlebotominae – sandflies), Hemiptera (Triatominae – kissing bugs),
<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>Mechanism</th>
<th>Example</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hirudin-like inhibitors</td>
<td>Single domain of ~ 7 kDa, N-terminal globular core stabilised by three disulfide bridges, C-terminal long, extended tail</td>
<td>Fast, tight-binding, competitive inhibition, N-terminal inhibits active site noncanonically, C-terminal binds to exosite-I</td>
<td>Hirudin</td>
<td>Hirudo medicinalis</td>
<td>(23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Similar to hirudin but C-terminal binds to exosite-II</td>
<td>BmGTI</td>
<td>Boophilus microplus</td>
<td>(113)</td>
</tr>
<tr>
<td>Kunitz-type proteinase inhibitors</td>
<td>Single Kunitz domain is ~ 7 kDa, Two tandem Kunitz domains, Soft ticks inhibitors: distorted reactive-site loop, lack of basic P1 residue, Hard ticks inhibitors: typical reactive-site loop, with basic P1 residue</td>
<td>Slow, tight-binding, competitive inhibition, N-terminal Kunitz domain inhibits active site noncanonically, C-terminal Kunitz domain binds to exosite-I</td>
<td>Ornithodorin Savignin, Monobin, Amblin, Boophilin, Hemalin</td>
<td>Ornithodoros moubata, Ornithodoros savignyi, Argas monolakensis, Amblyomma hebraeum, Boophilus microplus, Haemaphysalis longicornis</td>
<td>(48), (49), (50), (45), (46), (47)</td>
</tr>
<tr>
<td>Kazal-type proteinase inhibitors</td>
<td>Single Kazal domain is ~ 6 kDa, Multiple non-classical Kazal domains in tandem, Typically isolated as two tandem domains proteins but cDNA sequence showed multiple-domains precursors</td>
<td>Slow, tight-binding, competitive inhibition, N-terminal Kazal domain inhibits active site canonically, C-terminal Kazal domain binds to exosite-I</td>
<td>Rhodniin Dipetalogastin Infestin-I–2</td>
<td>Rhodnius proluxus Dipetalogaster maximus Triatoma infestans</td>
<td>(52), (53), (54), (55)</td>
</tr>
<tr>
<td>Lipocalin family</td>
<td>Single domain of ~ 16 kDa, Eight-stranded (A-B-C-D-E-F-G-H) with [beta]-barrel and a central ligand-binding pocket, Directional inversion in B and C strands compared to typical lipocalin topology</td>
<td>Binds to exosite-I only</td>
<td>Triabin</td>
<td>Triatoma pallidipennis</td>
<td>(59), (60)</td>
</tr>
<tr>
<td>Anophelin &amp; thrombostasin</td>
<td>Single domain of ~ 7 to 9 kDa, No cysteines, Acidic segment in the middle of the molecule</td>
<td>Slow, tight-binding, competitive inhibition, Inhibits active site and exosite-I, Lack of kinetic information, Inhibits active site, Exosites binding not determined</td>
<td>Anophelin</td>
<td>Anopheles albimanus</td>
<td>(61)</td>
</tr>
<tr>
<td>Madanin &amp; chimadanin</td>
<td>Single domain of ~ 7 kDa, No cysteines, Containing a 11- residues acidic segment in the middle</td>
<td>Binds to exosite-I only, Inhibits active site, Exosites binding not determined</td>
<td>Madanin 1 &amp; 2</td>
<td>Haemaphysalis longicornis</td>
<td>(64)</td>
</tr>
<tr>
<td>Antistasin-like inhibitor</td>
<td>Cysteine-rich domain of ~ 7 kDa, Containing a 26-residues segment with conserved cysteines (6) and disulfide linkages, Domain typically repeated in tandem</td>
<td>Inhibits active site, Exosites binding not determined</td>
<td>Theromin</td>
<td>Theromyzon tessulatum</td>
<td>(68)</td>
</tr>
<tr>
<td>Granulin-like inhibitor</td>
<td>Single granulin domain is ~ 6 kDa, Conserved cysteines (12) and disulfide linkages, Domain typically repeated in tandem</td>
<td>Inhibits active site, Exosites binding not determined</td>
<td>Leech granulin</td>
<td>Hirudo nipponia</td>
<td>(117)</td>
</tr>
<tr>
<td>TTI</td>
<td>Short peptide of ~ 4 kDa, No cysteines, Inhibits active site, Exosites binding not determined</td>
<td>TTI, Glossina morisitans morisitans</td>
<td>TTI</td>
<td>(118)</td>
<td></td>
</tr>
<tr>
<td>NTI-I</td>
<td>Short peptide of ~ 3 kDa, Non-competitively inhibits active site</td>
<td>NTI-I, Hyalomma damedorii</td>
<td>NTI-I</td>
<td>(119)</td>
<td></td>
</tr>
<tr>
<td>Microphilins</td>
<td>Short peptide of ~ 1.7 kDa, No cysteines, Inhibits active site, Exosites binding not determined</td>
<td>Microphilin (2 isoforms), Boophilus microplus</td>
<td>Microphilin</td>
<td>Boophilus microplus</td>
<td>(111)</td>
</tr>
<tr>
<td>BmAP</td>
<td>High molecular weight of 60 kDa, Inhibits active site, Binds to at least one exosite</td>
<td>BmAP, Boophilus microplus</td>
<td>BmAP</td>
<td>Boophilus microplus</td>
<td>(112)</td>
</tr>
<tr>
<td>BmGTI</td>
<td>26 kDa, Binds to exosite-I only, Also enhances APC activity</td>
<td>BmGTI, Boophilus microplus</td>
<td>BmGTI</td>
<td>Boophilus microplus</td>
<td>(113)</td>
</tr>
</tbody>
</table>
Phthiraptera (Anoplura – sucking lice) and Siphonaptera (fleas), as well as some annelids in the subclass of Hirudinea (leeches), parasitic nematodes such as hookworms, and even mammals (vampire bats). Physiologically, the duration (e.g. seconds in mosquitoes to months in hookworms), behavior (obligatory or facultative) and mechanisms (e.g. pool-feeding/telmophages in ticks or capillary-feeding/solenophages in mosquitoes) of their blood feeding habits differ. However, they all face the common physical, mechanical and chemical defenses of their hosts, including the skin and vessel walls, as well as the haemostatic, inflammatory and immunological responses. In order to obtain the enormous amount of blood required (relative to their body weight), it is essential for haematophagous animals to overcome these barriers with potent pharmacological agents that are capable of attenuating those physiological responses of their hosts (18, 19). These agents include vasodilators, anticoagulants, anti-platelets, immunosuppressors and anti-inflammatory compounds (18–22).

### Exogenous anticoagulants from haematophagous animals

Over the years, a large number of exogenous anticoagulants from haematophagous animals have been identified, although not all of them have been characterised in detail (4, 20, 21). These anticoagulants target blood coagulation proteinases to prevent clot formation during the ingestion and digestion of blood meals. Unlike physiological inhibitors of blood coagulation proteinases, which mainly comprise two groups (serpin and Kunitz), enormous molecular diversity can be observed in the exogenous anticoagulants from haematophagous animals. Based on the mechanism of action, these exogenous anticoagulants from haematophagous animals can be broadly classified as: a) thrombin inhibitors; b) FXa inhibitors; c) extrinsic tenase complex (ETC) inhibitors; d) intrinsic tenase complex (ITC) inhibitors; and e) contact system protein inhibitors. In this review, we provide an overview on the structure, function and mechanism of exogenous anticoagulants from haematophagous animals and rationalise their molecular diversity.

#### Thrombin inhibitors

**Hirudin**

The most well-known example of thrombin inhibitor, hirudin, was isolated more than 50 years ago from the periphrangial glands of the medicinal leech *Hirudo medicinalis* (Table 1). Hirudin is a 65-residue protein (~7 kDa) which specifically inhibits thrombin (23). Many hirudin isoforms with minor variations in the primary structures were subsequently reported (23, 24). This family of inhibitors was also isolated from other species of leeches (25, 26) (Table 1). Hirudin binds thrombin with a Kᵢ value of 22 fM. The residue Tyr64 of hirudin is sulfated and desulfated hirudin binds to thrombin 10 times weaker, with a Kᵢ of 207 fM (27). Hirudin becomes a slow-binding inhibitor at high ionic strength solutions (0.2 and above), highlighting the importance of electrostatic interactions in the complex formation (27, 28).

Three-dimensional structures of hirudin were determined using nuclear magnetic resonance (NMR) spectroscopy (29, 30), and its structures complexed with thrombin were determined using X-ray crystallography (31–35). The N-terminal domain of hirudin (residues 1–48) folds into a globular unit stabilised by three disulfide bridges, and the C-terminal domain (residues 49–65) assumes a long and extended conformation (Fig. 1A). The first three residues on hirudin N-terminus bind to a hydro-

<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>Mechanism</th>
<th>Example</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variegin</td>
<td>• Short peptide of 3 to 4 kDa&lt;br&gt;• No cysteines&lt;br&gt;• Flexible conformation&lt;br&gt;• C-terminal sequence highly identical to that of hirudin</td>
<td>• Fast, tight-binding, competitive inhibitor&lt;br&gt;• N-terminus drives fast binding&lt;br&gt;• Middle segment inhibits active site canonically&lt;br&gt;• C-terminus binds exosite-I</td>
<td>Variegin (multiple isoforms)</td>
<td>Amblyomma variegatum</td>
<td>(69)</td>
</tr>
<tr>
<td>Others</td>
<td>• Lack of detailed structural information</td>
<td>• Slow, tight-binding, competitive inhibitor&lt;br&gt;• Lack of functional characterisation</td>
<td>Calcaratin (14 kDa)&lt;br&gt;Unnamed (45 kDa)&lt;br&gt;Tabanin (7 kDa)&lt;br&gt;Simulidin (11 kDa)&lt;br&gt;Crude extract</td>
<td>Americanin (12 – 16 kDa)&lt;br&gt;Crude extract&lt;br&gt;Crude extract</td>
<td>Amblyomma americanum&lt;br&gt;Boophilus calcaratus&lt;br&gt;Anopheles stephensi&lt;br&gt;Tabanus bovinus&lt;br&gt;Simulium vittatum&lt;br&gt;17 species of flies (genus: tabanus)&lt;br&gt;3 species of mosquitoes (genus: Anopheles)&lt;br&gt;Simulium argus&lt;br&gt;Ponstrongylus megistus</td>
</tr>
</tbody>
</table>
phobic pocket at the active site of thrombin in a non-canonical form (i.e. in the opposite direction of natural substrates such as fibrinogen), forming a short parallel β-pleated sheet with residues Ser214 – Gly216 of thrombin [chymotrypsinogen numbering system (36)]. In contrast, a canonical inhibitor runs in an anti-parallel direction with respect to residues Ser214 – Gly216 of thrombin and possesses a basic P1 residue occupying the acidic S1 site [nomenclature: substrate/inhibitor residues are numbered from the P1-P1 ‘prime’ scissile bond toward the N-terminus and C-terminus respectively. Corresponding substrate/inhibitor binding pockets on the proteinase are numbered accordingly, with ‘S’ replacing ‘P’ (37)]. This primary specificity pocket (S1) on the hirudin-bound thrombin is not occupied, differing from those of canonical inhibitors. The N-terminal amino group interacts with thrombin catalytic residues through hydrogen bonds. The C-terminal domain of hirudin is disordered in NMR structures (29, 30) but binds in ordered, extended conformation to the thrombin exosite-I in crystal structure (31–35). The thrombin exosite-I is flanked by two loops (Phe34 – Leu41 and Lys70 – Glu80) that are rich in basic residues (34). The hirudin C-terminus, rich in acidic residues, is inserted into exosite-I through specific electrostatic interactions. In addition, hydrophobic contacts also make significant contributions to the interaction (32–34). The specific, tight-binding nature of hirudin is thus a result of the extensive contacts in both the active site and exosite-I of thrombin (Fig. 2A).

Figure 1: Thrombin inhibitors. A) Hirudin (PDB: 1HRT): the N-terminal core is stabilised by three disulfide bridges and the C-terminal tail is in a long and extended conformation. B) Haemadin (PDB: 1E0F): the N-terminal core is stabilised by three disulfide bridges and the C-terminal tail is in a long and extended conformation. C) Boophilin (PDB: 2ODY): the first of the two tandem Kunitz domains contains a typical reactive-site loop (arrow). D) Ornithodorin (PDB: ITOC): the first of the two tandem Kunitz domains contains a distorted reactive-site loop (arrow). E) Rhodniin (PDB: 1TBQ): the first of the two tandem Kazal domains contains a typical reactive-site loop (arrow). F) Triabin (PDB: 1AVG): the eight stranded β-barrel fold.
Haemadin

Haemadin was isolated from the Indian leech Haemadipsa sylvestris (Table 1) (38). Although haemadin and hirudin share low sequence similarity, they exhibit a common three-dimensional fold (39). With 57 residues, haemadin is slightly smaller than hirudin. Haemadin is a slow and tight-binding inhibitor of thrombin, with $K_i = 210 \text{fM}$ (38). Similar to hirudin, haemadin has a globular N-terminal core stabilised by three disulfide bridges with an extended, acidic C-terminal tail (Fig. 1B). The first three N-terminal residues non-canonically bind to the active site of thrombin, similar to the N-terminus of hirudin. Interestingly, the acidic C-terminus of haemadin (residues 41 to 57) binds to the thrombin exosite-II instead of exosite-I (Fig. 2B) (39, 40).

Thrombin exosite-II is a highly basic surface (more so than exosite-I), situated on the opposite side of the exosite-I. Exosite-II is also the heparin-binding site of thrombin (41). The distinct surfaces targeted by haemadin (exosite-II) and hirudin (exosite-I), despite overall similarity in their 3D structures, makes an interesting comparison of the molecular diversity observed in the anticoagulants from haematophagous animals: both are peptide inhibitors of thrombin from leeches, but each targets distinct sites.

Kunitz-type thrombin inhibitors

The Kunitz-type inhibitors [with bovine pancreatic trypsin inhibitor (BPTI) as a typical example] are one of the most exten-
sively studied families of serine proteinase inhibitors (42). A typical Kunitz-type domain has a reactive-site loop which binds and runs antiparallel to the enzyme active site residues Ser214 – Gly216. The basic P1 residue of the inhibitor binds to the S1 specificity pocket of the enzyme, similar to the natural substrates (canonical inhibition) (43). This structural fold is commonly found in anticoagulants from ticks. Kunitz-type inhibitors identified from two separate families of ticks (Ixodidae – hard ticks and Argasidae – soft ticks) appear to belong to two different protein subclasses, based on their sequences. This provides evidence for the independent evolution of anticoagulant adaptations of blood-feeding behaviors in the hard and soft ticks (44).

Comparing sequences only, Kunitz-type thrombin inhibitors from hard ticks have a reactive-site loop with normal topology (Fig. 1C). Molecules in this group include amblin from Amblyomma hebraeum (45), boophilin from Boophilus microplus (46) and haemalin from Haemaphysalis longicornis (47) (Table 1). These molecules have two tandem Kunitz domains. They typically have a lower affinity for thrombin [amblin Kᵢ = 20 nM (45), boophilin Kᵢ = 1.8 nM (46)] compared to the soft tick Kunitz-type inhibitors (see below). These molecules were initially thought to bind to the active site of thrombin with their first Kunitz domain in a canonical fashion. However, the recently reported crystal structure of the boophilin-thrombin complex revealed that the first Kunitz domain of boophilin non-canonically binds to the active site of thrombin, despite the presence of a reactive-site loop with normal topology (Fig. 2C) (46).

Kunitz-type thrombin inhibitors from soft ticks include orni-thodorin from Ornithodoros moubata (48), savignin from Orni-thodoros savignyi (49, 50) and monobin from Argas monolakensis (51) (Table 1). Kinetically, savignin (Kᵢ = 4.89 pM) (49) and monobin (Kᵢ = 7 pM) (51) are slow, tight-binding, competitive inhibitors of thrombin. They are atypical Kunitz-type inhibitors as they have two highly distorted tandem Kunitz domains (Fig. 1D). This distortion is due to the placement of a disulfide bridge, and the two domains are connected by an extended linker peptide. In addition to the distorted reactive-site loop, the basic P1 residue is also absent. As a result, this group of molecules non-canonically inhibit thrombin, as shown by the crystal structure of the thombin-ornithodorin (48). The first of the two Kunitz domains of orni-thodorin binds to the thrombin active site. The linker peptide and the second Kunitz domain bind to the thrombin exosite-I (Fig. 2D). The distorted reactive-site loop is not in contact with thrombin. Instead, the first three residues of the N-terminal Kunitz domain run parallel to thrombin Ser214 – Gly216, making several hydrophobic contacts with the thrombin active site, similar to hirudin. The other residues in the first domain of ornithodorin also make contacts with the thrombin 60-loop (Leu59 – Asn62) and autolysis loop (Leu144 – Gly150) near the active site. Interactions between the exosite-I of thrombin and the second Kunitz domain of ornithodorin are mainly mediated through ionic pairings and are strengthened by hydrophobic interactions (48).

Kazal-type thrombin inhibitors

One other commonly encountered serine proteinase inhibitor – the Kazal-type inhibitor (42) – is also utilised as the structural scaffold template for exogenous anticoagulants. Kazal-type thrombin inhibitors include rhodnin from Rhodnius prolixus (52, 53), dipetalogastin from Dipetalogaster maximus (54) and infestin from Triatoma infestans (55) (Table 1). Typically, these molecules contain multiple non-classical Kazal domains and bind to thrombin in a slow, tight-binding, competitive mode. The first and second cysteines of the non-classical Kazal domains are separated by one or two residues. In contrast, seven or eight spacer residues are found in the classical Kazal domains. The Kᵢ for rhodnin, dipetalogastin (domain 3–4) and infestin (domain 1–2) are 0.2 pM (52), 0.05 pM (54) and 25 pM (55), respectively.

### Table 1: Inhibitory Specificities of Kazal-Type and Kunitz-Type Inhibitors

<table>
<thead>
<tr>
<th>Domains</th>
<th>Kᵢ (nM)</th>
<th>Trypsin</th>
<th>Thrombin</th>
<th>FXIIa</th>
<th>Plasmin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1R 2R 3R</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 3 4</td>
<td>5.2</td>
<td>0.8</td>
<td>59.2</td>
<td>0.078</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>3.1</td>
<td>0.025</td>
<td>NI</td>
<td>NI</td>
<td>NI</td>
</tr>
<tr>
<td>3 3 3</td>
<td>3.3</td>
<td>NI</td>
<td>18.2</td>
<td>0.067</td>
<td>0.4</td>
</tr>
<tr>
<td>1 1 1</td>
<td>11</td>
<td>NI</td>
<td>53</td>
<td>0.128</td>
<td>2.1</td>
</tr>
</tbody>
</table>

**Figure 3: Differential specificities showed by different combinations of infestin domains.** The cDNA of Kazal-type thrombin inhibitor infestin is translated into seven tandem domains. The putative site of post-translational cleavage is between the Ala-Glu, present between the signal peptide and domain 1R, between the domains 1R and 2R, between the domains 3R, and 1 and between the domains 2 and 3 (indicated by red arrows). Matured proteins representing the domains 1R, 1–2 and 3–4 were isolated. Combinations of different domains showed differential specificities towards serine proteinases (italic Kᵢ indicates that the binding to that enzyme is the strongest). The domain 1R is not an anticoagulant; instead it inhibits neutrophil elastase, subtilisin A and chymotrypsin. The presence of domain 4 appeared to be responsible for FXIIa specificity. The domain 1–2 is specific for thrombin. ND: not determined; NI: non inhibitory.
Interestingly, these molecules are typically isolated as proteins with two tandem domains (Fig. 1E), although their cloned cDNA are usually translated into additional domains [dipetalogastin – six domains (54); infestin – seven domains (56)]. The crystal structure of the rhodniin-thrombin complex showed a different binding mechanism than the Kunitz-type inhibitors. In Kazal-type inhibitors, the first domain canonically binds to the thrombin active site. The second Kazal domain (along with inter-domains linker) binds to the thrombin exosite-I. The reactive-site loop of rhodniin has a P1 His (52, 53). In contrast to the non-canonical inhibitors, the imidazole side chain of the P1 His of rhodniin is inserted into the thrombin S1 pocket (53). The thrombin 60- and autolysis loops also make contacts with the first Kazal domain. Although the second Kazal domain binds to the thrombin exosite-I, the interaction appears to exclude the reactive-site loop (Fig. 2E). Only two pairs of residues directly form salt bridges, but the second Kazal domain and the linker peptide have an overall positive charge which might help to target the molecule to the negatively charged surface on exosite-I (53).

The roles of the additional domains of Kazal-type inhibitors were investigated in infestin. Combinations of different domains showed differential specificities towards blood coagulation proteinases (Fig. 3). Domain 1–2 (55) is most specific for thrombin, whereas the presence of domain 4 (in infestin domains 1–4, 3–4 and 4) (55, 57) switch the specificity towards FXIIa. It should also be noted that the post-translational processing mechanisms of these molecules are yet to be fully elucidated. However, the infestin domains are postulated to be processed between Ala-Glu, which does not show significant homology to other known pro-thrombin inhibitors (61). A slow, tight-binding, competitive inhibitor of thrombin, anophelin, has a unique amino acid sequence which does not show significant homology to other known proteins (except thrombostasin, which will be discussed below). Significantly, anophelin contains no cysteines but is rich in acidic residues in the middle region (Fig. 1A) (61). The acidic region possibly targets thrombin exosite-I, similar to the action of

**Lipocalin-like thrombin inhibitor**

The kissing bug *Triatoma pallidipennis* utilises the lipocalin fold as a scaffold for anticoagulants (Table 1). Although the lipocalin family of proteins has low sequence similarity, they all share a characteristic fold of an eight-stranded (A-B-C-D-E-F-G-H), anti-parallel β-barrel and a central ligand-binding pocket. This family of proteins displays a wide variety of functions (58). Triabin, isolated from *Triatoma pallidipennis* (59), is the only thrombin inhibitor belonging to this family. Triabin binds to thrombin exosite-I (but not to the active site) in an equimolar ratio, with a K_i of 3 pM (59). Structurally, triabin deviates slightly from the typical lipocalin topology due to a directional inversion of the B and C strands, creating an up-up-down-down topology in the first four strands of the β-barrel (Fig. 1F). In contrast to the hirudin C-terminus, which is inserted into a deep cleft formed by the surface loops in exosite-I of thrombin, the β-barrel structure of triabin interacts with a relatively flat surface near the end of the cleft. The contacts are mainly mediated through hydrophobic interactions, covering a larger area of exosite-I (Fig. 2F) (60).

**Anophelin**

Isolated and cloned from the salivary glands of the mosquito *Anopheles albimanus*, anophelin, a 6.5 kDa protein, belongs to a unique class of thrombin inhibitors (61) (Table 1). Anophelin is a slow, tight-binding, competitive inhibitor of thrombin with a K_i of 5.87 pM. Like many other exogenous thrombin inhibitors, anophelin binds to both the active site and exosite-I of thrombin (62). However, anophelin has a unique amino acid sequence which does not show significant homology to other known proteins (except thrombostasin, which will be discussed below). Significantly, anophelin contains no cysteines but is rich in acidic residues in the middle region (Fig. 1A) (61). The acidic region possibly targets thrombin exosite-I, similar to the action of

---

**Figure 4: Sequences of anophelin, thrombostasin, madanins and chimadanin.** A) The alignment of anophelin and thrombostasin showed limited sequence identities. The identical residues are coloured red, while similar residues are coloured blue. Both proteins are rich in acidic residues, marked by the symbol @. The identities between the two proteins are concentrated in the acidic region in the middle of the molecules and a short basic stretch in the C-terminus. B) The sequence alignment of madanins and chimadanin also showed limited identity. The identical residues are coloured red, while similar residues are coloured blue. Similar to anophelin and thrombostasin, the proteins are rich in acidic residues, marked by the symbol @. The proteins also bear a similar arrangement of residues (a cluster of acidic residues in the middle along with a short stretch of basic residues near the C-terminus). Chimanadin has an extended C-terminus, compared to madanins.
the hirudin C-terminus. However, the disulfide-stabilised hirudin/Kunitz/Kazal-type domains are not present in anophelin (due to the lack of cysteines), suggesting that a distinct structural feature is used to inhibit the thrombin active site.

**Thrombostasin**

Thrombostasin was isolated from the horn fly *Haematobia irritans* (Table 1). It is an 81-residue protein that inhibits the thrombin active site. The primary sequence of thrombostasin contains no cysteines and shows limited similarity to anophelin. Like anophelin, there is a cluster of acidic residues in the middle (all 18 acidic residues fall within positions 10 to 48) and a short stretch of basic residues near the C-terminus (Fig. 4A) (63).

**Madanins and chimadanin**

Madanins (isoforms 1 and 2) (64) and chimadanin (65) are two groups of 7 kDa proteins identified in the salivary gland cDNA libraries of the hard tick *Haemaphysalis longicornis* (Table 1). Recombinant madanins bind only to thrombin exosite-I without affecting its amidolytic activity (commonly used to probe the function of thrombin active site) (64). In contrast, recombinant chimadanin inhibits the function of thrombin active site; however, the binding to either exosite was not investigated (65). The translated amino acid sequence of madanins and chimadanin are without cysteines and show low similarity with each other, except for an acidic region in the middle of the molecules and a basic stretch near the C-terminus (Fig. 4B). Neither proteins show significant homology to other proteins. However, a similar arrangement of residues (a cluster of acidic residues in the middle and a short stretch of basic residues near the C-terminus) is seen in both proteins and is also observed in anophelin and thrombostasin. The C-terminus of chimadanin is extended, compared to the madanins. A detailed comparison between madanin and chimadanin could reveal interesting structure-function information concerning thrombin inhibition. It is possible that the binding of madanin to thrombin exosite-I is through the acidic region, which is also conserved in chimadanin. Thus, chimadanin may bind exosite-I, although this interaction was not investigated. The region in chimadanin that binds to the active site of thrombin could reside in the non-conserved N-terminus or in the extended C-terminus.

**Antistasin-like thrombin inhibitor**

In addition to the hirudin-like fold, leeches also utilise the antistasin-like scaffolds to derive thrombin inhibitors [antistasin is a FXa inhibitor (66, 67); see below for details]. Theromin, an antistasin-like thrombin inhibitor isolated from *Theromyzon tessulatum*, inhibits the thrombin active site with a $K_i$ of 12 fM (68) (Table 1). Overall sequence homology between theromin and antistasin is low, but crucially, a 26-residue segment with conserved cysteines and disulfide pattern (three pairs of disulfide bridges) is present in both. Antistasin is a single chain 119-residue protein with two tandem repeats (66, 67), while theromin is a disulfide-linked homodimer of 67-residue chains (proteolytically processed tandem repeats).

---

**Figure 5: FXa and extrinsic tenase complex (ETC) inhibitors.**

A) the FXa inhibitor TAP (PDB: 1KIG): a single Kunitz-type domain containing a distorted reactive-site loop B) the FXa inhibitor NAP5 (PDB: 2P3F): a single Ascaris-type domain C) the FXa inhibitor antistasin (PDB: 1SKZ): two tandem antistasin-like domains D) the ETC inhibitor NAPc2 (PDB: 2H9E): a single Ascaris-type domain.
Variegin
Recently, we characterised a novel, fast and tight-binding thrombin inhibitor, variegin, isolated from the tropical bont tick *Amblyomma variegatum* (69) (Table 1). Despite its small size (32 residues) and flexible structure (lack of secondary structures), variegin binds thrombin with strong affinity (native variegin $K_i = 10.4\ pM$, synthetic variegin $K_i = 146\ pM$) and high specificity. In an extended conformation, variegin binds to thrombin exosite-I with its C-terminus and to the active site with its middle segment, utilising its N-terminus for fast binding kinetics. Remarkably, the variegin C-terminus (FDFEAIPEEYLDDES) contains a stretch of sequence that is almost identical to that of hirudin C-terminus (GDFFEEIPPEYLEDQ), demonstrating an interesting case of structural convergence in proteins of these unrelated animals. However, the N-termini of both proteins are distinct. Variegin inhibits the thrombin active site through a substrate-like mechanism (canonical inhibition) and hence is cleaved by thrombin (Lys as the P1 residue) (69).

Our results show that variegin is structurally and functionally similar to the human-designed, bivalent thrombin inhibitor, bivalirudin (14). Interestingly, while the development of bivalirudin represents successful rational drug design, variegin demonstrates the ability of nature to produce similar ‘designs’ through evolution. Moreover, nature’s version of bivalirudin (variegin) is a more potent inhibitor, and its inhibitory activity is largely retained after cleavage by thrombin. Variegin thus appears to retain advantages of hirulog (i.e. clearance via proteolysis instead of through the renal route as in the case of hirudin), but with improved potency and possibly a longer duration of action.

**FXa inhibitors**

**Kunitz-type FXa inhibitors**
One of the main classes of FXa inhibitors is the atypical, non-canonical Kunitz-type inhibitors from the soft ticks including tick anticoagulant peptide (TAP) from *Ornithodoros moubata* (70) and FXa-inhibitor (FXal) from *Ornithodoros savignyi* (71) (Table 2). In contrast to the tandem Kunitz domains arrangement observed in Kunitz-type thrombin inhibitors, TAP and FXal contain a single domain (Fig. 5A). Kinetically, both are slow, tight-binding, competitive inhibitors of FXa. Native and recombinant TAP have $K_i$ values of 0.588 nM and 0.18 nM, respectively (70, 72), while native FXal has a $K_i$ of 0.83 nM (71). The crystal structure of TAP in complex with FXa (73) showed a similar non-canonical mode of active site inhibition as observed in the thrombin-hirudin (31–34) and thrombin-ornithodorin complexes (48). The first three N-terminal residues make multiple contacts with the FXa active site and catalytic triad. In addition, TAP interacts with some of the residues on the Na⁺-binding site (Arg222 and Lys224) and autolysis loops (Arg143, Glu146, Lys147 and Arg149) of FXa (Fig. 6A) (73).

**Ascaris-type FXa inhibitors**
The Ascaris family of serine proteinase inhibitors are characterised by 10 cysteine residues forming a unique disulfide pattern in a single domain (74). A group of Ascaris-type FXa inhibitors (75–84 residues) were identified in the hookworms *Ancylostoma caninum* (NAP5/6 or AcAP5/6 and NAPc2/3/4 or AcAPc2/3/4) (75–77) and *Ancylostoma ceylanicum* (AceAP1) (78) (Table 2). NAP5/6 (77) and AceAP1 (78) bind to the FXa active site, whereas NAPc2/3/4 bind to an FXa exosite (76, 77, 79). Binding of NAPc2/3/4 to FXa facilitates their inhibition of the FVIIa-TF complex (76, 77), thus are considered as extrinsic tenase complex (ETC) inhibitors (see below for details). Detailed studies on NAP5 showed that it inhibits the FXa active site competitively, with a $K_i$ of 43 PM (77). NAP5 (Fig. 5B) canonically binds to the FXa active site through the reactive-site loop that possesses a P1 Arg (Fig. 6B). In addition to the active site, NAP5 interacts with residues on the Na⁺-binding site (Arg222) and autolysis loop (Arg143, Lys147, Arg150, Gln151), similar to TAP (80). Interestingly, the C-terminus of NAP5 also interacts with a novel FXa exosite (this exosite partially overlaps with the FXa heparin binding exosite) in a symmetry-related FXa molecule (in the crystal) (80). This exosite interaction is similar to that observed in NAPc2-FXa complex (see below for details) (79). Whether this observation is an artefact due to crystal packing or has any physiological consequences remains to be confirmed.

Despite the structural similarity between NAP5 and AceAP1, both act through different mechanisms of action. Mechanistically, NAP5 is a competitive inhibitor of FXa amidolytic activity while AceAP1 is a two-site, partial non-competitive inhibitor. AceAP1 needs an exosite (yet to be identified) on FXa to which it binds with a lower affinity (700 nM) for full inhibition of the active site (affinity is 2 nM). Moreover, the FXa-AceAP1 complex binds to the FVIIa-TF complex in the same way as the FXa-NAPc2 complex, while NAP5 is devoid of such activity (78, 81).

**Antistasin-like FXa inhibitors**
FXa inhibitors with an antistasin-like domain include antistasin isolated from *Haementeria officinalis* (66, 67), ghilanten isolated from *Haementeria ghiliani* (82) and therostasin isolated from *Theromyzon tessulatum* (83) (Table 2). A 119-residue protein, antistasin consists of two tandem domains containing the 26-residue antistasin-like signature. Each domain in antistasin contains 10 cysteines, forming five intra-domain disulfide bridges (Fig. 5C). It is a slow, tight-binding, competitive inhibitor ($K_i = 0.3–0.6\ nM$). The inhibition of FXa active site is through the canonical reactive-site loop residing in the N-terminal domain (Arg as the P1 residue) (84). The crystal structure of antistasin was modelled in complex with FXa (85). Other than inhibition of the active site through the canonical reactive-site loop, the complex model suggests possible binding to the Na⁺-binding loop (Arg222, Lys223 and Lys224) of FXa (85). Overall sequence similarity between therostasin and antistasin is low, although the domain signature, the canonical reactive-site loop and the P1 Arg are all conserved. Compared to antistasin, therostasin is more potent ($K_i = 34\ pM$), smaller (82 residues), and has fewer cysteines (six residues). However, it does not display the similar tandem domains pattern observed in antistasin (83).

**Serpin family FXa inhibitors**
Serpins are a superfamily of 45–55 kDa proteins that inhibit their target enzymes by irreversibly locking the proteinases in a covalent acyl-enzyme intermediate (86). The crude salivary gland
extract of female yellow fever mosquitos (Aedes aegypti) was found to inhibit the FXa active site through a reversible, non-competitive mechanism (87). The only anti-FXa fraction in the extract, named anticoagulant-factor Xa (AFXa, 54 kDa), was isolated and cloned from the salivary glands of the mosquito (Table 2). Its primary sequence shows high similarities with other serpins (e.g. plasminogen activator inhibitor-2). Post-translational modifications of AFXa are likely to be important for its activity (four N-linked glycosylation sites are present). However, compared to typical serpins, AFXa has a shorter reactive-site loop and different hinge residues. If the activity of AFXa is the same as that of the salivary glands extract fraction (reversible, non-competitive inhibition of FXa), it would be interesting to investigate its differences with the physiological serpin (antithrombin), which also inhibits FXa but through a distinct mechanism (irreversible, heparin-dependant, competitive inhibition) (88).

Uncompetitive FXa inhibitors from Hyalomma

Two uncompetitive inhibitors of FXa amidolytic activity were isolated from Hyalomma truncatum (K_i = 0.69 nM) (89) and Hyalomma dromedarii (K_i = 134 nM) (90) (Table 2). Both proteins have similar masses (17 kDa and 15 kDa, respectively). As they bind to the enzyme-substrate complex but not to the enzyme

Figure 6: Interactions between FXa and its inhibitors. The interface residues between FXa and its inhibitors are mapped. On FXa, the active site surfaces are coloured green. The active site includes the active site pocket, Na+ binding loop and autolysis loop. The extended exosite surface including the heparin binding exosite is coloured orange. On the inhibitors, the active site-targeting residues are coloured magenta; the exosite-targeting residues are coloured blue. A) the FXa-TAP complex (PDB: 1KIG) B) the FXa-NAP5 complex (PDB: 2P3F) C) the FXa-NAPc2 complex (PDB: 2H9E): NAPc2 binds to FXa as scaffold to inhibit the FVIIa-TF complex and is classified as an extrinsic tenase complex inhibitor.
Table 2: FXa inhibitors from haematophagous animals.

<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>Mechanism</th>
<th>Example</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kunitz-type proteinase inhibitors</td>
<td>Single Kunitz domain of ~ 6 to 7 kDa</td>
<td>Slow, tight-binding, competitive inhibition</td>
<td>TAP</td>
<td>Omithodoras moubata</td>
<td>(70, 73)</td>
</tr>
<tr>
<td></td>
<td>Distorted reactive-site loop, lack of basic P1 residue</td>
<td>N-terminal Kunitz domain inhibits active site non-canonically</td>
<td>FXaI</td>
<td>Omithodoras savignyi</td>
<td>(71)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-terminal Kunitz domain binds to exosite-I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ascaris-type proteinase inhibitors</td>
<td>Single domain of ~ 9 to 11 kDa</td>
<td>Competitive inhibition</td>
<td>NAP5 (AcAP5)</td>
<td>Ancylostoma canium</td>
<td>(75–77)</td>
</tr>
<tr>
<td></td>
<td>Conserved cysteine residues (10) and disulfide linkages</td>
<td>Inhibits active site canonically</td>
<td>NAP6 (AcAP6)</td>
<td>Ancylostoma canium</td>
<td>(75–77)</td>
</tr>
<tr>
<td></td>
<td>Reactive-site loop with basic P1 residue</td>
<td></td>
<td></td>
<td></td>
<td>(78)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two-site partial non-competitive inhibition</td>
<td>AceAPI</td>
<td>Ancylostoma ceylanicum</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antistasin-like inhibitors</td>
<td>Cysteine-rich domain of ~ 7 kDa</td>
<td>Inhibits active site canonically</td>
<td>Antistasin</td>
<td>Haementeria officinalis</td>
<td>(66, 67)</td>
</tr>
<tr>
<td></td>
<td>Containing a 26-residues segment with conserved cysteines (6) and disulfide linkages</td>
<td>Interactions with exosites not reported</td>
<td>Ghilanten</td>
<td>Haementeria ghilianii</td>
<td>(82)</td>
</tr>
<tr>
<td></td>
<td>Domain typically repeated in tandem (except therostasin)</td>
<td></td>
<td>Therostasin</td>
<td>Theromyzon tessulatum</td>
<td>(83)</td>
</tr>
<tr>
<td>Atypical serpins</td>
<td>Single domain of ~ 45 – 55 kDa</td>
<td>Mechanically different from typical serpins</td>
<td>AFXa</td>
<td>Aedes aegypti</td>
<td>(87, 88)</td>
</tr>
<tr>
<td></td>
<td>Three β-sheets and eight or nine α-helices</td>
<td>Only example so far showed reversible, non-competitive inhibition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Typically glycosylated</td>
<td>Post-translational modification important for activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shorter reactive-site loop and different hinge residues compared to typical serpins</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salp family</td>
<td>~ 9 to 14 kDa</td>
<td>Recombinant salp14 inhibits FXa active site</td>
<td>Salp14</td>
<td>Ixodes scapularis</td>
<td>(129)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombinant salp9pac is not active</td>
<td>Salp9pac</td>
<td>Ixodes scapularis</td>
<td>(129)</td>
</tr>
<tr>
<td>Hemerythrin family</td>
<td>Single domain of ~ 15 kDa</td>
<td>Inhibits active site</td>
<td>Lefaxin</td>
<td>Haementeria depressa</td>
<td>(130)</td>
</tr>
<tr>
<td></td>
<td>Has a topology consisting of four helix bundles</td>
<td>Maximal inhibitory activity at low concentration (&lt;1 mM) of CaCl₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draculin</td>
<td>Glycoprotein of ~ 90 kDa</td>
<td>Slow, non-competitive, tight binding inhibitor of active site</td>
<td>Draculin</td>
<td>Desmodus rotundus</td>
<td>(131)</td>
</tr>
<tr>
<td>Uncompetitive inhibitors from Hyalomma</td>
<td>~ 15 to 17 kDa</td>
<td>Uncompetitive inhibitors of active site</td>
<td>Unnamed</td>
<td>Hyalomma truncatum</td>
<td>(89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unnamed</td>
<td>Hyalomma dromedani</td>
<td>(90)</td>
</tr>
<tr>
<td>Others</td>
<td>Lack of detailed structural information</td>
<td>Lack of functional characterisation</td>
<td>Unnamed</td>
<td>Amblyomma americanum</td>
<td>(132)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unnamed</td>
<td>Culicoides variipennis sonorensis</td>
<td>(133)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unnamed</td>
<td>Simulium vittatum</td>
<td>(134)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unnamed</td>
<td>Haementeria ghiliani</td>
<td>(135)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unnamed</td>
<td>3 species of mosquitoes</td>
<td>(126)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Crude extract</td>
<td>3 species of black flies</td>
<td>(127)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inhibiting prothrombinase complex by attenuating FV activity</td>
<td>Unnamed</td>
<td>Simulium vittatum</td>
<td>(136)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Crude extract</td>
<td>Dermacentor andersoni</td>
<td>(137)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Crude extract</td>
<td>Triatoma infestans</td>
<td>(128)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inhibiting prothrombinase complex by unknown mechanism</td>
<td>Unnamed</td>
<td>Rhipicephalus appendiculatus</td>
<td>(138)</td>
</tr>
</tbody>
</table>
alone, their detailed characterisation could possibly reveal novel FXa exosites which strongly modulate the active site function of substrate-bound FXa.

Extrinsic tenase complex (ETC) inhibitors

Kunitz-type ETC inhibitors

The two main classes of exogenous ETC inhibitors act through a similar but not identical mechanism as the physiological ETC inhibitor, tissue factor pathway inhibitor (TFPI). TFPI is a Kunitz-type inhibitor with three tandem domains, binding to both FXa and FVIIa-TF to form a quaternary complex (91). The hard tick *Ixodes scapularis* contains two different ETC inhibitors with Kunitz scaffolds (Table 3). One of them, *ixolaris* (15.7 kDa), possesses two tandem Kunitz domains and does not bind to the FXa active site, in contrast to TFPI (three domains and binds the FXa active site). It was hypothesised that the second Kunitz domain binds first to FX/Xa before binding to the FVIIa-TF complex via the first Kunitz domain (92). *Ixolaris* binds to FX and FXa with affinities between 0.5–10 nM (93). Factor X/Xa residues that are involved in binding to *ixolaris* are on a surface that largely overlaps with their heparin binding proexosite/exosite (93, 94). In addition, the binding of *ixolaris* to FX (93) and FXa (94) impaired their interactions with FVIIa and prothrombin, respectively. Much less information is available for the inhibition of the FVII-TF complex by the FX/FXa-*ixolaris* complex. Since *ixolaris* is a Kunitz-type inhibitor, it is likely that the reactive-site loop on the first Kunitz domain binds to the FVIIa active site. However, the putative P1 residue is Glu, instead of a preferred basic residue (92). The other ETC inhibitor, *penthalaris*, has five tandem Kunitz domains, compared to only two in *ixolaris*. *Penthalaris* uses FX or FXa as scaffold to inhibit the FVIIa-TF complex in the same way as *ixolaris*. The contribution of the three additional Kunitz domains to the interaction/function is not yet clear (95).

### Table 3: Extrinsic tenase complex inhibitors from haematophagous animals.

<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>Mechanism</th>
<th>Example</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kunitz-type protease inhibitors</td>
<td>Single Kunitz domain of ~ 6 to 7 kDa</td>
<td>Similar to TFPI</td>
<td><em>ixolaris</em></td>
<td>Ixodes scapularis, Ixodes [k2]</td>
<td>(92–94)</td>
</tr>
<tr>
<td></td>
<td>Multiple Kunitz domains in tandem</td>
<td>One Kunitz domain binds FX/FXa heparin binding pro-exosite/exosite as scaffold</td>
<td><em>Penthalaris</em></td>
<td><em>Ancyclostoma canium</em></td>
<td>(95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Another Kunitz domain inhibits FVIIa-TF complex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forms inactive extrinsic tenase-inhibitor quaternary complex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ascaris-type protease inhibitors</td>
<td>Single domain of ~ 9 to 11 kDa</td>
<td>Binds FXa exosite as scaffold</td>
<td>NAPc2/AcAPC2</td>
<td><em>Ancyclostoma canium</em></td>
<td>(76, 77)</td>
</tr>
<tr>
<td></td>
<td>Conserved cysteine residues (10) and disulfide linkages</td>
<td>Inhibitor-FXa complex binds FVIIa-TF complex</td>
<td>NAPc3/AcAPC3</td>
<td><em>Ancyclostoma canium</em></td>
<td>(76, 77)</td>
</tr>
<tr>
<td></td>
<td>Reactive-site loop with basic P1 residue</td>
<td>Forms inactive extrinsic tenase-inhibitor quaternary complex</td>
<td>NAPc4/AcAPC4</td>
<td><em>Ancyclostoma canium</em></td>
<td>(76, 77)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Forms inactive extrinsic tenase-inhibitor quaternary complex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Binding to FXa is through exosite and active site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>Lack of detailed structural information</td>
<td>Lack of functional characterisation</td>
<td>Crude extract</td>
<td>Dermacentor andersoni</td>
<td>(137)</td>
</tr>
</tbody>
</table>
Intrinsic tenase complex (ITC) inhibitor

**Lipocalin family ITC inhibitor**

Nitrophorin-2 (or Prolixin-s, ~20 kDa), isolated from the kissing bug *Rhodnius prolixus* (101), is an ITC inhibitor which belongs to the lipocalin family (102, 103). Four nitric oxide/histamine transporting molecules with high sequence similarities (nitrophorin 1 to 4) were characterised from the kissing bug and all possess vasodilatory (through nitric oxide release) and antihistaminic (through histamine binding) activities. Interestingly, only nitrophorin-2 has a strong anticoagulant activity, mediated through specific protein-protein interactions. Nitrophorin-2 binds specifically to FIX/FXa with $K_i \sim 13$ nM, most likely by targeting the Gla-domain of FIX/FXa. This binding interferes with FIX activation (by both the FVIIa-TF complex and FXIa) and with FIXα activity in the ITC (102). Through comparison of the anticoagulant with nitrophorin-1 and 4 (non-anticoagulants), the surfaces on nitrophorin-2 involving FIX/FXa binding have been suggested to be within its B-C and E-F strands as well as its C-terminus (104).

**Contact system protein inhibitors**

**Kunitz-type contact system inhibitors**

Haemaphysalin (16 kDa, two tandem Kunitz domains), from the hard tick *Haemaphysalis longicornis* (105), directly binds to both FXII and HMWK (high molecular weight kinogen) in the presence of Zn$^{2+}$ and interferes with their associations with activating surfaces. Amidolytic activity of FXIIα and kallikrein are not affected by the inhibitor. As a result of haemaphysalin binding, reciprocal activations of FXII and prekallikrein are disrupted and initiation of classical intrinsic pathway of coagulation is inhibited (105). Similar to haemaphysalin, hamadarin from the mosquito *Anopheles stephensi* has an identical mechanism of inhibition on contact system proteins. Hamadarin also has a mass of 16 kDa, although no sequence information is available to ascertain the identity of hamadarin as a Kunitz-type inhibitor (106).

A group of Kunitz-type inhibitors isolated from *Boophilus microplus* were reported to inhibit the amidolytic activity of plasma kallikrein. BmTI-A (15 kDa) (107) and BmTI-2 (17 kDa) (108) have two tandem domains, while BmTI-D (8 kDa) has a single Kunitz domain (108). They inhibit plasma kallikrein with $K_i$ values of 120 nM, 48 nM and 12 nM, respectively. However, they also inhibit other serine proteinases such as trypsin, chymotrypsin and neutrophil elastase (107, 108). Similarly, a few inhibitors from *Rhizophorus sanguineus*, such as RsTIQ2 (12 kDa, two Kunitz domains) and RsTIQ7 (8 kDa, single Kunitz domain), inhibit amidolytic activity of plasma kallikrein in addition to trypsin, neutrophil elastase and plasmin (109).

**Molecular diversity**

Successful blood-feeding strategies are crucial for the survival of haematophagous animals. They belong to a wide variety of different lineages and have evolved over hundreds of millions of years, developing enormous diversity in their blood-feeding habits, mechanisms, hosts and the molecular armory for anticoagulation. The molecular diversity among the anticoagulants can be discussed under two categories: (1) functional divergence among closely related proteins, and (2) functional convergence of structurally unrelated proteins.

**Functional divergence among closely related proteins**

A number of anticoagulants, which specifically target different coagulation proteinases or complexes, share select sets of common structural scaffolds. New functions or specificities are achieved through gene duplications and accumulation of mutations on common molecular scaffolds. In general, the same sets of scaffolds are frequently found in closely related species, whereas different sets of scaffolds are found in phylogenetically distant species. For example, the Kunitz-type proteinase inhibitors scaffold is common among the ticks, the hirudin-like scaffold is exclusively found in the leeches, and the Ascaris-type inhibitors scaffold is dominant in the hookworms. Frequently, more than one scaffold is utilised within a single family or order of haematophagous animals. For instance, anticoagulants from the leeches include the hirudin-like and the antistasin-like scaffolds. Similarly, anticoagulants with the Kazal-type and lipocalin scaffolds are found in the kissing bugs. Thus, each time a certain lineage of animal independently adapts to haematophagy, a different set of protein scaffolds may be utilised as anticoagulants, multiplying the molecular diversity of these compounds.

Within a single family of scaffold, a wide range of functional divergences is common. For example, a single Kunitz domain (108) may be functional in trypsin inhibition, a single domain (48 nM) in kallikrein inhibition, and a single domain (12 nM) in both serum proteinases, such as trypsin, neutrophil elastase, and plasmin. A number of anticoagulants, which specifically target different coagulation proteinases or complexes, share select sets of common structural scaffolds. New functions or specificities are achieved through gene duplications and accumulation of mutations on common molecular scaffolds. In general, the same sets of scaffolds are frequently found in closely related species, whereas different sets of scaffolds are found in phylogenetically distant species. For example, the Kunitz-type proteinase inhibitors scaffold is common among the ticks, the hirudin-like scaffold is exclusively found in the leeches, and the Ascaris-type inhibitors scaffold is dominant in the hookworms. Frequently, more than one scaffold is utilised within a single family or order of haematophagous animals. For instance, anticoagulants from the leeches include the hirudin-like and the antistasin-like scaffolds. Similarly, anticoagulants with the Kazal-type and lipocalin scaffolds are found in the kissing bugs. Thus, each time a certain lineage of animal independently adapts to haematophagy, a different set of protein scaffolds may be utilised as anticoagulants, multiplying the molecular diversity of these compounds.

<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>Mechanism</th>
<th>Example</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
</table>
| Lipocalin family | Single domain of ~16 kDa  
Eights-stranded (A-B-C-D-E-F-G-H)antiparallel β-barrel and a central ligand-binding pocket  
Also functions as vasodilator and anti-histamine  
β binds to FIX/FXa Gla-domain  
Inhibits both FIX activation and FXa activity in intrinsic tenase complex | Nitrophorin-2 (prolixin-s)  
Rhodnius prolixus | (101–104) | |
| Others | Lack of detailed structural information  
Inhibits intrinsic tenase complex activity by attenuating FVIII activity | Crude extract  
Triatoma infestans | (128) | |

Table 4: Intrinsic tenase complex inhibitors from haematophagous animals.
can be gained through divergent evolution. In ticks, the Kunitz-type scaffold diverged into thrombin inhibitors (ornithodorin and boophilin), FXa inhibitors (TAP), ETC inhibitors (ixolaris) and contact system proteins inhibitors (haemaphysalin). This functional divergence occurs through both changes in the interface residues and the addition of the tandem domains. At times, these anticoagulants interact with similar regions of closely related target proteinases yet evolved distinct molecular mechanisms to achieve specificity. Studying each set of anticoagulants with the same scaffold helps us to identify specific functional sites. For example, the autolysis loop of thrombin and FXa are targeted by the Kunitz-type inhibitors (ornithodorin-thrombin and TAP-FXa). The autolysis loops, although conserved at similar positions on all coagulation proteinases, are distinct in many aspects, modulating their interactions with substrates (110). Thus, comparing and contrasting the structural features of these two sets of interactions could reveal information that allows design of specific inhibitors for different coagulation proteinases by targeting the factor-specific features of the autolysis loops.

Similarly, the Ascaris-type scaffold is utilised for FXa inhibitors (NAP5), ETC inhibitors (NAPc2) and an inhibitor inhibiting both FXa and ETC (AceAP1). It can be observed that functionally distinct molecules (NAP5 and NAPc2) can be derived from a single structural family (Ascaris-type inhibitor) in one species (Ancyclostoma caninum), while the two functions are combined on a single molecule (AceAP1) in another species (Ancyclostoma ceylanium).

Variations on a scaffold can also result in molecules (hirudin vs. haemadin) which inhibit the same target enzyme (thrombin) but which bind to distinct surfaces (exosite-I vs. exosite-II). In addition, structurally similar inhibitors isolated from different leech species show adaptations to different hosts. Hirudin-like molecules isolated from other leeches, such as bufrudin (25) and hirullin (26) from Hirudinaria manillensis, show distinct structural features. The C-terminal sulfotyrosine of hirudin is not conserved in bufrudin (25), and hirullin was found to be glycosylated at Thr45 (isoform P6) or Thr50 (isoform P18) (26). Hirudinaria manillensis is primarily a mammalian parasite, while Hirudo medicinalis is an amphibian parasite. Thus bufrudin and hirullin may be tailored to inhibit mammalian coagulation proteins more efficiently. Characterisations of these molecules are important for the understanding of structure-function relationships, particularly the role of post-translational modifications of hirudin.

Target alterations by domain duplications are also demonstrated by the Kazal-type inhibitors infestins. Derived from the same precursor, specificities of infestins rely on post-translational processing of a broad specificity inhibitor (infestin 1–4) which can be cleaved into one specific for thrombin (infestin 1–2) and one specific for FXIIa (infestin 3–4) (55, 57) (Fig. 3). Investigation into the processing mechanisms could answer interesting questions on the regulation of anticoagulants expression profiles in the saliva.

### Functional convergence of structurally unrelated proteins

Haematophagy has evolved independently many times, and different molecular scaffolds are used to achieve the inhibition of a limited number of key enzymes in the coagulation cascade. A single target proteinase is often inhibited by distinct families of inhibitors, but which bind to distinct surfaces (exosite-I vs. exosite-II). In addition, structurally similar inhibitors isolated from different leech species show adaptations to different hosts. Hirudin-like molecules isolated from other leeches, such as bufrudin (25) and hirullin (26) from Hirudinaria manillensis, show distinct structural features. The C-terminal sulfotyrosine of hirudin is not conserved in bufrudin (25), and hirullin was found to be glycosylated at Thr45 (isoform P6) or Thr50 (isoform P18) (26). Hirudinaria manillensis is primarily a mammalian parasite, while Hirudo medicinalis is an amphibian parasite. Thus bufrudin and hirullin may be tailored to inhibit mammalian coagulation proteins more efficiently. Characterisations of these molecules are important for the understanding of structure-function relationships, particularly the role of post-translational modifications of hirudin.

Table 5: Contact system proteins inhibitors from haematophagous animals.

<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>Mechanism</th>
<th>Examples</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
</table>
| **Kunitz-type proteinase inhibitors** | ● Single Kunitz domain is ~ 7 kDa  
● Some of the inhibitors have two tandem Kunitz domains (~ 16 kDa) | ● Binds directly to both FXIIa and HMWK  
● Inhibits associations of FXIIa and HMWK with activating membranes  
● Identical with haemaphysalin | Haemaphysalin  
Haemaphysalis longicornis | (105) |
|                           |                                               |                                                                           | Hamadarin  
Anopheles stephensi | (106) |
|                           |                                               |                                                                           | BmTI-A/D/2  
Boophilus microplus | (107, 108) |
|                           |                                               |                                                                           | RsTI-Q2/Q7  
Rhipicephalus sanguineus | (109) |
| **Kazal-type proteinase inhibitors** | ● Single Kazal domain is ~ 6 kDa  
● Multiple non-classical Kazal domains  
● cDNA sequence showed multiple-domains precursors | ● Inhibits FXIIa amidolytic activity  
● Specificity towards different proteinase changes with different tandem repeats | Infestin 1–4  
Triatoma infestans | (55, 57) |
|                           |                                               |                                                                           | Infestin 3–4  
Triatoma infestans | (55, 57) |
|                           |                                               |                                                                           | Infestin 4  
Triatoma infestans | (55, 57) |
| **Antistasin-like inhibitors** | ● Cysteine-rich domain of ~ 5 kDa  
● Containing a 26-residue segment with conserved cysteines and disulfide linkages | ● Inhibits amidolytic activity of plasma kallikrein  
● Also inhibits tissue kallikrein and trypsin | Piguamerin  
Hirudo nipponia | (139) |
proteins. In some cases, multiple families of proteins that target the same proteinase are found in a single species. At least four different classes of thrombin inhibitors were isolated from *Boophilus microplus*: 1) boophilin (46); 2) microphilin (111); 3) BmAP (112); and 4) BmGTI (113) (Table 1). Thus functional convergence is seen in different protein scaffolds. So far, there are 14 different structural classes of thrombin inhibitors and eight classes of FXa inhibitors isolated from haematophagous animals (described above). Frequently, different structural features can recognise the same molecular surfaces on the target proteinase. For instance, simultaneous binding of the thrombin exosite-I and active sites are observed in the hirudin-like (23, 25, 114), Kunitz-type (46, 48), Kazal-type (52–55), anaphelin (61) and variegin (69) thrombin inhibitors. Similarly, both Kunitz-type (70, 73) and Ascaris-type NAP5 (80) inhibitors target the active site, Na+-binding loop and autolysis loop of FXa. Thus, starting from a wide variety of structural scaffolds and through numerous rounds of evolutionary ‘iterations’, consensus ‘hot-spots’ on the molecular surfaces of proteinases are targeted efficiently. Such hotspots are likely to be the ‘vulnerable’ target sites for the development of potent and specific inhibitors of these proteinases. In addition, the search for the ‘best’ structural fits for inhibition has been extensively explored by nature. Much valuable information about these molecular surfaces, particularly the exosites, is gained from careful analyses of exogenous anticoagulants. High specificity and hence minimal side effects can often be achieved by targeting such exosites (115).

Structurally distinct molecules target the same molecule with different mechanisms of action. For example, diversity in structure (at least eight structural classes) enabled different inhibition mechanisms – competitive (67, 70, 73, 77, 80, 84, 85), noncompetitive (87, 88), partial non-competitive (78, 81), and uncompetitive (89, 90) – of a single target (FXa). Detailed structure-function relationships studies of these proteins will provide valuable information for the design of novel FXa inhibitors.

**Future prospects**

Haematophagous animals include an estimated 15,000 species of arthropods (~400 genera) (18) and a large number of leeches and hookworms. It has been postulated that independent adaptations of blood-feeding behavior have evolved at least six times within the haematophagous arthropods alone (18). Only a small number of haematophagous animals (~50, based on articles listed on PubMed database) have been studied to date for their anticoagulants. Considering both the structural and functional diversity of the anticoagulants that have been characterised so far, the molecular diversity in exogenous anticoagulants from haematophagous animals is potentially massive. Haematophagous animals are a ready source of antithrombotic agents, and the study of their exogenous anticoagulants offers many advantages. For example, hirudin has both served as an effective antithrombotic directly (23) and has inspired the design of a smaller and more widely used agent, bivalirudin (116). These studies contribute to our basic understanding of the molecular interactions and mechanisms involved in the inhibition of coagulation proteinases or complexes. Thus, the exploration of novel exogenous anticoagulants from these animals provides a great opportunity in the search for novel antithrombotic agents.

**Acknowledgements**

We thank Dr. Stephen Mackessy for proofreading and his comments on the manuscript.

**References**

8. Wu KK, Matijevic-Aleksic N. Molecular aspects of proteinases or complexes. Thus, the exploration of novel exogenous anticoagulants from these animals provides a great opportunity in the search for novel antithrombotic agents.

**Future prospects**

Haematophagous animals include an estimated 15,000 species of arthropods (~400 genera) (18) and a large number of leeches and hookworms. It has been postulated that independent adaptations of blood-feeding behavior have evolved at least six times within the haematophagous arthropods alone (18). Only a small number of haematophagous animals (~50, based on articles listed on PubMed database) have been studied to date for their anticoagulants. Considering both the structural and functional diversity of the anticoagulants that have been characterised so far, the molecular diversity in exogenous anticoagulants from haematophagous animals is potentially massive. Haematophagous animals are a ready source of antithrombotic agents, and the study of their exogenous anticoagulants offers many advantages. For example, hirudin has both served as an effective antithrombotic directly (23) and has inspired the design of a smaller and more widely used agent, bivalirudin (116). These studies contribute to our basic understanding of the molecular interactions and mechanisms involved in the inhibition of coagulation proteinases or complexes. Thus, the exploration of novel exogenous anticoagulants from these animals provides a great opportunity in the search for novel antithrombotic agents.

**Acknowledgements**

We thank Dr. Stephen Mackessy for proofreading and his comments on the manuscript.

**References**

8. Wu KK, Matijevic-Aleksic N. Molecular aspects of...
53. van de Locht A, et al. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodinoid in complex with thrombin. EMBO J 1995; 14: 5149–5157.
71. Ibrahim MA, et al. Factor X (FXa) inhibitor from the nymphs of the camel tick Hyalomma dromedarii.
Koh et al. Anticoagulants from hematophagous animals


