Isolation, Sequence Analysis, and Cloning of Haemadin
AN ANTICOAGULANT PEPTIDE FROM THE INDIAN LEECH*

Karl-Hermann Strube, Burkhard Kröger‡, Siegfried Bialojan, Marcus Otte§, and Johannes Dött¶

From the Department of Biotechnology, BASF Aktiengesellschaft, D-6700 Ludwigshafen and the §Institute of Biochemistry, Technische Hochschule, D-6100 Darmstadt, Germany

A slow, tight-binding inhibitor of thrombin with an apparent molecular mass of about 5 kDa has been isolated from Haemadipsa sylvestris, an Indian leech of the family of Haemadipsidae. The inhibitory activity, called haemadin, is thrombin specific since it does not inhibit other proteases like trypsin, chymotrypsin, factor Xa, or plasmin. NH2-terminal amino acid sequence analysis (residues 1–45) does not reveal any homology to known serine protease inhibitors, including the thrombin-specific inhibitor hirudin. The haemadin cDNA cloned by polymerase chain reaction techniques codes for a polypeptide of 57 amino acid residues preceded by 20 residues of a signal peptide sequence. A synthetic gene coding for the mature haemadin was expressed in Escherichia coli. Recombinant haemadin displays a similar inhibition constant and specific activity as its natural counterpart. Although there is no obvious sequence identity between haemadin and hirudin, both proteins seem to share common mechanisms for thrombin inhibition.

Animals depending on a diet of ingested fresh blood have a vital interest in mechanisms that interfere with the coagulation system of the blood donor. A variety of different principles have been described and characterized from blood-sucking vertebrates (1), insects (2, 3), and leeches (4, 5) reviewed in Ref. 5). Their salivas contain anticoagulant compounds that maintain the blood in a fluid state during the uptake and in their guts. Leeches seem to have evolved several potent anticoagulant activities that inhibit the coagulation cascade at different steps. The salivas of the Mexican leech Haementeria officinalis, as well as the giant leech Haementeria ghilianii, display activities that inhibit coagulation factor Xa (6, 7). Decorsin, isolated from Macrobdella decora, is an antagonist of platelet membrane glycoprotein IIb-IIIa that interferes with platelet aggregation (8).

The most prominent natural anticoagulant is hirudin, a 65-amino-acid residue polypeptide with high specificity toward thrombin, isolated from the medicinal leech, Hirudo medicinalis (9). The observed inhibition constant for the thrombin-hirudin complex of about 10−14 M is the result of a combined interaction at two different sites on thrombin (10, 11). The acidic residues in the carboxyl-terminal tail of hirudin interact with positively charged residues in thrombin’s fibrinogen-binding exosite, and amino acid residues from the amino-terminal part make contacts with the active site cleft on thrombin. The individual functional parts of the inhibitor exhibit binding affinities to thrombin in the nanomolar to micromolar range, but the additive contributions result in the tight-binding inhibitor-proteinase complex. Recently, hirudin-like proteins from other leeches have been characterized (12, 13). Due to their high degree of homology, one would not expect another inhibition mechanism.

We focused in our study on the Indian leech, Haemadipsa sylvestris. In contrast to hirudinaria, H. sylvestris is a land-living leech. Since bites of this hemathophageous leech result in prolonged bleeding times (5) we expected these animals to display inhibitory activity on the hemostatic process. Initial experiments indicated that this observation is due to thrombin inhibition. The present paper describes the isolation and characterization of a thrombin-specific inhibitor, called haemadin. On the basis of amino acid sequences, it appeared attractive to us to clone the corresponding cDNA to test the recombinant inhibitor in vivo models of thrombotic diseases as well as to get a better understanding of thrombin inhibition mechanisms.

EXPERIMENTAL PROCEDURES

Materials
Indian leeches (H. sylvestris Blanchard) were obtained from Animalpharma (Erlangen, Germany). Bovine α-thrombin was purchased from Sigma, and human α-thrombin was isolated as described previously (14) and active site titrated as described previously (15); purity was >98%. Plasmin was from Calbiochem, FVIIa from Novo, trypsin and chymotrypsin from Boehringer Mannheim, and FXa from New England Biolabs. Chromogenic and fluorogenic substrates were purchased from Kabi Vitrum (S-2251, S-2238, S-2222, S-2586), Pentapharm (FVIIa-substrate) and Bachem (Tos-Gly-Pro-Arg-AMC). Recombinant hirudin was supplied by BASF AG. PCR components were obtained from Perkin-Elmer Cetus, the protein fusion and purification system pMAL-p2 from New England Biolabs (GeneExpress), and preformed tricine gels were from Anamed, Benheim, Germany.

Methods
Purification of Haemadin—About 200 g of leeches were homogenized for 10 min at 4 °C in buffer A (20 mM sodium phosphate, pH 7.4) using an Ultraturrax blender. Alternatively, 200–500 leeches were decapitated, and their heads were homogenized in the same way. The homogenates were cleared by centrifugation at 6,000 × g and the pellet resuspended in buffer A. After a second, identical centrifugation step, both supernatants were combined and 2 volumes of buffer B (50 mM Tris-HCl, pH 8.5) were added. The protein extract was applied to a source of the American Society for Biochemistry and Molecular Biology, Inc.
to a Q-Sepharose anion-exchange column (Pharmacia/LKB) and washed with buffer B. Bound proteins were eluted using a linear gradient of 0-1 M sodium chloride in buffer B. Fractions containing thrombin inhibitory activity were pooled, diluted with 2 volumes of buffer C (20 mM sodium phosphate, pH 7.5), and applied to a thrombin-Sepharose affinity column. Thrombin-Sepharose was prepared by coupling 10,000 NIH U bovine thrombin (Sigma) to 6.6 g of CNBr-activated Sepharose 4B (Pharmacia/LKB) according to the manufacturer's guidelines. The affinity column was washed with buffer C, followed by buffer C containing 500 mM sodium chloride. Specifically bound proteins were eluted with buffer D (100 mM glycine- HCl, pH 2.8), and the pH of the eluates was immediately adjusted to neutral with 1 M Tris- HCl, pH 8.5. The peak fraction was cleaved with factor Xa according to the manufacturer's guidelines with a 1:650 (w/w) ratio of FXa/affinity column. After 16 h at room temperature, released proteins were separated by mono- Q anion-exchange chromatography using the same buffer system as described above. Fractions obtained were tested for trypsin inhibitory activity, Thrombin-Sepharose and affinity chromatography on thrombin-coupled CNBr-activated Sepharose 4B (Pharmacia/LKB) according to the manufacturer's guidelines. The column was washed with buffer C, followed by buffer containing 500 mM sodium chloride. Specifically bound proteins were eluted with buffer D (100 mM glycine-HCl, pH 2.8), and the pH of the eluates was immediately adjusted to neutral with 1 M Tris-HCl, pH 8.5. The peak fraction was cleaved with factor Xa according to the manufacturer's guidelines with a 1:650 (w/w) ratio of FXa/affinity column. After 16 h at room temperature, released proteins were separated by mono- Q anion-exchange chromatography using the same buffer system as described above. Fractions obtained were tested for trypsin inhibitory activity, and the corresponding chromogenic sub- strate were incubated, and the reaction was started by addition of 10.4 pm enzyme (final concentration). The kinetic con- stants $K_i$, $K_m$, and $V_{max}$ were determined according to the theory of non-linear regression analysis of the data as previously described (18, 19).

Isolation of RNA and Construction of cDNA Library—Total RNA was isolated from whole leeches by the guanidinium thiocyanate method according to the manufacturer's guidelines (Stratagene). Poly(A) RNA was isolated by oligo(dT) affinity separation with the PolyATtract system (Promega). A cDNA library was constructed in lambda Uni-ZAP XR according to the manufacturer's guidelines (Stratagene). The complexity of the library was 1.8 $\times$ 10^9 by amplification by standard procedures (24).

PCR Cloning of the Haemadin cDNA—First strand cDNA was synthesized from 5 µg of total RNA with 1 µg of primer YXT (5'-CGAGGAGGATGTCGACCGGACC(T)$_{18}$-3'), using SuperScript reverse transcriptase (Gibco-BRL). The cDNA was purified from excess primers by elution with GeneClean (BiolOl, La Jolla, CA) and PCR amplified with oligonucleotide A (5'-GGAATGCATGCATGCTGCGAAGCGAC(T)$_{18}$-3') with N = C, G, T, or A; G; Y = C or T, and oligonucleotide Y (5'-CGAGGAGGATGTCGACCGGACC(T)$_{18}$-3'). PCR products were separated by gel electrophoresis on 1% gel containing 0.5% agarose/1% Triton X-100, and applied onto a 16% Tricine gel containing 0.5% fibrinogen and 2 mNIHU/ml thrombin. After 2-8 h thrombin inhibitory activity was visualized by clear areas of non-polymerized fibrin.

Amidolytic Assays and Kinetic Analysis—Protein concentration in eluates collected from different columns was determined using the Micro BCA Test-Kit (Pierce) according to the instruction manual.

To monitor haemadin purification, thrombin inhibitory activity was determined in a chromogenic assay. 100 µl of sample were preincubated at 37°C for 10 min with 100 µl of thrombin (0.4 NIHU/ml). 50 µl of 25 mM chromogenic substrate S-2238 (Kabi Vitrum) were added. After 5-10 min the reaction was stopped by addition of 20 µl of glacial acetic acid, and the change in optical density was measured at 405 nm.

Analogous assays were performed using trypsin, plasmin, chymotrypsin, FXa and FVa, and the corresponding chromogenic substrates S-2238, S-2251, S-2296, S-2222, and FVIIa-substrate, respectively.

Slow-binding and tight-binding inhibition assays of human $\alpha$-thrombin were performed at 25°C as previously described (18, 19). Trypsin (5 µg/ml) and Arg-Arg-AMC was used as substrate at a final concentration of 50 µM ($K_i$, $K_m$, $V_{max}$) were determined according to the theory of non-linear regression analysis of the data as previously described (18, 19).

Isolation of RNA and Construction of cDNA Library—Total RNA was isolated from whole leeches by the guanidinium thiocyanate method according to the manufacturer's guidelines (Stratagene). Poly(A) RNA was isolated by oligo(dT) affinity separation with the PolyATtract system (Promega). A cDNA library was constructed in lambda Uni-ZAP XR according to the manufacturer's guidelines (Stratagene). The complexity of the library was 1.8 $\times$ 10^9 by amplification by standard procedures (24).

PCR Cloning of the Haemadin cDNA—First strand cDNA was synthesized from 5 µg of total RNA with 1 µg of primer YXT (5'-CGAGGAGGATGTCGACCGGACC(T)$_{18}$-3'), using SuperScript reverse transcriptase (Gibco-BRL). The cDNA was purified from excess primers by elution with GeneClean (BiolOl, La Jolla, CA) and PCR amplified with oligonucleotide A (5'-GGAATGCATGCATGCTGCGAAGCGAC(T)$_{18}$-3') with N = C, G, T, or A; G; Y = C or T, and oligonucleotide Y (5'-CGAGGAGGATGTCGACCGGACC(T)$_{18}$-3'). PCR products were separated by gel electrophoresis on 1% agarose/TBS gel containing 0.5% fibrinogen and 2 mNIHU/ml thrombin. After 2-8 h thrombin inhibitory activity was visualized by clear areas of non-polymerized fibrin.

Amidolytic Assays and Kinetic Analysis—Protein concentration in eluates collected from different columns was determined using the Micro BCA Test-Kit (Pierce) according to the instruction manual.

To monitor haemadin purification, thrombin inhibitory activity was determined in a chromogenic assay. 100 µl of sample were preincubated at 37°C for 10 min with 100 µl of thrombin (0.4 NIHU/ml). 50 µl of 25 mM chromogenic substrate S-2238 (Kabi Vitrum) were added. After 5-10 min the reaction was stopped by addition of 20 µl of glacial acetic acid, and the change in optical density was measured at 405 nm.

Analogous assays were performed using trypsin, plasmin, chymotrypsin, FXa and FVa, and the corresponding chromogenic substrates S-2238, S-2251, S-2296, S-2222, and FVIIa-substrate, respectively.

Slow-binding and tight-binding inhibition assays of human $\alpha$-thrombin were performed at 25°C as previously described (18, 19). Trypsin (5 µg/ml) and Arg-Arg-AMC was used as substrate at a final concentration of 50 µM ($K_i$, $K_m$, $V_{max}$) were determined according to the theory of non-linear regression analysis of the data as previously described (18, 19).
A Thrombin-specific Inhibitor from H. sylvestris

**Fig. 1.** Preparative rHPLC separation of haemadin. Chromatography was carried out on a C-4 Hi-pore rHPLC column using 0.1% trifluoroacetic acid/water (solvent A) and 0.1% trifluoroacetic acid/acetonitril (solvent B) as eluents. Eluates of the thrombin-Sepharose were separated with a linear gradient slope of 0.5%/min and peaks were fractionated manually as indicated by bars. Dashed peaks (A–J) indicate fractions with thrombin inhibitory activity.

**Fig. 2.** SDS-gel electrophoresis of purified haemadin. Samples were pretreated with Tricine gel sample buffer and heated for 5 min at 95 °C. A, Coomassie Brilliant Blue stain of the 16% Tricine gel; lane 1, r-hirudin (25 pg); lane 2, haemadin (5 pg); lane 3, molecular mass standards (16.9, 14.4, 8.15, 6.2, and 2.5 kDa). B, zymographic analysis of haemadin. Tricine gels were shaken for 15 min in 1% Triton X-100 to remove excess SDS and put on a preformed fibrin plate. After incubation for 2–8 h at 37 °C, antithrombin activity was visualized by clear areas in the polymerized fibrin layer.

**J** with inhibitory activity could be isolated by this procedure.

Homogeneity of the major rHPLC peak fraction (peak F, Fig. 1) was shown on SDS-Tricine-polyacrylamide gels, and the molecular weight was estimated by comparison with standard proteins. A single band of about 5 kDa was stained with Coomassie Brilliant Blue as shown in Fig. 2A. To prove identity of the stained protein band with the isolated inhibitory activity, a novel, very sensitive overlay technique was performed. Therefore, an agarose gel containing human fibrinogen and thrombin was poured and overlayed with a detergent-treated Tricine gel, equivalent to the one shown in Fig. 2A. After 2–8 h, a clear band indicating thrombin inhibitory activity was visible that corresponded directly to the Coomassie-stained protein (Fig. 2B).

Thrombin specificity of the inhibitor was further revealed by determination of its activity toward other serine proteases. In these assays we tested plasmin, trypsin, chymotrypsin, FVIIa, and FXa with their corresponding chromogenic substrates. No inhibition by haemadin was detectable up to a 300-fold molar excess of haemadin over the respective protease (data not shown).

The NH$_2$-terminal amino acid sequences of the polypeptides with inhibitory activities that we found in several rHPLC peaks were determined by gas-phase sequencing. The amino acid sequences obtained from peaks B, E, F, and G (see Fig. 1) share the same 20 NH$_2$-terminal amino acid residues. From the major fraction F, we were able to determine a continuous stretch of 40 amino acid residues with a few undetermined positions. To verify these ambiguous amino acid positions we performed peptide mapping. This sequencing strategy allowed us to determine the following stretch of 45 NH$_2$-terminal amino acid residues of haemadin: NH$_2$-IRFGMGKVPCPDGEVYTCDGEGKICLYGQSCNDG-QCSDPKPSS... Database searches with this amino acid sequence did not reveal any significant similarities with other protease inhibitors indicating that haemadin represents a new type of thrombin inhibitor.

cDNA Cloning of Haemadin—For the evaluation of the therapeutic potential of the new thrombin inhibitor, larger amounts of this protein are necessary. Since the natural source is limiting, we proceeded into production of haemadin by recombinant DNA technology. For the molecular cloning of the haemadin cDNA, we followed a rapid amplification of cDNA ends-like strategy described previously (25) that we modified after Fritz et al. (26). In brief, two subsequent PCRs with nested oligonucleotide primers were used to amplify specific sequences from an oligo(dT)-primed first strand cDNA. The oligo(dT) primer used in first strand synthesis (YXT) is extended by 5' sequences that correspond to two separate but overlapping oligonucleotides (X, Y).

The NH$_2$-terminal amino acid sequence of haemadin was reverse translated and the deduced DNA sequence scanned for oligonucleotide sequences with low codon degeneracy. Two oligonucleotide guessmers, according to regions determined by amino acid residues 4 through 11 and 19 through 26 were synthesized (A and B, respectively). The calculated complex-
ities were 256 and 384 different sequences for mixture A and B, respectively.

First strand cDNA, primed with oligo(YXT) from *H. sylvestris* total RNA, was used as the template for "PCR amplification 1" with the primer pair B/X. The result of this amplification as checked on agarose gels was a smear of heterogeneously sized DNAs as shown in Fig. 3A, lane 1. The region corresponding to fragment sizes between 100 and 1000 bp was cut into eight individual gel slices, and aliquots of each were used as templates in subsequent PCRs with the nested primer pair B/X. Five out of eight amplifications showed a single band of about 280 bp (Fig. 3B, lanes 1–8). To verify this "PCR amplification 2" product as a haemadin cDNA fragment, this fragment was subcloned to generate plasmid pBS-HSTI.

Sequence analysis showed an open reading frame following oligonucleotide B that extends through the known haemadin sequence (Fig. 4A). It was completed by 12 additional amino acid residues, followed by a stop codon. The size of the 3' untranslated region was 75 nucleotides followed by a poly(A) tail. The consensus signal for polyadenylation AATAAA precedes the poly(A) tract by 21 nucleotides.

In order to clone the complete cDNA, we synthesized cDNA from *H. sylvestris* mRNA that we cloned unidirectionally into lambda Uni-ZAP XR arms. Oligonucleotide C, derived from opposite strand sequences downstream of the haemadin coding region (position 181–158, opposite strand, Fig. 4A) was synthesized, and a PCR reaction was carried out in combination with the reverse primer (Stratagene) present on the Uni-ZAP phage arms. As the template we took one phage lysate from the amplified leech library. The PCR resulted in a predominant DNA fragment of about 500 bp (Fig. 3C, lane 1). This amplification product was subcloned to generate plasmid pBS-HSTI and sequenced.

The haemadin cDNA sequence, as it was present in the cloned PCR product (C/reverse), is shown in Fig. 4B. The sequence starts with the first cDNA nucleotide following the cloning linker. The first ATG start codon leads into an uninterrupted reading frame of 77 amino acid residues. The first 20 amino acid residues most probably represent the signal peptide since they are not present in the mature protein and display the general consensus of a signal peptide (27). In addition, computer analysis of haemadin's hydrophobicity clearly characterizes this region as a prominent hydrophobic stretch.

**Heterologous Expression of Haemadin—** A commercially available *E. coli* system, GeneExpress, was chosen for the heterologous expression of haemadin. The pMAL-p2 vector (Fig. 5A) allows inducible periplasmic expression and therefore increases the probability for correctly folded recombinant proteins. To account for possible differences in the codon usage, a synthetic gene was made by oligonucleotide primer B (boxed) and extends through an open reading frame, a putative AATAAA polyadenylation signal (in italics) and a putative poly(A) tail, terminated by 20 residues that represent the complement of oligonucleotide X (boxed). The predicted amino acid sequence is shown in *three-letter code*. The region known from peptide sequence analysis is shown in *boldface letters*. Panel B, nucleotide and deduced amino acid sequence of the complete coding region of haemadin. The sequence starts at the 5' end of the cDNA and proceeds through the entire coding region for haemadin. The cDNA sequence is terminated by 24 nucleotides that represent the complement of oligonucleotide C (boxed). The predicted amino acid sequence is shown in *three-letter code*. A putative signal peptide is indicated by *italic letters*.
under the control of the isopropyl-β-D-thiogalactoside-inducible Pmal promoter. The coding region for the maltose-binding protein (malE) is preceded by a signal peptide (sp) and continues through the indicated tetrapeptide recognition sequence Ile-Glu-Gly-Arg for protease factor Xa. Directional cloning in Xmn I/BamHI leads to in frame fusion with haemadin (inset). Panel B, comparison between natural and recombinant haemadin. Coomassie Brilliant Blue staining of 15% Tricine gels. M, for molecular weight standards see legend to Fig. 2; lane 1, purified maltose-binding protein/haemadin fusion peptide before cleavage; lane 2, fusion peptide after cleavage with factor Xa; lane 3, purified recombinant haemadin; lane 4, purified natural haemadin.

**Table I**

**Kinetic constants for the inhibition of human α-thrombin by haemadin and r-haemadin**

<table>
<thead>
<tr>
<th>Ionic strength</th>
<th>NaCl</th>
<th>$K_e^a$</th>
<th>$k_{on}$</th>
<th>$k_{off}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>0.124</td>
<td>83.3 ± 36.9</td>
<td>ND$^a$</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>0.224</td>
<td>99.3 ± 26.8</td>
<td>1.24 ± 0.08</td>
<td>1.23 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>0.324</td>
<td>95.1 ± 11.0</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>0.224</td>
<td>210.0 ± 62.0$^a$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ From tight-binding inhibition experiments.

$^b$ From slow-binding inhibition experiments.

$^c$ ND, not determined.

$^d$ Performed with r-haemadin.

proteins were separated by mono-Q anion-exchange chromatography. By this procedure we generated about 25 mg of r-haemadin. Natural haemadin and r-haemadin seem to be identical as judged by SDS-PAGE analysis (Fig. 5B). Purified r-haemadin was sequenced and the intact NH$_2$ terminus was found to be independent on ionic strength. The $K_w$ was found to be in frame fusion with malE-thrombin.

**Interaction between Human α-Thrombin and Haemadin**

From tight-binding inhibition experiments, a 1:1 stoichiometry of thrombin-haemadin complexes was found. Steady-state velocities obtained at different inhibitor concentrations ([I] = 0.13 to 4.5 $\times$ $[E_0]$; $[S] = 50 \mu M$) were fitted by non-linear regression analysis as previously described (18, 19). The dissociation constant $K_I = 99 \pm 26 \times 10^{-15}$ M is in excellent agreement with the $K_w$ determined from slow-binding inhibition experiments (Table I). As an increase in the value of $K_I$ with an increase of the ionic strength of the buffer would indicate an important contribution of ionic interactions for haemadin-thrombin complex formation we performed tight-binding inhibition experiments. The results for haemadin, however, indicated that the $K_w$ of the thrombin-haemadin interaction is independent on ionic strength (Table I). In addition, the inhibition constant of r-haemadin was determined in tight-binding inhibition experiments. The $K_I$ of 210 ± 62 $\times 10^{-15}$ M is in agreement with that of the natural inhibitor indicating proper folding of r-haemadin in the malE-haemadin fusion protein.

To investigate the rate constants for thrombin-haemadin complex formation, we used the slow-binding kinetic approach. Data of a set of progress curves obtained with 10.4 pm enzyme for eight inhibitor concentrations ([I] = 10$^0$ to $60 \times [E_0]$) yielded values for the apparent first-order rate constant $k_{app}$. From the slope of a plot of $k_{app}$ versus [I] the dissociation rate constant $k_{off} = 20.0 \pm 3.0 \times 10^{-8}$ s$^{-1}$ was determined which yields the dissociation constant $K_I = k_{on}/k_{off} = 161.3 \pm 15.6 \times 10^{-15}$ M$^{-1}$. These data are consistent with $k_{off} = 1.24 \times 10^{-4}$ s$^{-1}$ obtained using the $K_w$ from tight-binding and the $k_{on}$ from slow-binding inhibition experiments.

**DISCUSSION**

In this paper we described the purification and cloning of a thrombin-specific inhibitor from *H. sylvestris* and its interaction with human α-thrombin. Although there is no apparent sequence homology of haemadin neither to serine protease inhibitors obeying the standard mechanism nor to hirudin and related inhibitors from blood sucking leeches, we like to compare this inhibitor with hirudin since both appear to exhibit common structural features, especially, the highly acidic carboxyl-terminal tail.

We determined the amount of thrombin inhibitory activity in homogenates of whole leeches to be about 60–85 NIHU/g leech. Since 5–10 leeches equal 1 g, the total antithrombin activity found per leech is much lower than that described for the well known thrombin inhibitor hirudin from *H. medicinalis* (5). Nevertheless, with 7,700 NIHU/mg protein based on calculations with material eluted from the thrombin-Sepharose, the specific activity of haemadin is in the range of that determined for hirudin.

The apparent molecular mass of haemadin determined by SDS-PAGE was about 5 kDa which correlates well to the molecular mass of 5.34 kDa calculated from the amino acid composition. To our knowledge, haemadin seems to be the smallest, thrombin-specific inhibitor to date (12, 13) which displays an inhibition constant as low as $10^{-33}$ M$^{-1}$.

The final HPLC step resulted in the detection of different thrombin inhibitory protein peaks. The components of the
main peaks were analyzed on SDS-PAGE and by NH₂-terminal sequence analysis. Although all of them displayed identical apparent molecular masses and identical NH₂ termini, the origin of different isoforms remains unclear. The differences could be due to minor amino acid exchanges in other parts of the molecule, making homogeneity, as it is known when the hirudin prepared from leech homogenates (28-31). Since a genomic blot analysis revealed only a single gene coding for haemadin (data not shown), the most probable explanation is the presence of subspecies or varieties of H. selystris in the input leech material. In a further approach the analysis of other cDNA clones might lead to their identification.

If haemadin forms complexes with α-thrombin in a hirudin-like manner, the following binding sites could be postulated.

(i) Crystal structure data of the thrombin complex with hirudin variant 2 indicate that the amino-terminal residues of hirudin stick into the active side pocket of thrombin (10). Hydrophobic interactions of hirudin residues Ile¹ and Tyr³ with the S2 subite and the ary1-binding site of thrombin, respectively, seem to be established, but for Thr² no direct interaction to amino acid residues within the S1 specificity pocket could be determined (10). Haemadin provides a positively charged arginine residue at this position and might therefore be able to form a salt bridge to Asp³ of thrombin. Interestingly, a V2R mutant improved the inhibitory properties of r-hirudin significantly (32), and an arginine residue in position 2 is also found in a natural hirudin-like protein from the leech H. mansfienšsis (13). Phe² of haemadin then could interact with the aryl-binding site of thrombin. A phenylalanine residue at this position seems to make more favorable interactions with thrombin than a tyrosine since the Y3F hirudin mutant has already been shown to display a 2-fold lower Kᵢ value for the thrombin-hirudin interactions (33).

(ii) Six of the 12 carboxyl-terminal amino acid residues by the CDNA analysis are identical with the hirudin carboxyl terminus. In hirudin, the carboxyl-terminal region binds to a surface groove of α-thrombin that has been termed the anion-binding exosite (17-19). The exosite is important for the recognition of macromolecular substrates like its interaction with fibrinogen (34, 35) and the endothelial cell surface receptor thrombomodulin (36). Recently, Vu and coworkers (37) proposed a thrombin receptor to interact with thrombin's exosite by a stretch of 11 amino acid residues containing 6 acidic residues. Despite these homologies haemadin lacks a tyrosine in its carboxyl-terminal part which in other cDNA clones might lead to their identification. The origin of different isoforms remains unclear. The difference in conformational changes in the presence of subspecies or varieties of H. selystris is a prerequisite for molecular modeling of low molecular weight thrombin inhibitors.

Structural as well as kinetic evidence provided by this study suggest distinct dissimilarities between haemadin and hirudin, and it will be of interest to perform further studies on the thrombin-haemadin interaction. Heterologous expression of r-haemadin will allow us to produce larger amounts of the inhibitor for pharmacological studies as well as for the crystallization with thrombin. Determination of the three-dimensional structure will provide us with a better understanding of thrombin inhibition mechanisms which is a prerequisite for molecular modeling of low molecular weight thrombin inhibitors.

Acknowledgements—We thank T. Brugger, S. Dollt, G. Fickartz, H. Heumann, D. Hölderle, A. Kiesel, M. Kern, K. Vogel, and U. Ruffer for excellent technical assistance, Dr. E. Fink for providing leeches, and Dr. T. Friedrich for valuable comments and discussion.

REFERENCES