



Identification and functional characterization of multiple haemadins and an oligomeric decorsin in the Asian land leech *Haemadipsa interrupta*

Christian Müller¹ · Dana Sponholz¹ · Céline Tolksdorf² · Bernhard H. Rauch² · Sebastian Kvist³

Received: 19 August 2024 / Accepted: 5 November 2024 / Published online: 21 November 2024
© The Author(s) 2024

Abstract

Haematophagous leeches rely on a broad variety of bioactive compounds to secure a sufficient blood meal from their vertebrate prey. Both the primary (platelet aggregation) and secondary (blood coagulation) haemostasis are the main targets of action. The platelet aggregation inhibitor decorsin was first described in the North American leech, *Macrobdella decora* Say, 1824, whereas the bivalent thrombin inhibitor hirudin was originally identified in the European medicinal leech, *Hirudo medicinalis* Linnaeus, 1758. Hirudin blocks both the catalytic site and the fibrinogen-binding site (exosite I) of thrombin. Haemadin of the Indian land leech *Haemadipsa sylvestris* Blanchard, 1894, is also a highly efficient bivalent thrombin inhibitor but blocks exosite II of thrombin. So far, only the archetypal form of haemadin from *H. sylvestris* has been purified and functionally characterized, and two putative haemadins have been identified in the salivary transcriptome of *Haemadipsa interrupta* Moore, 1835, a terrestrial leech inhabiting mainly the Malayan peninsula. Here, we describe the identification of ten additional putative haemadins in the transcriptomic data set of *H. interrupta*, first generated by another study. Furthermore, we identified a putative oligomeric decorsin, which represents the first finding of this anticoagulant in a haemadipsid leech. Both the putative decorsin and a selection of haemadins were expressed, purified, and functionally characterized. The putative haemadins displayed a broad spectrum of thrombin-inhibitory potencies, whereas the putative oligomeric decorsin was indeed a weak inhibitor of platelet aggregation.

Keywords Haemadin · Decorsin · Blood coagulation · Platelet aggregation · Hematophagous leeches

Handling Editor: Julia Walochnik

✉ Christian Müller
christian.mueller@uni-greifswald.de

Dana Sponholz
dana.sponholz@uni-greifswald.de

Céline Tolksdorf
celine.tolksdorf@uni-oldenburg.de

Bernhard H. Rauch
bernhard.rauch@uni-oldenburg.de

Sebastian Kvist
sebastian.kvist@nrm.se

¹ Animal Physiology, Zoological Institute and Museum, University of Greifswald, 17489 Greifswald, Germany

² Pharmacology and Toxicology, University Medicine Oldenburg, Carl Von Ossietzky University Oldenburg, 26129 Oldenburg, Germany

³ Research Division, Swedish Museum of Natural History, 104 05 Stockholm, Sweden

Introduction

Members of the leech genus *Haemadipsa* Tennent, 1859, are terrestrial leeches and inhabit tropical and sub-tropical rainforests across South and Southeast Asia (Borda and Siddall 2010). They are about 15–50 mm in length (Lai et al. 2011; Gąsiorek and Różycka 2017), possess three jaws (Sawyer 1986), are haematophagous, and mainly feed on vertebrate hosts (Fogden and Proctor 1985; Hanya et al. 2019) including humans (Stammers 1950; Wilson and Eisenberg 1982). Based on their feeding preferences, members of the genus *Haemadipsa* have been successfully applied as a source for iDNA isolation and subsequent mammalian biodiversity screening (Schnell et al. 2012; Drinkwater et al. 2021; Wiltling et al. 2022; Fahmy et al. 2023). During a blood meal, they can ingest up to 13 times their initial body weight (Fogden and Proctor 1985). As described for other haematophagous leeches like *Hirudo medicinalis* Linnaeus, 1758, *Haemadipsa* leeches secrete bioactive compounds during

feeding, in order to secure sufficient intake of blood. However, comparably little is known so far on the exact composition of the saliva that is secreted.

Strube et al. (1993) described haemadin from the Indian land leech *Haemadipsa sylvestris* Blanchard, 1894. Haemadin is a highly effective inhibitor of thrombin and resembles hirudin, the archetype of leech-derived thrombin inhibitors, in sequence and structure. Both factors are bivalent inhibitors of thrombin and block the catalytic site of the protease via their N-terminal segments. But whereas the N-terminal segment of hirudin comprises five amino acid residues, the respective counterpart in haemadin displays a length of nine amino acid residues. However, in both cases, an aromatic amino acid residue (tyrosine or phenylalanine) that is essential for function (Lazar et al. 1991; Strube et al. 1993; Acquasaliente et al. 2023) is located at position 3. In addition, the acidic C-terminal tail of haemadin does not interact with the exosite I (the fibrinogen-binding site) of thrombin (as is the case for hirudin) but instead blocks the exosite II (the heparin-binding site) (Richardson et al. 2000, 2002). Recent investigations revealed clear evidence that the interaction between haemadin and exosite II is only transient in nature and that the C-terminal tail prefers to bind to the exosite I of a nearby thrombin molecule instead (Acquasaliente et al. 2023).

Kvist et al. (2014) presented a salivary transcriptome of *Haemadipsa interrupta* Moore, 1935, a leech that is distributed across the Malayan peninsula. The data set was screened for the presence of sequences that may encode putative anticoagulants, and several promising candidates were identified including antistasin, guamerin, manillase, bdellin, eglin C, LDTI, and LAPP, respectively. In addition, two coding sequences of putative haemadins/hirudins were described. The first of these two sequences displayed a higher degree of sequence similarity to haemadin, whereas the second was more similar to hirudins (Kvist et al. 2014). However, none of the putative anticoagulants was functionally characterized. The first aim of the present study was hence to test whether or not the putative hirudins/haemadins of *H. interrupta* are indeed thrombin inhibitors.

The uptake and long-term storage of a blood meal by haematophagous leeches critically rely on the presence of different anticoagulants in their saliva, and inactivation or loss of the respective coding genes may have detrimental effects on survival. Consequently, the genomes of *H. medicinalis* and *Hirudinaria manillensis* Lesson, 1842, contain several copies of hirudin and hirudin-like factor (HLF) genes (Müller et al. 2016; Lukas et al. 2019). The second aim of the present study was hence to analyze whether or not the transcriptome of *H. interrupta* may contain additional sequences of putative hirudins/haemadins.

Interestingly, one of the putative haemadins of *H. interrupta* contains an RGD motif that is located within the acidic

C-terminal tail of the molecule. The RGD motif is known to mediate the binding of the respective protein to cellular receptors of the integrin family (D'Souza et al. 1991), including to the glycoprotein IIb/IIIa complex on platelets (Calvete 1995). The presence of a RGD motif is a typical feature of decorsins, a type of leech-derived factor that was first described in the North American leech *Macrobdella decora* Say, 1824 (Seymour et al. 1990) and subsequently also in *Limnobdella mexicana* Blanchard, 1893 (Pfordt et al. 2022). Hirudins and decorsins share conserved structural and genomic features (Min et al. 2010; Müller et al. 2019). In contrast to hirudins, decorsins do not inhibit thrombin but are inhibitors of platelet aggregation. The RGD motif mentioned above is located between the 5th and the 6th cysteine residue within the central globular domain of decorsins, but not within the C-terminal tail as observed for the first putative haemadin of *H. interrupta*. Nevertheless, it might be possible that the factor is indeed an inhibitor of platelet aggregation. The third aim of the present study was hence to evaluate whether or not the presence of an RGD motif within the C-terminal tail of a putative haemadin facilitates the inhibition of platelet aggregation.

All hirudins, haemadins, and decorsins described so far contain only one central globular domain and are hence “monomeric” molecules. The existence of a dimeric hirudin-like molecule (named tandem-hirudin, TH) was confirmed only recently (Lukas et al. 2022), and the authors developed a model for the generation of putative oligomeric hirudin-like molecules (multiplication hypothesis). However, so far, no representative of hirudin-like molecules with more than two repeats of the central globular domain could be identified. The fourth aim of the present study was hence to evaluate whether or not *H. interrupta* may encode dimeric or even oligomeric hirudin-like molecules.

Materials and methods

Transcriptomic data

Transcriptomic data of *Haemadipsa interrupta* were obtained from GenBank sequence read archive number SRX246952 (Kvist et al. 2014).

Bioinformatics tools

Basic Local Alignment Search Tool (BLAST) searches were performed using either the respective NCBI web portal (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) or BioEdit v7.2.5 (Hall 1999) and customized settings (word size and expected threshold values) for search algorithm parameters.

Multiple sequence alignments were generated using either ClustalX 2.1 (Larkin et al. 2007) or the CLC

Sequence Viewer software package v8.0 (QIAGEN, Aarhus, Denmark) at default settings. Alignments were exported as msf-files and further processed using GeneDoc v2.7 (Nicholas and Nicholas 1997). Signal peptide sequences were predicted using the Phobius web portal (Käll et al. 2007) and SignalP6.0 (Teufel et al. 2022).

Gene synthesis

cDNA fragments of putative haemadins/hirudins and the oligomeric decorsin were generated using the gene synthesis service of Synbio Technologies (Monmouth Junction, NJ, USA).

Amplification and cloning of haemadin/hirudin and decorsin cDNAs

For the amplification of putative haemadins/hirudins and the oligomeric decorsin, primers were derived from the respective transcriptome database sequences. A list of all primers that were used in the study is provided in Supplementary Information Table S1. PCR reactions were performed using Q5 high-fidelity DNA polymerase (New England Biolabs, Frankfurt a. M., Germany), fragments of relevant sizes were purified and cloned, and their sequences were determined.

Expression, purification, processing, and quantification of putative haemadins/hirudins and decorsins

The detailed protocol to clone cDNAs as well as to express, purify, and quantify the respective putative haemadins/hirudins and decorsin was described in a couple of recent publications (Müller et al. 2016, 2020a, b; Pfordt et al. 2022). Briefly, we applied an expression and purification system developed by QIAGEN (Hilden, Germany). The pQE30Xa vector encodes a factor Xa protease recognition site located between the His-tag coding region at the 5' side and the multiple cloning site at the 3' side. A subsequent factor Xa protease treatment cleaves off the His-tag and results in a recombinant protein that is devoid of any vector-derived amino acid residues at the N-terminus. Molar concentrations of final protein solutions were calculated by dividing the absorbance at 280 nm by the molar absorption coefficient according to the equation $\epsilon = (nW \times 5500) + (nY \times 1490) + (nC \times 125)$ (Gill and von Hippel 1989; Pace et al. 1995). Representative gel images that illustrate the entire expression and purification process for three recombinant factors are provided in Supplementary Information Fig. S1.

Blood coagulation assays

To verify the thrombin-inhibitory potency of putative haemadins/hirudins, we performed the thrombin time test (TT; reference range 16.8–21.4 s) using a BFT II analyzer (Siemens Healthcare, Erlangen, Germany). All steps were carried out according to the manufacturer's instructions. Protein samples were diluted with buffer to reach final concentrations in the reaction assays of 3.2 $\mu\text{mol/l}$ or 0.32 $\mu\text{mol/l}$, respectively. The desired amount of substrate was directly transferred into the cuvette immediately before the plasma was added. Dade® Ci-Trol® 1 (Siemens Healthcare, Erlangen, Germany) was used as standardised human plasma. The incubation of reaction mixtures was carried out at 37.4 °C. Measurements that exceeded 300 s were stopped and considered as complete inhibition of clot formation. Coagulation tests were performed in three technical replicates.

Platelet aggregation assays

All assays were performed with human blood samples that were obtained from a healthy volunteer (CM) after written informed consent and approval from the institutional ethics committee. The blood collection, blood preparation, and subsequent experiment followed the procedure as already described in Lukas et al. (2019) and Pfordt et al. (2022) with few modifications. Briefly, 10 ml of venous blood was taken from the antecubital vein using an S-Monovette® (Sarstedt, Nümbrecht, Germany) prefilled with citrate buffer. The first centrifugation step of the blood collection tube was performed at 200 g for 20 min. After centrifugation, the supernatant (platelet-rich plasma, PRP) was transferred, and the remaining blood was centrifuged again for 10 min at 2000 g. The supernatant was dedicated as platelet-poor plasma (PPP), transferred, and used as a reference value for maximal platelet aggregation. Measurements were performed using a TA-8 V aggregometer (Diagnostica Stago S.A.S., Asnières-sur-Seine, France). The snake venom-derived platelet aggregation inhibitors tirofiban and eptifibatid (Sigma-Aldrich, Taufkirchen, Germany) were used as positive controls. PRP was pre-incubated with the respective test and control compounds (final concentration 3.2 $\mu\text{mol/l}$) or buffer for 1 min at 37 °C. For the measurement, PRP was then transferred into cuvettes and stimulated with ADP (200 $\mu\text{mol/l}$; Hart Biologicals, Hartlepool, UK; final concentration 5 $\mu\text{mol/l}$) after 1 min of runtime. The final volume of each cuvette was 250 μL of diluted PRP. All experiments were performed at 37 °C over a time period of 10 min. Maximal aggregation in percentage and maximum slope of the curve in percentage per minute were calculated as quantitative output parameters (Zhou and Schmaier 2005). Platelet aggregation tests were performed in two to four technical replicates.

Results

Identification of putative haemadins/hirudins in *H. interrupta*

We thoroughly re-investigated the raw reads of *H. interrupta* (SRX246952) and confirmed the observations of Kvist et al. (2014). Furthermore, we were able to identify the coding sequences of ten additional putative haemadin/hirudin sequences. The lengths of all 12 putative haemadins/hirudins (named Hint_V1–V12) varied considerably between 39 and 66 amino acid residues (without signal peptides), and the isoelectric points (pI values) covered a range between 3.97 and 8.60. A summary of biochemical properties and the N-terminal sequences of all factors are provided in Table 1.

The degrees of amino acid sequence identity/ similarity/gaps between haemadin of *H. sylvestrus* and the putative haemadins/hirudins of *H. interrupta* varied between

77/84/0% (Hint_V1) and 18/30/28% (Hint_V2 = contig_c1987_1). Within the putative haemadins/hirudins of *H. interrupta*, Hint_V1 and Hint_V5 were almost identical (94/96/0% sequence identity/similarity/gaps), whereas Hint_V2 was distantly related to all the other factors (12–20/20–28/13–44% sequence identity/similarity/gaps). A multiple sequence alignment of all putative haemadin/hirudin variants of *H. interrupta* is shown in Fig. 1a; a compilation of all pairwise sequence comparisons can be found in Supplementary Information Table S2. Five factors (Hint_V3 = contig_c7044_1, _V4, _V10, _V11, and _V12) contained the RGD motif within the C-terminal tail (see Fig. 1a). Notably, three factors (Hint_V1, _V3, and _V5) contained a phenylalanine (F) residue at position 3 of the mature protein (marked with an arrow in Fig. 1a). An aromatic amino acid residue (tyrosine, phenylalanine or tryptophan) at position 3 is essential for binding of haemadin/hirudins to the active center of thrombin (Lazar et al. 1991; Strube et al. 1993; Acquasaliente et al. 2023).

Table 1 Molecular and structural properties of hirudins of *H. medicinalis* (Hmed_HV1) and *L. mexicana* (Lmex_HV2), haemadin of *H. sylvestrus*, and of putative haemadins of *H. interrupta* (Hint_V1–12)

Factor	read	pI	length in aa	MW	N-terminus	RGD-motif
Hmed_HV1		4,12	65	6968.54	VV Y TD	-
Lmex_HV2		4,64	66	7502.19	VH F MP	-
Haemadin		4.07	57	6257.86	IR F GMGKVP	-
Hint_V1*	4668	3.97	57	6209.82	VR F GMGNVP	-
Hint_V2*	105860	4.24	56	6672.27	AR I F F PK	-
Hint_V3*	22944	4.54	66	7326.21	VR F GMGNVP	+
Hint_V4*	86849	4.51	54	5762.36	VRTQ I DDAPP	+
Hint_V5	88009	4.11	57	6233.80	VR F GMGNVP	-
Hint_V6	77515	6.24	47	5360.17	QE I TQTKK	-
Hint_V7	100720	4.29	52	5990.76	VR S EMKD	-
Hint_V8	19306	4.51	40	4286.86	MA S GAQ	-
Hint_V9	47621	4.94	39	4541.28	MM P EFN F P	-
Hint_V10	23816	8.60	53	5871.62	M R SQ M NV P K	+
Hint_V11	52567	6.04	53	5652.36	F D SQ M NV P K	+
Hint_V12	33695	6.11	60	6906.68	V R S Y C Y E E E L P E	+

The read numbers refer to the respective reads in SRX246952. The sequences of the N-termini are provided, and the presence (+) or absence (-) of the RGD motif within the C-terminal tail is indicated. The aromatic amino acid residues tyrosine (Y) or phenylalanine (F) within the N-termini are highlighted in yellow. Asterisks indicate the factors that were selected for further functional analyses

pI isoelectric point, MW molecular mass

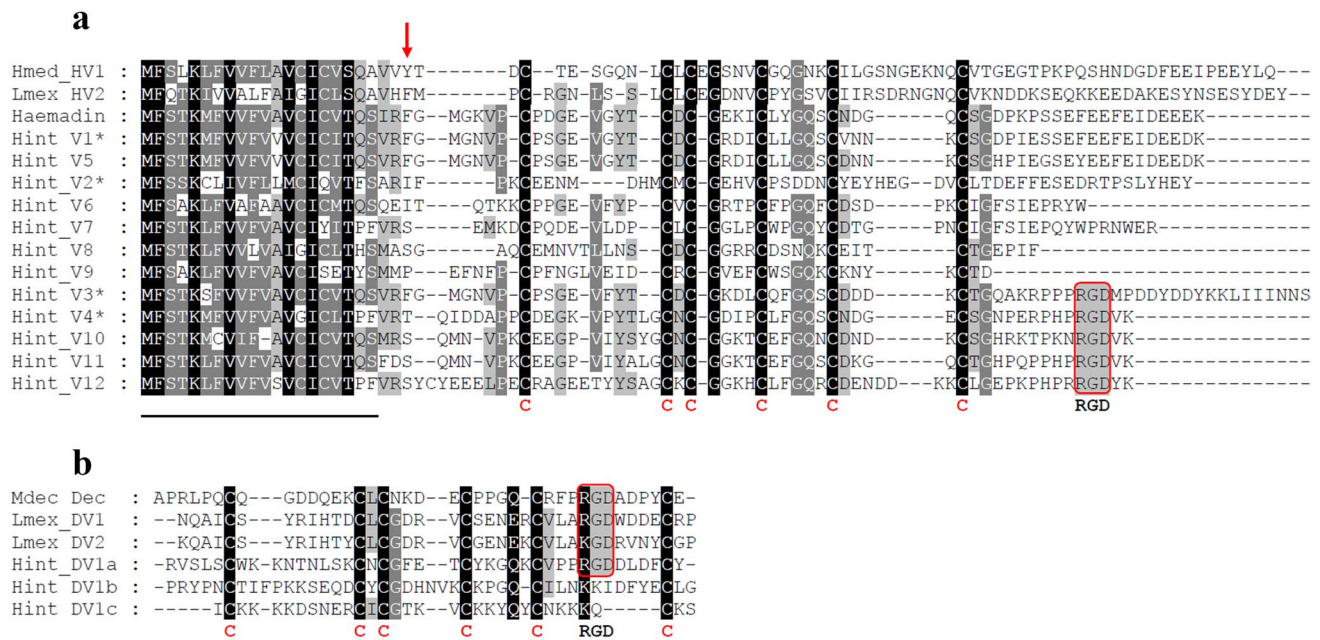


Fig. 1 Multiple amino acid sequence alignments. **a** Hirudin variants HV1 of *H. medicinalis* (Hmed_HV1) and HV2 of *L. mexicana* (Lmex_HV2), haemadin of *H. sylvestrus* and putative haemadin variants Hint_V1–12 of *H. interrupta*. **b** Decorsin of *M. decora* (Mdec_Dec), decorsin variants 1 and 2 of *L. mexicana* (Lmex_DV1+2), and the three globular domains of the putative oligomeric ornatin of *H. interrupta* (Hint_DV1a-c). A black background indicates fully

conserved residues; a gray background indicates partially conserved residues. The six conserved cysteine residues giving rise to the three-dimensional structure are marked in bold and red. The canonical RGD/KGD motif in decorsins is marked in bold and boxed. The signal peptide sequences are underlined. Asterisks indicate the factors that were selected for further functional analyses. Abbreviations are used according to the IUPAC code

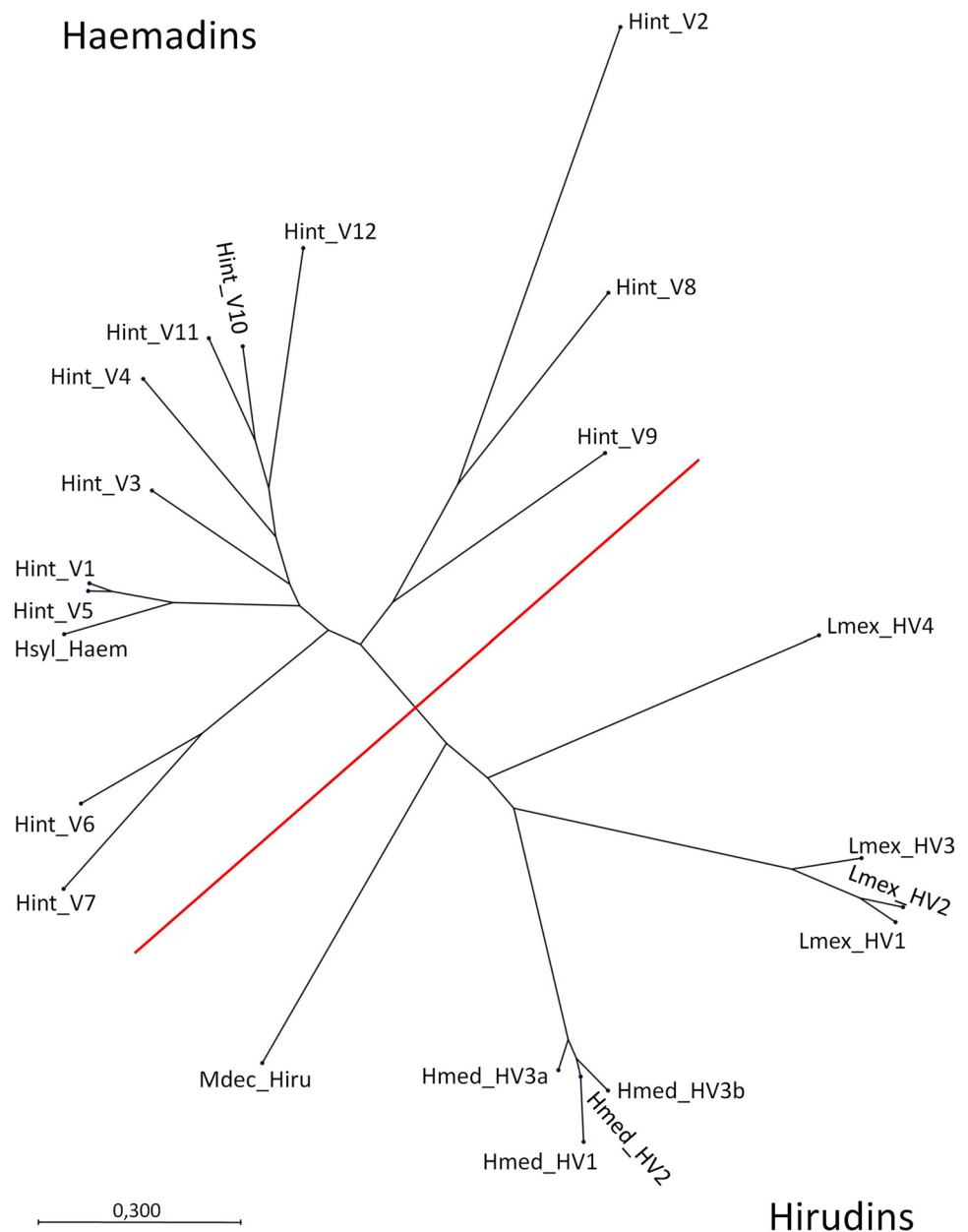
To further evaluate the relationships between the putative haemadins/hirudins of *H. interrupta* we performed phylogenetic investigations and included haemadin of *H. sylvestrus* and hirudins of *H. medicinalis*, *M. decora* and *L. mexicana* into our analysis. The resulting unrooted phylogenetic tree displays a clear separation between hirudins and (putative) haemadins of *H. interrupta* including Hint_V2 (= contig_c1987_1) (see Fig. 2). Interestingly, the RGD motif containing haemadin sequences (Hint_V3, _V4, and _V10-12) nicely group together.

Identification of a putative oligomeric decorsin in *H. interrupta*

To our surprise, the re-analysis of the data revealed not only the presence of a great diversity of putative haemadin sequences in the salivary transcriptome of *H. interrupta*, but also a sequence of a putative oligomeric decorsin. As already stated in the Introduction section, decorsins are inhibitors of platelet aggregation and contain an RGD motif that is located between the 5th and 6th cysteine residues within the central globular domain. *H. interrupta* expresses a putative decorsin (named Hint_DV1) that is composed of a short N-terminus of five amino acid

residues in length followed by three repeats of the (central) globular domain (named Hint_DV1a-c, respectively) and an elongated C-terminal tail. The complete amino acid sequence of Hint_DV1 and a schematic drawing that illustrates its putative secondary structure are provided in Supplementary Information Fig. S2. The composition of Hint_DV1 with three repeats of the (central) globular domain leads to the presence of 18 cysteine residues within the molecule. Hint_DV1 can be designated as an oligomeric decorsin, and this is to our knowledge the first description of a member of the hirudin superfamily that is composed of more than two repeats of the central globular domain. Interestingly, only the first globular domain repeat (Hint_DV1a) contains an RGD motif between the 5th and the 6th cysteine residue (see Fig. 1b and Supplementary Information Fig. S2). The degrees of sequence identity/similarity/ gaps between the decorsins of *M. decora*, *L. mexicana*, and Hint_DV1a of *H. interrupta* are rather low and within a range of 24–30/34–38/9%, respectively. A multiple amino acid sequence alignment of decorsins and the three repeats of the (central) globular domain of Hint_DV1 (Hint_DV1a-c) is shown in Fig. 1b, a compilation of all pairwise sequence comparisons can be found in Supplementary Information Table S3.

Fig. 2 Unrooted phylogenetic tree based on amino acid sequences including the signal peptides of haemadins and hirudins generated using the neighbor-joining algorithm. Hsyl_Haem = haemadin of *H. sylvestrus*; Hint_V1–12 = putative haemadins of *H. interrupta*; Mdec_Hiru = hirudin of *M. decora*; Lmex_V1–4 = hirudin variants 1–4 of *L. mexicana*; Hmed_HV1–3 = hirudin variants 1–3 of *H. medicinalis*. The scale bar indicates substitutions per site. The red bar illustrates the separation of the haemadin and the hirudin clades



Functional characterization I: thrombin inhibition

Some, but not all putative haemadins of *H. interrupta* displayed structural and biochemical features that are mandatory for thrombin inhibition. Of special importance in that context is the composition of the N-terminus with a special emphasis on the first three amino acid residues, the acidic composition of the C-terminal tail, and an overall isoelectric point of about 4. Based on these criteria (see Table 1 for details), we decided to functionally characterize the following selection of putative haemadins of *H. interrupta*: Hint_V1 (most similar to haemadin), Hint_V2 (= contig c1987_1), Hint_V3 (= contig c7044_1, RGD motif), and Hint_V4 (RGD motif). All factors were successfully

expressed, purified, processed, and tested in the thrombin time test. Both Hint_V1 and Hint_V3 were strong inhibitors of thrombin, whereas Hint_V2 displayed a weak, but measurably inhibitory potency. Hint_V4, however, did not inhibit the activity of thrombin (see Fig. 3).

Functional characterization II: platelet aggregation

To measure the inhibition of platelet aggregation we selected the haemadin variants Hint_V1 (no RGD motif), Hint_V3, and Hint_V4 (both with RGD motifs) of *H. interrupta*. The snake venom-derived platelet aggregation inhibitors tirofiban and eptifibatid were used as positive controls. As

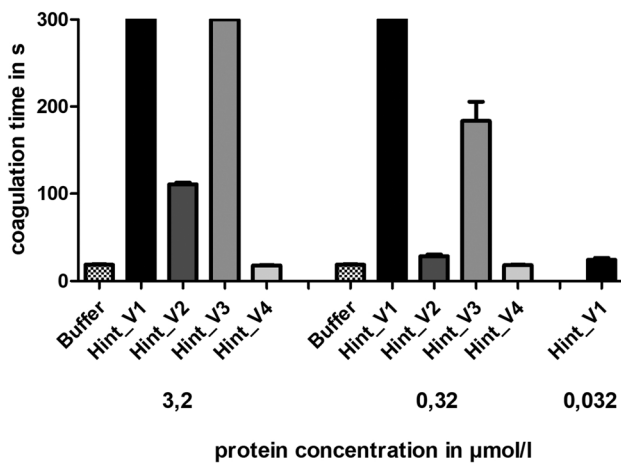


Fig. 3 Standard blood coagulation assays of putative haemadin variants of *H. interrupta* (Hint_V1–4) using the thrombin time test (TT). Results are the mean of three independent measurements

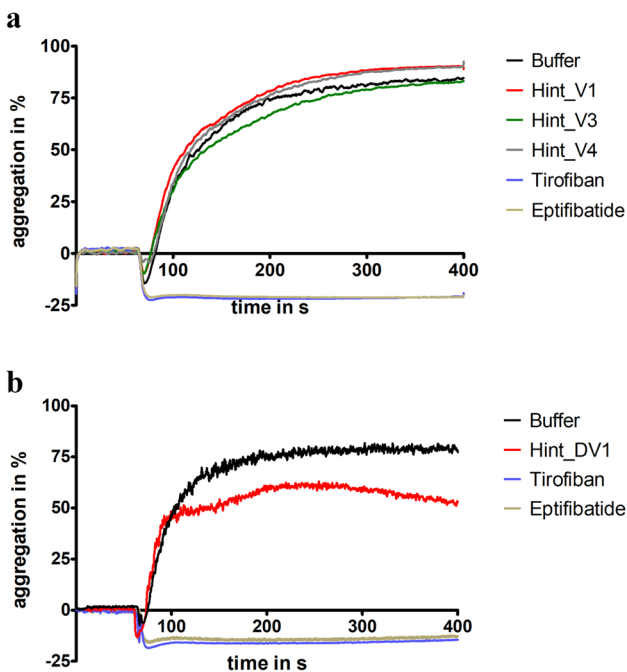


Fig. 4 Standard platelet aggregation assays of putative haemadin variants and the putative oligomeric decorsin of *H. interrupta*. **a** Platelet aggregation curves of putative haemadins Hint_V1, Hint_V3, and Hint_V4. **b** Platelet aggregation curves of the putative oligomeric decorsin Hint_DV1. Eptifibatide and tirofiban were used as positive control substances for the complete inhibition of aggregation; buffer was used as a negative control. Platelet aggregation was induced by the addition of ADP to a final concentration of 5 µmol/l. Results are the mean of two independent measurements

can be seen in Fig. 4a, none of the haemadins of *H. interrupta* displayed an inhibitory effect on platelet aggregation.

In addition, we also expressed, purified, and tested the putative oligomeric decorsin of *H. interrupta*. The inhibitory

effect of Hint_DV1 on platelet aggregation was less pronounced compared to tirofiban and eptifibatide, but clearly detectable (see Fig. 4b).

Discussion

Haemadipsa interrupta is a terrestrial leech that feeds on the blood of vertebrate hosts. Unsurprisingly, the analysis of a salivary gland transcriptome of the leech revealed the expression of a broad variety of putative anticoagulants, among them two representatives of the haemadin/hirudin-type of thrombin inhibitors (Kvist et al. 2014). However, a thorough re-investigation of the data set uncovered the presence of several additional putative haemadins/hirudins in addition to a putative oligomeric decorsin. The variety of putative anticoagulants in *H. interrupta* is obviously much higher than originally assumed. In the present study, we focused on haemadins/hirudins and did not search for other types of anticoagulants (e.g. antistasins), but it might well be that our observations represent only the tip of the iceberg. Our results on the great diversity of putative haemadins/hirudins in *H. interrupta* are in line with similar observations in other haematophagous leeches including *H. medicinalis* (Müller et al. 2016), *Poecilobdella (Hirudinaria) manillensis* Lesson, 1842 (Müller et al. 2017), *Whitmania pigra* Blanchard, 1887 (Müller et al. 2022), and *Limnobdella mexicana* (Pfordt et al. 2022) and hold true also for other classes of major anticoagulants in leech saliva like the hirustasin/antistasin family (Iwama et al. 2021; Liu et al. 2023). However, a solely sequence-based prediction of a function is highly problematic and may be misleading. For that reason, we performed both coagulation and platelet aggregation assays on a selection of the putative haemadins/hirudins in *H. interrupta* to verify their biological function.

The archetypes of haemadin and hirudin are both highly effective bivalent inhibitors of thrombin (Strube et al. 1993; Bichler and Fritz 1991), yet their modes of action are slightly different. Whereas the acidic C-terminal tail of hirudin blocks exosite I of thrombin (Grütter et al. 1990), the tail of haemadin interacts with exosite II (Richardson et al. 2000). Both factors block the catalytic centre of thrombin by inserting their N-termini into the active site cleft. Of special importance in that context are the first three amino acid residues, especially the third residue has to be an aromatic amino acid residue (Lazar et al. 1991; Strube et al. 1993; Acquasaliente et al. 2023). Notably, only three out of the 12 putative haemadins/hirudins of *H. interrupta* (namely Hint_V1, Hint_V3, and Hint_V5, see Table 1) contain a phenylalanine residue at position 3, Hint_V2 at position 4. We also found that Hint_V1 and Hint_V3 were highly potent thrombin inhibitors, whereas the inhibitory effect of Hint_V2 was much reduced, but still observable. Hint_V4 contains a

threonine residue at position 3 and displayed no inhibitory potency at all. A comparable substitution of phenylalanine (or tyrosine) with threonine at position 3 in hirudin variant HV1 of *H. medicinalis* increases the K_i value of the respective recombinant factor by more than three orders of magnitude (Lazar et al. 1991). We have not yet functionally tested the remaining putative haemadins/hirudins of *H. interrupta*, but based on the structural and biochemical features of the respective factors (see Table 1) and the degrees of sequence identities/similarities (see Table S1), it seems reasonable that only Hint_V5 is a thrombin inhibitor. The biological functions of Hint_V6-V12 remain obscure; they may not be haemadins (in a narrow sense) but rather haemadin-like factors (a comparable term was recently introduced for the hirudin-like factors, HLFs; Müller et al. 2019).

Haemadin and hirudin do not only differ in their binding modes to thrombin (see above), but also in the length of their N-termini: haemadin comprises nine amino acid residues in front of the first cysteine residue, hirudin only five. The lengths of the respective N-termini of Hint_V1-V12 differ between six and 12 amino acid residues (see Table 1 for details). The highly potent thrombin-inhibitory variants Hint_V1 and Hint_V3 comprise nine residues (identical to haemadin), the less effective variant Hint_V2 has six residues, and the inactive variant Hint_V4 has 10 residues. It would be interesting to see whether or not a modification of the N-terminus (e.g., removal of the residues 6–9 in Hint_V1) would alter the thrombin-inhibitory potency. Acquasaliente et al. (2023) constructed a deletion mutant of haemadin, named haemanorm, a 29-amino acid peptide containing the amino acid residues 1–9 and 41–57 of haemadin. Haemanorm is a highly effective bivalent thrombin inhibitor that blocks both the active site and, surprisingly, the exosite I of thrombin (Acquasaliente et al. 2023).

Interestingly, neither the variable lengths of the N-termini in Hint_V1-V12 nor the presence/absence of the RGD motif within the C-terminal tail altered the allocation of the factors in phylogenetic analysis: all haemadins/haemadin-like factors formed a clade that was clearly separated from the hirudin clade (see Fig. 2). The question whether or not the haemadins/haemadin-like factors nonetheless belong to the hirudin superfamily crucially depends on their gene structures. All hirudins and hirudin-like factors analyzed so far share a common gene structure that is composed of four exons and three introns (Scacheri et al. 1993; Müller et al. 2017, 2022). The gene structures of haemadins and haemadin-like factors remain to be determined. It is, however, reasonable to assume that their genes share the same basic structure.

Decorsins (and the structurally and functionally equivalent ornatins; Mazur et al. 1991) are inhibitors of platelet aggregation and belong to the hirudin superfamily as well (Müller et al. 2019). They are characterized by the

presence of RGD (or in some cases KGD; see Fig. 1b) motifs that are located within the central globular domains between the 5th and the 6th cysteine residue at the apex of an exposed hairpin loop and facilitate the binding to integrin receptors (Krezel et al. 2000). Five out of the 12 haemadins/haemadin-like factors of *H. interrupta* (Hint_V3, Hint_V4, and Hint_V10-12) contain RGD motifs that are located within the C-terminal tail (see Fig. 1a and Table 1). The tail is highly flexible in solution (Folkers et al. 1989), and the RGD motif within the tail may hence come in contact and bind to integrin receptors as well. However, none of the RGD motif-containing factors displayed any evidence of an inhibitory effect on platelet aggregation (see Fig. 4a). It might be that the binding of an RGD motif located in the tail region of the protein is of low affinity and thus unstable, and/or those additional structural components within the central globular domain are necessary for binding.

In contrast, the oligomeric decorsin Hint_DV1 displayed a weak but detectable inhibitory effect on platelet aggregation. The inhibitory effect was much less pronounced compared to both the positive controls tirofiban and eptifibatid and the decorsins of *L. mexicana* and *Haementeria vizottoi* Castro, 1971 (Pfordt et al. 2022). However, our results represent the very first evidence at all for the biological activity of a hirudin-like factor that is composed of more than one central globular domain. Tandem hirudin (TH) of *P. manillensis* was described just recently (Lukas et al. 2022), but no biological function could be assigned to the factor. Furthermore, the identification of Hint_DV1 supports the hypothesis of hirudin multiplication (Lukas et al. 2022). Again, it would be highly interesting to determine and analyze the gene structure of Hint_DV1. The comparably low inhibitory potency of our preparation of Hint_DV1 could well be due to the imperfect formation of the secondary structure of the protein during heterologous expression in a bacterial host. The structures of the central globular domains of all hirudins, HLFs, haemadins, and decorsins entirely depend on the correct formation of the three disulfide bonds. The folding of hirudin is a complex process (Chatrenet and Chang 1993) and perturbations may result in the formation of scrambled and hence inactive isoforms (Chatrenet and Chang 1992; Chang et al. 1995). The situation is even more complicated in the case of Hint_DV1 which contains 18 cysteine residues instead of six and requires the correct formation of nine disulfide bonds.

The identification of such a broad variety of haemadins and haemadin-like factors in addition to an oligomeric decorsin in the land leech *H. interrupta* certainly expands our knowledge of that particular class of leeches and may also help to better understand the complex biology of haematophagous leeches in general.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00436-024-08404-w>.

Acknowledgements The authors are thankful to the colleagues who generated and provided the transcriptome dataset explored in the present study. In addition, we would like to thank all members of the Animal Physiology Working Group at the University of Greifswald for their help and support. A special thanks goes to Jan-Peter Hildebrandt and Undine Lauf for their critical reading of the manuscript.

Author contribution Christian Müller conceived the ideas, designed the methodology, and performed the sequence-based investigations; Christian Müller, Dana Sponholz, and Céline Tolksdorf performed the experiments; Christian Müller and Sebastian Kvist analyzed the data and drafted the manuscript; Christian Müller and Bernhard H. Rauch supervised the experimental work.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability Original sequence data are deposited in GenBank under sequence read archive number SRX246952. The respective read numbers are provided within the manuscript.

Declarations

Ethical approval The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the University Medicine Oldenburg, Carl von Ossietzky University Oldenburg (ethics vote 2024–034).

Consent to participate and consent for publication Written informed consent was obtained from all subjects involved in the study.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acquasaliente L, Pierangelini A, Pagotto A, Pozzi N, De Filippis V (2023) From haemadin to haemanorm: synthesis and characterization of full-length haemadin from the leech *Haemadipsa sylvestris* and of a novel bivalent, highly potent thrombin inhibitor (haemanorm). *Protein Sci* 32(12):e4825. <https://doi.org/10.1002/pro.4825>
- Bichler J, Fritz H (1991) Hirudin, a new therapeutic tool? *Ann Hematol* 63(2):67–76. <https://doi.org/10.1007/BF01707275>
- Borda E, Siddall ME (2010) Insights into the evolutionary history of Indo-Pacific bloodfeeding terrestrial leeches (Hirudinida: Arhynchobdellida: Haemadipsidae). *Invert System* 24(5):456–472. <https://doi.org/10.1071/IS10013>

- Calvete JJ (1995) On the structure and function of platelet integrin $\alpha_{IIb}\beta_3$, the fibrinogen receptor. *Proc Soc Exp Biol Med* 208(4):346–360. <https://doi.org/10.3181/00379727-208-43863a>
- Chang JY, Schindler P, Chatrenet B (1995) The disulfide structures of scrambled hirudins. *J Biol Chem* 270(20):11992–11997. <https://doi.org/10.1074/jbc.270.20.11992>
- Chatrenet B, Chang JY (1992) The folding of hirudin adopts a mechanism of trial and error. *J Biol Chem* 267(5):3038–3043. [https://doi.org/10.1016/S0021-9258\(19\)50691-3](https://doi.org/10.1016/S0021-9258(19)50691-3)
- Chatrenet B, Chang JY (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. *J Biol Chem* 268(28):20988–20996. [https://doi.org/10.1016/S0021-9258\(19\)36883-8](https://doi.org/10.1016/S0021-9258(19)36883-8)
- D'Souza SD, Ginsberg MH, Plow EF (1991) Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends Biochem Sci* 16(7):246–250. [https://doi.org/10.1016/0968-0004\(91\)90096-e](https://doi.org/10.1016/0968-0004(91)90096-e)
- Drinkwater R, Jucker T, Potter JHT, Swinfield T, Coomes DA, Slade EM, Gilbert MTP, Lewis OT, Bernard H, Struebig MJ, Clare EL, Rossiter SJ (2021) Leech blood-meal invertebrate-derived DNA reveals differences in Bornean mammal diversity across habitats. *Mol Ecol* 30(13):3299–3312. <https://doi.org/10.1111/mec.15724>
- Fahmy M, Andrianoely D, Wright PC, Hekkala E (2023) Leech-derived iDNA complements traditional surveying methods, enhancing species detections for rapid biodiversity sampling in the tropics. *Environmental DNA* 5(6):1557–1573. <https://doi.org/10.1002/edn3.474>
- Fogden SCL, Proctor J (1985) Notes on the feeding of land leeches (*Haemadipsa zeylanica* Moore and *H. picta* Moore) in Gunung Mulu National Park, Sarawak. *Biotropica* 17(2):172–174. <https://doi.org/10.2307/2388511>
- Folkers PJ, Clore GM, Driscoll PC, Dodt J, Köhler S, Gronenborn AM (1989) Solution structure of recombinant hirudin and the Lys-47 - Glu mutant: a nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing study. *Biochemistry* 28(6):2601–2617. <https://doi.org/10.1021/bi00432a038>
- Gąsiorek P, Różycka H (2017) Feeding strategies and competition between terrestrial *Haemadipsa* leeches (Euhirudinea: Arhynchobdellida) in Danum Valley rainforest (Borneo, Sabah). *Folia Parasitol* 64:031. <https://doi.org/10.14411/fp.2017.031>
- Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319–326. [https://doi.org/10.1016/0003-2697\(89\)90602-7](https://doi.org/10.1016/0003-2697(89)90602-7)
- Grütter MG, Priestle JP, Rahuel J, Grossenbacher H, Bode W, Hofsteenge J, Stone SR (1990) Crystal structure of the thrombin-hirudin complex: a novel mode of serine protease inhibition. *EMBO J* 9(8):2361–2365. <https://doi.org/10.1002/j.1460-2075.1990.tb07410.x>
- Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41(2):95–98
- Hanya G, Morishima K, Koide T, Otani Y, Hongo S, Honda T, Okamura H, Higo Y, Hattori M, Kondo Y, Kurihara Y, Jin S, Otake A, Shiroishi I, Takakuwa T, Yamamoto H, Suzuki H, Kajimura H, Hayakawa T, Suzuki-Hashido N, Nakano T (2019) Host selection of hematophagous leeches (*Haemadipsa japonica*): implications for iDNA studies. *Ecol Res* 34(6):842–855. <https://doi.org/10.1111/1440-1703.12059>
- Iwama RE, Tessler M, Siddall ME, Kvist S (2021) The origin and evolution of antistasin-like proteins in leeches (Hirudinida, Clitellata). *Genome Biol Evol* 13(1):evaa242. <https://doi.org/10.1093/gbe/evaa242>
- Käll L, Krogh A, Sonnhammer ELL (2007) Advantages of combined transmembrane topology and signal peptide prediction - the Phobius web server. *Nucleic Acids Res* 35(Web Server issue):W429–432. <https://doi.org/10.1093/nar/gkm256>
- Krezel AM, Ulmer JS, Wagner G, Lazarus RA (2000) Recombinant decorsin: dynamics of the RGD recognition site. *Protein Sci* 9(8):1428–1438. <https://doi.org/10.1110/ps.9.8.1428>

- Kvist S, Brugler MR, Goh TG, Giribet G, Siddall ME (2014) Pyrosequencing the salivary transcriptome of *Haemadipsa interrupta* (Annelida: Clitellata: Haemadipsidae): anticoagulant diversity and insight into the evolution of anticoagulation capabilities in leeches. *Invertebr Biol* 133(1):74–98. <https://doi.org/10.1111/ivb.12039>
- Lai Y-T, Nakano T, Chen J-H (2011) Three species of land leeches from Taiwan, *Haemadipsa rjukjuana* comb. n., a new record for *Haemadipsa picta* Moore, and an updated description of *Tritetrab-della taiwana* (Oka). *ZooKeys* 139:1–22. <https://doi.org/10.3897/zookeys.139.1711>
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>
- Lazar JB, Winant RC, Johnson PH (1991) Hirudin: amino-terminal residues play a major role in the interaction with thrombin. *J Biol Chem* 266(2):685–688. [https://doi.org/10.1016/S0021-9258\(17\)35224-9](https://doi.org/10.1016/S0021-9258(17)35224-9)
- Liu Z, Zhao F, Huang Z, Hu Q, Meng R, Lin Y, Qi J, Lin G (2023) Revisiting the Asian buffalo leech (*Hirudinaria manillensis*) genome: focus on antithrombotic genes and their corresponding proteins. *Genes* 14:2068. <https://doi.org/10.3390/genes14112068>
- Lukas P, Wolf R, Rauch BH, Hildebrandt J-P, Müller C (2019) Hirudins of the Asian medicinal leech, *Hirudinaria manillensis*: same same, but different. *Parasitol Res* 118(7):2223–2233. <https://doi.org/10.1007/s00436-019-06365-z>
- Lukas P, Melikian G, Hildebrandt J-P, Müller C (2022) Make it double: identification and characterization of a Tandem-Hirudin from the Asian medicinal leech *Hirudinaria manillensis*. *Parasitol Res* 121(10):2995–3006. <https://doi.org/10.1007/s00436-022-07634-0>
- Mazur P, Henzel WJ, Seymour JL, Lazarus RA (1991) Ornatin: potent glycoprotein IIb-IIIa antagonists and platelet aggregation inhibitors from the leech *Placobdella ornata*. *Eur J Biochem* 202(3):1073–1082. <https://doi.org/10.1111/j.1432-1033.1991.tb16472.x>
- Min G-S, Sarkar IN, Siddall ME (2010) Salivary transcriptome of the North American medicinal leech. *Macrobodella Decora* *J Parasitol* 96(6):1211–1221. <https://doi.org/10.1645/GE-2496.1>
- Müller C, Mescke K, Liebig S, Mahfoud H, Lemke S, Hildebrandt J-P (2016) More than just one: multiplicity of hirudins and hirudin-like factors in the medicinal leech, *Hirudo medicinalis*. *Mol Genet Genomics* 291:227–240. <https://doi.org/10.1007/s00438-015-1100-0>
- Müller C, Haase M, Lemke S, Hildebrandt J-P (2017) Hirudins and hirudin-like factors in Hirudinidae: implications for function and phylogenetic relationships. *Parasitol Res* 116(1):313–325. <https://doi.org/10.1007/s00436-016-5294-9>
- Müller C, Lukas P, Lemke S, Hildebrandt J-P (2019) Hirudin and decorsins of the North American medicinal leech *Macrobodella decora*: gene structure reveals homology to hirudins and hirudin-like factors of Eurasian medicinal leeches. *J Parasitol* 105(3):423–431. <https://doi.org/10.1645/18-117>
- Müller C, Lukas P, Böhmert M, Hildebrandt J-P (2020a) Hirudin or hirudin-like factor - that is the question: insights from the analyses of natural and synthetic HLF variants. *FEBS Lett* 594:841–850. <https://doi.org/10.1002/1873-3468.13683>
- Müller C, Lukas P, Sponholz D, Hildebrandt J-P (2020b) The hirudin-like factors HLF3 and HLF4 - hidden hirudins of European medicinal leeches. *Parasitol Res* 119:1767–1775. <https://doi.org/10.1007/s00436-020-06697-1>
- Müller C, Wang Z, Hamann M, Sponholz D, Hildebrandt J-P (2022) Life without blood: molecular and functional analysis of hirudins and hirudin-like factors of the Asian non-hematophagous leech *Whitmania pigra*. *J Thromb Haemost* 20(8):1808–1817. <https://doi.org/10.1111/jth.15762>
- Nicholas KB, Nicholas Jr HB (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author
- Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* 4:2411–2423. <https://doi.org/10.1002/pro.5560041120>
- Pfördt V, Kalatehjari P, Tolksdorf C, Rauch BH, Müller C (2022) Go West: hirudins and decorsin/ornatin-like platelet aggregation inhibitors in two representatives of American hematophagous leeches. *Parasitologia* 2:313–325. <https://doi.org/10.3390/parasitologia2040026>
- Richardson JL, Kröger B, Hoeffken W, Sadler JE, Pereira P, Huber R, Bode W, Fuentes-Prior P (2000) Crystal structure of the human α -thrombin-haemadin complex: an exosite II-binding inhibitor. *EMBO J* 19(21):5650–5660. <https://doi.org/10.1093/emboj/19.21.5650>
- Richardson JL, Fuentes-Prior P, Sadler JE, Huber T, Bode W (2002) Characterization of the residues involved in the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. *Biochemistry* 41(8):2535–2542. <https://doi.org/10.1021/bi011605q>
- Sawyer RT (1986) Leech biology and behaviour. Oxford University Press/Clarendon Press, Oxford
- Scacheri E, Nitti G, Valsasina B, Orsini G, Visco C, Ferrera M, Sawyer RT, Sarmientos P (1993) Novel hirudin variants from the leech *Hirudinaria manillensis*. Amino acid sequence, cDNA cloning and genomic organization. *Eur J Biochem* 214(1):295–304. <https://doi.org/10.1111/j.1432-1033.1993.tb17924.x>
- Schnell IB, Thomsen PF, Wilkinson N, Rasmussen M, Jensen LRD, Willerslev E, Bertelsen MF, Gilbert MTP (2012) Screening mammal biodiversity using DNA from leeches. *Curr Biol* 22(8):R262–263. <https://doi.org/10.1016/j.cub.2012.02.058>
- Seymour JL, Henzel WJ, Nevins B, Stults JT, Lazarus RA (1990) Decorsin. A potent glycoprotein IIb-IIIa antagonist and platelet aggregation inhibitor from the leech *Macrobodella decora*. *J Biol Chem* 265(17):10143–10147. [https://doi.org/10.1016/S0021-9258\(19\)38791-5](https://doi.org/10.1016/S0021-9258(19)38791-5)
- Stammers FMG (1950) Observations on the behaviour of land-leeches (genus *Haemadipsa*). *Parasitology* 40(3–4):237–246. <https://doi.org/10.1017/s0031182000018084>
- Strube K-H, Kröger B, Bialojan S, Otte M, Dodt J (1993) Isolation, sequence analysis and cloning of haemadin. An anticoagulant peptide from the Indian leech. *J Biol Chem* 268(12):8590–8595. [https://doi.org/10.1016/S0021-9258\(18\)52916-1](https://doi.org/10.1016/S0021-9258(18)52916-1)
- Teufel F, Armenteros JJA, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol* 40(7):1023–1025. <https://doi.org/10.1038/s41587-021-01156-3>
- Wilson HR, Eisenberg GHG (1982) Land leeches of the genus *Haemadipsa* (Haemadipsoidea: Haemadipsidae) I. Conditions essential to laboratory colonization. Letterman army institute of research institute report no. 121
- Wilting A, Nguyen TV, Axtner J, Nguyen A, Schmidt A, Le M, Nguyen AHQ, Rawson BM, Tilker A, Fickel J (2022) Creating genetic reference datasets: indirect sampling of target species using terrestrial leeches as sample “collectors.” *Environ DNA* 4(2):311–325. <https://doi.org/10.1002/edn3.256>
- Zhou L, Schmaier AH (2005) Platelet aggregation testing in platelet-rich plasma: description of procedures with the aim to develop standards in the field. *Am J Clin Pathol* 123(2):172–183. <https://doi.org/10.1309/y9ec-63rw-3xg1-v313>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

All authors have read and agreed to the published version of the manuscript.