



# Make it double: identification and characterization of a Tandem-Hirudin from the Asian medicinal leech *Hirudinaria manillensis*

Phil Lukas<sup>1</sup> · Georgij Melikian<sup>1</sup> · Jan-Peter Hildebrandt<sup>1</sup> · Christian Müller<sup>1</sup>

Received: 13 July 2022 / Accepted: 19 August 2022 / Published online: 25 August 2022  
© The Author(s) 2022

## Abstract

Haematophagous leeches express a broad variety of secretory proteins in their salivary glands, among them are hirudins and hirudin-like factors. Here, we describe the identification, molecular and initial functional characterization of Tandem-Hirudin (TH), a novel salivary gland derived factor identified in the Asian medicinal leech, *Hirudinaria manillensis*. In contrast to the typical structure of hirudins, TH comprises two globular domains arranged in a tandem-like orientation and lacks the elongated C-terminal tail. Similar structures of thrombin inhibitors have so far been identified only in kissing bugs and ticks. Expression of TH was performed in both cell-based and cell-free bacterial systems. A subsequent functional characterization revealed no evidence for a thrombin-inhibitory potency of TH.

**Keywords** Hirudin · Hirudin-like factors · *Hirudinaria manillensis* · Blood coagulation · Medicinal leeches · Tandem-Hirudin

## Introduction

Haematophagous animals, no matter whether insects, ticks, leeches or even mammals, face the same fundamental challenges during their blood meals on vertebrate hosts: staying undetected, getting enough food and preventing premature haemostasis (Ware and Luck 2017). For the latter, a broad arsenal of bioactive compounds has been developed by nature including inhibitors of platelet aggregation (primary haemostasis) and inhibitors of blood coagulation (secondary haemostasis) (Ware and Luck 2017; Cherniack 2011; Klotz et al. 2014; Oka et al. 1989; Li et al. 2018; Manship et al. 2012). Putative targets of blood coagulation inhibitors are factor Xa (Stuart–Prower factor) and factor IIa (thrombin) (Koh and Kini 2010, 2009). Hirudin is probably the best known natural direct thrombin inhibitor and was originally identified in the European medicinal leech, *Hirudo*

*medicinalis* (Markwardt 1955, 1970). However, several other thrombin inhibitors have been described in leeches (haemadin), ticks (boophilin, ornithodorin and varieglin), kissing bugs (rhodniin and triabin) and mosquitoes (anophelein), respectively (Strube et al. 1993; Macedo-Ribeiro et al. 2008; Friedrich et al. 1993; Mans et al. 2002; Oka et al. 1989; Koh et al. 2007; Noeske-Jungblut et al. 1995). Hirudin and haemadin share a common structure with a short N-terminal part, one central globular domain stabilized by three disulfide bridges and an elongated C-terminal tail (Grütter et al. 1990; Richardson et al. 2000). In contrast, boophilin, ornithodorin and rhodniin show a tandem-like structure with two globular domains connected by a short linker and without an elongated tail (Macedo-Ribeiro et al. 2008; Friedrich et al. 1993; Mans et al. 2002). Strikingly, the globular domains are stabilized by three disulfide bonds as well, yet in a completely different way: Cys1-Cys2, Cys3-Cys5 and Cys4-Cys-6 in hirudin and haemadin; Cys1-Cys6, Cys2-Cys4 and Cys3-Cys5 in boophilin and ornithodorin; and Cys1-Cys5, Cys2-Cys4 and Cys3-Cys6 in rhodniin, respectively (Grütter et al. 1990; Richardson et al. 2000; Macedo-Ribeiro et al. 2008; Friedrich et al. 1993; Mans et al. 2002). A tandem-like structure has been described for leech-derived factors so far only for members of the antistasin-family like antistasin itself and ghilanten (Whitlow et al.

Handling Editor: Una Ryan

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

<sup>1</sup> Animal Physiology, Zoological Institute and Museum, University of Greifswald, Felix-Hausdorff-Str. 1, D-17489 Greifswald, Germany

1999). Both are inhibitors of factor Xa (Blankenship et al. 1990; Dunwiddie et al. 1989). We have recently described the identification and functional characterization of a novel class of leech-derived compounds, the hirudin-like factors (HLFs), in members of the genera *Hirudo* and *Hirudinaria* (Lukas et al. 2019; Müller et al. 2016, 2017, 2019a). HLFs comprise structural features characteristic to hirudin (e.g. six cysteine residues within a central globular domain and a common gene structure composed of four exons and three introns), but may considerably differ in biochemical properties like molecular weight (MW) and isoelectric point (pI value) (Müller et al. 2017). Some of the HLFs show thrombin-inhibitory potencies as high as hirudin (e.g. HLF-11long of *Hirudo orientalis*), whereas others have reduced inhibitory capacities (e.g. HLF5 and HLF8 of *Hirudinaria manillensis*) or a very low or even no detectable potencies at all (e.g. HLF1D of *Hirudo verbana* and HLF6 of *Hirudinaria manillensis*) (Lukas et al. 2019; Müller et al. 2019a). All these factors were identified prior to the availability of genomic or transcriptomic data by either cDNA library screening or targeted and non-targeted PCR approaches. As an unexpected outcome of such a reaction, we identified a cDNA that encodes a yet unknown putative salivary gland-derived factor of *Hirudinaria manillensis* with two hirudin-like globular domains arranged in a tandem-like orientation. We therefore named this factor Tandem-Hirudin (TH). The structure of TH resembles those of the thrombin inhibitors of ticks and kissing bugs mentioned above, and we hence derived the hypothesis that TH might be an inhibitor of thrombin as well. In the present study, we describe the structure and our attempts to express and functionally characterize TH.

## Material and methods

### Genotyping of animals and tissue preparation

The biological material used in this study (specimen of *Hirudinaria manillensis* and salivary gland preparations) was already described by Müller et al. (2017). Information on genotyping data and GenBank accession numbers can be obtained from the respective publication.

### Sequence analysis

Nucleotide and amino acid sequence alignments were generated using the CLS Sequence Viewer software package v8.0 (CLC bio, Aarhus, Denmark) with the following parameters: gap open cost, 5.0; gap extension cost, 2.0; end gap cost, free. 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 server (Käll et al. 2007) and SignalP6.0 (Teufel et al. 2022).

### Isolation and cloning of a Tandem-Hirudin cDNA

Tissue preparation, RT-PCR and primer design were already described in detail elsewhere (Müller et al. 2017). For bacterial expression, the cDNA of TH without its putative signal peptide coding sequence was cloned into the bacterial expression vector pQE30Xa (Qiagen, Hilden, Germany). In addition, a plasmid for cell free protein synthesis of TH was generated as well (see below).

### Bacterial expression, sulfitolysis and refolding of TH

The *E. coli* strain GM1674 (Palmer and Marinus 1994) was used for bacterial expression of TH. Expression and purification of TH were performed as described by Lukas et al. (2019). However, the successful downstream processing of TH required a partial sulfitolysis of the protein prior to the factor Xa-cleavage. The sulfitolysis was performed by modified methods obtained from Lundblad (1995) and Grant (2017). Tandem-Hirudin was first dialysed overnight against 0.5 L of a sodium sulfite buffer (0.05 mol/l sodium sulfite, 0.2 mol/l cysteine-HCl, 8 mol/l urea, 0.1 mol/l Tris-HCl, pH 7.0) at room temperature in a 5 ml microdialysis capsule (QuixSep, Roth, Karlsruhe, Germany) and a dialysis membrane with a molecular mass cut off (MWCO) of 5,000 Da (Roth, Karlsruhe, Germany). After dialysis the capsule was exposed to air for 4 h at room temperature. For refolding, the capsule was placed in 1 L of refolding buffer (50 mmol/l Tris-HCl, 400 mmol/l sucrose, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, pH 8.0, modified after Moghadam et al. (2015) and dialysed for another 48 h with one buffer replacement after 24 h. A fresh volume of refolding buffer was degassed and enriched with 3 mmol/l glutathione under nitrogen atmosphere. The capsule was incubated for another 4 h at room temperature and constant stirring. After incubation, the nitrogen was replaced with air, and 0.3 mmol/l glutathione disulfide was added, and the capsule incubated for 20 h at 4 °C. A final dialysis was performed against reaction buffer (20 mmol/l Tris/HCl, 100 mmol/l NaCl, 2 mmol/l CaCl<sub>2</sub>, pH 8.0) for 24 h at 4 °C and repeated for another 24 h. Factor-Xa cleavage was performed for 48 h at room temperature (Lukas et al. 2019; Müller et al. 2017).

### Cell-free synthesis of TH in bacterial cell lysates

For the cell-free synthesis of TH in bacterial cell lysates, the NEBExpress® Cell-free *E. coli* Protein Synthesis System (New England Biolabs, Frankfurt a. M., Germany) was applied. Expression of proteins was achieved by a coupled transcription (T7 RNA polymerase supplied with the kit)

and translation process. The coding sequences for TH as well as hirudin variant HV1 and HLF5 (both factors served as positive controls) were cloned into the expression vector pET-28a(+) (Merck Millipore, Darmstadt, Germany) using the unique *NcoI* restriction site at position 296 to generate an appropriate start codon for translation. As a consequence, the N-terminal amino acid sequence of hirudin variant HV1 was changed from VVYTDC to MVYTDC. The N-terminal amino acid sequence of TH was accordingly modified from VCTVC to MVYTDC. All further steps of the synthesis reactions followed the instructions and recommendations as outlined by the manufacturer. Briefly, 250 ng of plasmid DNA in a total reaction volume of 50  $\mu$ l were incubated for 4 h at 37 °C in an orbital shaker (ThermoMixer C, Eppendorf, Hamburg, Germany). The empty vector pET-28a(+) and the DHFR-His control template (supplied with the kit) were used for control purposes in separate protein synthesis reactions. To support a proper disulfide bond formation the PURExpress® Disulfide Bond Enhancer (New England Biolabs, Frankfurt a. M., Germany) was added. Aliquots of all synthesis reactions were analysed by SDS-PAGE on a 20% gel.

### Coagulation assays

To verify the putative anti-coagulatory activity of TH, three blood coagulation assays were performed: the activated partial thromboplastin time test (aPTT; reference range 22.7–28.9 s), the prothrombin time test (PT; reference range 10.7–13.7 s) and the thrombin time test (TT; reference range 16.8–21.4 s) using a BFT II Analyzer (Siemens Healthcare, Erlangen, Germany). All steps followed the instructions outlined by the manufacturer. For the coagulation assays, the protein samples were either diluted with dialysis buffer to reach final concentrations of 50  $\mu$ mol/l or 5  $\mu$ mol/l, respectively, or used in the highest concentration possible. The desired amount of substrate was transferred directly 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 lasted up to 300 s were stopped and declared as a complete inhibition of coagulation. Clot formation was additionally controlled by eye to exclude technical errors. All samples were measured immediately after cell-free synthesis. TH expressed in bacteria was kept in aliquots and thawed on ice immediately before the measurements.

### Platelet aggregation assay

The platelet aggregation assays were performed as already described in detail by Lukas et al. (2019).

For TH expressed and purified from bacteria, 10  $\mu$ l of protein with a concentration of 75  $\mu$ mol/l were transferred into the cuvettes, and 180  $\mu$ l PRP were added, resulting in a final concentration of 3.8  $\mu$ mol/l. Ten microlitre of dialysis buffer served as the control. Eptifibatide in an equimolar amount was used as a positive control for inhibition of platelet aggregation. Output parameters were calculated as described (Lukas et al. 2019).

### Trypsin activity assay

The putative effects of TH on the trypsin activity were evaluated by a chromogenic assay slightly modified after Erlanger et al. (1961). Briefly, wells on a transparent 96 well plates (Sarstedt, Nümbrecht, Germany) were prefilled with 22.4  $\mu$ l of either dialysis buffer (negative control) or sample solution to give a final concentration of 0.4 or 3.8  $\mu$ mol/l. Afterwards, 9.6  $\mu$ l of a standard trypsin solution used for cell culture purposes (2.5 g/l, PAN-Biotech, Aidenbach, Germany) was added. Equimolar amounts of the soybean trypsin inhibitor (SBTI, AppliChem, Darmstadt, Germany) were used as a positive control. The samples were incubated for 5 min at 37 °C followed by the addition of 167  $\mu$ l of substrate solution (1 mmol/l N  $\alpha$ -benzoyl-L-arginine-4-nitroanilide hydrochloride (BAPA) in 0.01% dimethyl sulfoxide (DMSO), 50 mmol/l Tris-HCl, 20 mmol/l CaCl<sub>2</sub>, pH = 8.2). Every 30 s, the absorbance at a wavelength of  $\lambda = 410$  nm was measured over a period of 5 min in a standard plate reader (Tecan, Crailsheim, Germany). Calculation of trypsin activity followed an adapted equation of Sørensen et al. (1999):

$$U/mg = (A_{410}/min \cdot GV \cdot VF) / (\epsilon \cdot d \cdot PV \cdot mg/ml).$$

( $A_{410}/min$  is the alteration of the extinction at  $\lambda = 410$  nm per minute,  $GV$  is the whole sample volume in ml,  $VF$  is the dilution factor,  $\epsilon$  is the molar attenuation coefficient in  $cm^2/\mu$ mol,  $d$  is the thickness of the cuvette in cm,  $PV$  is the amount of trypsin solution in ml, and  $mg/ml$  is the concentration of trypsin solution).

## Results

### Identification and characterization of a TH encoding cDNA

Isolation of cDNA of TH followed the same general protocol as already described for the identification of recently reported HLFs (Müller et al. 2016, 2017, 2019a). Extracts of RNA derived from two *Hirudinaria manillensis* individuals were analysed by reverse transcription polymerase chain reaction (RT-PCR) with a forward primer that spans a portion of the coding region of the signal peptide and an oligo dT-primer. Amplified fragments were cloned, and sequences of individual clones were determined. Sequence analyses of

two clones revealed inserts with lengths of 572 bp (clone 1) or 517 bp (clone 2), respectively, including a polyA-stretch. Clone 1 contains an additional stretch of 56 nucleotides immediately before the polyA tail but is otherwise identical to clone 2. Sequence data were deposited in GenBank under Acc. No. MZ361731 and MZ361732. Both sequences contain open reading frames of 306 bp (including the stop codon) that encode a putative protein of 101 amino acid residues in length including a signal peptide sequence. The derived amino acid sequences revealed the presence of 13 cysteine residues. With exception of the first, all cysteine residues were arranged in a manner (position and spacing) typical for the central globular core domain of hirudins and HLFs (Fig. 1).

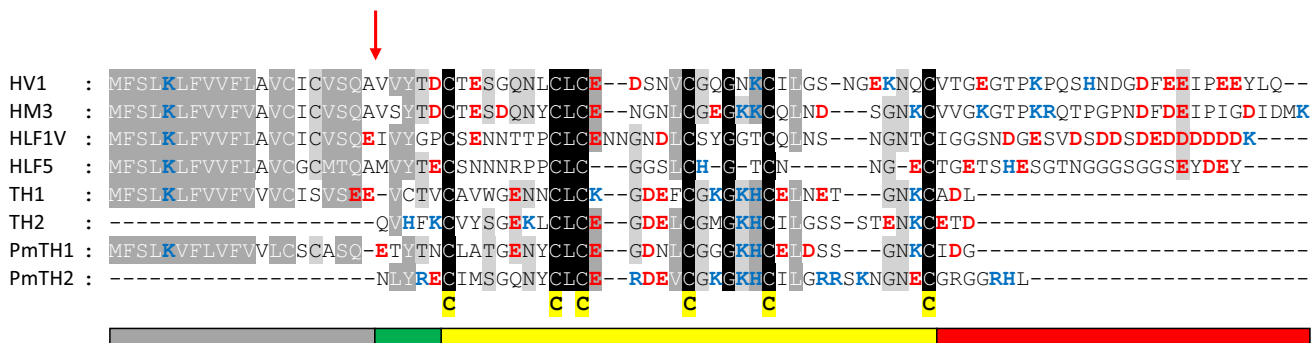
In addition, the arrangement of cysteine residues clearly pointed to a tandem-like duplication of the central globular domain in the putative new protein, and the factor was hence termed Tandem-Hirudin (TH). Detailed comparison of the central globular domains (Cys1-Cys6) of representatives of hirudins and HLFs with the two domains of TH (TH1

and TH2) revealed degrees of sequence identity/similarity between 29–65% and 38–78%, respectively (Table 1).

### Molecular properties of TH

Hirudins and HLFs of *Hirudo spec.* and *Hirudinaria manillensis* contain highly conserved signal peptide sequences of 20 amino acid residues in length. The putative signal peptide sequence of TH almost perfectly fits into that scheme (see Fig. 1); however, cleavage site predictions using several online software tools led to heterogeneous outcomes, most probably due to the presence of the two glutamic acid residues at positions 19 and 20.

Hirudins and HLFs with thrombin-inhibiting activity exhibit several characteristic molecular features including an acidic pI value of about 3.2–4.6 (Müller et al. 2019a, 2017). The low value is mainly caused by several acidic amino acid residues within the elongated C-terminal tail, but the central globular domain (Cys1-Cys6) as well has to be acidic for inhibitory activity (Müller et al. 2019a). Analyses



**Fig. 1** Multiple sequence alignment of hirudin-variant 1 (HV1) and hirudin-like factor HLF1V from *Hirudo medicinalis*, hirudin-variant HM1 and HLF5 from *Hirudinaria (Poecilobdella) manillensis* and domain-I (TH1) and domain-II (TH2) of Tandem-Hirudins TH and PmTH (a putative TH encoded by *Poecilobdella manillensis*). Black background indicates conserved residues; grey background indicates similar residues. Acidic amino acid residues are labelled in red, basic

amino acid residues in blue. A red arrow indicates the signal peptide cleavage site. The six conserved cysteine residues giving rise to the three dimensional structure are marked in bold and labelled in yellow. Boxes indicate the structural and functional domains of hirudin HV1: signal peptide (grey), N-terminus (green), central globular domain (yellow) and elongated C-terminal tail (red). Abbreviations are used according to the IUPAC code

**Table 1** Comparison of the central globular domains (Cys1-Cys6) of hirudins (hirudin variants HV1 of *Hirudo medicinalis* and HM3 of *Hirudinaria manillensis*) and hirudin-like factors (HLF1V of *Hir-*

*udo medicinalis* and HLF5 of *Hirudinaria manillensis*) with the two domains of TH (TH1 and TH2) and PmTH (PmTH1 and PmTH2). Displayed are degrees of amino acid sequence identity/similarity

	HM3	HLF1V	HLF5	TH1	TH2	PmTH1	PmTH2
HV1	58/73	41/55	29/50	41/52	52/67	50/64	59/64
HM3		47/61	24/44	50/71	50/52	62/75	51/62
HLF1V			52/55	38/55	38/50	44/58	37/48
HLF5				34/43	29/38	34/46	31/40
TH1					55/64	65/78	54/65
TH2						64/70	56/70
PmTH1							57/74

of the TH1 and TH2 domains revealed a length of 32 or 34 amino acid residues and pI values of 5.59 and 4.96, respectively (Table 2).

Beside the central globular domain and the elongated C-terminal tail, hirudins and HLFs comprise a short N-terminal stretch of five amino acid residues immediately in front of the first cysteine residue at position 6 including a highly conserved tyrosine (or phenylalanine) residue at position 3 (see Fig. 1). The N-terminus of TH comprises either four (VCTV; assuming a signal peptide of 20 amino

acid residues) or five residues (EVCTV; assuming a signal peptide of 19 amino acid residues). In both cases, the N-terminus of TH does not fit very well to the consensus sequence and lacks the conserved tyrosine or phenylalanine residue (Müller et al. 2017; Braun et al. 1988; Wallace et al. 1989; Vindigni et al. 1994; Betz et al. 1992). Strikingly, TH is the first natural representative of the hirudin family known so far that contains a tryptophan residue (see Fig. 1).

The structure of TH has not yet been determined. Structure modelling using i-TASSER (Roy et al. 2010; Yang et al. 2015) confirmed the formation of two separated globular domains, each stabilized by three disulfide bridges. According to the model, the first cysteine residue of TH at position 2 is not involved in the formation of disulfide bridges. Figure 2 illustrates the putative structure of TH in comparison to the archetype hirudin variant HV1.

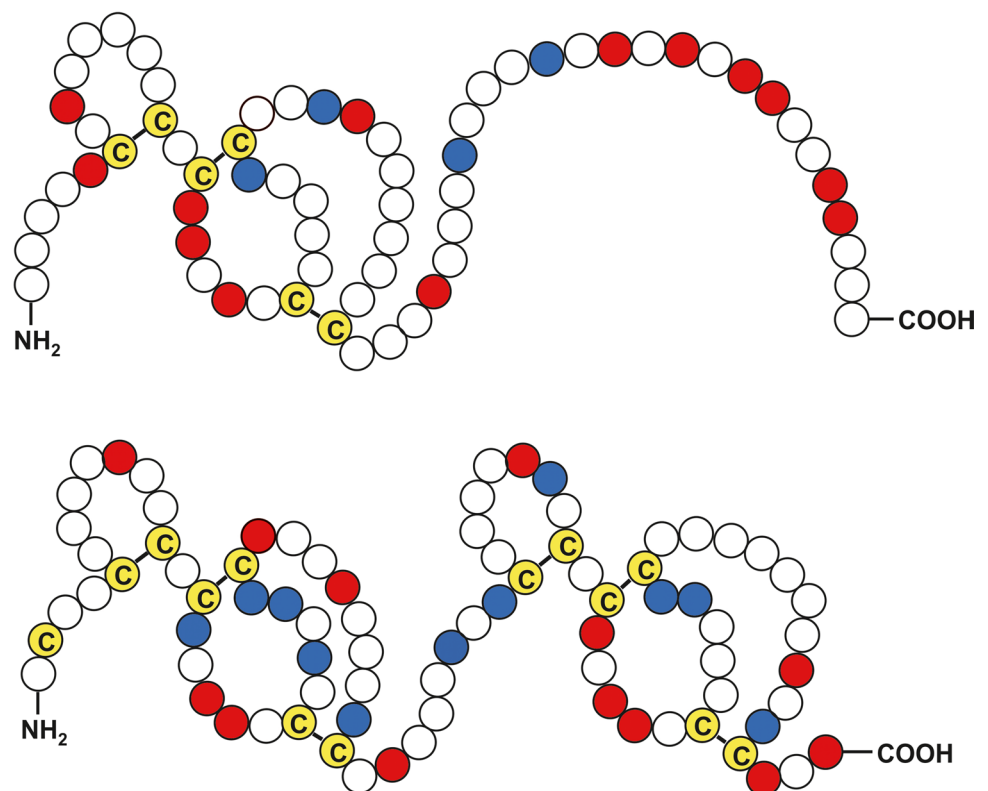
**Table 2** Lengths, numbers of acidic, basic amino acid residues, predicted isoelectric points of globular domains of hirudins (hirudin variants HV1 of *Hirudo medicinalis*, HM3 of *Hirudinaria manillensis*), hirudin-like factors (HLF1V of *Hirudo medicinalis*, HLF5 of *Hirudinaria manillensis*), tandem hirudins TH (TH1, TH2), PmTH (PmTH1, PmTH2)

Factor	Length C1–C6	Acidic/basic	pI value
HV1	34	4/2	4.41
HM3	32	5/3	4.51
HLF1	34	3/0	3.57
HLF5	25	½	6.70
TH1	32	5/5	5.59
TH2	34	5/4	4.96
PmTH1	32	5/3	4.50
PmTH2	35	4/7	8.33

### Bacterial expression and purification of TH

Bacterial expression of TH (r-TH) was performed in *E. coli* strain GM1674 according to the protocol already described earlier (Lukas et al. 2019). r-TH was expressed in large amounts; however, most of the protein could be found in inclusion bodies and only minor amounts in the cytosolic fraction. Inclusion bodies were purified and resolved in denaturing buffers containing 8 mol/l urea for further purification by His-tag affinity chromatography (Lukas et al.

**Fig. 2** Schematic representation of HV1 (upper part) and TH (lower part). Acidic amino acid residues are labelled in red, basic amino acid residues in blue and cysteine residues in yellow. Putative disulfide bonds are indicated by bold lines



2019). During dialysis at pH 8.0, the protein solution partly precipitated, and only the soluble fraction was used for further processing. However, the subsequent cleavage of the N-terminal His-tag by factor Xa digestion largely failed. Various attempts have been made to remove the remaining His-tag including the addition of up to 0.05% SDS and 0.5 mol/l urea during the cleavage reaction, increase of reaction temperature up to 37 °C and the irreversible immobilization of His-tag-TH on Co(III)-IDA (Wegner and Spatz 2013; Wegner et al. 2016), but they largely failed, too. Finally, partial sulfitolysis as described in the “Material and Method” section solved the problem. For control purposes, purification and processing by sulfitolysis was also done for hirudin-variant 1 (r-HV1).

### Functional characterisation of bacterially synthesized r-TH

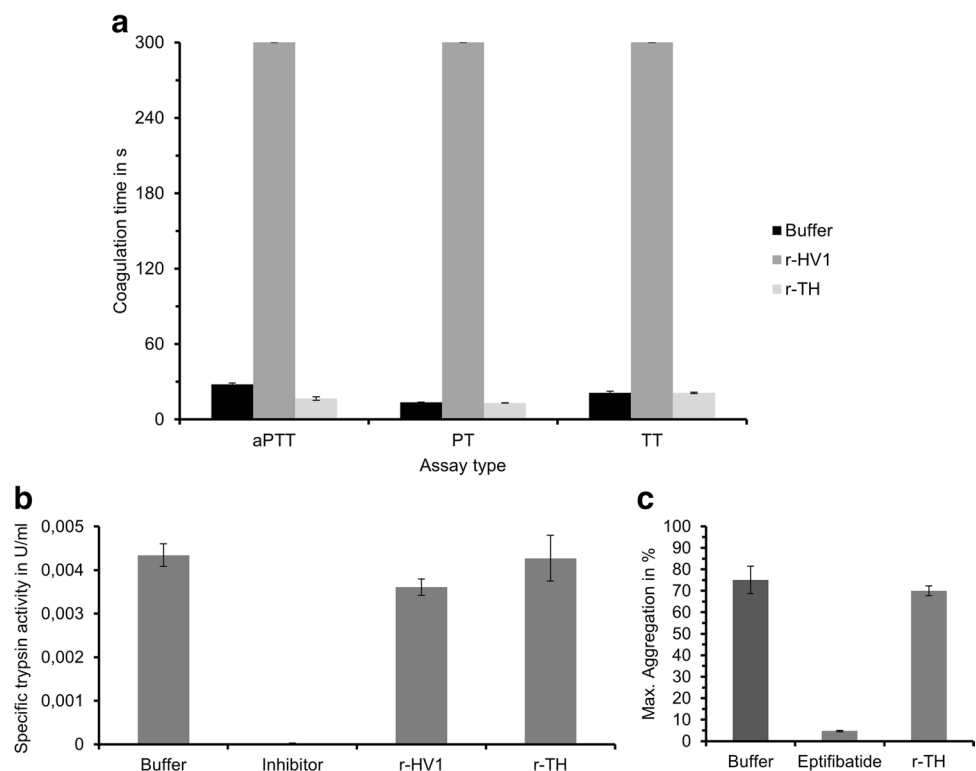
Functional characterisation of bacterially synthesized r-TH included blood coagulation assays as well as measurements of trypsin activity and platelet aggregation. In the blood coagulation assays, r-TH was compared to equimolar amounts of r-HV1 as a positive and buffer as the negative control. No anti-coagulatory effect of r-TH could be observed in all three assays (activated partial thromboplastin time, prothrombin time and thrombin time, respectively), while r-HV1 completely inhibited blood coagulation (Fig. 3A). In addition, no inhibitory effects of r-TH and r-HV1 could be observed for

either the trypsin activity or the platelet aggregation assays (Fig. 3B and C).

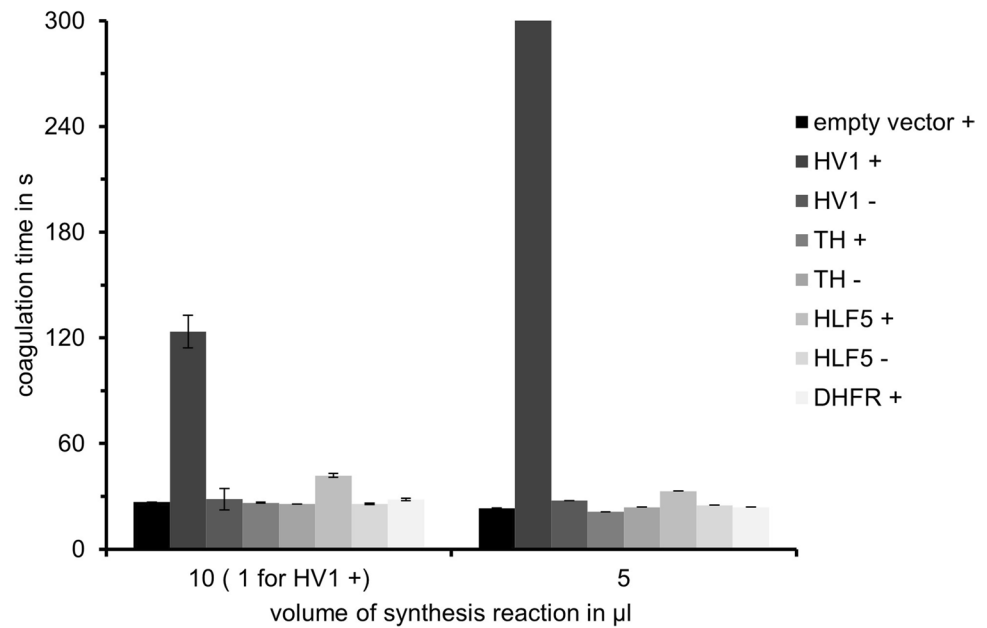
### Cell-free synthesis of TH in bacterial cell lysates

To circumvent the procedural problems with the bacterial cell-based expression and purification of TH described above, we applied a bacteria-based cell-free protein synthesis protocol. N-terminally modified (incorporation of a methionine residue) variants of TH and hirudin variant HV1 as well as native HLF5 (a factor that naturally comprises a methionine residue at the N-terminus; see Fig. 1) were synthesized in lysates of *E. coli*. Synthesis reactions were performed with or without the addition of a disulfide bond enhancer. Aliquots of all synthesis reactions were analysed in a SDS-PAGE and tested in the thrombin time coagulation assay. As can be seen in Fig. 4, the cell lysate that contained hirudin variant HV1 revealed a very high thrombin-inhibitory potency. The potency, however, was almost entirely dependent on the presence of the disulfide bond enhancer in the synthesis reaction. For HLF5, the coagulation test revealed a comparably low, but clearly detectable inhibition of thrombin. Again, the effect was dependent on the presence of the disulfide bond enhancer in the synthesis reaction. For TH, however, no inhibition could be measured, even if the disulfide bond enhancer was present in the synthesis reaction (Fig. 4).

**Fig. 3** Functional characterisation of bacterially synthesized r-TH and hirudin variant 1 (r-HV1) in blood coagulation (A), trypsin activity (B) and platelet aggregation (C) assays. Measurements for blood coagulation were cancelled after 300 s and considered as a complete inhibition of coagulation. Concentrations used in the assays were 3.2 µmol/l (A), 0.4 µmol/l (B) and 3.8 µmol/l (C), respectively. aPTT indicates the activated partial thromboplastin time test, PT the prothrombin time test and TT the thrombin time test



**Fig. 4** Standard blood coagulation assays using the thrombin time test (TT) of cell-free synthesized TH, hirudin HV1 and HLF5 with (+) or without (–) the addition of disulfide bond enhancer during the synthesis reaction in bacterial cell lysates. Lysates containing either an empty vector or a vector containing the coding region of the dihydrofolate reductase (DHFR) served as the negative controls. Aliquots of synthesis reactions used for the measurements were either 5  $\mu$ l ( $n=1$ , right) of 10  $\mu$ l (1  $\mu$ l for HV1) ( $n=3$ , left)



## Discussion

More than 100 different proteins are predicted to occur in leech saliva (Baskova et al. 2004); however, most of these factors still remain uncharacterized or even unidentified. For only a few leech saliva components, functional data on one hand and molecular data on the other hand have already been assigned (Hildebrandt and Lemke 2011; Baskova and Zavalova 2001). Among them, hirudin is probably the best known and so far the only one that found its way from bench to bedside (Organisation to Assess Strategies for Ischemic Syndromes (OASIS-2) Investigators 1999; Greinacher and Lubenow 2001; Chuang et al. 2001; Frame et al. 2010; Nar 2012). Hirudin is one of the most effective natural inhibitors of thrombin, and its presence seems to be restricted to members of the order of *Hirudinida* (Müller et al. 2019a; Scacheri et al. 1993; Baskova et al. 1983; Scharf et al. 1989). Functional counterparts of hirudin in terms of thrombin inhibition, however, are present in a broad range of haematophagous animals like ticks, kissing bugs and mosquitoes as well (Koh and Kini 2010, 2009). These factors are analogous to hirudin and hence very likely the result of convergent evolution. Some of these non-leech-derived thrombin inhibitors comprise a notable structure as they are composed of two repeated domains that are arranged in a tandem-like order (Macedo-Ribeiro et al. 2008; Friedrich et al. 1993; Mans et al. 2002). Similar tandem-like structure are described for members of the leech antistasin-family including antistasin itself and ghilanten (Whitlow et al. 1999; Kim et al. 2007). Both are inhibitors of factor Xa. We describe and both functionally and molecularly characterize a novel and

so far unique leech-derived factor of *Hirudinaria manilensis*, called Tandem-Hirudin (TH), that combines the core structural features of hirudin with the tandem-like arrangement found in leech antistasins and thrombin-inhibitors of ticks and kissing bugs.

## Structural properties of TH

The Tandem-Hirudin is composed of two central globular hirudin-like domains connected by a short linker sequence and contains a putative N-terminal signal peptide sequence. The presence of a signal peptide sequence was supported by all appropriate software tools; however, the exact processing side was somewhat uncertain.

The remaining N-terminal tail of TH differs both in length (four instead of five residues) and composition from respective sequences of hirudin(s) and HLFs (Müller et al. 2017). The N-terminus of hirudin blocks the reactive site of thrombin, and alterations (deletions and/or substitutions) within the N-terminus may greatly alter its thrombin-inhibitory potency (Müller et al. 2019a; Braun et al. 1988; Wallace et al. 1989; Betz et al. 1992). However, there is no strict consensus sequence of the N-terminus. But so far, all investigations concordantly confirmed that an aromatic amino acid residue at position 3 (Y3, F3 or W3) is crucial for thrombin inhibition (Lazar et al. 1991; Müller et al. 2019a). TH completely lacks a respective residue within the N-terminus. Even more strikingly, TH lacks the elongated C-terminal tail that is characteristic for all hirudins (Salzet 2000). The C-terminal tail of hirudin blocks the fibrinogen-binding site of thrombin, the so-called exosite 1 (Maraganore and Fenton 1990; Xin et al. 2015; Priestle et al. 1993). In boophilin,

ornithodorin and rhodniin, the C-terminal second globular domains adopt these functions (Macedo-Ribeiro et al. 2008; Mans et al. 2002; Friedrich et al. 1993). Whether or not the same is true for TH remains to be proven. At least, the biochemical properties of TH do not seem to be prohibitive. Hirudins and HLFs with thrombin-inhibitory potencies (recently termed unorthodox hirudins) comprise an overall pI value of about 3.2–4.6, but functionality requires both an acidic central globular domain and an acidic C-terminal tail (Müller et al. 2019a; Lukas et al. 2019). The pI value of TH is about 5.03, and the respective value for the first domain is 5.59 and for the second 4.96 (Table 2). Taken together, TH displays some unique structural characteristics that are not common in hirudin(s) and HLFs and hence raised questions about its functionality.

### Genetics of TH

The genes of hirudins and HLFs comprise a common structure of four exons and three introns (Scacheri et al. 1993; Müller et al. 2017, 2019a). The first exon encodes the signal peptide, the second and the third exon encode the short N-terminal stretch and the central globular domain, whereas the fourth exon encodes the elongated C-terminal tail. Despite intensive efforts, we failed to identify and characterize the gene encoding TH in our specimen of *Hirudinaria manillensis*. However, when thoroughly analysing the genome data of *Hirudinaria (Poecilobdella) manillensis* (Guan et al. 2020; GenBank Acc. No. SRX4286737), we identified a segment of about 18,200 bp in length on contig00065 (GenBank Acc. No. JADEYA01000065, position 1,28,3800 to 1,302,000) that contained the genes of four putative hirudins/HLFs and the gene of a putative TH (PmTH, see also Fig. 1). The respective segment was annotated and deposited as a Third Party Annotation (TPA) in GenBank under Acc. No. BK059527. Figure 5 summarizes the genetic structure of the respective segment in the *Hirudinaria (Poecilobdella) manillensis* genome. The putative PmTH gene comprises six exons and

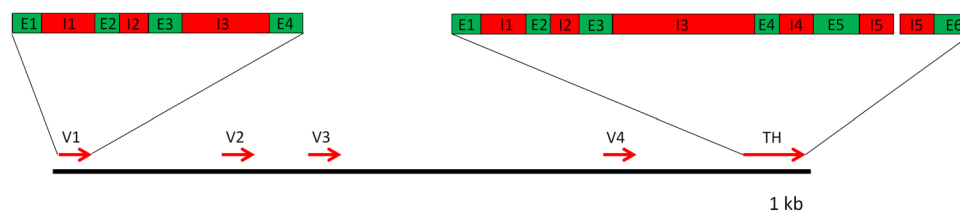
five introns, whereby the exons 4 and 5 exhibit the same structure as the exons 2 and 3 and are hence most likely the result of a duplication event.

The exact composition of exons 2 and 3 supports such a duplication (or even multiplication) scenario. Exon2 of hirudin and the HLF genes starts with the second nucleotide of a triplet that codes for the first amino acid residue of the mature factors (the first nucleotide of the respective triplet is provided by exon 1). Exon 3 ends with the first nucleotide of a triplet that is completed by the first two nucleotides of exon 4 (in hirudin and HLF genes) or by the duplicated exon 2 (= exon 4) in PmTH. Further rounds of such exon duplication events would lead to the formation of molecules with multiple hirudin-like domains. A scheme visualizing the multiplication hypothesis is shown in Fig. 6.

Comparable scenarios of multiplication have been described for members of the antistasin-family in *Hydra* sp. (Holstein et al. 1992) or in the disc abalone *Haliotis discus discus* (Nikapitiya et al. 2010).

An alternative way to generate a TH-encoding mRNA would be trans-splicing. RNA trans-splicing combines two separate pre-mRNA molecules to form a chimeric non-co-linear RNA molecule (Lei et al. 2016). So far, trans-splicing has been described in several groups of eukaryotes including trypanosomes, nematodes, insects and even mammals, but no yet in leeches. However, the exploitation of such a mechanism would have the potential to greatly enhance the variability of hirudins and HLFs in these animals.

The presence of TH genes outside the genus *Hirudinaria (Poecilobdella)* remains to be proven. In our own analysis, we did not find any evidence for TH genes in the genomes of leech species like *Hirudo medicinalis* (Kvist et al. 2020; Babenko et al. 2020), *Whitmania pigra* (Tong et al. 2022) or *Helobdella robusta* (Simakov et al. 2013). The identification of additional TH and/or OH genes is a prerequisite for any further investigations on the origin of TH genes in an evolutionary context.

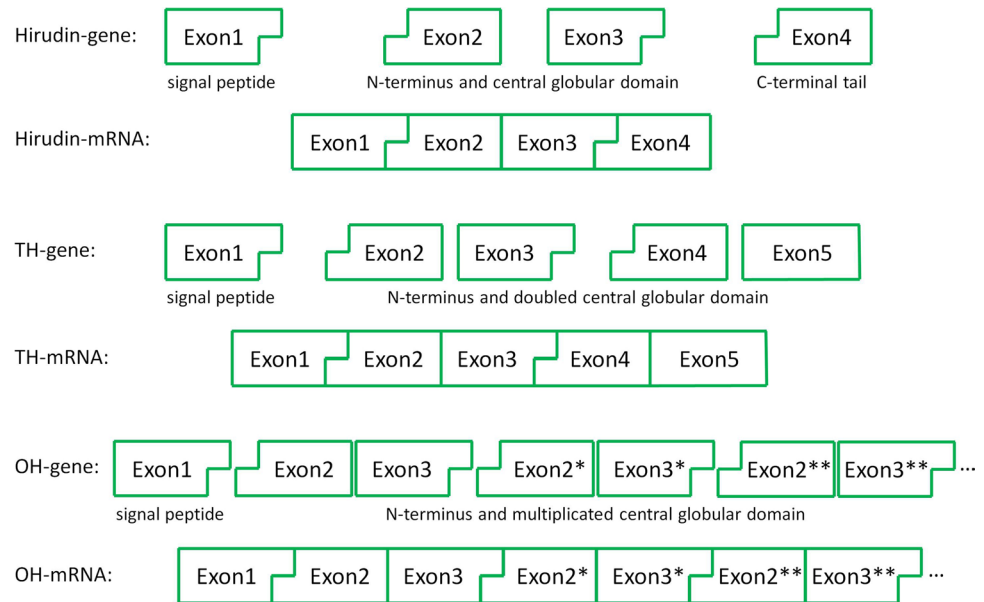


**Fig. 5** Schematic representation of a segment on genome contig00065 of *Hirudinaria (Poecilobdella) manillensis* comprising four putative hirudin/HLF genes (Hiru\_V1–V4) and one putative tandem hirudin gene (PmTH). Exons (E) are labelled in green, and introns (I) are labelled in red. Red arrows indicate position, size and orienta-

tion of the respective genes. The upper boxes illustrate the structures of two genes (left: putative hirudin/HLF Hiru\_V1; right: putative tandem hirudin PmTH). Exons (E) are labelled in green, and introns (I) are labelled in red. The bar indicates a length of 1 kb



**Fig. 6** Schematic drawing visualizing the multiplication hypothesis. TH indicates the tandem hirudin, OH a putative molecule with multiple hirudin-like domains. Exon-overlapping codons are illustrated by an angled connection



## Synthesis and functional characterization of TH

We have multiple times successfully used bacterial expression systems for the large-scale synthesis of His-tagged hirudin(s) and HLFs that were subsequently purified by affinity chromatography and functionally validated (Müller et al. 2016, 2019a, 2019b; Lukas et al. 2019). For TH, this approach failed. TH was very well over-expressed by the bacteria, but was almost completely found in insoluble inclusion bodies. Intense efforts have been made to reconstitute functionally active TH. Whereas the His-tag of TH could be successfully removed, the final TH did not show anti-coagulatory potency or any other inhibitory effects (Fig. 3). Two possible scenarios could explain these results. First, TH does not comprise these activities at all. Second, TH may not be in its correct and hence functional conformation (Chang et al. 1995). As described above, TH comprises 13 cysteine residues. Six pairs of these residues are supposed to be involved in the formation of disulfide bridges, three within the first and three within the second repeat of the tandem globular domains (see Fig. 2). The bacterial expression and in vitro processing of TH by sulfitolysis and refolding may likely have not resulted in the correct formation of all these several disulfide bridges (Thomas and Baneyx 1996; Baneyx and Mujacic 2004; García-Fruitós et al. 2016; Mukherjee and Gupta 2016). Mathematical calculations reveal a number of 135,135 different combinations of disulfide bridges based on 13 cysteine residues (and one remaining free SH group) (Combet et al. 2000). Clearly, not all of the calculated combinations are sterically possible, but nevertheless, the number indicates the very low probability of a correct folding simply by chance (Zhang and Snyder

1989). To improve the probability of correct folding, we decided to use cell-free protein synthesis approaches.

To our knowledge, the bacterial cell lysate-based approach has never been before applied for the synthesis of hirudins in particular or leech saliva components in general. Our data clearly indicate that both hirudin variant HV1 of *Hirudo medicinalis* and HLF5 of *Hirudinaria manillensis* were synthesized in biologically active forms (see Fig. 4). The observed thrombin-inhibitory potency, however, was completely dependent on the presence of a disulfide bond enhancer in the synthesis reaction, once again highlighting the crucial importance of correct disulfide bond formation and folding of hirudins and HLFs for their biological activities. For TH, no inhibitory effect on thrombin activity could be observed, even when the disulfide bond enhancer had been present in the synthesis reaction mixture.

Further large-scale synthesis and purification approaches will hopefully result in higher yields of TH to perform additional functional assays to confirm the present results and/or identify the actual biological target of TH. In addition, it is without any doubt of great interest to determine the structure of TH in a complex with its biological target.

## Summary

The current study describes the identification and partially successful synthesis of a putative secretory protein of *Hirudinaria manillensis*. The protein comprises a structure that has not yet been described for members of the hirudin superfamily: two central globular domains of hirudin are arranged in a tandem-like way. The protein was hence termed Tandem-Hirudin (TH). Several functional tests including

coagulation assays, platelet aggregation assay and a trypsin activity assay did not yet reveal the biological target of TH.

**Acknowledgements** We want to thank Bethany Sawyer (Biopharm Ltd, Swansea, UK) for providing the *Hirudinaria* specimen used in this study. Platelet aggregation assays were performed with the help and support of Robert Wolf and Prof. Dr. Bernhard H. Rauch (Department of Pharmacology, University Medicine Greifswald). In addition, we want to thank all members of our groups for their permanent support. Finally, we are thankful to two unknown reviewers for their helpful comments and recommendations.

**Author contribution** PL and CM conceived the ideas and designed the methodology. PL, GM and CM performed the experiments. CM and PL analysed the data and drafted the manuscript. JPH supervised the practical and theoretical works. All authors contributed to the preparation of the final manuscript and gave approval for publication.

**Funding** Open Access funding enabled and organized by Projekt DEAL. Phil Lukas received a doctoral stipend from the Friedrich-Ebert-Stiftung, Germany.

**Data availability** All relevant sequence data were deposited in GenBank for public access.

**Code availability** Not applicable.

## Declarations

**Ethics approval** We declare that the experiments described in this paper comply with the current laws in Germany. All applicable international, national and/or institutional guidelines for the care and use of animals were followed. All procedures in our studies involving human participants (platelet aggregation experiments) were approved by the local ethics committee and performed according to the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Conflict of interest** 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

- Babenco VV, Podgorny OV, Manuvera VA, Kasianov AS, Manolov AI, Grafkskaia EN, Shirokov DA, Kurdyumov AS, Vinogradov DV, Nikitina AS, Kovalchuk SI, Anikanov NA, Butenko IO, Pobeguts OV, Matyushkina DS, Rakitina DV, Kostryukova ES, Zgoda VH, Baskova IP, Trukhan VM, Gelfand MS, Govorun VM, Schiöth HB, Lazarev VN (2020) Draft genome sequences of *Hirudo medicinalis* and salivary transcriptome of three closely related medicinal leeches. *BMC Genomics* 21(1):331. <https://doi.org/10.1186/s12864-020-6748-0>
- Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* 22(11):1399–1408. <https://doi.org/10.1038/nbt1029>
- Baskova IP, Zavalova LL (2001) Proteinase inhibitors from medicinal leech *Hirudo medicinalis*. *Biochemistry (mosc)* 66(7):869–883. <https://doi.org/10.1023/A:1010223325313>
- Baskova IP, Cherkesova DU, Mosolov VV (1983) Hirudin from leech heads and whole leeches and “pseudo-hirudin” from leech bodies. *Thromb Res* 30(5):459–467. [https://doi.org/10.1016/0049-3848\(83\)90180-9](https://doi.org/10.1016/0049-3848(83)90180-9)
- Baskova IP, Zavalova LL, Basanova AV, Moshkovskii SA, Zgoda VG (2004) Protein profiling of the medicinal leech salivary gland secretion by proteomic analytical methods. *Biochemistry (mosc)* 69(7):770–775. <https://doi.org/10.1023/b:biry.0000040202.21965.2a>
- Betz A, Hofsteenge J, Stone SR (1992) Interaction of the N-terminal region of hirudin with the active-site cleft of thrombin. *Biochemistry* 31(19):4557–4562. <https://doi.org/10.1021/bi00134a004>
- Blankenship DT, Brankamp RG, Manley GD, Cardin AD (1990) Amino acid sequence of ghilanten: anticoagulant-antimetastatic principle of the south American leech. *Haementeria Ghiliani* *Biochem Biophys Res Comm* 166(3):1384–1389. [https://doi.org/10.1016/0006-291X\(90\)91020-S](https://doi.org/10.1016/0006-291X(90)91020-S)
- Braun PJ, Dennis S, Hofsteenge J, Stone SR (1988) Use of site-directed mutagenesis to investigate the basis for the specificity of hirudin. *Biochemistry* 27(17):6517–6522. <https://doi.org/10.1021/bi00417a048>
- 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>
- Cherniack EP (2011) Bugs as drugs, part two: worms, leeches, scorpions, snails, ticks, centipedes, and spiders. *Altern Med Rev* 16(1):50–58
- Chuang P, Parikh C, Reilly RF (2001) A case review: anticoagulation in hemodialysis patients with heparin-induced thrombocytopenia. *Am J Nephrol* 21(3):226–231. <https://doi.org/10.1159/000046252>
- Combet C, Blanchet C, Deléage GC, G, (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci* 25(3):147–150. [https://doi.org/10.1016/s0968-0004\(99\)01540-6](https://doi.org/10.1016/s0968-0004(99)01540-6)
- Erlanger BF, Kokowsky N, Cohen W (1961) The preparation and properties of two new chromogenic substrates of trypsin. *Arch Biochem Biophys* 95:271–278. [https://doi.org/10.1016/0003-9861\(61\)90145-x](https://doi.org/10.1016/0003-9861(61)90145-x)
- Frame JN, Rice L, Bartholomew JR, Whelton A (2010) Rationale and design of the PREVENT-HIT study: a randomized, open-label pilot study to compare desirudin and argatroban in patients with suspected heparin-induced thrombocytopenia with or without thrombosis. *Clin Ther* 32(4):626–636. <https://doi.org/10.1016/j.clinthera.2010.04.012>
- Friedrich T, Kroger B, Bialojan S, Lemaire HG, Hoffken HW, Reuschenbach P, Otte M, Dödt J (1993) A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. *J Biol Chem* 268(22):16216–16222. [https://doi.org/10.1016/S0021-9258\(19\)85408-X](https://doi.org/10.1016/S0021-9258(19)85408-X)
- García-Fruitós E, González-Montalbán N, Morell M, Vera A, Ferraz RM, Arís A, Ventura S, García-Fruitós E, González-Montalbán N, Morell M, Vera A, Ferraz RM, Arís A, Ventura S, Villaverde A (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Fact* 4:27. <https://doi.org/10.1186/1475-2859-4-27>
- Greiner A, Lubenow N (2001) Recombinant hirudin in clinical practice: focus on lepirudin. *Circulation* 103(10):1479–1484. <https://doi.org/10.1161/01.cir.103.10.1479>

- 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>
- Guan DL, Yang J, Liu YK, Li Y, Mi D, Ma LB, Wang ZZ, Xu SQ, Qiu Q (2020) Draft Genome of the Asian buffalo leech *Hirudinaria manillensis*. *Front Genet* 10:1321. <https://doi.org/10.3389/fgene.2019.01321>
- Hildebrandt JP, Lemke S (2011) Small bite, large impact—saliva and salivary molecules in the medicinal leech. *Hirudo Medicinalis* *Naturwissenschaften* 98(12):995–1008. <https://doi.org/10.1007/s00114-011-0859-z>
- Holstein TW, Mala C, Kurz E, Bauer K, Greber M, David CN (1992) The primitive metazoan *Hydra* expresses antistasin, a serine protease inhibitor of vertebrate blood coagulation: cDNA cloning, cellular localisation and developmental regulation. *FEBS Lett* 309(3):288–292. [https://doi.org/10.1016/0014-5793\(92\)80791-e](https://doi.org/10.1016/0014-5793(92)80791-e)
- Kim H, Chu TT, Kim DY, Kim DR, Nguyen CM, Choi J, Lee JR, Hahn MJ, Kim KK (2007) The crystal structure of guamerin in complex with chymotrypsin and the development of an elastase-specific inhibitor. *J Mol Biol* 76(1):184–192. <https://doi.org/10.1016/j.jmb.2007.11.089>
- Klotz SA, Dorn PL, Mosbacher M, Schmidt JO (2014) Kissing bugs in the United States: risk for vector-borne disease in humans. *Environ Health Insights* 18(2):49–59. <https://doi.org/10.4137/EHI.S16003>
- Koh CY, Kini RM (2009) Molecular diversity of anticoagulants from haematophagous animals. *Thromb Haemost* 102(3):437–453. <https://doi.org/10.1160/TH09-04-0221>
- Koh CY, Kazimirova M, Trimmell A, Takac P, Labuda M, Nuttall PA, Kini RM (2007) Variagin, a novel fast and tight binding thrombin inhibitor from the tropical bont tick. *J Biol Chem* 282(40):29101–29113. <https://doi.org/10.1074/jbc.M705600200>
- Kvist S, Manzano-Marin A, de Carle D, Trontelj P, Siddall ME (2020) Draft genome of the European medicinal leech *Hirudo medicinalis* (Annelida, Clitellata, Hirudiniformes) with emphasis on anticoagulants. *Sci Rep* 10(1):9885. <https://doi.org/10.1038/s41598-020-66749-5>
- 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)
- Lei Q, Li C, Zuo Z, Huang C, Cheng H, Zhou R (2016) Evolutionary insights into RNA trans-splicing in vertebrates. *Genome Biol Evol* 8(3):562–577. <https://doi.org/10.1093/gbe/evw025>
- Li B, Gou M, Han J, Yuan X, Li Y, Li T, Jiang Q, Xiao R, Li Q (2018) Proteomic analysis of buccal gland secretion from fasting and feeding lampreys (*Lampetra morii*). *Proteome Sci* 16:9. <https://doi.org/10.1186/s12953-018-0137-5>
- Lukas P, Wolf R, Rauch BH, Hildebrandt JP, 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>
- Lundblad RL (1995) *Techniques in protein modification*. CRC Press, Boca Raton, Florida
- Macedo-Ribeiro S, Almeida C, Calisto BM, Friedrich T, Mentele R, Stürzebecher J, Fuentes-Prior P, Pereira PJB (2008) Isolation, cloning and structural characterisation of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. *PLoS One* 3(2):e1624. <https://doi.org/10.1371/journal.pone.0001624>
- Mans BJ, Louw AI, Neitz AWH (2002) Amino acid sequence and structure modeling of savignin, a thrombin inhibitor from the tick. *Ornithodoros Savignyi* *Insect Biochem Mol Biol* 32(7):821–828. [https://doi.org/10.1016/S0965-1748\(01\)00169-2](https://doi.org/10.1016/S0965-1748(01)00169-2)
- Manship BM, Walker AJ, Jones LA, Davies AJ (2012) Blood feeding in juvenile *Paragnathia formica* (Isopoda: Gnathiidae): biochemical characterization of trypsin inhibitors, detection of anticoagulants, and molecular identification of fish hosts. *Parasitology* 139(6):744–754. <https://doi.org/10.1017/S0031182011002320>
- Markwardt F (1955) Untersuchungen über Hirudin. *Naturwissenschaften* 42:537–538. <https://doi.org/10.1007/BF00630151>
- Markwardt F (1970) Hirudin as an inhibitor of thrombin. *Meth Enzymol* 19:924–932. [https://doi.org/10.1016/0076-6879\(70\)19082-3](https://doi.org/10.1016/0076-6879(70)19082-3)
- Moghadam M, Ganji A, Varasteh A, Falak R, Sankian M (2015) Refolding process of cysteine-rich proteins: chitinase as a model. *Rep Biochem Mol Biol* 4(1):19–24
- Mukherjee J, Gupta MN (2016) Paradigm shifts in our views on inclusion bodies. *Curr Biochem Eng* 3(1):47–55. <https://doi.org/10.2174/2212711902666150302212051>
- Müller C, Mescke K, Liebig S, Mahfound H, Lemke S, Hildebrandt JP (2016) More than just one: multiplicity of hirudins and hirudin like factors in the medicinal leech. *Hirudo Medicinalis* *Mol Genet Genomics* 291(1):227–240. <https://doi.org/10.1007/s00438-015-1100-0>
- Müller C, Haase M, Lemke S, Hildebrandt JP (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 L, Böhmert M, Hildebrandt JP (2019) Hirudin or hirudin-like factor - that is the question. *FEBS Lett* 594(5):841–850. <https://doi.org/10.1002/1873-3468.13683>
- Müller C, Lukas P, Lemke S, Hildebrandt JP (2019) Hirudin and decorsins of the North American medicinal leech *Macrob- della 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>
- Nar H (2012) The role of structural information in the discovery of direct thrombin and factor Xa inhibitors. *Trends Pharmacol Sci* 33(5):279–288. <https://doi.org/10.1016/j.tips.2012.03.004>
- Nikapitiya C, De Zoysa M, Oh C, Lee Y, Ekanayake PM, Whang I, Choi CY, Lee JS, Lee J (2010) Disk abalone (*Haliotis discus discus*) expresses a novel antistasin-like serine protease inhibitor: molecular cloning and immune response against bacterial infection. *Fish Shellfish Immunol* 28(4):661–671. <https://doi.org/10.1016/j.fsi.2009.12.027>
- Noeske-Jungblut C, Haendler B, Donner P, Alagon A, Possani L, Schleuning WD (1995) Triabin, a highly potent exosite inhibitor of thrombin. *J Biol Chem* 270(48):28629–28634. <https://doi.org/10.1074/jbc.270.48.28629>
- Oka K, Ohtaki N, Yasuhara T, Nakajima T (1989) A study of mosquito salivary gland components and their effects on man. *J Dermatol* 16(6):469–474. <https://doi.org/10.1111/j.1346-8138.1989.tb01587.x>
- Organisation to assess strategies for ischemic syndromes (OASIS-2) investigators (1999) Effects of recombinant hirudin (lepirudin) compared with heparin on death, myocardial infarction, refractory angina, and revascularisation procedures in patients with acute myocardial ischaemia without ST elevation: a randomised trial. *Lancet* 353(9151):429–439. [https://doi.org/10.1016/S0140-6736\(98\)07332-2](https://doi.org/10.1016/S0140-6736(98)07332-2)
- Palmer BR, Marinus MG (1994) The *dam* and *dcm* strains of *Escherichia coli* - a review. *Gene* 143(1):1–12. [https://doi.org/10.1016/0378-1119\(94\)90597-5](https://doi.org/10.1016/0378-1119(94)90597-5)
- Priestle JP, Rahuel J, Rink H, Tones M, Grütter MG (1993) Changes in interactions in complexes of hirudin derivatives and human alpha-thrombin due to different crystal forms. *Protein Sci* 2(10):1630–1642. <https://doi.org/10.1002/pro.5560021009>
- Richardson JL, Kroeger B, Hoeffken W, Sadler JE, Pereira P, Huber R, Bode W, Fuentes-Prior P (2000) Crystal structure of the human alpha-thrombin-haemadin complex: an exosite

- II-binding inhibitor. *EMBO J* 19(21):5650–5660. <https://doi.org/10.1093/emboj/19.21.5650>
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5(4):725–738. <https://doi.org/10.1038/nprot.2010.5>
- Salzter M (2000) Anticoagulants and inhibitors of platelet aggregation derived from leeches. *FEBS Lett* 492(3):187–192. [https://doi.org/10.1016/S0014-5793\(01\)02212-8](https://doi.org/10.1016/S0014-5793(01)02212-8)
- Scacheri E, Nitti G, Valsasina B, Orsini G, Visco C, Ferrera M, Sawyer T, 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>
- Scharf M, Engels J, Tripiet D (1989) Primary structures of new 'iso-hirudins.' *FEBS Lett* 255(1):105–110. [https://doi.org/10.1016/0014-5793\(89\)81070-1](https://doi.org/10.1016/0014-5793(89)81070-1)
- Simakov O, Marletaz F, Cho S-J, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo D-H, Larsson T, Lv J, Arendt D, Savage R, Osoegawa K, de Jong P, Grimwood J, Chapman JA, Shapiro H, Aerts A, Otiillar RP, Terry AY, Boore JL, Grigoriev IV, Lindberg DR, Seaver EC, Weisblat DA, Putnam NH, Rokhsar DS (2013) Insights into bilaterian evolution from three spiralian genomes. *Nature* 493(7433):526–531. <https://doi.org/10.1038/nature11696>
- Sørensen H, Sørensen S, Bjerregaard C, Michaelsen S (1999) Chromatography and capillary electrophoresis in food analysis. Royal Society of Chemistry, Cambridge, U.K.
- Strube KH, 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>
- Thomas GJ, Baneyx FO (1996) Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing Heat-shock proteins. *J Biol Chem* 271(19):11141–11147. <https://doi.org/10.1074/jbc.271.19.11141>
- Tong L, Dai S-X, Kong D-J, Yang P-P, Tong X, Tong X-R, Bi X-X, Su Y, Zhao Y-Q, Liu Z-C (2022) The genome of medicinal leech (*Whitmania pigra*) and comparative genomic study for exploration of bioactive ingredients. *BMC Genomics* 23(1):76. <https://doi.org/10.1186/s12864-022-08290-5>
- Vindigni A, De Filippis V, Zanotti G, Visco C, Orsini G, Fontana A (1994) Probing the structure of hirudin from *Hirudinaria manillensis* by limited proteolysis Isolation, characterization and thrombin-inhibitory properties of N-terminal fragments. *Eur J Biochem* 226(2):323–333. <https://doi.org/10.1111/j.1432-1033.1994.tb20056.x>
- Wallace A, Dennis S, Hofsteenge J, Stone SR (1989) Contribution of the N-terminal region of hirudin to its interaction with thrombin. *Biochemistry* 28(26):10079–10084. <https://doi.org/10.1021/bi00452a030>
- Wegner SV, Spatz JP (2013) Cobalt(III) als stabiles und inertes Vermittlerion zwischen Nitrilotriessigsäure und Proteinen mit His6-Tag. *Angew Chem* 125(29):7743–7746. <https://doi.org/10.1002/ange.201210317>
- Wegner SV, Schenk FC, Spatz JP (2016) Cobalt (III)-mediated permanent and stable immobilization of histidine-tagged proteins on NTA-functionalized surfaces. *Chemistry* 22(9):3156–3162. <https://doi.org/10.1002/chem.201504465>
- Whitlow M, Arnaiz DO, Buckman BO, Davey DD, Griedel B, Guilford WJ, Koovakkat SK, Liang A, Mohan R, Phillips GB, Seto M, Shaw KJ, Xu W, Zhao Z, Light DR, Morrissey MM (1999) Crystallographic analysis of potent and selective factor Xa inhibitors complexed to bovine trypsin. *Acta Crystallogr D Biol Crystallog* 55(Pt 8):1395–1404. <https://doi.org/10.1107/S0907444999007350>
- Xin X, Liu W, Sun L (2015) Improving the bioactivity of rHirudin with boronophenylalanine site-specific modification. *Mol Med Rep* 11(5):3774–3779. <https://doi.org/10.3892/mmr.2015.3194>
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12(1):7–8. <https://doi.org/10.1038/nmeth.3213>
- Zhang RM, Snyder GH (1989) Dependence of formation of small disulfide loops in two-cysteine peptides on the number and types of intervening amino acids. *J Biol Chem* 264(31):18472–18479. [https://doi.org/10.1016/S0021-9258\(18\)51490-3](https://doi.org/10.1016/S0021-9258(18)51490-3)
- Dunwiddie C, Thornberry NA, Bull HG, Sardana M, Friedman PA, Jacobs JW, Simpson E (1989) Antistasin, a leech-derived inhibitor of factor Xa. Kinetic analysis of enzyme inhibition and identification of the reactive site. *J Biol Chem* 264(28):16694–16699. [https://doi.org/10.1016/S0021-9258\(19\)84761-0](https://doi.org/10.1016/S0021-9258(19)84761-0)
- Grant GA (2017) Modification of cysteine. *Curr Protoc Protein Sci*:15.1.1–15.1.23. [10.1002/cpps.22](https://doi.org/10.1002/cpps.22)
- 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>
- Koh CY, Kini RM (2010) Thrombin Inhibitors from haematophagous animals. In: Kini R, Clemetson K, Markland F, McLane M, Morita T (eds) *Toxins and Hemostasis*. Springer, Dordrecht. pp 239–254. [https://doi.org/10.1007/978-90-481-9295-3\\_15](https://doi.org/10.1007/978-90-481-9295-3_15)
- Maraganore JM, Fenton JW (1990) Thrombin inhibition by synthetic hirudin peptides. In: Liu CY, Chien S (eds) *Fibrinogen, thrombolysis, coagulation, and fibrinolysis*. Advances in Experimental Medicine and Biology, vol 281. Springer, Boston, MA. pp 177–183. [https://doi.org/10.1007/978-1-4615-3806-6\\_17](https://doi.org/10.1007/978-1-4615-3806-6_17)
- Nicholas KB, Nicholas Jr HB (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author
- Ware FL, Luck MR (2017) Evolution of salivary secretions in haematophagous animals. *Biosci Horiz* 10:hzw015. <https://doi.org/10.1093/biohorizons/hzw015>

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