Hirudin or hirudin-like factor - that is the question: insights from the analyses of natural and synthetic HLF variants

Christian Müller [iD], Phil Lukas [iD], Michel Böhmert and Jan-Peter Hildebrandt [iD]

Animal Physiology and Biochemistry, Zoological Institute and Museum, University of Greifswald, Germany

The hirudin-like factor 1 (HLF1) of *Hirudo medicinalis* belongs to a new class of leech-derived factors. In previous investigations, HLF1 did not exhibit anticoagulatory activities. Here, we describe the analysis of natural and synthetic variants of HLF1 and HLF-Hyb, a yet uncharacterized member of the HLF family. Modifications within the N terminus of HLF1 have a strong impact on its activity. Some variants of HLF1 exhibit thrombin-inhibiting activity comparable to hirudins, whereas others have reduced or no activity. The analyses of HLF-Hyb variants revealed a strong impact of the central globular domain on activity. Our results indicate a comparable mode of action of hirudins and thrombin-inhibiting HLF variants. Finally, we propose and discuss criteria for classifying hirudins and HLFs.

Keywords: blood coagulation; Hirudin; Hirudin-like factors; *Hirudo medicinalis*; medicinal leeches

The saliva of hematophagous leeches comprises a complex mixture of bioactive molecules [1–4]. To date, merely a handful of these factors are functionally characterized and only the thrombin inhibitor hirudin found its way to clinical application [5,6]. The hirudin-like factors (HLFs) represent a recently described subclass of compounds derived from the salivary glands of medicinal leeches of the genera *Hirudo* and *Hirudinaria* [7,8]. 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 such as molecular weight and isoelectric point (pI value). In previous works, we have described the purification and functional characterization of HLF1 (originating from *Hirudo medicinalis*) [7] and HLF5, HLF6, and HLF8 (originating from *Hirudinaria manillensis*) [9], respectively. While HLF1 and HLF6 did not exhibit any measurable anticoagulatory activity, both HLF5 and

HLF8 did. However, the antithrombin activity of HLF5 and HLF8 was markedly lower compared to hirudins of the genera *Hirudo* and *Hirudinaria*, but was in a range of activity compared to a hirudin variant that is actually expressed in the North American leech *Macrobdella decora* (hirudin$_{Mdec}$) [10]. Based on these observations, we have recently proposed to use the terms 'full hirudin', 'unorthodox hirudins', and 'HLFs' to also linguistically differentiate between factors with high, medium, or low-to-zero thrombin-inhibitory activity.

Differences of proteins in their biological activities are very likely caused by structural differences. And indeed, there are pronounced differences in the primary structure (amino acid sequence) of full hirudins, unorthodox hirudins, and HLFs. Common theme are the six conserved cysteine residues, but features such as the length and composition of the C-terminal tail, the number and position of positively (hence basic) or negatively (hence acidic) charged amino acid residues,

or the overall pI value (isoelectric point) differ in a wide range [8]. Interestingly, all full and also all unorthodox hirudins that were functionally tested so far and that exhibited an anticoagulatory and antithrombin activity (justifying their classification as hirudins) comprise a pI value of about 4.1–4.6. These pI values are the result of an excess of acidic amino acid residues compared to basic amino acid residues, especially within the elongated C-terminal tail of these factors. Analyses of the thrombin–hirudin interactions based on cocrystallographic data revealed several hydrophilic and hydrophobic interactions of the C-terminal tail with the surface of thrombin, especially at the fibrinogen binding site (exosite 1) of thrombin [11,12]. In contrast, the N-terminal amino acid residues of hirudin plug into the active site of thrombin and hence block its catalytic activity [13,14]. As a result, hirudin is a so-called bivalent thrombin inhibitor [15,16]. Alterations of the N-terminal five amino acid residues or changes within the C-terminal tail greatly modify the inhibitory potency of hirudin [17–19]. The central globular domain of hirudin is stabilized by the formation of three disulfide bonds, and activity of the hirudin molecule completely depends on the correct formation of all bonds [20–22]. However, the effects of changes of other amino acid residues within the globular domain of hirudin are by far less well analyzed, compared to investigations of the N- and C-terminal tails [23]. Incorporation of an RGD motif by replacement of Ser_{32}-Asp_{33}-Gly_{34} with the tripeptide Arg-Gly-Asp between cysteine residues 5 and 6 of hirudin variant HV1 did not alter its antithrombin activity [24]. The RGD motif acts as a recognition site for integrin binding [25] and facilitates the inhibition of platelet aggregation by the leech factors ornatin [26], decorsin [27], and RGD–hirudin [28].

A cDNA encoding the HLF1 was first purified from a cDNA library of *Hirudo* sp. [7]. Later on, the cDNAs and genes of several HLF1 variants could be identified in specimens of *H. medicinalis*, *Hirudo verbana*, and *Hirudo orientalis*, respectively [8]. HLF1 is of about the same length (62 aa) as the hirudin variants (65–66 aa) and comprises a C-terminal tail that contains a large number (13) of acidic amino acid residues. In addition, basic amino acid residues are almost completely absent, with a lysine residue at the very C-terminal end as the only exception. The N terminus of HLF1 has a length of four amino acid residues (counted up to the first cysteine residue), with a tyrosine residue at position 2. In hirudins, a respective tyrosine residue is present at position 3 [29]. The several variants of HLF1 differ in the length of the C-terminal tail and in the actual number of acidic amino acid

residues within (ranging from 13 to 16). Additionally, they differ in the number (4 or 5) and the composition of the N-terminal amino acid residues [8]. Strikingly, all HLF1 variants with five amino acid residues at the N terminus comprise a tyrosine residue at position 3.

The aim of the present study was to evaluate the putative influence of the N terminus of HLF1 on its antithrombin activity. Furthermore, we wanted to characterize two additional natural HLFs of *H. medicinalis*, namely HLF2 and HLF-Hyb. HLF1 and HLF2 differ markedly in their pI values: 3.2 for HLF1 and 6.7 for HLF2, mainly caused by the different ratios of acidic versus basic amino acid residues (16 : 1 in HLF1, 7 : 8 in HLF2). These differences are especially prominent within the C-terminal tail, but are furthermore spread over the entire molecule including the N terminus and the central globular domain. HLF-Hyb represents a somehow curious example of HLFs, as it is a hybrid composed of both HLF1 and HLF2. It comprises the N-terminal part (including the central globular domain) of HLF2 and the elongated acidic C-terminal tail of HLF1. Consequently, both its pI value (about 4.2) and the acidic versus basic amino acid residue ratio (14 : 7) are right in between. Hence, HLF-Hyb might be a very good candidate to disentangle the putative influence of different parts of HLF molecules on their functionality.

Materials and methods

Genotyping of animals and tissue preparation

The biological material used in this study (specimens of *H. medicinalis*, *H. verbana*, and *H. orientalis* and salivary gland preparations) was already described by Müller *et al.* [7,8]. Information on genotyping data and GenBank accession numbers can be obtained from the same publications.

Overexpression and purification of His-tagged HLFs and hirudins

The procedure to clone cDNAs encoding HLFs and hirudins to overexpress and purify the respective proteins was previously described in detail [7]. Briefly, we applied a system developed by Qiagen (Hilden, Germany). The pQE30Xa vector encodes a factor Xa protease recognition site between the His-tag coding region on the 5′ side and the multiple cloning site on the 3′ side. Factor Xa protease treatment cleaves off the His-tag and results in a recombinant protein that is free of any vector-derived amino acids at the N terminus.

Partial cDNAs of HLFs were cloned into pQE30Xa in a way that the first amino acid of the respective HLF

(without the signal sequence) was located directly adjacent to the factor Xa protease cleavage site. The cDNA sequences of interest were amplified using appropriate primer pairs and Phusion™ High-Fidelity DNA polymerase (Thermo Scientific, Schwerte, Germany). For the overexpression, pQE30Xa clones containing inserts encoding the HLF variants of interest were transformed into appropriate *Escherichia coli* strains. Two flasks, each containing 500 mL of LB medium with ampicillin, were inoculated with 10 mL of a preculture. From the start of inoculation, optical densities were determined in a regular frequence. At an $OD_{600} = 0.5$, the expression of hirudin and HLF variants was induced by adding IPTG to a final concentration of 1 mM. After 4 h of expression, cells were harvested, the pellet was carefully resuspended in binding buffer (20 mM Tris/HCl, 500 mM NaCl, 5 mM imidazole, pH 7.9), and the cells were sonicated using a Sonopuls homogenizer (Bandelin, Berlin, Germany). After centrifugation for 1 h at 4 °C and 4500 r.p.m. (~ 3900 *g*) in a Labofuge 400R (Thermo Scientific), the supernatant was carefully removed, filtered through a filter with a pore size of 0.45 μm, and loaded on a self-packed column containing Ni–IDA His-Bind® resin (Merck, Darmstadt, Germany). Washing and elution steps were performed as recommended by the manufacturer of the resin. Equal volumes of every fraction were analyzed by SDS/PAGE on 20% gels. Prior to the treatment with factor Xa protease, fractions of interest were dialyzed twice for 24 h at 4 °C against a 100-fold excess of reaction buffer (20 mM Tris/HCl, 100 mM NaCl, 2 mM CaCl$_2$, pH 8.0) in a dialysis membrane with a MWCO of 5000 (Roth, Karlsruhe, Germany). The final volume was approximately 10 mL.

Factor Xa protease treatment and purification

The treatment of fusion proteins containing the factor Xa protease recognition sequence consisted of three steps: (a) factor Xa protease cleavage, (b) removal of factor Xa protease, and (c) cleanup of the digested protein. All steps were performed as recommended by the manufacturer (Qiagen). Purity of recombinant hirudins and HLFs was confirmed by SDS/PAGE on 20% gels. Molar concentrations of protein solutions were calculated by dividing the absorbance at 280 nm by the molar absorption coefficient according to the equation $\varepsilon = (nW \times 5500) + (nY \times 1490) + (nC \times 125)$ [30,31].

Blood coagulation assays

To verify the putative antithrombotic activity of purified natural and synthetic HLFs, the thrombin time test (TT, reference range 16.8–21.4 s) was performed 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 diluted with dialysis buffer to reach final concentrations of 50 and 5 μM, respectively. The desired amount of substrate was transferred directly into the cuvette immediately before the plasma was added. Dade® Ci-Trol® 1 (Siemens Healthcare) was used as standardized 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 clot formation.

Generation of synthetic HLF variants

The chimeric HLF variant HLF-Hyb2a was generated using the gene synthesis service of BioCat GmbH (Heidelberg, Germany). All other HLF variants were generated by PCR using appropriate primers to incorporate alterations in the nucleotide sequences that lead to the desired alterations in the amino acid sequences. Table 1 summarizes the origin (natural or synthetic) of all factors that are described and analyzed in the study.

Results

The main aim of the present study was to characterize variants of HLFs in their ability to negatively influence the thrombin activity and hence the blood coagulation cascade. Both naturally occurring and genetically modified variants of the HLF1 were tested in respective

Table 1. Origin (natural or synthetic) and N-terminal amino acid sequences of all factors described and analyzed in the study. Acidic residues are marked in red, basic residues are marked in blue, and neutral aliphatic residues are marked in green.

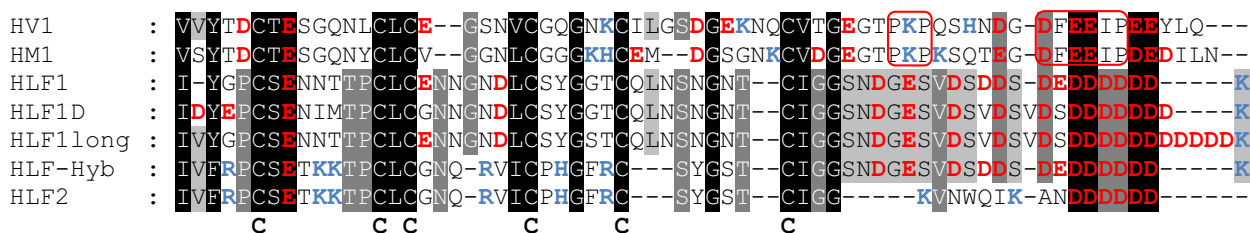| Factor	Origin	N terminus
HV1	Natural	VVYTD
HV2	Natural	ITYTD
HLF1	Natural	I-YGP
HLF2	Natural	IVFRP
HLF1V	Synthetic	IVYGP
VLF1long	Natural	IVYGP
HLF1D	Natural	IDYEP
HLF1D-DG	Synthetic	IDYGP
HLF1D-VE	Synthetic	IVYEP
HLF1D-VG	Synthetic	IVYGP
HLF-Hyb	Natural	IVFRP
HLF-HybG	Synthetic	IVFGP
HLF-HybD	Synthetic	IVFRD
HLF1V-YR	Synthetic	IVYRP
HLF1V-FR	Synthetic	IVFRP
HLF1V-FG	Synthetic	IVFGP
HLF1V-FRD	Synthetic	IVFRD
HLF-Hyb2a	Synthetic	IVYGP
HLF-Hyb2b	Synthetic	IVFRP

coagulation assays. In addition, the HLF2, hybrids between HLF1 and HLF2 (HLF-Hyb and HLF-Hyb2), and variants thereof were analyzed.
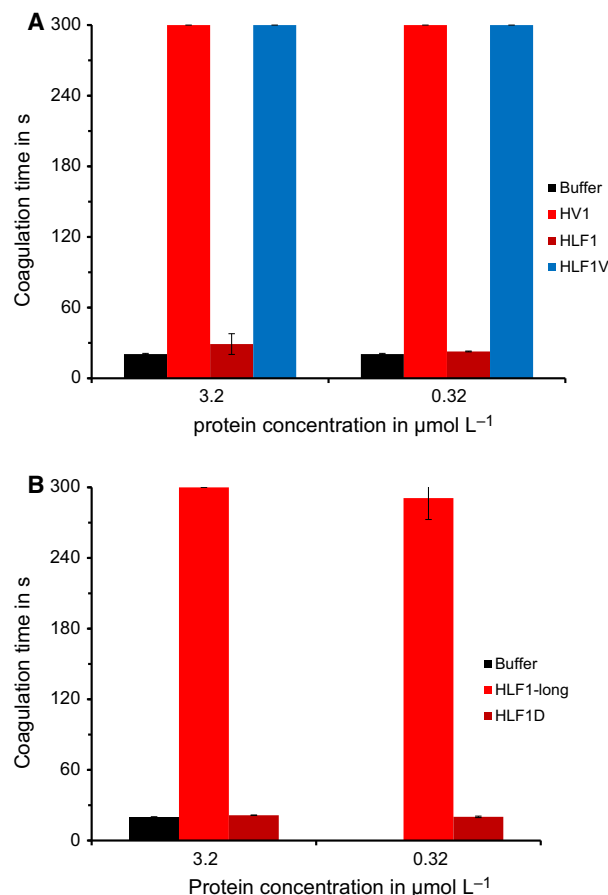
HLF1: effects of the N terminus

HLF1 of *H. medicinalis* was already previously analyzed and did not exhibit anticoagulatory activity when tested in the aPTT assay [7]. Comparison of natural HLF1 variants derived from specimen of *H. medicinalis*, *H. verbana,* and *H. orientalis*, respectively, revealed differences in the length of the C-terminal tail and in position 2 of the N terminus. Whereas all HLF1 variants of *H. medicinalis* identified so far displayed a gap at the respective position, the HLF1 variants of *H. verbana* contained either a gap, an aspartic acid residue, or a valine residue. Specimen of *H. orientalis* expressed variants with either a gap or a valine residue [8]. Figure 1 summarizes examples of the different HLF1 variants identified in *H. medicinalis* (HLF1), *H. verbana* (HLF1D), and *H. orientalis* (HLF1long).

In order to evaluate the effect of a valine residue at position 2, we genetically modified the corresponding cDNA of HLF1 of *H. medicinalis* and inserted an additional codon (GTT) at the respective position. The resulting synthetic factor was termed HLF1V and displayed a very strong antithrombin activity almost identical to hirudin variant HV1 of *H. medicinalis*. In contrast, HLF1 had a minor, but detectable, effect at only the highest concentration tested (Fig. 2A).

To further analyze the significance of position 2 in HLF1, we tested the natural variants HLF1D (containing an aspartic acid residue at position 2) of *H. verbana* and HLF1long (containing a valine residue at position 2 and a stretch of additional five aspartic acid residues at the C-terminal tail; see Fig. 1) of *H. orientalis*, respectively. HLF1D had no detectable

Fig. 2. Standard blood coagulation assays using the TT. (A) Hirudin variant HV1 and HLF variants HLF1 and HLF1V of *H. medicinalis*. (B) HLF variants HLF1-long of *H. orientalis* and HLF1D of *H. verbana*. n = 3, error bars indicate SD.

antithrombin activity even at the highest concentration. HLF1long in contrast had almost the same activity as HLF1V (Fig. 2B).

Fig. 1. Multiple sequence alignment of hirudin variants HV1 of *H. medicinalis* and HM1 of *H. manillensis* and HLF variants HLF1, HLF1V, HLF1D, HLF2, and HLF-Hyb. The alignments were generated using the CLS Sequence Viewer software package v8.0 (CLC bio, Aarhus, Denmark). Black background indicates conserved residues; gray background indicates similar residues. The six conserved cysteine residues giving rise to the three-dimensional structure are marked in bold, acidic amino acid residues are marked in red, and basic amino acid residues are marked in blue. The PKP and DFxxIP motifs are boxed. Abbreviations are used according to the IUPAC code.

Site-directed mutagenesis of HLF1D

The observation that HLF1D did not show any anti-coagulatory activity, but HLF1V and HLF1long did, raised the question, what particular amino acid residue might be responsible for the observed effects. HLF1D on the one hand and HLF1V and HLF1long on the other hand differ in two positions within the first five amino acid residues (before the first cysteine residue at position 6): position 2 (D versus V) and position 4 (E versus G) (Fig. 1). We genetically modified the cDNA of HLF1D to construct the following synthetic HLF1D-variants E4G (HLF1D-DG), D2V (HLF1D-VE), and E4G+D2V (HLF1D-VG) (Table 1). HLF1D-DG displayed only a small effect at the highest concentration, prolonging the coagulation time for a few seconds (just above the upper tolerance limit of the assay). In contrast, HLF1D-VG (with an N terminus identical to HLF1V and HLF1long; Table 1) had a very strong effect, almost comparable to HLF1V and HLF1long. The effect of HLF1D-VE was less pronounced, but still much higher compared to both HLF1 and HLF1D (Fig. 3).

The chimeric HLF: HLF-Hyb

HLF-Hyb comprises the N-terminal part of HLF2 (amino acids 1–38) and the C-terminal part of HLF1 (amino acids 39-59) (Fig. 1). We expressed and tried to purify both HLF2 and HLF-Hyb. Unfortunately, HLF2 almost completely precipitated during dialysis, most likely caused by the close proximity of the buffer pH value (8.0) and the pI value of HLF2 (6.7). HLF-Hyb exhibited only a weak inhibitory activity even

when tested at the highest possible concentration of 3.2 µM (Fig. 4B).

HLF-Hyb: effects of the central globular domain

HLF-Hyb differs from HV1 and HLF1V both in the presence of a basic amino acid residue at position 4 (R4) (Fig. 1) and in the length and charge of the central globular domain (Table 2). To disentangle the importance of both structural components, we constructed two N-terminal variants of HLF-Hyb: HLF-HybG and HLF-HybD. In the first variant, R4 was replaced by a glycine residue (R4G), and in the second variant, P5 was replaced by an aspartic acid residue (P5D). In both cases, the positive charge of the arginine residue R4 is either removed or counteracted. However, both variants of HLF-Hyb did show a detectable, but only weak, antithrombin activity (Fig. 4B). Consequently, the central globular domain
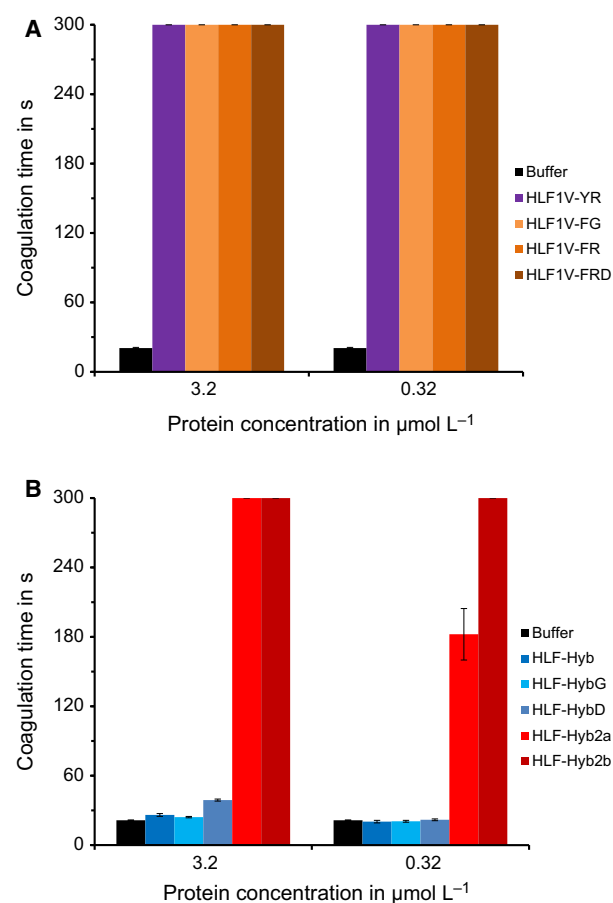
Fig. 4. Standard blood coagulation assays using the TT. (A) HLF variants HLF1V-YR, HLF1V-FG, HLF1V-FR, and HLF1V-FRD. (B) HLF variants HLF-Hyb, HLF-HybG, HLF-HybD, HLF-Hyb2a, and HLF-Hyb2b. n = 3, error bars indicate SD.

Fig. 3. Standard blood coagulation assays using the TT of HLF variant HLF1D, HLF1D-DG, HLF1D-VE, and HLF1D-VG. n = 3, error bars indicate SD.

Table 2. Molecular properties of hirudin variant HV1 and HLF variants HLF1 and HLF2 of *H. medicinalis*. C1-C6 indicates the number of amino acid residues between the cysteine residues 1 and 6 (including C1 and C6). The ratio acidic/basic indicates the number of acidic and basic amino acid residues within that region.

Factor	Length C1-C6	Acidic/Basic	pI value
HV1	34	4/2	4.41
HLF1	34	3/0	3.57
HLF2/HLF-Hyb	30	1/5	8.68

of HLF-Hyb is most likely responsible for its low inhibitory activity on thrombin.

To further support this assumption, we followed two complementary experimental approaches. First, we stepwise modified the N terminus of HLF1V and constructed four additional variants of HLF1V by replacing the glycine residue at position 4 with an arginine residue (G4R, HLF1V-YR) or by replacing the tyrosine residue at position 3 with a phenylalanine residue (Y3F, HLF1V-FG) in combination with changes at position 4 (Y3F + G4R, HLF1V-FR) and position 5 (Y3F + G4R + P5D, HLF1V-FRD), respectively. Hereby, the variant Y3F + G4R (HLF1V-FR) exactly mimics the N terminus of HLF2 (Table 1). Second, we constructed two reversed HLF1-HLF2 hybrids that contain the C terminus of HLF2 fused to the core domain of HLF1 in combination with either the N terminus of HLF1V (HLF-Hyb2a) or HLF2 (HLF-Hyb2b), respectively (Fig. 5 and Table 3). All synthetic N-terminal variants of HLF1V and both reversed HLF1-HLF2 hybrids showed a high antithrombin activity even at the lowest concentration tested (0.32 μM; Fig. 4A,B).

Discussion

Biology

So far, we have verified the expression of different HLF1 cDNAs in three species of the genus *Hirudo*: *H. medicinalis*, *Hirudo verbanda*, and *H. orientalis*, respectively [8]. These variants vary both in length (due to different numbers of repetitive aspartate residues in the C-terminal tail) and the composition of the N terminus. Whereas Y2 variants are expressed in all three species, the HLF1D variant could be identified in *H. verbana* only, and functionally active V2Y3 variants seem to be absent in *H. medicinalis*.

HLF1long (as a representative of the natural V2Y3 variants) exhibits an antithrombin activity comparable to the hirudin variants of *Hirudo* sp. and *H. manillensis*. It seems possible that these active HLF1 variants represent a functional redundancy with respect to thrombin inhibition. Worth mentioning in that context is the observation that *H. medicinalis* comprises at least four hirudin genes: one for HV1, one for HV2, and two for HV3. In contrast, both *H. verbana* and *H. orientalis* each comprise only two hirudin genes: HV1 and HV3 genes in *H. verbana* and two HV3 genes in *H. orientalis* [8]. One might speculate that the presence and expression of active HLF1 variants compensates for the underrepresentation of hirudin variants. Such an underrepresentation may have quantitative (gene dosis effects) or qualitative (target specificity on thrombins of different putative leech host species) implications. The putative biological functions of the Y2 and HLF1D variants (which lack anticoagulatory activities) remain to be determined.

Structure

The N termini of hirudins have been examined by mutational analyses of HV1 of *H. medicinalis* [19,32,33] and HM1 of *H. manillensis* [34,35]. These investigations were mainly carried out by substitutions or chemical modifications of particular amino acid residues of interest. To our best knowledge, there is no mutant of hirudin with a deletion of either the first or the second amino acid residue (thereby shifting the tyrosine residue Y3 to position 2: Y2 mutant) for a direct comparison of such a mutant with HLF1. However,
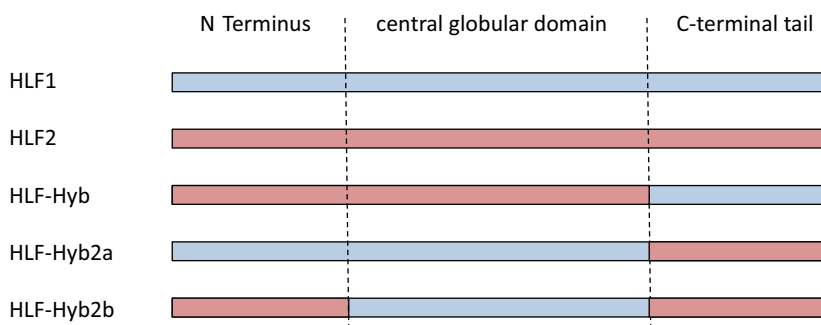
Fig. 5. Schematic representation of the N termini, the central globular domains, and the C-terminal tails of HLF1 (light blue), HLF2 (light red), and the chimeric variants HLF-Hyb, HLF-Hyb2a, and HLF-Hyb2b.

Table 3. Composition and activity of HLF1-HLF2-hybrids. The asterisk indicates a Y3F substitution in the N terminus of HLF-HybG.

Variant	N terminus	Central domain	C terminus	Inhibitory activity
HLF-Hyb	HLF2	HLF2	HLF1	Very low
HLF-HybG	HLF1*	HLF2	HLF1	Very low
HLF1V-FR	HLF2	HLF1	HLF1	Very high
HLF-Hyb2a	HLF1	HLF1	HLF2	High
HLF-Hyb2b	HLF2	HLF1	HLF2	Very high

addition of a methionine residue at the N terminus and therefore shifting the tyrosine residue Y3 to position 4 strongly decrease the activity of hirudin [36]. In addition, the hirudin mutant Y3A displays a more than hundred-fold increase in its Ki value compared to wild-type hirudin [19,37]. The tyrosine residue at exactly the position 3 is obviously crucial for full hirudin activity and can only be replaced by similar aromatic amino acid residues such as phenylalanine (Y3F) or tryptophan (Y3W) without negatively affecting the activity [34,35]. It is hence very likely that a Y2 mutant of hirudin will lack antithrombin activity comparable to HLF1 (Fig. 2).

In HV1, a substitution of the valin residue at position 2 with a glutamate residue (V2E) strongly reduces the affinity to thrombin [32]. Our data clearly indicate a similar effect of an acidic amino acid residue at position 2 in HLF1D, too. However, the substitution D2V in HLF1D (HLF1D-VE) does not completely restore full antithrombin activity (Fig. 3), most likely due to the acidic glutamate residue E4. Again, to our knowledge there is no directly comparable mutant (e.g., T4E) of hirudin documented. Whereas the substitution of E4 with a glycine residue alone (HLF1D-DG: E4G) has almost no effect on thrombin activity, the combination of both substitutions (HLF1D-VG: D2V and E4G) eventually restores full antithrombin activity (Fig. 4).

Replacements of the tyrosine residue at position 3 with a phenylalanine residue in HLF1 (HLF1-FG: Y3F) or the neutral glycine residue at position 4 with a basic arginine residue (HLF1V-YR: G4R) did not negatively affect the antithrombin activities of the respective HLF1V-variants (Fig. 4A). Whereas Y3F is a functionally similar amino acid residue replacement, G4R is not. Basic amino acid residues have been introduced so far only at positions 1 and 2 in HV1 (V1R, V1K, V2R, V2K) [19] and HM1 (S2R) [35], but occur in natural hirudins as well (e.g., R2 in hirudin P6 of *H. manillensis*) [38]. With the exception of V1K in HV1, they do not negatively affect the antithrombin activity, but may even enhance it (S2R in HM1). Our

data indicate that a basic amino acid residue at position 4 is not prohibitive for high activity, whether or not it is accompanied by a replacement of the apolar proline residue at position 5 with an acidic aspartate residue (P5D). Taken together, our data point to very similar mechanisms of interaction between the N termini of both hirudins and HLF1V, respectively, and the catalytic center of thrombin.

The C-terminal tail of hirudin blocks the exosite 1, the fibrinogen binding site, of thrombin [14,39,40]. This interaction is mainly due to several ionic and nonionic interactions between particular amino acid residues in both hirudin and thrombin [18,37,41]. Of particular importance within this context are the numerous acidic amino acid residues (7 out of 29 residues in HV1, 9 out of 27 residues in HM1) [17]. Interestingly, HLF1 (13 out of 24) and HLF1D (12 out of 26) contain even more acidic amino acid residues in their tails, but exhibited no (when tested for HLF1 with the aPPT test, Ref. [7]) or a very low (when tested with the more sensitive TT, Fig. 2) antithrombin activity. However, high activity could easily be restored by only a few modifications at the N termini (see above). Hence, an elongated acidic tail *per se* does not guarantee a significant inhibitory potency of a hirudin or a hirudin-like factor, but needs a functional N terminus as a molecular counterpart. The HLFs HLF5 and HLF8 of *H. manillensis* [9] and the hirudin of *M. decora* (hirudin_{Mdec}) [10,42] support this assumption. In all cases, these factors comprise C-terminal tails with a generally acid character (five acidic amino acid residues out of 23 in HLF5, eight out of 24 in HLF8, and six out of 26 in hirudin_{Mdec}, respectively), but with otherwise widely different primary structures compared to hirudins and HLF1. The N termini, however, perfectly fit into the general scheme [7]. We therefore conclude that the N termini of hirudins and HLFs are strongly restricted in their variability, whereas the C-terminal tails are much less constrained, as long as they comprise a general acidic character.

Compared to the termini, the core domain of hirudin got much less attention in terms of functional analysis. Only a few amino acid residues were substituted, and none of these alterations caused major changes in activity [23]. Our data, however, indicate a crucial role of the core domain for the activity of HLFs (and probably hirudins, too) as well. HLF-Hyb, a natural chimeric factor, comprises the N-terminal part (including the core domain) of HLF2 (amino acids 1–38) and the C-terminal part of HLF1 (amino acids 39–59) (Fig. 1). The factor exhibited only a very weak activity even when tested at the highest concentration of 3.2 μM (Fig. 4A). This lack of activity

cannot be explained by the general biochemical characteristics of the molecule (both the molecular mass of 6.4 kDa and the pI value of 4.16 are well within the range of hirudins) or the composition of either the N-terminal five amino acid residues or the C-terminal tail of HLF-Hyb: When combined with the core domain of HLF1, the resulting factor HLF-Hyb2b exhibits a very high antithrombin activity (Fig. 4B). In contrast, a synthetic hybrid factor (HLF-HybG) that comprises the termini of HLF1V-FG fused to the core domain of HLF2 lacks activity as well (Fig. 4B). The only reasonable explanation for these observations is a prohibitive effect of the core domain of HLF2. To our best knowledge, this is the first indication of such a strong influence of the globular domain of HLFs on the antithrombin activity of the whole molecule. To date, we can only speculate about the molecular determinants. The core domain of HLF2 (and HLF-Hyb) is a little shorter (30 amino acid residues) compared to both HV1 und HLF1 (34 amino acid residues), but more strikingly, the numbers of acidic versus basic residues and consequently the pI values markedly differ (Table 2). The core domain of HLF2 is basic (pI = 8.68), whereas the core domains of HV1 and HLF1 are acidic (pI = 4.41 and 3.57, respectively). These differences may alter the three-dimensional structure and/or the surface charge of the core domains and hence influence the interactions with the thrombin molecule. In both conditions, the correct sterical orientations of the N terminus and the C-terminal tail in the complex with thrombin may be enabled or hampered.

Nomenclature

In the present paper, we describe the functional characterization of both natural and synthetic variants of HLFs that exhibit antithrombin activities that can be almost zero, be in the medium range, or, for the first time, even be as high as the activities of hirudins. The question arises: What makes a hirudin a hirudin? Historically, the term 'hirudin' was introduced by Franz [43] and Jacoby [44] and simply refers to the genus of the respective source, the European medicinal leech, *H. medicinalis*. A more advanced definition of hirudin may be based on a combination of biological, functional, structural, and evolutionary properties. In that context, the term 'hirudin' in a broad sense might refer to secreted leech-derived factors (biology) with an antithrombin (and hence anticoagulatory) activity (function), a molecular mass of 5–7 kDa combined with a of signature of six cysteine residues forming three intramolecular disulfide bridges (structure) and a

common ancestor (evolution). Such a definition would include all hirudins such as HV1 and HM1, but also HLF1V, HLF5, HLF8, and Hirudin_{Mdec}. However, it would also include hemadin of *Haemadipsa sylvestris*, a factor that is functionally different from hirudins in a way that it blocks the exosite 2 of thrombin [45]. In contrast, the definition would not include HLF1 due to its almost complete lack of antithrombin activity. As described above, HLF1 and HLF1V only differ in the presence of one amino acid residue. An alternative definition of 'hirudin' might primarily focus on the historical perspective. In that context, the term 'hirudin' in a narrow sense refers to a factor that comprises antithrombin activity combined with the structural features of HV1 and HV2 of *H. medicinalis*, including the PKP and DFxxIP motifs in addition to the six conserved cysteine residues (Fig. 1). However, such a definition negates the evolutionary perspective because it would include factors such as HV1 and HM1, but exclude, for example, HLF1V and Hirudin_{Mdec}. As already mentioned above, HV1, HM1, HLF1, and Hirudin_{Mdec} comprise an identical gene structure and hence share a common ancestor and consequently the same evolutionary origin [8,10]. We do not have a solution that answers all questions and includes all possible aspects of the problem, but for now, we propose the following terminology:

- 'classical hirudins' are the hirudins in a narrow sense, for example, HV1 and HM1
- 'unorthodox hirudins' are factors that lack the additional motifs of classical hirudins (as mentioned above), but comprise antithrombin activity, for example, HLF5, HLF8, HLF1V, and Hirudin_{Mdec}
- 'HLFs' are factors that comprise the key structural features of hirudins, but lack the antithrombin activity, for example, HLF1 and HLF6.

Acknowledgements

Author contributions

CM conceived the ideas and designed the methodology; CM, PL, and MB purified leech factors and measured their activities; CM analyzed 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.

Ethical 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.

References

1 Ascenzi P, Amiconi G, Bode W, Bolognesi M, Coletta M and Menegatti E (1995) Proteinase inhibitors from the European medicinal leech *Hirudo medicinalis*: structural, functional and biomedical aspects. *Mol Aspects Med* **16**, 215–313.

2 Baskova IP and Zavalova LL (2001) Proteinase inhibitors from the medicinal leech *Hirudo medicinalis*. *Biochemistry (Mosc.)* **66**, 703–714.

3 Baskova IP, Kostrjukova ES, Vlasova MA, Kharitonova OV, Levitskiy SA, Zavalova LL, Moshkovskii SA and Lazarev VN (2008) Proteins and peptides of the salivary gland secretion of medicinal leeches *Hirudo verbana*, *H. medicinalis*, and *H. orientalis*. *Biochemistry (Mosc.)* **73**, 315–320.

4 Hildebrandt JP and Lemke S (2011) Small bite, large impact – saliva and salivary molecules in the medical leech, *Hirudo medicinalis*. *Naturwiss* **98**, 995–1008.

5 Becker RC and Cannon CP (1994) Hirudin: Its biology and clinical use. *J Thromb Thrombolysis* **1**, 7–16.

6 Greinacher A and Warkentin TE (2008) The direct thrombin inhibitor hirudin. *Thromb Haemost* **99**, 819–829.

7 Müller C, Mescke K, Liebig S, Mahfoud H, Lemke S and 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.

8 Müller C, Haase M, Lemke S and Hildebrandt J-P (2017) Hirudins and hirudin-like factors in *Hirudinidae*: implications for function and phylogenetic relationships. *Parasitol Res* **116**, 313–325.

9 Lukas P, Wolf R, Rauch BH, Hildebrandt J-P and Müller C (2019) Hirudins of the Asian medicinal leech, *Hirudinaria manillensis*: same same, but different. *Parasitol Res* **118**, 2223–2233.

10 Müller C, Lukas P, Lemke S and Hildebrandt J-P (2019) Hirudin and decorsins of the North American medicinal leech *Macrobdella decora*: gene structure reveals homology to hirudins and hirudin-like factors

of Eurasian medicinal leeches. *J Parasitol* **105**, 423–431.

11 Chang JY, Ngai PK, Rink H, Dennis S and Schlaeppi JM (1990) The structural elements of hirudin which bind to the fibrinogen recognition site of thrombin are exclusively located within its acidic C-terminal tail. *FEBS Lett* **261**, 287–290.

12 Rydel TJ, Ravichandran KG, Tulinsky A, Bode W, Huber R, Roitsch C and Fenton JW (1990) The structure of a complex of recombinant hirudin and human α-thrombin. *Science* **249**, 277–280.

13 Grütter MG, Priestle JP, Rahuel J, Grossenbacher H, Bode W, Hofsteenge J and Stone SR (1990) Crystal structure of the thrombin-hirudin complex: a novel mode of serine protease inhibition. *EMBO J* **9**, 2361–2365.

14 Rydel TJ, Tulinsky A, Bode W and Huber R (1991) Refined structure of the hirudin-thrombin complex. *J Mol Biol* **221**, 583–601.

15 Bichler J and Fritz H (1991) Hirudin, a new therapeutic tool? *Ann Hematol* **63**, 67–76.

16 Warkentin TE (2004) Bivalent direct thrombin inhibitors: hirudin and bivalirudin. *Best Pract Res Clin Haematol* **17**, 105–125.

17 Betz A, Hofsteenge J and Stone SR (1991a) Ionic interactions in the formation of the thrombin-hirudin complex. *Biochem J* **275** (Pt 3), 801–803.

18 Betz A, Hofsteenge J and Stone SR (1991b) Role of interactions involving C-terminal nonpolar residues of hirudin in the formation of the thrombin-hirudin complex. *Biochemistry* **30**, 9848–9853.

19 Betz A, Hofsteenge J and Stone SR (1992) Interaction of the N-terminal region of hirudin with the active-site cleft of thrombin. *Biochemistry* **31**, 4557–4562.

20 Chatrenet B and Chang JY (1992) The folding of hirudin adopts a mechanism of trial and error. *J Biol Chem* **267**, 3038–3043.

21 Chatrenet B and Chang JY (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. *J Biol Chem* **268**, 20988–20996.

22 Chang JY (1994) Controlling the speed of hirudin folding. *Biochem J* **300** (Pt 3), 643–650.

23 Betz A, Hopkins PC, Le Bonniec BF and Stone SR (1994) Contribution of interactions with the core domain of hirudin to the stability of its complex with thrombin. *Biochem J* **298** (Pt 2), 507–510.

24 Mo W, Zhang YL, Chen HS, Wang LS and Song HY (2009) A novel hirudin derivative characterized with anti-platelet aggregations and thrombin inhibition. *J Thromb Thrombolysis* **28**, 230–237.

25 Ruoslahti E and Pierschbacher MD (1987) New perspectives in cell adhesion: RGD and integrins. *Science* **238**, 491–497.

26 Mazur P, Henzel WJ, Seymour JL and Lazarus RA (1991) Ornatins: potent glycoprotein IIb-IIIa antagonists

and platelet aggregation inhibitors from the leech *Placobdella ornata*. *Eur J Biochem* **202**, 1073–1082.

27 Krezel AM, Wagner G, Seymour-Ulmer J and Lazarus RA (1994) Structure of the RGD protein decorsin: conserved motif and distinct function in leech proteins that affect blood clotting. *Science* **264**, 1944–1947.

28 Liu X, Yan X, Mo W, Song H and Dai L (2005) 1H NMR assignment and secondary structure of recombinant RGD-hirudin. *Magn Reson Chem* **43**, 956–961.

29 Scharf M, Engels J and Tripier D (1989) Primary structures of new 'iso-hirudins'. *FEBS Lett* **255**, 105–110.

30 Gill SC and von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **182**, 319–326.

31 Pace CN, Vajdos F, Fee L, Grimsley G and Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* **4**, 2411–2423.

32 Wallace A, Dennis S, Hofsteenge J and Stone SR (1989) Contribution of the N-terminal region of hirudin to its interaction with thrombin. *Biochemistry* **28**, 10079–10084.

33 Lazar JB, Winant RC and Johnson PH (1991) Hirudin: amino-terminal residues play a major role in the interaction with thrombin. *J Biol Chem* **266**, 685–688.

34 De Filippis V, Vindigni A, Altichieri L and Fontana A (1995) Core domain of hirudin from the leech *Hirudinaria manillensis*: chemical synthesis, purification, and characterization of a Trp3 analog of fragment 1–47. *Biochemistry* **34**, 9552–9564.

35 De Filippis V, Quarzago D, Vindigni A, Di Cera E and Fontana A (1998) Synthesis and characterization of more potent analogues of hirudin fragment 1–47 containing non-natural amino acids. *Biochemistry* **37**, 13507–13515.

36 Loison G, Findeli A, Bernard S, Nguyen-Juilleret M, Marquet M, Riehl-Bellon N, Carvallo D, Guerra-Santos L, Brown SW, Courtney M *et al.* (1988) Expression and secretion in *S. cerevisiae* of biologically active leech hirudin. *Bio/Technology* **6**, 72–77.

37 Huang Y, Zhang Y, Zhao B, Xu Q, Zhou X, Song H, Yu M and Mo W (2014) Structural basis of RGD-hirudin binding to thrombin: Tyr3 and five C-terminal residues are crucial for inhibiting thrombin activity. *BMC Struct Biol* **14**, 26.

38 Steiner V, Knecht R, Boernsen KO, Gassmann E, Stone SR, Raschdorf F, Schlaeppi JM and Maschler R (1992) Primary structure and function of novel O-glycosylated hirudins from the leech *Hirudinaria manillensis*. *Biochemistry* **31**, 2294–2298.

39 Maraganore JM, Chao B, Joseph ML, Jablonski J and Ramachandran KL (1989) Anticoagulant activity of synthetic hirudin peptides. *J Biol Chem* **264**, 8692–8698.

40 Fenton JW 2nd, Villanueva GB, Ofosu FA and Maraganore JM (1991) Thrombin inhibition by hirudin: how hirudin inhibits thrombin. *Haemostasis* **21** (Suppl 1), 27–31.

41 Braun PJ, Dennis S, Hofsteenge J and Stone SR (1988) Use of site-directed mutagenesis to investigate the basis for the specificity of hirudin. *Biochemistry* **27**, 6517–6522.

42 Min GS, Sarkar IN and Siddall ME (2010) Salivary transcriptome of the North American medicinal leech, *Macrobdella decora*. *J Parasitol* **96**, 1211–1221.

43 Franz F (1903) Ueber den die Blutgerinnung aufhebenden Bestandtheil des medicinischen Blutegels. *Naunyn Schmiedeberg's Arch Exp Path Pharmakol* **49**, 342–366.

44 Jacoby C (1904) Über Hirudin. *Dtsch Med Wochenschr* **30**, 1786–1794.

45 Richardson JL, Kröger B, Hoeffken W, Sadler JE, Pereira P, Huber R, Bode W and Fuentes-Prior P (2000) Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. *EMBO J* **19**, 5650–5660.