Tyrosine Sulfation Modulates Activity of Tick-Derived Thrombin Inhibitors

Robert E. Thompson,[1‡] Xuyu Liu,[1‡] Jorge Ripoll-Rozada,[2,3‡] Noelia Alonso-García,[2,3] Benjamin L. Parker,[4] Pedro José Barbosa Pereira,[2,3] Richard J. Payne[1]

[1]School of Chemistry, The University of Sydney, Sydney, NSW 2006, Australia

[2]IBMC - Instituto de Biologia Molecular e Celular, Universidade do Porto, 4200-135 Porto, Portugal

[3]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal

[4]Charles Perkins Centre, The University of Sydney, NSW 2006, Australia

[‡] These authors contributed equally

Correspondence and requests for materials should be addressed to Richard J. Payne (email: richard.payne@sydney.edu.au)

ABSTRACT

Madanin-1 and chimadanin are two small cysteine-free thrombin inhibitors that facilitate blood feeding in the tick *Haemaphysalis longicornis*. Here we report a hitherto undiscovered post-translational modification of these two proteins – tyrosine sulfation – that is critical for potent anti-thrombotic and anticoagulant activity. Inhibitors produced in baculovirus-infected insect cells displayed heterogeneous sulfation of two tyrosine residues within each of the proteins. One-pot ligation-desulfurization chemistry enabled access to homogeneous samples of all possible sulfated variants of the proteins. Tyrosine sulfation of madanin-1 and chimadanin proved crucial for thrombin inhibitory activity, with the doubly sulfated variants three orders of magnitude more potent than the unmodified inhibitors. The three-dimensional structure of madanin-1 in complex with thrombin revealed a unique mode of inhibition, with the sulfated tyrosine residues binding to the basic exosite II of the protease. The importance of tyrosine sulfation within this family of thrombin inhibitors, together with their unique binding mode, paves the way for the development of anti-thrombotic drug leads based on these privileged scaffolds.

Blood-feeding organisms, including leeches, flies, mosquitos and ticks, produce an arsenal of anti-immune and anticoagulant salivary proteins to neutralize the biological defences of their hosts.[1-4] Vertebrate haemostatic mechanisms, namely blood clotting, constitute a major obstacle to meal acquisition and digestion by blood-feeding parasites. To date, a number of proteins from hematophagous invertebrates have been identified that target thrombin, a central proteinase in the blood clotting cascade. Thrombin contributes to many stages of the clotting process, however, the enzyme is best characterized in the proteolytic cleavage of a soluble protein known as fibrinogen, to the insoluble protein fibrin, that initiates the formation of an insoluble clotted mesh known as the hemostatic plug. Thrombin-targeting inhibitors produced by blood-feeding organisms typically display potent inhibition of the enzyme, often targeting both the active site of the serine proteinase as well as the basic exosite I (or fibrinogen-binding exosite) required for efficient recognition of many of its physiological substrates (Fig. 1A).[5-10] For example, hirudin (from the medicinal leech *Hirudo medicinalis*) efficiently blocks thrombin procoagulant activity by establishing strong electrostatic interactions between its acidic and unstructured C-terminal tail and the exosite I of thrombin, while the globular N-terminal region blocks the active site of the enzyme.[5] Sulfation of a tyrosine (Y) residue within the acidic tail is a native post-translational modification (PTM) that increases the affinity of hirudin towards thrombin by more than 10-fold.[1,11]

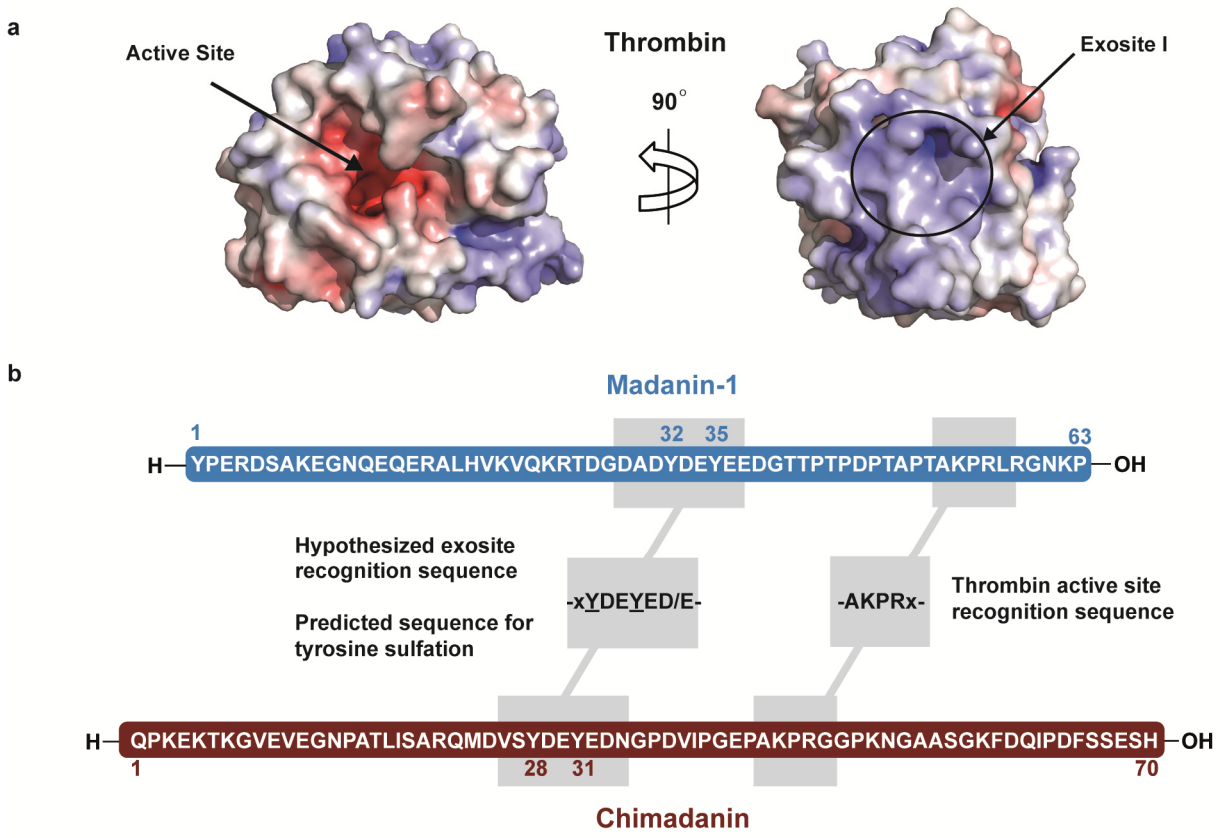

Figure 1. Madanin-1 and chimadanin: thrombin inhibitors from the bush tick *Haemaphysalis longicornis.* **a,** Representation of human thrombin apo structure (PDB entry 3U69)[12] showing (left panel) the active site cleft, and

rotated 90º around y (right panel) the exosite I. **b,** Amino acid sequences of madanin-1 and chimadanin from the bush tick *Haemaphysalis longicornis,* with predicted recognition sequences[13] and sites for tyrosine sulfation highlighted.

While a number of other disulfide-rich natural anticoagulants have been identified that target thrombin,[1,3] several structurally unrelated small cysteine-free thrombin inhibitors from hematophagous invertebrates have recently been shown to inhibit this crucial enzyme.[14-19] These molecules do not display any significant sequence similarity, and their molecular mechanisms of action are also diverse. This is highlighted by the three-dimensional structures of the complexes formed between thrombin and the substrate-like variegin (MEROPS family I74) or the unique proteolysis-resistant anophelin (MEROPS family I77).[9,10] Two other small and flexible specific thrombin inhibitors, madanin-1 and chimadanin (MEROPS families I53 and I72, respectively), have been identified in the bush tick *Haemaphysalis longicornis*.[17,18] The 60-residue madanin-1 was functionally characterized using recombinant protein produced in a prokaryotic system.[13] Unmodified madanin-1 exhibited an inhibition constant (*K*i) for bovine thrombin of 55.6 ± 5.5 nM and prolonged the thrombin-induced clotting time of human plasma (thrombin time, TT) by two-fold at a concentration of 5 μM. In striking contrast to the highly potent inhibitory activity of other proteinaceous inhibitors of thrombin, madanin-1 behaves as a cleavable competitive inhibitor, and loses affinity to thrombin upon proteolysis.[13] It is conceivable, however, that critical PTMs (unavailable in traditional prokaryotic expression systems) may enhance the anticoagulant activity of this family of inhibitors. Indeed, with the notable exception of hirudin, the presence of PTMs in thrombin inhibitors from hematophagous animals has not been determined, despite cases in which the native inhibitor is more potent than recombinant or synthetic forms, e.g. TTI from the fly *Glossina morsitans*, that may point toward activity modulation by the presence of unidentified PTMs.[20]

Upon inspection of the madanin-1 and chimadanin amino acid sequences we noted the presence of two tyrosine residues within conserved acidic stretches (Fig. 1B), which constitute good candidates for post-translational sulfation by comparison to conserved tyrosine sulfation sites found within a number of eukaryotic proteins.[21] We hypothesized that specific sulfation at these sites would likely enhance binding to regions of basic electrostatic potential (exosites) on the surface of thrombin, thus leading to increased anticoagulant activity. In this study we thoroughly address these hypotheses by: 1) identifying sulfation sites within madanin-1 and chimadanin *via* heterologous expression in insect cells, 2) accessing homogeneously and site-specifically modified madanin-1 and chimadanin proteins through chemical synthesis, and 3) characterizing the mode of thrombin inhibition of the sulfated inhibitors by X-ray crystallography. Herein, we demonstrate a dramatic increase in the bioactivity of these small anti-thrombotic proteins following introduction of tyrosine sulfation, and that madanin-1 inhibits thrombin by occupying the active site and exosite II of the enzyme. Therefore, we describe the identification of a new mechanistic class of cysteine-free thrombin inhibitors.

RESULTS

Madanin-1 and chimadanin undergo specific tyrosine sulfation *in vivo*

To date, the identification and functional characterization of the vast majority of thrombin inhibitors from blood-feeding organisms has proven possible only using recombinant DNA technology, due to the paucity of proteinaceous material that can be isolated from natural sources. While unambiguously establishing the primary sequence of anticoagulants, this approach does not allow for analysis of PTMs that may modulate protein activity. In both madanin-1 and chimadanin, two tyrosine residues (Y32 and Y35 in madanin-1, and Y28 and Y31 in chimadanin) are found in the vicinity of conserved acidic amino acids, making them likely substrates for the trans-Golgi tyrosylprotein sulfotransferase enzymes (TPST-1 and TPST-2) responsible for specific tyrosine sulfation of proteins traversing the secretion pathway.[21-23] In order to examine this possibility, codon-optimized sequences encoding madanin-1 and chimadanin fused N-terminally to the honeybee mellitin signal sequence to robustly direct the recombinant proteins to the secretion pathway were expressed in *Trichopulsia ni* High Five insect cells. Following expression, the cell media containing the secreted proteins was analyzed by liquid chromatography coupled to tandem mass spectrometry. High resolution mass analysis of chimadanin identified the major expression product containing two sulfates with approximately 30% mono-sulfated and 10% non-sulfated with and without a pyroglutamate residue on the N-terminus (Supplementary Fig. 1 and Supplementary Table 1). Tandem mass spectral analysis of the disulfated isoform with higher collisional dissociation localized the modifications to Y28 and Y31, consistent with our predictions. Analysis of expression medium containing madanin-1 revealed a highly heterogeneous mixture bearing both *O*-linked glycosylation and sulfation. In this instance, the non-sulfated variants were the major species present, with approximately 25% comprising of the two monosulfated and disulfated variants (Supplementary Fig. 2 and Supplementary Table 2). Tandem mass analysis with electron transfer dissociation (ETD) was performed to pinpoint the modification sites and demonstrated that sulfation occurred on tyrosine residues Y32 and Y35 (rather than on O-glycans).

Synthesis of madanin-1 and chimadanin bearing homogeneous sulfation patterns

Having determined that both madanin-1 and chimadanin are sulfated when expressed in eukaryotic cells, we were next interested in characterizing the effect of tyrosine sulfation at both putative modification sites of madanin-1 and chimadanin on thrombin inhibitory activity. The heterogeneous nature of the expressed sulfoproteins, together with the small quantities obtained through expression in insect cells prompted us to explore total chemical synthesis as a means to access each of the three possible sulfoforms of both madanin-1 and chimadanin in homogeneous form. Madanin-1 sulfoforms were assembled in the N- to C-terminal direction *via* a one-pot kinetically-controlled ligation-desulfurization strategy that we have reported previously[24] (see Supplementary Methods for full synthetic details). This first step involved preparation of three peptide fragments by Fmoc-strategy solid-phase peptide synthesis (SPPS)

followed by stepwise assembly of the full-length polypeptide sequences by native chemical ligation. The synthesis began with ligation of the peptide **1** (corresponding to N-terminal residues 1-28) activated as a S-trifluoroethyl thioester with one of peptides **2, 3** or **4**, spanning the acidic stretch of madanin-1 (residues 29-47) bearing an unactivated S-propionate ethyl ester thioester moiety on the C-terminus and a nucleophilic β-SH aspartic acid[25] functionality on the N-terminus (see Supplementary Methods for full synthetic details). Depending on the desired madanin-1 sulfoform targeted, this central peptide was prepared with sulfotyrosine residues site specifically installed at Y32 (**2**), Y35 (**3**) or Y32 and Y35 (**4**). The highly acid labile phenolic sulfate ester functionality in these sulfopeptides was protected as neopentyl (nP) sulfate esters to enable assembly by Fmoc-SPPS, with the knowledge that this protecting group could be cleaved in the presence of nucleophilic reagents.[21,26] Importantly, this synthetic strategy allowed us to modify only the central peptide thioester fragment in the synthetic route to access each of the three possible sulfoforms of madanin-1. The first chemoselective ligation between peptide thioester **1** and peptide thioester **2**, **3** or **4** was performed under standard native chemical ligation conditions [6 M guanidine hydrochloride (Gnd.HCl), 200 mM Na_2HPO_4, 50 mM *tris*-(2-carboxyethyl)phosphine (TCEP), pH 7.4-7.5] in the absence of an external thiol additive and was complete in 1 h (as judged by HPLC-MS analysis) to afford **5-7**. Serendipitously, these reaction conditions also led to deprotection of the neopentyl sulfate esters, with the unprotected sulfotyrosine moieties completely stable to the ensuing conditions. Trifluoroethanethiol (TFET) was added directly to crude **5-7** to convert the unreactive S-propionate ethyl ester thioester to an activated S-trifluoroethyl thioester acyl-donor. From here, peptide **8**, corresponding to the C-terminus of madanin-1 (residues 48-60) and bearing an N-terminal cysteine residue, was added to assemble the full length polypeptide sequence of the differentially sulfated madanin-1 proteins **9-11**. Without purification, conversion of the non-native cysteine and β-SH aspartic acid residues in **9-11** to native alanine and aspartate, respectively, was achieved under radical desulfurization conditions[27] to afford the native madanin-1 sulfoproteins in crude form. Following a single reversed-phase HPLC purification (using 0.1 M ammonium acetate and acetonitrile as eluents to prevent acidolysis of the fragile tyrosine sulfate ester) this one-pot ligation-desulfurization sequence provided madanin-1 bearing single sulfotyrosine residues at residue Y32 (**12**, 28% isolated yield), Y35 (**13**, 54% isolated yield), or a doubly sulfated isoform with modifications at Y32 and Y35 (**14**, 40% isolated yield) on multi-milligram scales.

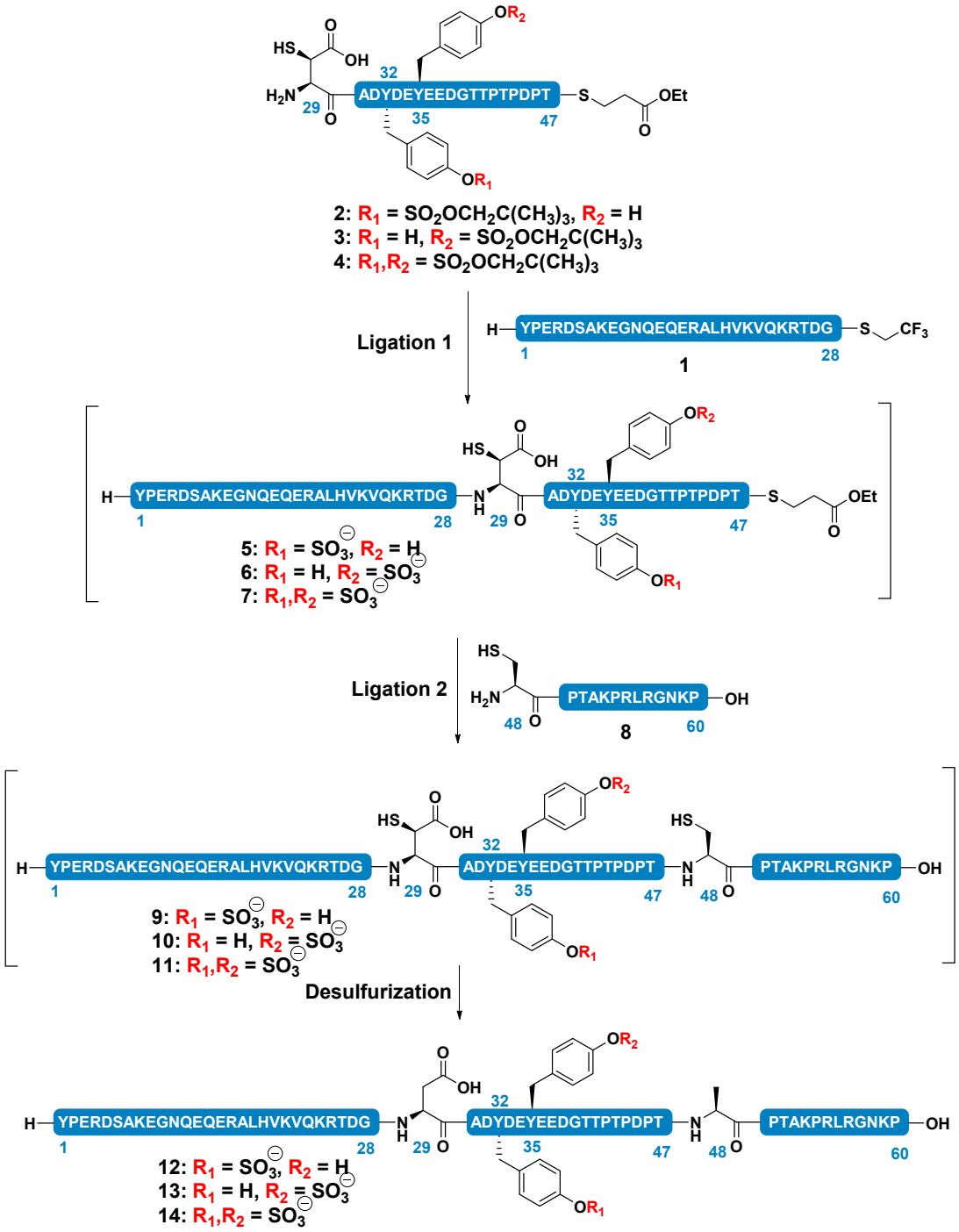

Figure 2. Assembly of homogeneously sulfated madanin-1 variants through one-pot kinetically-controlled ligation-desulfurization in the N- to C-direction. Conditions: **Ligation 1:** madanin-1(29-47) thioester (**2**, **3** or **4**, 1.0 equiv.), madanin-1(1-28) thioester **1** (1.1-1.2 equiv, 5 mM), 6 M Gnd.HCl, 200 mM Na_2HPO_4, 50 mM TCEP, pH 7.4-7.5, 1 h (nP protecting groups in **2-4** were labile under the ligation conditions and afforded the free aryl mono sulfate esters); **Ligation 2:** *in situ* addition of madanin-1(48-60) **8** (1.2-1.3 equiv.), 2 vol.% TFET, pH 7.4-7.5, 12 h; **Desulfurization:** degassed with argon, followed by addition of TCEP (200 mM), reduced glutathione (40 mM) and VA-044 (20 mM), 16 h, 37 °C. HPLC purification (eluent: 0.1 M NH_4OAc → MeCN) afforded sY32-monosulfated madanin-1 (**12**, 28% yield), sY35-monosulfated madanin-1 (**13**, 54% yield) and sY32,sY35 disulfated madanin-1 (**14**, 40% yield). NB: isolated yields are after 3 steps carried out in one-pot, square brackets around structures indicate that the compound was not isolated.

In a similar manner to the preparation of madanin-1 sulfoproteins **12-14** the synthesis of the three possible sulfoforms of the 70-residue protein chimadanin involved the ligation-based assembly of three synthetic peptide fragments: 1) a C-terminal peptide (corresponding to residues 42-70) bearing an N-terminal γ-(SSMe)-glutamate moiety[28] (**15**), 2) peptides (corresponding to residues 21-41) with variation in the sulfation pattern at Y28 and Y31 as well as a C-terminal S-propionate ethyl ester thioester and an N-terminal thiazolidine moiety (**16-18**), and 3) an N-terminal fragment **19** corresponding to residues 1-20, containing an N-terminal pyroglutamate residue (to prevent the formation of mixed species arising from acid-catalyzed deamination during synthetic manipulation). In contrast to madanin-1, the peptide fragments for chimadanin were assembled in the C-to-N direction in one-pot through two sequential ligation reactions, separated by a thiazolidine deprotection step and a final desulfurization reaction to afford the fully assembled proteins (see Supplementary Methods for full synthetic details). Briefly, C-terminal peptide **15** was reacted with sulfopeptide thioesters **16**, **17** or **18** under standard ligation conditions using TFET as a thiol additive[24] and, following completion of the reactions (as judged by HPLC-MS), *in situ* treatment with methoxyamine at pH 4.2 facilitated thiazolidine deprotection. Introduction of peptide thioester **19** in the presence of TFET led to assembly of the linear polypeptide chains of the differentially sulfated chimadanins. Finally, *in situ* desulfurization of the non-native cysteine and γ-(SH)-glutamate residues afforded the monosulfated sY28 (35% isolated yield) and sY31 (55% isolated yield) chimadanin proteins, as well as the doubly modified chimadanin sulfoform (31% isolated yield) after a single HPLC purification step.

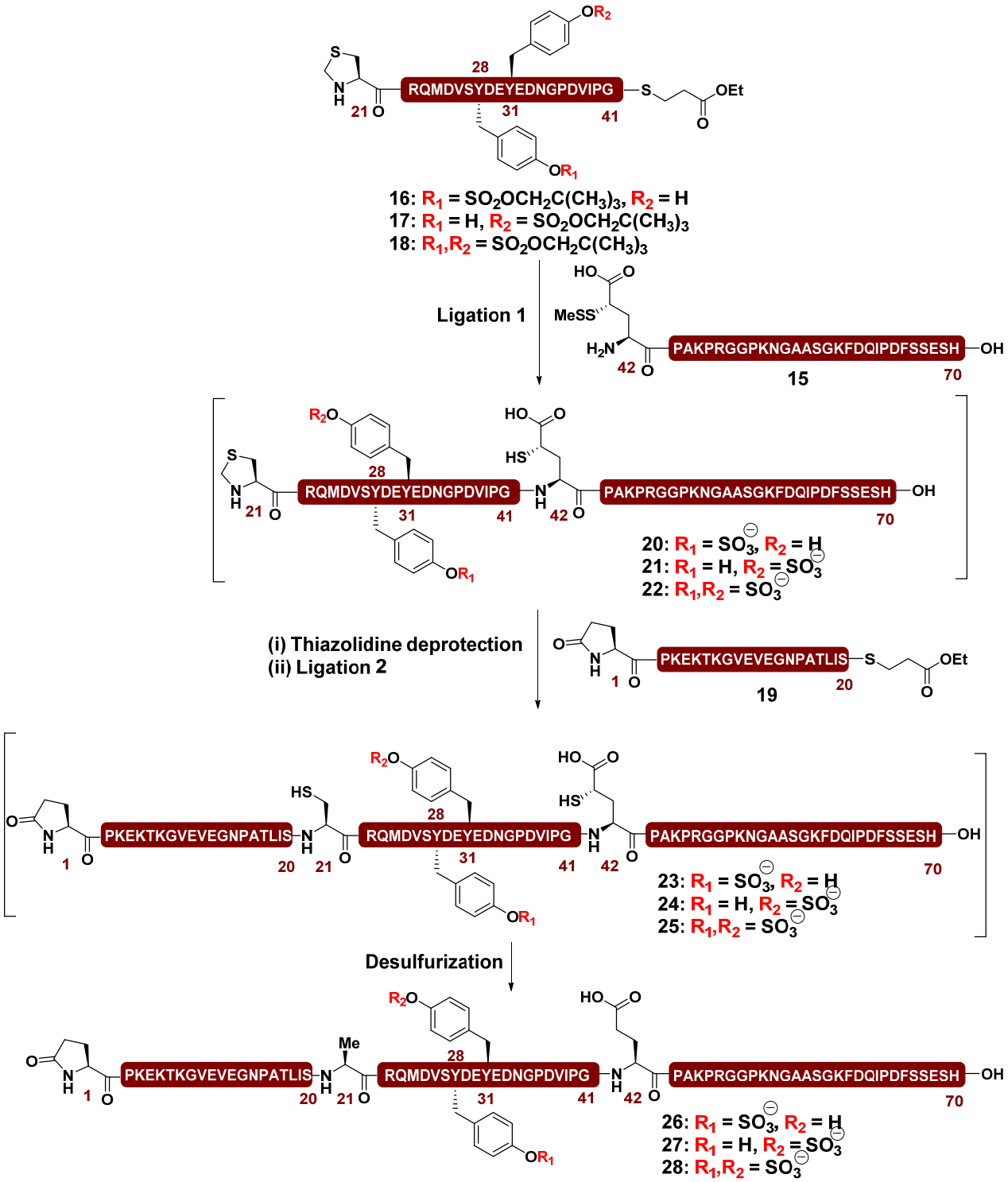

Figure 3. Assembly of differentially sulfated chimadanin proteins through one-pot sequential native chemical ligation reactions, followed by desulfurization. Conditions: **Ligation 1:** Peptide thioester (**16**, **17** or **18**, 1.0 equiv.), peptide **15** (1.1-1.2 equiv, 5 mM), 6 M Gnd.HCl, 100 mM Na_2HPO_4, 25 mM TCEP, 2 vol.% TFET, pH 6.8, 30 °C, 2 h (nP protecting groups in **16-18** were labile under the ligation conditions and afforded the free aryl mono sulfate esters); **Thiazolidine deprotection:** 0.2 M $MeONH_2$ pH 4.2, 30 °C, 2 h; **Ligation 2:** peptide thioester **19**, adjustment to pH 6.8, 2 vol.% TFET, 30 °C, 18 h; **Desulfurization:** degassed with argon, then addition of TCEP (200 mM), reduced glutathione (40 mM), VA-044 (20 mM), pH 6.2, 37 °C, 5 h. HPLC purification (eluent: 0.1 M NH_4OAc → MeCN) afforded sY28-monosulfated chimadanin (**26**, 35% yield), sY31-monosulfated chimadanin (**27**, 55% yield) and

disulfated chimadanin (**28**, 31% yield). NB: isolated yields are after 4 steps carried out in one-pot, square brackets around structures indicate that the compound was not isolated.

Position and extent of sulfation modulates anticoagulant activity

With the homogeneously sulfated madanin-1 (**12-14**) and chimadanin (**26-28**) proteins in hand, we next sought to address how the position and degree of sulfation within each sequence affected blood clotting and thrombin inhibitory activity. Towards this end, we first investigated the effect of sulfation of the proteins on the time required for thrombin-catalyzed clotting of blood plasma (thrombin time, TT, see Supplementary Methods for details of the assay). Unsulfated madanin-1 only prolonged TT weakly at concentrations up to 5 μM (28.7 ± 0.9 s at 5 μM vs. 20.3 ± 0.9 s in the absence of inhibitor; Fig. 4A), consistent with previous observations.[13] However, upon sulfation at Y32 (sulfoform **12**), 1 μM of madanin-1 prolonged TT to beyond 60 s. Interestingly, monosulfation at the other tyrosine sulfation site (Y35; sulfoform **13**) had a much weaker effect on the prolongation of TT by madanin-1, only slightly above the unsulfated sequence. In contrast, sulfation of both Y32 and Y35 (sulfoform **14**) transformed madanin-1 into a highly potent anticoagulant: an effect on TT similar to that of 5 μM unsulfated madanin-1 is observed at a 500-fold lower inhibitor concentration, and all higher concentrations tested prolonged TT to beyond 60 s (Fig. 4A). This suggests that there is cooperativity for these two sulfotyrosine sites for anticoagulant activity. Unsulfated chimadanin exhibited more potent anticoagulant activity than unsulfated madanin-1, extending the TT of blood plasma to >60 s at a concentration of 5 μM. Interestingly, monosulfation of chimadanin at Y28 (**26**) did not increase TT relative to the unsulfated inhibitor. In contrast, monosulfation of chimadanin at Y31 (**27**) led to a drastic increase in TT at concentrations as low as 100 nM. As with madanin-1, disulfation of chimadanin at Y28 and Y31 (**28**) had a profound effect on anticlotting activity, with a nearly 2-fold prolongation of TT at a concentration of 10 nM.

a

Madanin-1 Thrombin Times (s)

	Unsulfated	sY32 (**12**)	sY35 (**13**)	sY32, sY35 (**14**)
5 µM	28.7 ± 0.9	> 60	39.7 ± 0.4	> 60
1 µM		> 60	30.7 ± 0.4	> 60
0.1 µM	22.7 ± 0.4	31.0 ± 2.0	24.7 ± 0.4	> 60
0.01 µM	18.7 ± 0.4	22.7 ± 0.4	21.7 ± 2.4	28.7 ± 0.9

Chimadanin Thrombin Times (s)

	Unsulfated	sY28 (**26**)	sY31 (**27**)	sY28, sY31 (**28**)
5 µM	> 60	> 60	> 60	> 60
1 µM	27.7 ± 0.4	31.0 ± 0.6	> 60	> 60
0.1 µM	19.7 ± 0.4	23.0 ± 0.0	36.0 ± 1.0	> 60
0.01 µM	19.3 ± 0.4	22.0 ± 0.0	23.0 ± 0.0	36.0 ± 0.6

* In the absence of inhibitor the TT for plasma was 20.3 ± 0.9 s.

c

Inhibitor	K_i (nM)
Madanin-1	408.6 ± 26.8
Madanin-1 sY32 (**12**)	16.3 ± 0.8
Madanin-1 sY35 (**13**)	217.4 ± 21.5
Madanin-1 sY32, sY35 (**14**)	1.02 ± 0.038
Chimadanin	214.4 ± 21.5
Chimadanin sY28 (**26**)	72.8 ± 5.6
Chimadanin sY31 (**27**)	4.1 ± 0.2
Chimadanin sY28, sY31 (**28**)	0.41 ± 0.04

b

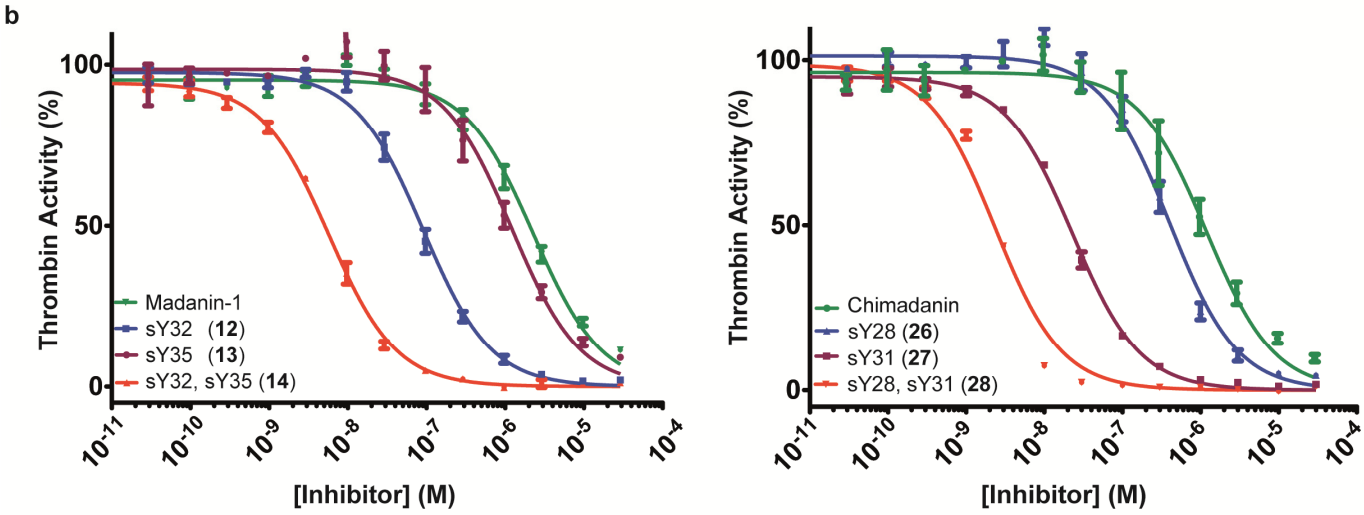

Figure 4. Tyrosine sulfation strongly modulates the anti-thrombin activity of madanin-1 and chimadanin. a, Thrombin times (TT), measuring the clotting time of human plasma in the presence of different concentrations of synthetic unsulfated madanin-1 and chimadanin and madanin-1 (**12-14**) and chimadanin (**26-28**) sulfoforms. TT values were not evaluated past 60 seconds. Errors depicted as standard deviation of the mean from three independent measurements. **b,** Dose-response curves for the inhibition of thrombin activity by increasing concentrations of synthetic unsulfated madanin-1 and chimadanin and madanin-1 (**12-14**) and chimadanin (**26-28**) sulfoforms in an amidolytic assay. Errors bars represent SEM values. **c,** Inhibition constants (K_i) for madanin-1 (**12-14**) and chimadanin (**26-28**) sulfoforms, determined by fitting the inhibited steady-state velocity data to the Morrison model. K_i values ± SEM values are provided.

Having observed that the inhibition of plasma coagulation is dramatically affected by sulfation of madanin-1 and chimadanin in a site-specific and cooperative manner, we next investigated the effect of tyrosine sulfation on the *in*

vitro inhibitory activity of madanin-1 and chimadanin. The amidolytic activity of thrombin against a chromogenic substrate was determined in the presence of a range of concentrations of each of the synthetic madanin-1 (**12-14**) and chimadanin (**26-28**) sulfoforms and inhibition constants (K_i) were determined (see Supplementary Methods for details of the assay, Fig. 4B for raw dose response curves and Fig. 4C for K_i values). Consistent with the TT data, sulfation of both madanin-1 and chimadanin increased the potency of thrombin inhibition in a site-specific manner. Unsulfated synthetic madanin-1 exhibited moderate potency against thrombin (K_i = 408.6 ± 26.8 nM) as was reported previously.[13] However, upon sulfation of madanin-1 at Y32 (in **12**), the potency of the inhibitor increased drastically (K_i = 16.3 ± 0.8 nM). Consistent with the TT studies, madanin-1 bearing a sulfotyrosine residue at position 35 (**13**) had only slightly improved inhibitory activity relative to the unsulfated protein (K_i = 217.4 ± 21.5 nM). Strikingly, doubly sulfated madanin-1 (**14**) exhibited extremely potent inhibition of thrombin (K_i = 1.02 ± 0.038 nM) almost three orders of magnitude more active than the unsulfated homologue. This result further highlights the cooperativity between the two sulfation sites to afford a high-affinity inhibitor of thrombin. Sulfation of chimadanin also drastically increased the inhibitory activity of the small protein. Unsulfated chimadanin exhibited moderate inhibition of thrombin activity, albeit two-fold more potent than madanin-1 (K_i = 214.4 ± 21.5 nM). Consistent with the TT results, monosulfation at Y28 (**26**) led to a modest improvement in thrombin inhibitory activity (K_i = 72.8 ± 5.6 nM), while monosulfation at Y31 (**27**) significantly improved the affinity for the enzyme (K_i = 4.1 ± 0.2 nM). As observed for the inhibition studies with madanin-1, doubly sulfated chimadanin (**28**) behaved as a potent thrombin inhibitor (K_i = 0.41 ± 0.04 nM), and again was almost three orders of magnitude more active than the unsulfated variant.

It was previously observed that recombinant unsulfated madanin-1 possesses two sites that are prone to proteolytic cleavage by thrombin (between K21-V22 and R54-L55).[13] We were therefore interested in assessing the rate at which both unsulfated madanin-1 and doubly sulfated madanin-1 **14** were processed by thrombin). Interestingly, while both unsulfated and sulfated madanin-1 were rapidly cleaved by thrombin between R54-L55 ($t_{1/2}$ ~ 15 min, Supplementary Fig. 24), cleavage between K21-V22 was slower for both inhibitors ($t_{1/2}$ ~ 4 h for unsulfated madanin-1, $t_{1/2}$ >> 24 h for doubly sulfated madanin-1, Supplementary Fig. 24). To ensure that neither processing events unduly affect thrombin inhibitory activity, we synthesized doubly sulfated madanin-1(1-54) and doubly sulfated madanin-1(22-54) (see Supplementary Methods for synthetic details), that represent the products of the proteolysis events. The inhibitory activity of these fragments (Ki = 5.96 ± 0.43 nM and 3.05 ± 0.22 nM respectively, Supplementary Fig. 17) were very similar to that of the full length inhibitor (Ki = 1.02 nM), indicating that the processing of madanin-1 by thrombin has little effect on the inhibitory activity of the sulfated polypeptide.

Madanin-1 exhibits a novel mechanism of thrombin inhibition

The cooperatively enhanced inhibition of thrombin by both doubly sulfated madanin-1 (**14**) and chimadanin (**28**) is striking, and points to strong recognition of these residues by one of the two positively charged exosites of thrombin.

The previously published three-dimensional structure of the complex between unsulfated madanin-1 and human thrombin revealed interpretable electron density only for a short segment of the inhibitor (^{M}A51-^{M}R54), occupying the non-primed specificity subsites within the active site cleft of thrombin (PDB entry 4BOH).[13] In order to understand the molecular details of thrombin recognition and inhibition by madanin-1, the complex between human α-thrombin and the synthetic disulfated madanin-1 (22-54) fragment was crystallized and its structure was determined at 1.6-Å resolution (Fig. 5 and Supplementary Table 3). Madanin-1 (22-54) fragment contains both the recognition sequence for thrombin cleavage[13] and the acidic stretch encompassing the two sites of sulfation. Further, this peptide retained potent inhibitory activity relative to the full-length inhibitor (*vide supra*), indicating that the mode of thrombin inhibition had been maintained (Supplementary Fig. 21). The final structural model comprises thrombin light-chain residues ^{T}A1B to ^{T}G14M and ^{T}I16 to ^{T}E247 from the heavy-chain (superscripted ^{T} and ^{M} denote the molecule – thrombin or madanin-1 – to which each residue belongs), with the only exception of residues ^{T}W148 to ^{T}V149C of the 149 insertion loop that are not defined in the electron density maps (residue numbering for thrombin is based on that of (chymo)trypsinogen). Comparison between the structures of free thrombin (PDB entry 3U69)[12] and the proteinase moiety of this complex revealed that only minor rearrangements occur upon madanin-1 binding (r.m.s.d. of 0.54Å for 279 aligned Cα atoms). The electron density maps were readily interpretable for madanin-1 residues ^{M}D31 to ^{M}R54 (Fig. 5A), but no significant density was detectable for the remaining N-terminal segment ^{M}V22-^{M}A30. Remarkably, madanin-1 runs along the surface of thrombin in an extended conformation, with its C-terminal region blocking the active site of the enzyme, while the sulfotyrosine-containing region of the inhibitor establishes specific interactions with the exosite II of thrombin (Fig. 5A). Similar to the full-length protein in solution, the thrombin-bound madanin-1 fragment doesn't display any secondary structure elements.[13]

The ^{M}A51-^{M}R54 segment of madanin-1 binds to the active site cleft of thrombin in a substrate-like manner as observed previously,[13] establishing extensive interactions with the proteinase (Fig. 5B). The P1 ^{M}R54 is embedded into the S1 pocket of thrombin, with its main-chain nitrogen and carbonyl oxygen establishing polar contacts with the carbonyl oxygen of ^{T}S214 and the main-chain nitrogen of ^{T}G193, respectively. In addition to several water-mediated contacts (e.g. between the carbonyl oxygen of ^{M}R54 and the side-chain of the catalytic ^{T}S195), the guanidinium group of ^{M}R54 forms a salt bridge with ^{T}D189, at the bottom of the S1 specificity pocket (Fig. 5B). The upstream residues ^{M}P53 and ^{M}K52 occupy the S2 and S3 subsites of thrombin, respectively, and ^{M}A51 contacts the carbonyl oxygen of ^{T}W96 and the side-chain hydroxyl of ^{T}Y60A through ordered solvent molecules. Further towards the N-terminus of madanin-1, the association between inhibitor and proteinase is ensured by a meshwork of exclusively water-mediated contacts involving residues ^{M}P49, ^{M}D45, ^{M}P42 and ^{M}T41 that extends to the aryl binding site of thrombin. A number of intramolecular interactions within madanin-1 (solvent-mediated between the carbonyl oxygen and side-chain hydroxyl of ^{M}T41 and the side-chain hydroxyl of ^{M}T43 and the carbonyl oxygen of ^{M}P42, respectively, as well as between the carbonyl oxygen and the side-chain carboxylate of ^{M}D45 and the main-chain nitrogen of ^{M}A48 and the side-chain

hydroxyl of ^{M}T47, respectively) account for the pronounced bend of the inhibitor that directs the N-terminal segment to the exosite II of thrombin.

The acidic ^{M}D31-^{M}G39 segment of madanin-1 binds to the positively charged exosite II of thrombin (Fig. 5C). The sulfate group of ^{M}Y35 directly interacts with the ε-amino group of ^{T}K240, while hydrophobic interactions between the aromatic ring of ^{M}Y35 and the side chains of ^{T}H91, ^{T}P92 and ^{T}R93 further anchor the sulfated residue in this pocket. Additionally, the sulfate group of ^{M}Y32 interacts directly with the ε-amino group of ^{T}K236 and with the main chain nitrogen of ^{T}R126 *via* a water molecule. The phenolic group of ^{M}Y32 slots between the side chains of ^{T}R126 and ^{T}K236 and is also stabilized by Van der Waals interactions with the side chain of ^{T}F232. In concert with these two strong electrostatic anchors, the side-chain carboxylate of ^{M}D31 forms a salt bridge with ^{T}R126, and the side chain of ^{M}D33 with ^{T}R101. Additionally, the main chain nitrogen and side-chain carboxylate of ^{M}D33 are hydrogen bonded to the carbonyl oxygen of ^{T}R233 and the carboxamide nitrogen of ^{T}N179, respectively. Furthermore, additional water-mediated interactions are established between the side-chains of ^{M}D33 and ^{T}H91, and between the carbonyl oxygens of ^{M}D33 and ^{T}R233. This extensive network of polar interactions, involving both sulfotyrosine residues and their acidic neighbors (^{M}D31 and ^{M}D33), accounts for the markedly increased affinity of sulfated madanin-1 towards thrombin.

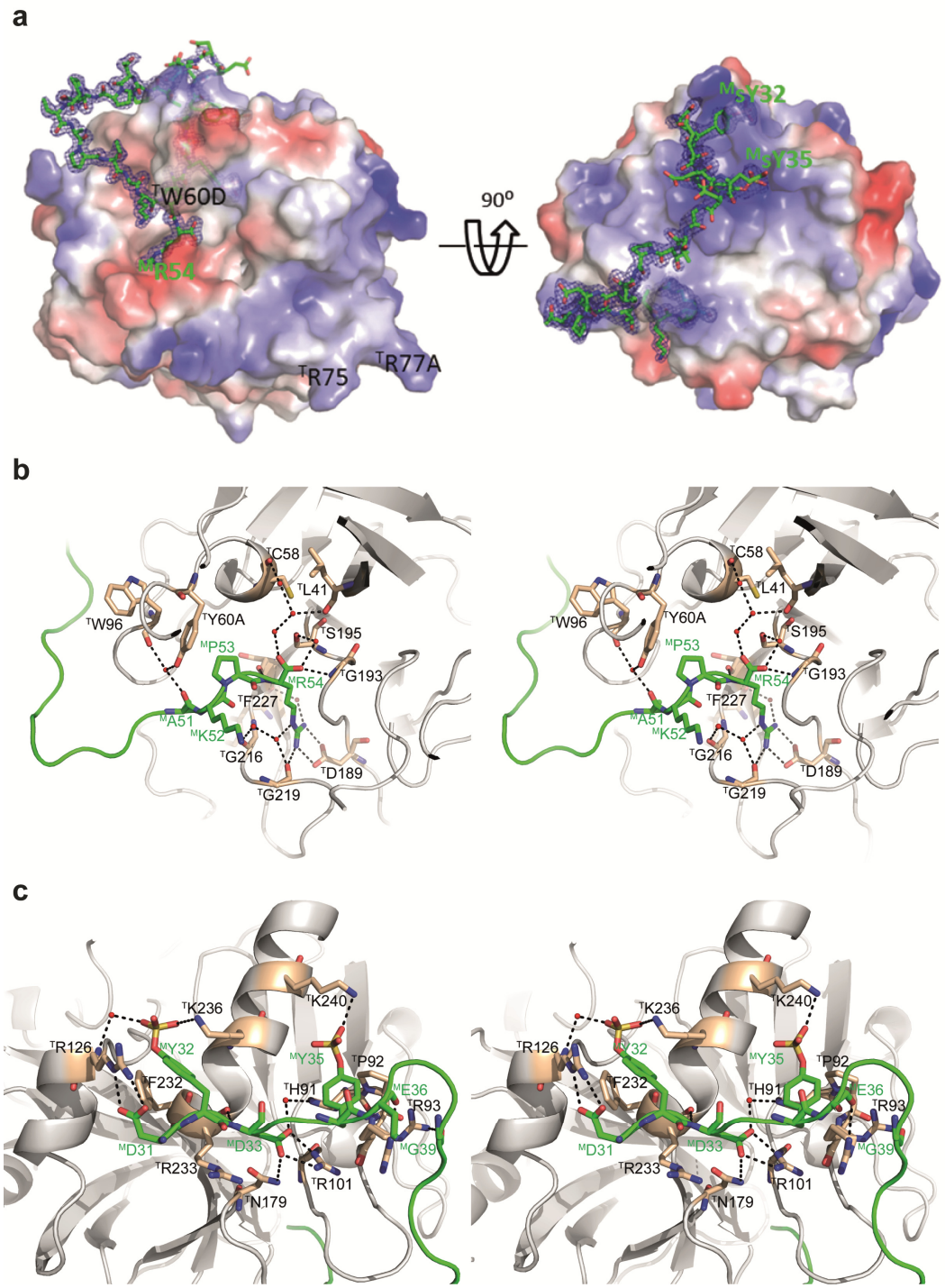

Figure 5. Madanin-1 blocks both the active site and the exosite II regions of human α-thrombin. a,The acidic ^{M}D31-^{M}G39 segment of madanin-1 (stick model with nitrogen atoms in blue, oxygen in red, and carbon in green) binds to the exosite II of thrombin (solid surface representation with positive surface electrostatic potential in blue and negative surface electrostatic potential in red), whereas the downstream ^{M}A51-^{M}R54 segment occupies the non-primed sites on the active site cleft of the proteinase. The 2Fo-Fc electron density map (1.0-σ contour) for madanin-1 is displayed as a blue mesh. The thrombin molecule is shown in the standard orientation for serine proteinases (left panel) and rotated 90º around x (right panel). **b,** Stereo close-up of the extensive network of interactions established between the C-terminal segment of madanin-1 (colored as in A) and the active site of the proteinase (gray cartoon with selected residues as sticks color-coded as madanin-1 except for carbon atoms (colored tan)). In particular, all interactions between residue ^{M}R54 and thrombin are represented. **c,** Stereo close-up of the interactions between the

N-terminal segment of madanin-1 (colored as in A) and the exosite II region of thrombin (colored as in B). Water molecules are represented as red spheres in B and C.

In order to further validate the specific exosite II-binding of sulfated madanin-1, we tested the potency of disulfated madanin-1 as an inhibitor of human γ-thrombin. This isoform of thrombin results from the autolytic or tryptic processing of the heavy chain of α-thrombin at two positions (between residues R75-Y76 and K149E-G150), disrupting the exosite I and dramatically reducing the affinity towards fibrinogen, while leaving the ability to process small chromogenic substrates virtually unaffected. [29,30] Additionally, the interaction between γ-thrombin and the exosite I-targeting inhibitor hirudin is severely affected, with a measured affinity six orders of magnitude lower than that observed for α-thrombin.[30] In agreement with the crystallographic model reported here, γ-thrombin was potently inhibited by disulfated madanin-1 **14** (K_i = 0.64 ± 0.05 nM) and chimadanin **28** (K_i = 12.49 ± 1.22 nM), thus supporting a binding mode independent of exosite I. Taken together, these data demonstrate that sulfation of madanin-1 and chimadanin significantly increases affinity for thrombin, specifically blocking the basic exosite II and the active site of the enzyme in a unique bivalent binding mechanism.

DISCUSSION

Hematophagous arthropods produce and secrete a rich source of bioactive proteins, however, biochemical and functional studies of the native proteins (as produced by a given organism) have been limited by not considering the potential presence and role of PTMs. In this study we have demonstrated that tyrosine sulfation of the tick-derived proteins madanin-1 and chimadanin dramatically enhances the thrombin inhibitory and anticoagulant activity of the proteins.

Recombinant production of madanin-1 and chimadanin in baculovirus-infected insect cells afforded heterogeneously sulfated proteins and pointed strongly towards sulfation of these thrombin inhibitors within the salivary glands of the *H. longicornis* tick. The three possible sulfated variants of both madanin-1 and chimadanin were efficiently synthesized in homogeneous form using one-pot ligation-desulfurization chemistry. Access to these synthetic proteins enabled the unequivocal demonstration that sulfation of two conserved tyrosine residues of madanin-1 and chimadanin leads to significant enhancement in antithrombotic and anticoagulant activity. Specifically, homogeneous doubly sulfated variants of madanin-1 (at Y32 and Y35) and chimadanin (at Y28 and Y31) were nearly three orders of magnitude more potent as inhibitors of thrombin relative to their unsulfated counterparts, with inhibition constants reaching the subnanomolar range. These inhibitory activities are much more consistent with the essential role of these anticoagulants for tick feeding and survival than those previously reported for unsulfated recombinant madanin-1.[13]

The structural features leading to the increased potency of disulfated madanin-1 were unveiled by X-ray crystallography. Unexpectedly, the disulfated sequence of madanin-1 is a strong binder of exosite II of thrombin, directing the downstream segment of the inhibitor towards the active site cleft of the proteinase, where it occupies the non-primed specificity subsites. Due to the significant sequence homology and similar placement of tyrosine sulfation sites in chimadanin, together with a comparable activity towards exosite I-disrupted γ-thrombin, we would predict that it shares the bidentate binding mode of madanin-1, specifically targeting the exosite II of thrombin, a unique property among these cysteine-less thrombin anticoagulants. Interestingly, the binding mode of the acidic stretch encompassing the sulfotyrosine residues is strikingly different from that of the C-terminal tail of leech-derived haemadin (PDB entry 1E0F),[31] but highly reminiscent of that of the physiologic substrate, platelet-derived glycoprotein Ibα (PDB entry 1P8V).[32] This substrate mimicry by madanin-1 (and chimadanin) parallels the fibrinogen-mimicking interactions displayed by the exosite I-targeting inhibitors such as hirudin[5] and variegin.[9] The exosite II recognition motif identified in this study is highly conserved in other related tick-derived proteins, including madanin-2 and the madanin-like proteins from *H. longicornis*, the analogously named madanins 1-4 from the hairy bont-legged tick (*Hyalomma marginatum rufipes*), haemathrin-1 and 2 from the *Haemaphysalis bispinosa* tick, and a chimadanin-like protein from the brown dog tick (*Rhipicephalus sanguineus*) (Fig. 6A). A consensus motif can be constructed from these sequences that possess the acidic, sulfotyrosine-containing exosite II-binding region, spatially separated from a

conserved active-site recognition sequence by a variable linker (Fig. 6B). Given the high sequence conservation of these proteins, it is highly likely that they share the unique bivalent binding mode described for madanin-1, and their inhibitory activity towards thrombin is dependent on the recognition of sulfotyrosine residues by the exosite II region of the proteinase (Fig. 6C). The key findings from this work prompt an extensive and systematic characterization of the role of PTMs on the activity and specificity (Supplementary Fig. 20) of other small cysteine-free anticoagulants through a combination of sulfoprotein total synthesis and biochemical interrogation. We anticipate that extensive knowledge of the structural mechanisms underpinning thrombin inhibition by these evolutionarily tuned molecules will create the blueprints for the next-generation of thrombin inhibitors.

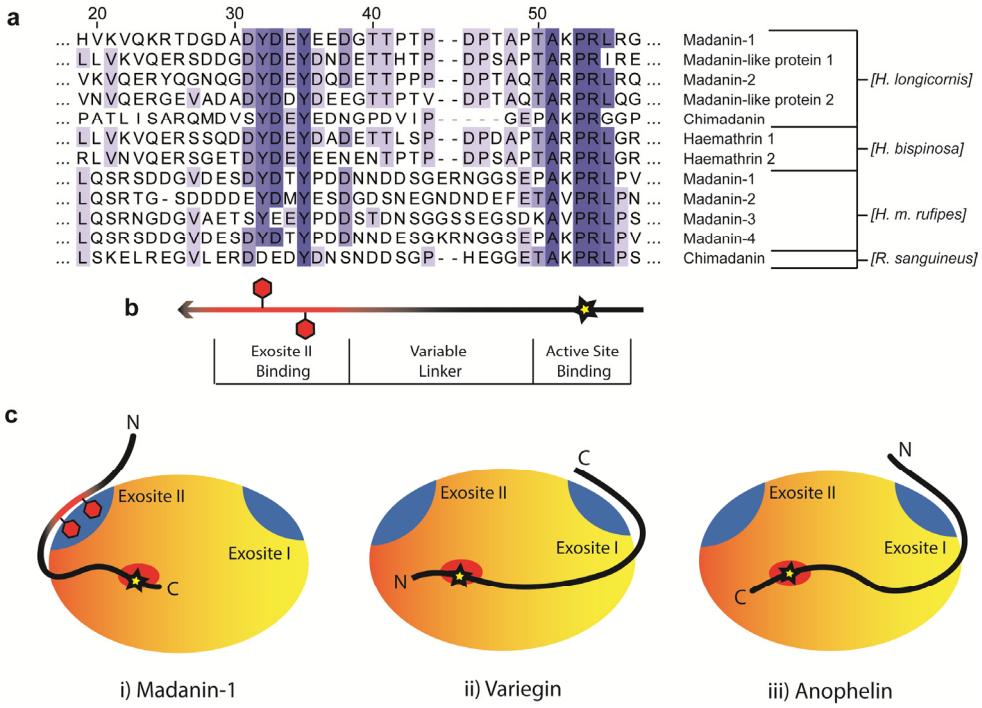

Figure 6. Madanin-1 and chimadanin belong to a new mechanistic family of exosite II binding inhibitors. a, Multiple amino acid sequence alignment of proposed exosite II-targeting proteins from other ixodid tick species: *Haemaphysalis longicornis* (UniProt entries: Q86FP9, Q4R1A2, Q86FP8, Q4R1A5, Q4R1A0), *Haemaphysalis bispinosa* (UniProt entries: A0A089VKV7, A0A089X5A5), *Hyalomma marginatum rufipes* (UniProt entries: E2J6S1, E2J6T8, E2J6R9, E2J6R0), and *Rhipicephalus sanguineus* (UniProt entry C9W1D6). Increased sequence conservation is depicted by deeper hues of blue background. Numbering is based on the mature *H. longicornis* madanin-1 sequence. **b,** Schematic representation of the consensus structure of exosite II-targeting inhibitors. Conserved sulfotyrosine residues are shown as red hexagons, and the active site binding region as a yellow star. **c,** Schematic representation of bivalent thrombin binding by three cysteine-less anticoagulants: i) madanin-1 and chimadanin-like exosite II-targeting inhibitors, ii) variegin, which binds both exosite I and the active site in a substrate-like orientation, and iii) the anophelins, which display a reverse binding mode to both exosite I and the active site.

Data availability statement

Data supporting the findings of this study are available within the article and its supplementary information file and from the corresponding author upon reasonable request.

References

1 Corral-Rodríguez, M. A., Macedo-Ribeiro, S., Pereira, P. J. B. & Fuentes-Prior, P. Leech-derived thrombin inhibitors: from structures to mechanisms to clinical applications. *J. Med. Chem.* **53**, 3847-3861 (2010).

2 Koh, C. Y. & Kini, R. M. Molecular diversity of anticoagulants from haematophagous animals. *Thromb. Haemost.* **102**, 437-453 (2009).

3 Corral-Rodriguez, M. A., Macedo-Ribeiro, S., Barbosa Pereira, P. J. & Fuentes-Prior, P. Tick-derived Kunitz-type inhibitors as antihemostatic factors. *Insect Biochem. Mol. Biol.* **39**, 579-595 (2009).

4 Kazimirova, M. & Stibraniova, I. Tick salivary compounds: their role in modulation of host defences and pathogen transmission. *Front. Cell. Infect. Microbiol.* **3**, 43 (2013).

5 Rydel, T. J. *et al.* The structure of a complex of recombinant hirudin and human α-thrombin. *Science* **249**, 277-280 (1990).

6 van de Locht, A. *et al.* Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. *EMBO J* **14**, 5149-5157 (1995).

7 van de Locht, A. *et al.* The ornithodorin-thrombin crystal structure, a key to the TAP enigma? *EMBO J.* **15**, 6011-6017 (1996).

8 Macedo-Ribeiro, S. *et al.* Isolation, cloning and structural characterisation of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. *PLoS ONE* **3**, e1624 (2008).

9 Koh, C. Y. *et al.* Crystal structure of thrombin in complex with S-variegin: insights of a novel mechanism of inhibition and design of tunable thrombin inhibitors. *PLoS ONE* **6**, e26367 (2011).

10 Figueiredo, A. C. *et al.* Unique thrombin inhibition mechanism by anophelin, an anticoagulant from the malaria vector. *Proc Natl Acad Sci U S A* **109**, E3649–E3658 (2012).

11 Hsieh, Y. S., Wijeyewickrema, L. C., Wilkinson, B. L., Pike, R. N. & Payne, R. J. Total Synthesis of Homogeneous Variants of Hirudin P6: A Post-translationally modified anti-thrombotic leech-derived protein. *Angew. Chem. Int. Ed.* **53**, 3947-3951 (2014).

12 Figueiredo, A. C. *et al.* Rational design and characterization of D-Phe-Pro-D-Arg-derived direct thrombin inhibitors. *PLoS ONE* **7**, e34354 (2012).

13 Figueiredo, A. C., de Sanctis, D. & Pereira, P. J. B. The tick-derived anticoagulant madanin is processed by thrombin and factor Xa. *PloS ONE* **8**, e71866 (2013).

14 Valenzuela, J. G., Francischetti, I. M. & Ribeiro, J. M. Purification, cloning, and synthesis of a novel salivary anti-thrombin from the mosquito *Anopheles albimanus*. *Biochemistry* **38**, 11209-11215 (1999).

15 Koh, C. Y. *et al.* Variegin, a novel fast and tight binding thrombin inhibitor from the tropical bont tick. *J. Biol. Chem.* **282**, 29101-29113 (2007).

16 Cappello, M. *et al.* Isolation and characterization of the tsetse thrombin inhibitor: a potent antithrombotic peptide from the saliva of *Glossina morsitans morsitans*. *Am. J. Trop. Med. Hyg.* **54**, 475-480 (1996).

17 Iwanaga, S. *et al.* Identification and characterization of novel salivary thrombin inhibitors from the ixodidae tick, *Haemaphysalis longicornis*. *Eur. J. Biochem.* **270**, 1926-1934 (2003).

18 Nakajima, C. *et al.* A novel gene encoding a thrombin inhibitory protein in a cDNA library from *Haemaphysalis longicornis* salivary gland. *J. Vet. Med. Sci.* **68**, 447-452 (2006).

19 Zhang, D., Cupp, M. & Cupp, E. Thrombostasin: purification, molecular cloning and expression of a novel anti-thrombin protein from horn fly saliva. *Insect Biochem. Mol. Biol.* **32**, 321-330 (2002).

20 Cappello, M. *et al.* Tsetse thrombin inhibitor: bloodmeal-induced expression of an anticoagulant in salivary glands and gut tissue of Glossina morsitans morsitans. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14290-14295 (1998).

21 Stone, M. J. & Payne, R. J. Homogeneous sulfopeptides and sulfoproteins: synthetic approaches and applications to characterize the effects of tyrosine sulfation on biochemical function. *Acc. Chem. Res.* **48**, 2251-2261 (2015).

22 Ouyang, Y. B. & Moore, K. L. Molecular cloning and expression of human and mouse tyrosylprotein sulfotransferase-2 and a tyrosylprotein sulfotransferase homologue in *Caenorhabditis elegans*. *J. Biol. Chem.* **273**, 24770-24774 (1998).

23 Ouyang, Y., Lane, W. S. & Moore, K. L. Tyrosylprotein sulfotransferase: purification and molecular cloning of an enzyme that catalyzes tyrosine O-sulfation, a common posttranslational modification of eukaryotic proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 2896-2901 (1998).

24 Thompson, R. E. *et al.* Trifluoroethanethiol: an additive for efficient one-pot peptide ligation− desulfurization chemistry. *J. Am. Chem. Soc.* **136**, 8161-8164 (2014).

25 Thompson, R. E., Chan, B., Radom, L., Jolliffe, K. A. & Payne, R. J. Chemoselective peptide ligation–desulfurization at aspartate. *Angew. Chem. Int. Ed.* **52**, 9723-9727 (2013).

26 Simpson, L. S., Zhu, J. Z., Widlanski, T. S. & Stone, M. J. Regulation of chemokine recognition by site-specific tyrosine sulfation of receptor peptides. *Chem. Biol.* **16**, 153-161 (2009).

27 Wan, Q. & Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **46**, 9248-9252 (2007).
28 Cergol, K. M., Thompson, R. E., Malins, L. R., Turner, P. & Payne, R. J. One-pot peptide ligation–desulfurization at glutamate. *Org. Lett.* **16**, 290-293 (2013).
29 Chang, T., Feinman, R. D., Landis, B. H. & Fenton, J. W., 2nd. Antithrombin reactions with alpha- and gamma-thrombins. *Biochemistry* **18**, 113-119 (1979).
30 Ascenzi, P. *et al.* Binding of hirudin to human alpha, beta and gamma-thrombin. A comparative kinetic and thermodynamic study. *J. Mol. Biol.* **225**, 177-184 (1992).
31 Richardson, J. L. *et al.* Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. *EMBO J* **19**, 5650-5660 (2000).
32 Dumas, J. J., Kumar, R., Seehra, J., Somers, W. S. & Mosyak, L. Crystal structure of the GpIbalpha-thrombin complex essential for platelet aggregation. *Science* **301**, 222-226 (2003).

Acknowledgements The research was supported by an Australian Research Council Future Fellowship (to R. J. P.) and by Australian Postgraduate Awards (to R. E. T. and X. L.) and a John Lamberton Research Scholarship (to R. E. T.) This work was funded in part by the European Social Fund through Programa Operacional Capital Humano (POCH) and by national funds through Fundação para a Ciência e a Tecnologia (Portugal) in the form of a postdoctoral fellowship SFRH/BPD/108004/2015 (to J. R.-R.). We acknowledge the ESRF for provision of synchrotron radiation facilities and thank the ESRF staff for help with data collection. Transnational Access to the High Throughput Crystallization Laboratory of the European Molecular Biology Laboratory Grenoble Outstation was supported by the European Community-Seventh Framework Program (FP7/2007-2013) Grant Protein Production Platform (PCUBE Agreement No. 227764).

Author Contributions R. E. T., X. L. and J. R-R. contributed equally to this work. R. E. T. and X. L. synthesized the thiol aspartate and glutamate amino acids, developed and employed the ligation methods for the assembly of differentially sulfated madanin-1 and chimadanin peptides and proteins and performed all the characterization experiments; J. R-R. carried out all baculovirus expression experiments, some of the thrombin inhibition experiments with the synthetic madanin-1 and chimadanin sulfoproteins and determined the structure of the thrombin-madanin-1 complex; N. A-G. performed some of the thrombin inhibition experiments with the synthetic madanin-1 and chimadanin sulfoproteins; B. L. P. performed the liquid chromatography-tandem mass spectrometry experiments on madanin-1 and chimadanin expressed in baculovirus; P. J. B. P. and R. J. P designed and directed the investigations and composed the manuscript together with the other authors.

Competing Financial Interests

The authors declare no competing financial interests.