Protein Synthesis

DOI: 10.1002/anie.201310777

Total Synthesis of Homogeneous Variants of Hirudin P6: A Post-Translationally Modified Anti-Thrombotic Leech-Derived Protein**

*Yves S. Y. Hsieh, Lakshmi C. Wijeyewickrema, Brendan L. Wilkinson, Robert N. Pike, and Richard J. Payne**

Dedicated to Professor Chi-Huey Wong on the occasion of his 65th birthday

Abstract: *Hirudin P6 is a leech-derived anti-thrombotic protein which possesses two post-translational modifications, O-glycosylation and tyrosine sulfation. In this study we report the ligation-based synthesis of a library of hirudin P6 proteins possessing homogeneous glycosylation and sulfation modifications. The nature of the modifications incorporated was shown to have a drastic effect on inhibition against both the fibrinogenolytic and amidolytic activities of thrombin and thus highlights a potential means for attenuating the biological activity of the protein.*

Post-translational modification (PTM) of proteins is a common mechanism to expand the structural diversity possible from the genome of an organism.[1] Two common PTMs include *O*-glycosylation of serine (Ser) and threonine (Thr) residues[2] and sulfation of tyrosine (Tyr) residues,[3] both of which are performed in the Golgi apparatus through the action of glycosyltransferases and tyrosylprotein sulfotransferases (TPSTs), respectively. Indeed, it is estimated that more than 50% of eukaryotic proteins are glycosylated and over 1% are Tyr sulfated. These two PTMs have been implicated in mediating a vast array of biological processes, including molecular recognition, cell differentiation, immune-regulation, and protein folding.[1,4] This has led to significant interest in glycosylated and sulfated proteins as therapeutics for numerous diseases.[4d] Despite the importance of these modifications, access to pure glycoproteins and sulfoproteins is extremely challenging. This is owing to the untemplated nature of the PTM process which is dictated by the relative activity of transferase enzymes, thus leading to heterogeneous mixtures of glycoforms[2,5] and sulfoforms[3b] that are often inseparable by chromatographic techniques. In addition, the acid labile nature of the phenolic sulfate ester leads to significant difficulties in purifying native sulfoproteins from biological sources with the modification intact.[3,6] Recently, chemical synthesis has emerged as a viable avenue for accessing homogeneous glycoproteins in order to interrogate the effect of glycan modifications on the activity of proteins.[2,5,7] Whilst a full-length sulfoprotein has not been prepared to date, it can be envisaged that these might also be accessed by synthetic means. Herein, we describe the efficient chemical synthesis of a library of homogeneously modified hirudin P6 (HP6) proteins, a 63 amino acid leech-derived protein that naturally possesses both *O*-glycosylation at Thr-43 and Tyr sulfation at Tyr-61 (Figure 1).[8] HP6 is produced within the parapharyngeal gland of *Hirudinaria manillensis*, with the isolated form exhibiting potent inhibitory activity against human thrombin, thus facilitating blood flow for the hematophagous activity of the organism.[8] HP6 consists of an N-terminal domain held together by three disulfide bonds and an acidic C-terminal domain, reminiscent of hirudin (HIRV1)[9] from the medicinal leech *H. medicinalis*. However, these two proteins share sequence homology of only 63%. HIRV1 has exhibited potent inhibition against human thrombin[9b,10] with several unmodified variants entering the clinic as anti-coagulant agents.[11] Whilst sulfation of HIRV1 has been reported to provide a significant increase in inhibitory activity against the fibrinogenolytic and amidolytic activities of thrombin, the effect of glycosylation or the combination of glycosylation and sulfation has not been interrogated on HIRV1 or HP6. Glycosylation of HP6 is

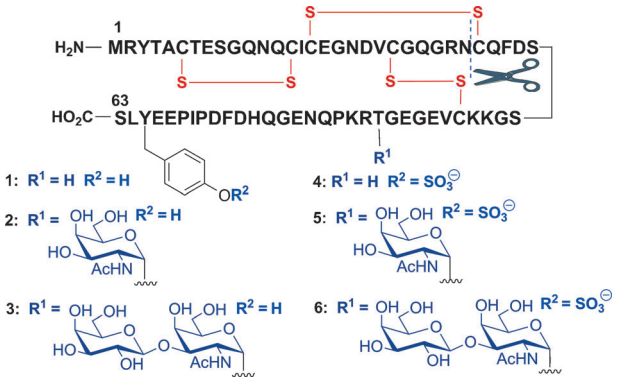


Figure 1. Structure of HP6 from *Hirudinaria manillensis* and target post-translationally modified HP6 proteins **1**–**6**.

[*] Y. S. Y. Hsieh, Dr. B. L. Wilkinson, Prof. R. J. Payne
School of Chemistry, The University of Sydney
Sydney, NSW 2006 (Australia)
E-mail: richard.payne@sydney.edu.au
Homepage: http://sydney.edu.au/science/chemistry/~payne/index
Dr. L. C. Wijeyewickrema, Prof. R. N. Pike
Department of Biochemistry and Molecular Biology, Monash University, VIC 3800 (Australia)

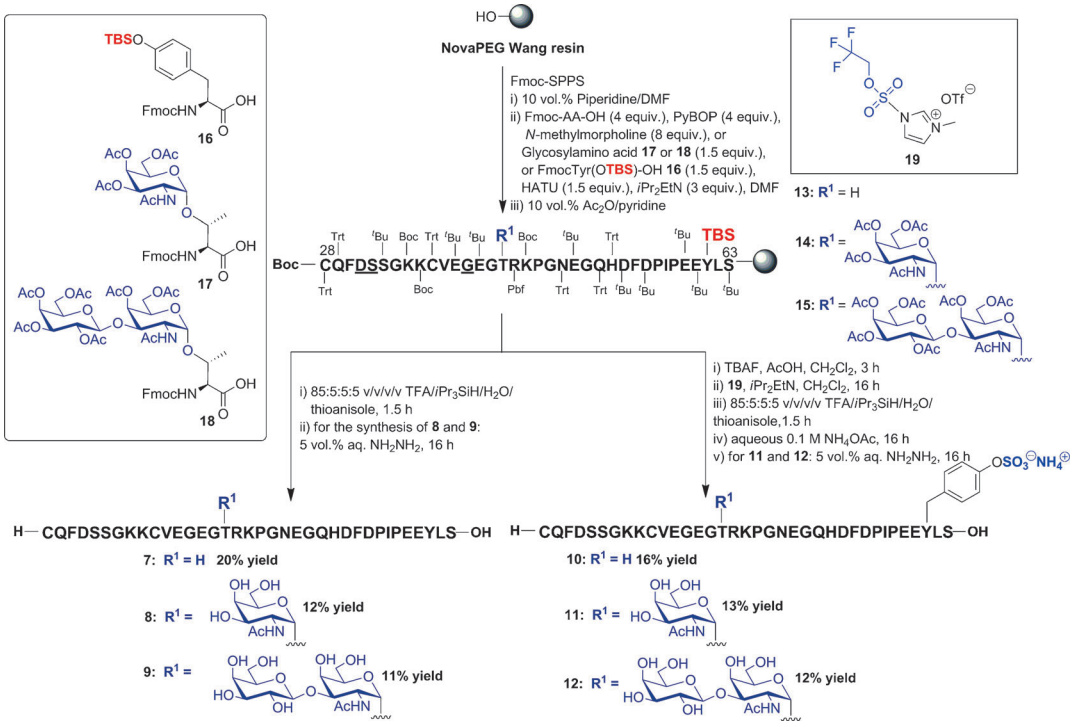
[**] The authors acknowledge the Australian Research Council Future Fellowship Scheme for research support (FT130100150). We thank Dr. Ben Crossett, Dr. Nick Proschogo and Leo Corcilius for technical assistance with high resolution mass spectrometry.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201310777.

reported to be of the mucin type, whereby α-linked *N*-acetylgalactosamine (α-GalNAc) or galactose(Gal)-β-1,3-GalNAc are appended at Thr-43.[8] The overarching goal of this study was therefore to employ chemical synthesis to access homogeneous variants of HP6 to probe the effect of Tyr sulfation and *O*-glycosylation on the anti-thrombotic activity of this protein.

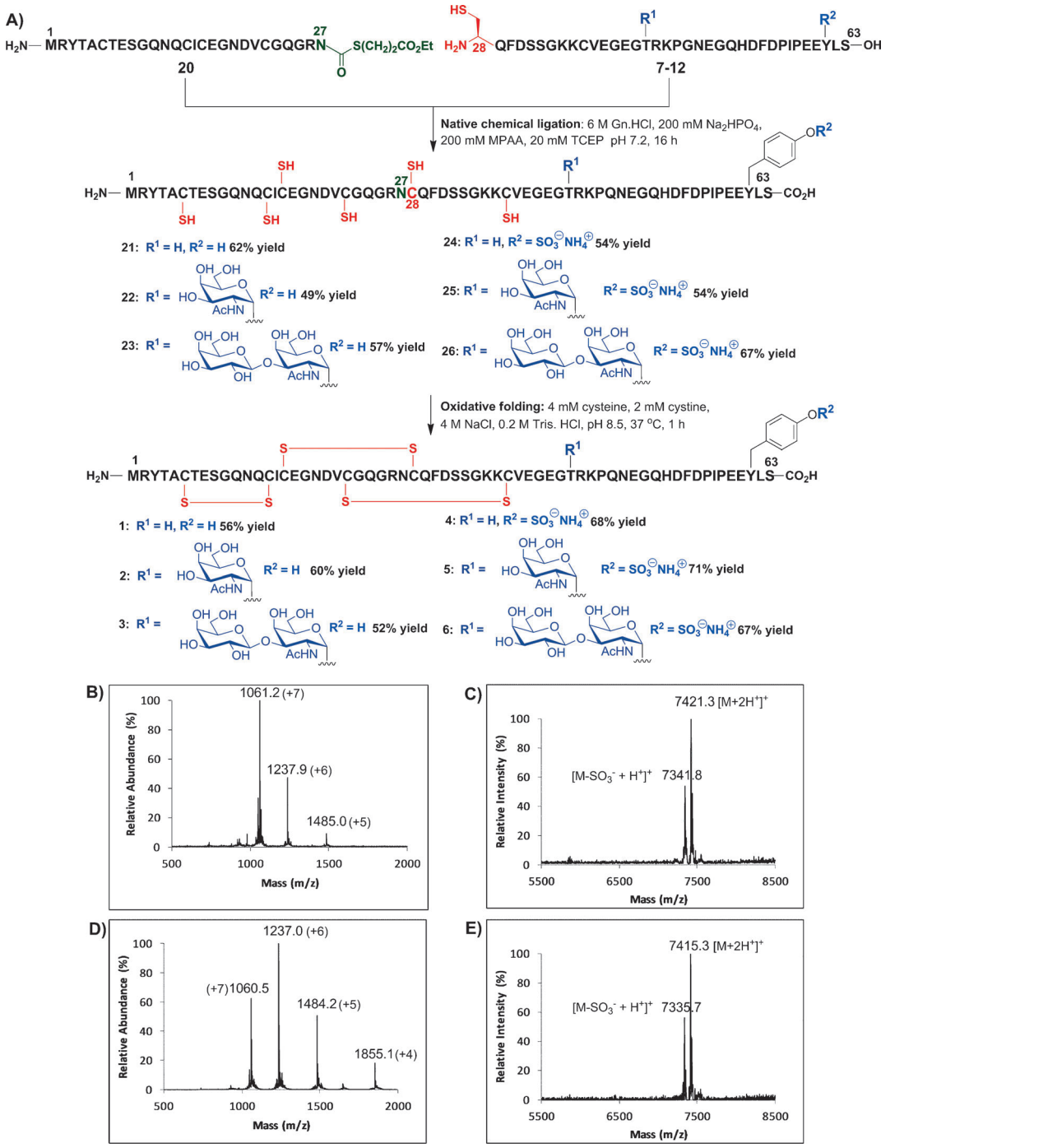
Six homogeneous variants of HP6 were targeted in this study, namely unmodified HP6 (**1**), GalNAc-derived HP6 (**2**), Gal-β-1,3-GalNAc-derived HP6 (**3**), as well as the Tyr-sulfated variants of **1**–**3** (**4**–**6**). HP6 possesses a total of six Cys residues that form three disulfide bridges in the native folded form of the protein, therefore offering several possible disconnection sites where native chemical ligation[12] could be employed. At the outset, we identified the peptide bond between Asn-27 and Cys-28 as a junction for the assembly of **1**–**6** (Figure 1). This strategy thus required the synthesis of six differentially glycosylated and sulfated peptide fragments (**7**–**12**) corresponding to HP6(28–63) and a single peptide thioester fragment representing HP6(1–27). The synthesis of **7**–**12** was achieved by using a divergent solid-phase sulfation strategy (see Supporting Information for details). Synthesis of resin-bound **13**–**15** commenced from NovaPEG Wang resin on a 150 μmol scale whereby the peptide chain was elongated to Arg-44, just prior to the first glycosylation site by Fmoc-SPPS, with the incorporation of Fmoc-Tyr(OTBS)-OH **16**[6a] at Tyr-61 (Scheme 1). At this stage, the resin was split into three 50 μmol batches to which unmodified Thr or glycosylamino acids **17** or **18**[13] were coupled. The chain was then elongated to provide the desired resin-bound peptides **13**–**15**. For this purpose a 2,4-dimethoxybenzyl(Dmb)-protected Gly residue was incorprated at Gly-40 and Fmoc-Asp(*t*Bu)-Ser($\Psi^{Me,Me}$pro)-OH at Asp-31–Ser-32 to improve the yield of the peptide syntheses by preventing en bloc peptide aggregation. Following complete elongation to **13**–**15**, each resin-bound (glyco)peptide was split into two equal 25 μmol batches to enable divergent solid-phase manipulations. The first batches of **13**–**15** were subjected to acidolytic cleavage of the peptide from the resin which occurred with concomitant removal of the side chain protecting groups, including the silyl ether from Tyr-61, and the N-terminal Boc-carbamate moiety. Following purification, the unglycosylated peptide fragment **7** was produced in 20% yield. The acetyl groups of glycopeptides derived from **14** and **15** were subsequently removed by hydrazinolysis. Purification then provided **8** and **9** in 12% and 11% yields, respectively, based on the original resin loading.

For the synthesis of sulfopeptides **10**–**12**, the TBS-ether was first chemoselectively removed from resin-bound **13**–**15** by treatment with buffered tetrabutylammonium fluoride (TBAF). Next, solid-phase sulfation was effected by treatment of the resin-bound peptides with trifluoroethylsulfuryl imidazolium sulfating agent **19**[14] and *N*,*N*-diisopropylethylamine. Trifluoroethyl (TFE) protection of the Tyr sulfate was chosen as it stabilizes the phenolic sulfate ester to the ensuing acidic cleavage from the resin, but can be subsequently removed by mild nucleophilic displacement. Indeed, cleavage with a trifluoroacetic acid (TFA)-based cocktail afforded the crude, TFE sulfate ester-protected (glyco)sulfopeptides which were subsequently subjected to treatment with 0.1 M ammonium acetate for 16 h to facilitate smooth deprotection of the



Scheme 1. Divergent solid-phase synthesis of differentially *O*-glycosylated and Tyr-sulfated HP6(28–63) fragments. TBS = *tert*-butyldimethylsilyl, Pbf = 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl, DS represents an Asp(*t*Bu)-Ser($\Psi^{Me,Me}$pro) residue, G represents an *N*-(2,4-dimethoxybenzyl) Gly residue.

TFE sulfate ester (Scheme 1). At this stage, unglycosylated sulfopeptide **10** was purified by reverse-phase HPLC using ammonium acetate and acetonitrile as the eluent (to prevent hydrolysis of the acid labile sulfate ester) and was produced in 16% yield after repetitive lyophilization from water to remove excess ammonium acetate. The crude unprotected sulfated glycopeptides were subsequently treated with aqueous hydrazine hydrate to remove the *O*-acetate protective groups on the glycans, and thus provide the desired sulfated glycopeptides **11** and **12** in 13% and 12% yield, respectively, following purification. Having successfully prepared the six C-terminal fragments, peptide thioester **20** was next synthesized using a side-chain anchoring strategy (Scheme 2, see Supporting Information).[15] With all requisite fragments in hand, differentially modified fragments **7**–**12** were submitted to native chemical ligation with peptide thioester **20** using 6 M



Scheme 2. A) Synthesis of differentially *O*-glycosylated and Tyr-sulfated variants of full-length HP6 proteins **1**–**6** by native chemical ligation and oxidative folding. B–E) Selected spectra (see Supporting Information for all data): B) ESI mass spectrum and C) MALDI-TOF mass spectrum of unfolded sulfated glycoprotein ligation product **26**; D) ESI mass spectrum and E) MALDI-TOF mass spectrum of folded sulfated HP6 glycoprotein **6**. NB: loss of sulfate observed due to the acidic matrix employed in MALDI-TOF experiments.

guanidine hydrochloride (Gn·HCl) and 200 mM Na_2HPO_4 at pH 7.2 as the ligation buffer with the addition of mercaptophenylacetic acid (MPAA) as an aryl thiol catalyst and tris(2-carboxyethyl)phosphine (TCEP) as a reducing agent. Each ligation was monitored by HPLC-MS and, gratifyingly, all reactions reached completion after 16 h (see Supporting Information). Following reverse-phase HPLC purification, all six full-length unfolded HP6 variants **21**–**26** were obtained in good yield (49–67 % yield) and excellent purity (see Supporting Information). With **21**–**26** in hand, all that remained to complete the synthesis of the native proteins was the oxidative folding to generate the three correctly aligned disulfide bonds; Cys6–Cys14, Cys16–Cys28 and Cys22–Cys37, predicted based on the sequence alignment with HIRV1.[8] The optimal procedure involved the use of a cysteine (4 mM) and cystine (2 mM) redox couple in the presence of NaCl (4 M final conc.) at a pH of 8.5 which led to folding of all six reduced HP6 variants after 1 h to form a single species as judged by HPLC-MS (see Supporting Information).[16] The correct disulfide connectivity was confirmed through digestion with thermolysin and trypsin followed by HPLC-MS analysis and MS/MS peptide sequencing of the resulting peptide fragments (see Supporting Information). Gratifyingly, the modified HP6 proteins **1**–**6** were isolated in good yields (52–71 %) following HPLC purification.

Having successfully completed the synthesis of the library of six HP6 proteins, we were now interested in interrogating the effect of the glycan and sulfate modifications on the inhibition of both the fibrinogenolytic and amidolytic activity of human thrombin. Inhibition of fibrinogen clot formation by **1**–**6** was measured spectrophotometrically as described previously (see Supporting Information for raw data).[17] Modification of HP6 had a drastic effect on inhibitory activity. Specifically, whilst unmodified HP6 **1** exhibited only moderate inhibition of the fibrinogenolytic activity ($IC_{50} = 4.97$ μM), addition of a monosaccharide (GalNAc) to Thr-43 provided a significant increase in activity, with an IC_{50} of 2.41 μM (Table 1). This effect was even more pronounced in HP6 variant **3**, bearing a disaccharide moiety appended at Thr-43, which exhibited an IC_{50} of 82 nM. It is conceivable that the carbohydrates are capable of forming favorable interactions with thrombin or, alternatively, inducing a conformation of HP6 whereby other C-terminal residues are capable of making additional interactions with the fibrinogen binding site of exosite I. Interestingly, truncated synthetic variants of HP6 (HP6 41–63) bearing identical glycan modifications did not demonstrate measurable inhibition of fibrinogenolytic activity ($IC_{50} > 5$ μM)[13b] thus underscoring the additional importance of the N-terminal domain. Incorporation of sulfation at Tyr-63 in sulfoprotein **4** led to a two orders of magnitude increase in inhibitory activity against the fibrinogenolytic activity ($IC_{50} = 5.80$ nM), consistent with the crucial role of sulfation in the anti-thrombotic activity of HIRV1. Intriguingly, in contrast to the increased potency observed upon *O*-glycosylation of **1**, the addition of glycans to **4** did not lead to an improvement in activity. Rather, **5** and **6** were almost an order of magnitude less potent ($IC_{50} = 40.6$ nM and 31.9 nM, respectively) suggesting that the presence of glycans is detrimental to the predicted ionic and hydrophobic interactions between the sulfated Tyr residue and the fibrinogen binding site.

Having demonstrated that the two PTMs were capable of modulating the fibrinogenolytic activity of thrombin, we next studied the effect on the amidolytic activity of the enzyme. Inhibition was measured in a fluorescence assay using a chromogenic peptide substrate as reported previously (Table 1).[19] As was observed with the effects on fibrinogenolytic activity, the PTMs had a drastic effect on amidolytic inhibitory activity. Unmodified HP6 again displayed the weakest activity, with an inhibition constant (K_i) of 14.1 nM. Addition of a single GalNAc unit to Thr-43 led to only a modest increase in activity ($K_i = 11.3$ nM), while the introduction of a Gal-β-1,3-GalNAc moiety to Thr-43 in **3** imparted a two-fold improvement in activity ($K_i = 6.38$ nM) when compared to unmodified HP6. This data suggests that the additional interactions or altered conformation induced by the presence of glycans, in particular the disaccharide moiety of **3**, contribute to a more favorable interaction between the N-terminal tail of HP6 and the thrombin active site. The most potent thrombin inhibitor of the six variants was sulfated and unglycosylated HP6 **4**, with a K_i of 820 pM. As was observed with the fibrinogenolytic assays, the addition of glycans to Thr-43 resulted in a significant reduction in inhibitory potency against the amidolytic activity of thrombin. Specifically, GalNAc-derived sulfoprotein **5** and Gal-β-1,3-GalNAc-derived sulfoprotein **6** exhibited inhibition constants of 6.80 nM and 7.63 nM, almost an order of magnitude less potent than **4**. These results again suggest that the interactions made by the sulfated Tyr moiety with exosite I residues of thrombin are crucial, either through correct positioning of the N-terminus of HP6 into the active site or by modifying the conformation of the active-site pocket for improved binding interactions with the N-terminal tail.[9b] Consistent with the trends observed for inhibition of fibrinogenolytic activity by **1**–**6**, it is possible that the addition of glycans to Thr-43 of HP6 leads to new binding interactions facilitating a structural change to HP6 or thrombin with a concomitant weakening of the binding interaction of the N-terminus of HP6 with the

Table 1: Inhibition of the fibrinogenolytic and amidolytic activities of human thrombin by homogeneously modified synthetic HP6 proteins.

Hirudin P6	IC_{50} against fibrinogenolytic thrombin activity [nM][a]	Inhibition constant (K_i) against amidolytic thrombin activity [nM][b]
1	4970	14.1 ± 0.61
2	2410	11.3 ± 0.17
3	82.2	6.38 ± 0.88
4	5.80	0.82 ± 0.04
5	40.6	6.80 ± 0.20
6	31.9	7.63 ± 0.19

[a] IC_{50} values calculated using a sigmoidal dose-response formula (GraphPad Prism). Standard errors for log IC_{50} values (from which IC_{50} values were derived) were all less than 10% of the overall values. [b] K_i values calculated using Easson–Stedman analysis, K_i values ± SEM are representative of at least two independent experiments. Positive control: HIRV1. $K_i = 20$ pM.[18]

active site. Taken together, these studies provide clear evidence that the pattern of PTMs, in this case *O*-glycosylation and Tyr sulfation, can significantly attenuate antithrombotic activity. While it is not yet possible to propose a precise molecular mechanism for the attenuated activity, it is clear that a complex interplay exists and that a given set of modifications can modulate inhibitory activity. Crystallographic and NMR spectroscopic studies should aid in the delineation of these effects and will be the subject of future work.

In summary, we have used a combination of divergent solid-phase methodologies and native chemical ligation to assemble six homogeneous post-translationally modified variants of HP6. To our knowledge these represent the first examples of full-length proteins bearing two different families of PTMs synthesized to date, as well as the first homogeneous sulfoproteins ever synthesized. The six proteins, which differed in the glycosylation and sulfation state at Thr-43 and Tyr-61, respectively, enabled a dissection of the effect of these modifications on the inhibition of both the fibrinogenolytic and amidolytic activities of thrombin. To our knowledge, modulation of inhibitory activity invoked by two different PTMs is an unprecedented phenomenon and provides an interesting avenue for future investigations into the detailed structure–activity relationships of other proteins that natively bear more than one type of PTM.

Received: December 12, 2013
Revised: February 5, 2014
Published online: March 11, 2014

Keywords: glycoprotein · glycosylation · ligation · sulfation · sulfoprotein

[1] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, Jr., *Angew. Chem.* **2005**, *117*, 7508 – 7539; *Angew. Chem. Int. Ed.* **2005**, *44*, 7342 – 7372.

[2] a) R. J. Payne, C. H. Wong, *Chem. Commun.* **2010**, *46*, 21 – 43; b) B. G. Davis, *Chem. Rev.* **2002**, *102*, 579 – 601.

[3] a) C. Seibert, T. P. Sakmar, *Pept. Sci.* **2008**, *90*, 459 – 477; b) E. Chapman, M. D. Best, S. R. Hanson, C. H. Wong, *Angew. Chem.* **2004**, *116*, 3610 – 3632; *Angew. Chem. Int. Ed.* **2004**, *43*, 3526 – 3548.

[4] a) P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, R. A. Dwek, *Science* **2001**, *291*, 2370 – 2376; b) S. R. Hanson, E. K. Culyba, T. L. Hsu, C. H. Wong, J. W. Kelly, E. T. Powers, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 3131 – 3136; c) E. K. Culyba, J. L. Price, S. R. Hanson, A. Dhar, C. H. Wong, M. Gruebele, E. T. Powers, J. W. Kelly, *Science* **2011**, *331*, 571 – 575; d) G. Walsh, R. Jefferis, *Nat. Biotechnol.* **2006**, *24*, 1241 – 1252; e) P. P. Wilkins, K. L. Moore, R. P. McEver, R. D. Cummings, *J. Biol. Chem.* **1995**, *270*, 22677 – 22680; f) W. S. Somers, J. Tang, G. D. Shaw, R. T. Camphausen, *Cell* **2000**, *103*, 467 – 479; g) M. Farzan, T. Mirzabekov, P. Kolchinsky, R. Wyatt, M. Cayabyab, N. P. Gerard, C. Gerard, J. Sodroski, H. Choe, *Cell* **1999**, *96*, 667 – 676; h) D. Sako, K. M. Comess, K. M. Barone, R. T. Camphausen, D. A. Cumming, G. D. Shaw, *Cell* **1995**, *83*, 323 – 331.

[5] C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408 – 4420.

[6] a) D. Taleski, S. J. Butler, M. J. Stone, R. J. Payne, *Chem. Asian J.* **2011**, *6*, 1316 – 1320; b) A. M. Ali, S. D. Taylor, *Angew. Chem.* **2009**, *121*, 2058 – 2060; *Angew. Chem. Int. Ed.* **2009**, *48*, 2024 – 2026; c) K. M. Koeller, M. E. B. Smith, C. H. Wong, *J. Am. Chem. Soc.* **2000**, *122*, 742 – 743.

[7] a) P. Wang, S. W. Dong, J. A. Brailsford, K. Iyer, S. D. Townsend, Q. Zhang, R. C. Hendrickson, J. Shieh, M. A. S. Moore, S. J. Danishefsky, *Angew. Chem.* **2012**, *124*, 11744 – 11752; *Angew. Chem. Int. Ed.* **2012**, *51*, 11576 – 11584; b) Y. Kajihara, N. Yamamoto, R. Okamoto, K. Hirano, T. Murase, *Chem. Rec.* **2010**, *10*, 80 – 100; c) P. Wang, S. Dong, J.-H. Shieh, E. Peguero, R. Hendrickson, M. A. S. Moore, S. J. Danishefsky, *Science* **2013**, *342*, 1357 – 1360.

[8] V. Steiner, R. Knecht, K. O. Bornsen, E. Gassmann, S. R. Stone, F. Raschdorf, J. M. Schlaeppi, R. Maschler, *Biochemistry* **1992**, *31*, 2294 – 2298.

[9] a) T. J. Rydel, K. G. Ravichandran, A. Tulinsky, W. Bode, R. Huber, C. Roitsch, J. W. Fenton 2nd, *Science* **1990**, *249*, 277 – 280; b) C. C. Liu, E. Brustad, W. Liu, P. G. Schultz, *J. Am. Chem. Soc.* **2007**, *129*, 10648 – 10649.

[10] S. R. Stone, J. Hofsteenge, *Biochemistry* **1986**, *25*, 4622 – 4628.

[11] a) A. Greinacher, N. Lubenow, *Circulation* **2001**, *103*, 1479 – 1484; b) M. A. Corral-Rodríguez, S. Macedo-Ribeiro, P. J. Pereira, P. Fuentes-Prior, *J. Med. Chem.* **2010**, *53*, 3847 – 3861.

[12] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776 – 779; b) S. B. H. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338 – 351.

[13] a) B. L. Wilkinson, R. S. Stone, C. J. Capicciotti, M. Thaysen-Andersen, J. M. Matthews, N. H. Packer, R. N. Ben, R. J. Payne, *Angew. Chem.* **2012**, *124*, 3666 – 3670; *Angew. Chem. Int. Ed.* **2012**, *51*, 3606 – 3610; b) Y. S. Y. Hsieh, D. Taleski, B. L. Wilkinson, L. C. Wijeyewickrema, T. E. Adams, R. N. Pike, R. J. Payne, *Chem. Commun.* **2012**, *48*, 1547 – 1549.

[14] A. Y. Desoky, J. Hendel, L. Ingram, S. D. Taylor, *Tetrahedron* **2011**, *67*, 1281 – 1287.

[15] a) S. Ficht, R. J. Payne, R. T. Guy, C. H. Wong, *Chem. Eur. J.* **2008**, *14*, 3620 – 3629; b) R. E. Thompson, K. A. Jolliffe, R. J. Payne, *Org. Lett.* **2011**, *13*, 680 – 683; c) L. R. Malins, K. M. Cergol, R. J. Payne, *ChemBioChem* **2013**, *14*, 559 – 563; d) L. R. Malins, R. J. Payne, *Org. Lett.* **2012**, *14*, 3142 – 3145.

[16] J. Y. Chang, *Biochem. J.* **1994**, *300*, 643 – 650.

[17] C. G. Binnie, B. W. Erickson, J. Hermans, *FEBS Lett.* **1990**, *270*, 85 – 89.

[18] J. G. Bieth, *Bull. Eur. Physiopathol. Respir.* **1980**, *16*, 183 – 197.

[19] N. M. Y. Ng, N. S. Quinsey, A. Y. Matthews, D. Kaiserman, L. C. Wijeyewickrema, P. I. Bird, P. E. Thompson, R. N. Pike, *Arch. Biochem. Biophys.* **2009**, *489*, 48 – 54.

