Glycopeptides

DOI: 10.1002/ange.201404904

Synthesis of the Highly Glycosylated Hydrophilic Motif of Extensins**

Akihiro Ishiwata, Sophon Kaeothip, Yoichi Takeda, and Yukishige Ito**

Abstract: *Extensin, the structural motif of plant extracellular matrix proteins, possesses a unique highly glycosylated, hydrophilic, and repeating Ser_1Hyp_4 pentapeptide unit, and has been proposed to include post-translational hydroxylation at proline residue and subsequent oligo-L-arabinosylations at all of the resultant hydroxyprolines as well as galactosylation at serine residue. Reported herein is the stereoselective synthesis of one of the highly glycosylated motifs, Ser(Galp$_1$)-Hyp(Araf$_4$)-Hyp(Araf$_4$)-Hyp(Araf$_3$)-Hyp(Araf$_1$). The synthesis has been completed by the application of 2-(naphthyl)methylether-mediated intramolecular aglycon delivery to the stereoselective construction of the Ser(Galp$_1$) and Hyp(Araf$_n$) fragments as the key step, as well as Fmoc solid-phase peptide synthesis for the backbone pentapeptide.*

Hydroxyproline-rich glycoproteins (HRGPs),[1] which are major structural components of plant extracellular matrices, are produced by extensive post-translational modifications of proline residues. They are first hydroxylated by prolyl 4-hydroxylases[2] and resultant hydroxyproline (Hyp) residues are glycosylated by L-arabinofuranosyltransferases (AFT).[3] These modifications are widespread in plants and essential for their developmental processes such as root hair growth.[4] Secreted peptide hormones of plant origin,[5] such as CLV3,[6] are modified in a similar manner. In addition to arabinofuranosylation, glycosylation of serine (Ser) residues has been found in extensins.[7] Glycosylation has been proposed to enhance their conformational rigidity and are important for molecular recognition[8] required for self-assembly of plant cell walls.[9]

Extensins are pivotal components of plant cell wall architectures, and are required for their self-assembly. Extensin monomers are bipartite in nature, thus consisting of hydrophobic and hydrophilic repeating motifs. Crosslinking of highly conserved Tyr-Xaa-Tyr motifs by oxidative coupling of their Tyr residues confers insolubility on extensins.[10,11] In contrast, the most common repeating motif of their hydrophilic region is the Ser-Hyp-Hyp-Hyp-Hyp pentapeptide modified by Ser α-D-galactopyranosylation [Ser(Galp_1)] and Hyp oligo-L-arabinofuranosylation [Hyp(Ara$f$)$_{1\text{-}4}$] (Figure 1).[12,13] The hydrophilic motif of a HRGP related to root hair growth[12] was proposed to consist of, from the N-to-C terminus, Ser(Galp_1), [Hyp(Araf_4)]$_2$, Hyp(Araf_3), and Hyp(Araf_1) residues. In Hyp(Ara$f_{1\text{-}3}$), all glycosides have been found to be β-linked, while Hyp(Araf_4)[14,15] has an α-L-Ara$f$ residue 1→3-linked to the Hyp(Araf_3). In addition to their structures and activities, biosynthetic as well as metabolic processes have been subjects of recent studies. For example, β-L-arabinofuranosidases (AFases)[16] and β-L-AFT[17] have been identified recently. Interestingly, HypBA1, one of the AFases, was indicated to be a cysteine glycosidase.[18]

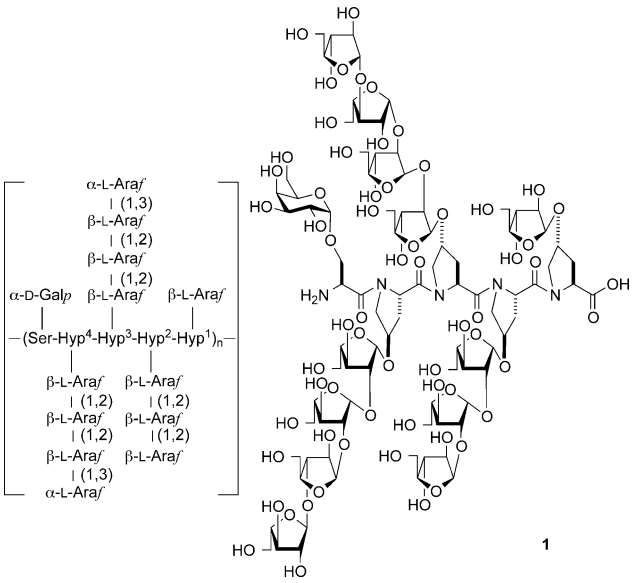


Figure 1. The hydrophilic repeating motif of the extensins **1**.

The hydrophilic motif **1** is synthetically challenging because they are extensively modified by oligosaccharides consisting of consecutive β-Araf*s*. Since stereoselective construction of β-Ara*f* glycosides is difficult to achieve because of its 1,2-*cis* nature, various approaches[19] based on direct[20] or intramolecular[21,22] glycosylation strategies have been examined by targeting extensin structure motifs.[23,24] In contrast, preparation of Ser(Galp_1) is intrinsically more straightforward and has been carried out in a conventional manner.[25]

For the synthesis of 1,2-*cis* glycosides, approaches based on intramolecular aglycon delivery (IAD)[26] have been

[*] Dr. A. Ishiwata, Dr. Y. Ito
Synthetic Cellular Chemistry Laboratory, RIKEN
2-1 Hirosawa, Wako, Saitama 351-0198 (Japan)
E-mail: aishiwa@riken.jp
yukito@riken.jp
Dr. S. Kaeothip, Dr. Y. Takeda, Dr. Y. Ito
ERATO glycotrilogy project, JST (Japan)

[**] We thank Dr. Hiroyuki Koshino (RIKEN Global Research Cluster) and his staff for technical help for technical help with ESI MS and CD, and Dr. Fumiaki Hayashi and Dr. Hui-ping Zhang (RIKEN Center for Life Science Technologies) for 900 MHz NMR measurements and are grateful to the Support Unit for Biomaterial Analysis (RIKEN Brain Science Institute) for tandem MS analysis. We also thank Ms. Akemi Takahashi for her kind technical assistance. This work was partly supported by Incentive Research Grant (2013) in RIKEN.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201404904.

employed successfully. Among a number of variants reported, 2-(naphthyl)methyl (NAP) ether mediated IAD has been shown to be most versatile,[27,28] and applied to the synthesis of CLV3[29] and *p*-nitrophenyl β-L-Ara*f*,[30] a valuable substrate for HypBA1. Herein we report the first synthesis of a hydrophilic repeating motif typical of extensins, Ser(Gal*p*$_1$)-Hyp-(Ara*f*$_4$)-Hyp(Ara*f*$_4$)-Hyp(Ara*f*$_3$)-Hyp(Ara*f*$_1$) (**1**). It features the extensive use of NAP-IAD for all 1,2-*cis* glycosides, including Ser(Gal*p*$_1$) and Hyp(Ara*f*$_n$) ($n = 1, 3, 4$). Subsequent Fmoc solid-phase peptide synthesis (Fmoc-SPPS) was carried out with suppression of diketopiperazine (DKP) formation.

Imaginary disconnection of the target structure led us to design fragments corresponding to Ser(Gal*p*$_1$) (**5**) and Hyp-(Ara*f*$_n$) ($n = 1, 3, 4$; **2**–**4**; Figure 2). Among them, **2** and **3** were prepared as previously reported[29] (Scheme 1). Namely, starting from **10**, oxidative mixed acetal (MA) formation with the donor **11** in the presence of 2,3-dichloro-5,6-dicyanobenzoquinone (DDQ) was followed by IAD mediated by MeOTf and 2,6-di-*tert*-butyl-4-methylpyridine (DTBMP) to afford the Hyp(Ara*f*$_1$) derivative **12** in a fully stereoselective manner. Then **12** was converted into **2** in a conventional manner.

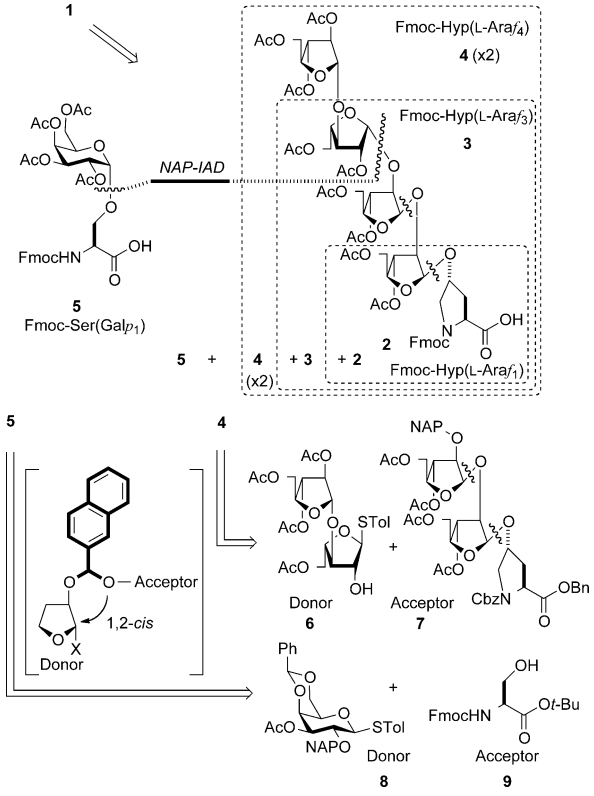
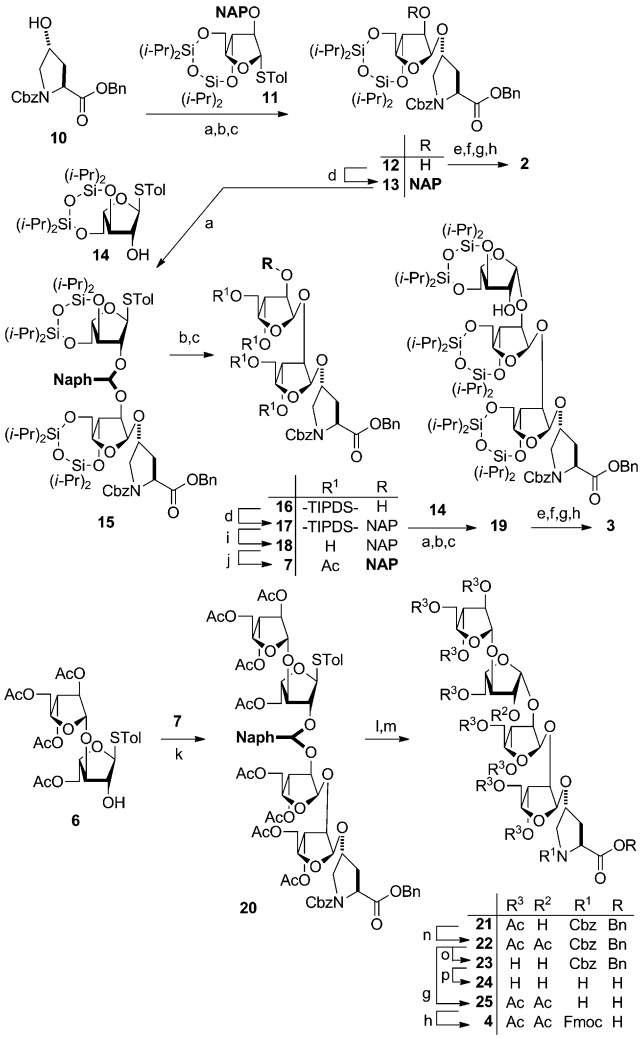


Figure 2. Synthetic plan for (oligo)saccharyl amino acid fragments of **1**. Cbz = benzyloxycarbonyl, Fmoc = 9-fluorenylmethyloxycarbonyl, Tol = 4-methylphenyl.



Scheme 1. Stereoselective synthesis of Hyp(Ara*f*$_{1\text{-}4}$) derivatives. a) DDQ, 4Å M.S., DCE, RT; b) MeOTf, DTBMP, 4Å M.S., DCE, 40°C, 48 h; c) TFA, $CHCl_3$, 0°C, 2 h, 70% (**12** from **10**), 73% (**16** from **13**), 51% (**19** from **17**); d) NAPBr, NaH, TBAI, DMF, −20°C, 6 h, 82% (**13**), 79% (**17**); e) TBAF, pyr/THF, 0°C, 1 h; f) Ac_2O, py, RT, 3 h; g) H_2, $Pd(OH)_2$, EtOAc/EtOH (2:1), RT, 8–16 h; h) FmocCl, DIPEA, CH_2Cl_2, RT, 5 h, 72% (**2** from **12**), 70% (**3** from **19**), 70% (**4** from **26**); i) TBAF, THF, 0°C, 1 h; j) Ac_2O, py, 89% from **17**; k) DDQ, 4Å M.S., DCE, RT, 74%; l) MeOTf, DTBMP, 4Å M.S., DCE, 40°C, 48 h; m) TFA, $CHCl_3$, 0°C, 0.5 h; n) Ac_2O, py, 84% from **20**; o) 0.1 M NaOH, MeOH, 0°C, 4 h; p) H_2, $Pd(OH)_2$, MeOH/H_2O/HOAc (30:10:1), RT, 15 h, 70% from **22**. M.S. = molecular seives, DCE = $(CH_2Cl)_2$, DDQ = 2,3-dichloro-5,6-dicyano-1,4-benzoquinone, DIPEA = $(iPr)_2NEt$, DTBMP = 2,6-di-*tert*-butyl-4-methylpyridine, Naph = 2-naphthyl, py = pyridine, TBAF = $(nBu)_4NF$, TBAI = $(nBu)_4NI$, Tf = trifluoromethanesulfonyl, TFA = trifluoroacetic acid, TIPDS = tetraisopropyldisiloxanylidene.

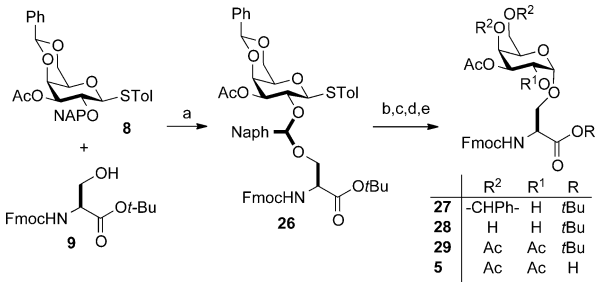
From the preceding study,[29] we learned that in order for the IAD to be most efficient, proper combination of donor and acceptor is critical. Specifically, the use of an NAP-masked acceptor in combination with a 2-*O*-unprotected donor is optimal for the construction of consecutive β-Ara*f* linkages. Accordingly, **12** was converted into the corresponding NAP ether **13**, which was coupled with the donor **14** through the MA **15** under standard NAP-IAD conditions. After conversion of **16** into **17**, chain elongation was carried out in a similar manner to give **3** from **17** through **19**.

For the preparation of the Hyp(Ara*f*$_4$) component **4**, a fragment coupling between the donor **6**[31] and the acceptor **7** was planned (Figure 2). The acceptor **7** was prepared from **17** through desilylation and acetylation of the resultant **18**

(Scheme 1). Coupling of **6** with **7** stereoselectively afforded **21** from the MA **20**, which was converted into **22**. Since, judging from NMR spectra, **22** was a mixture of rotamers, its homogeneity was confirmed after being converted into the free Hyp(Araf_4) **24** from **23**. The compound **22** was then converted into **4** by way of **25**.

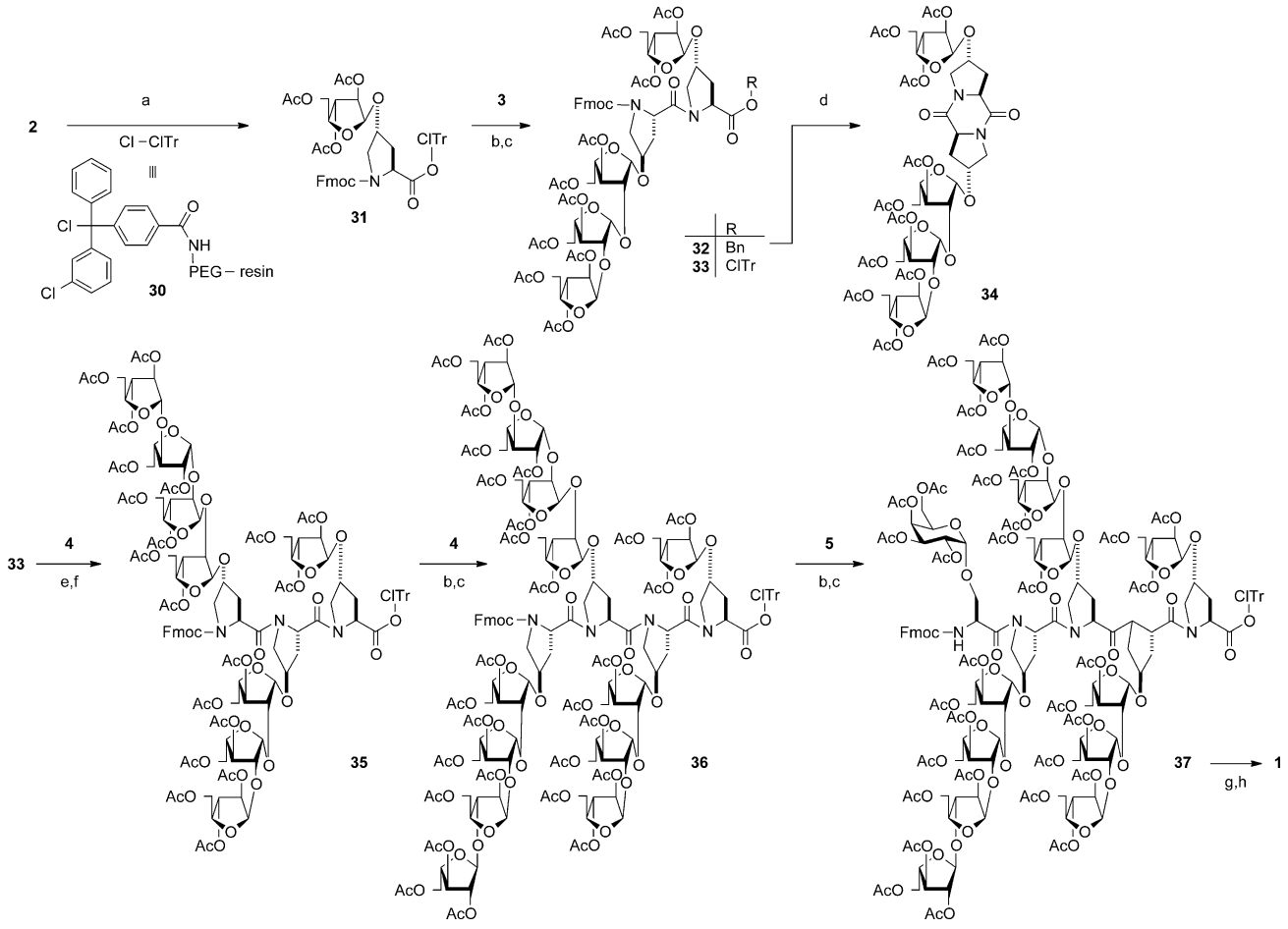
The Fmoc-Ser(α-Galp_1) derivative **5** was synthesized from the NAP-protected donor **8**[31] (Figure 2, Scheme 2), although **5** was also obtained predominantly by direct intermolecular glycosylation.[31] The donor **8** was treated with the Fmoc-Ser-O-*t*Bu **9** under oxidative mixed-acetal-formation conditions to give the MA **26**, which was converted into the expected α-galactosylserine derivative **29** in a highly stereoselective manner after mild acidic work-up (10% TFA in $CHCl_3$ at 0°C) of **27** and *O*-acetylation of the resultant triol **28**. Further acidic removal of the *tert*-butylester of **29** gave the desired Fmoc-Ser(Galp_1) derivative **5**.[25]

We initially planned to conduct solution-phase peptide synthesis using 1-[(1-(cyano-2-ethoxy-2-oxo-ethylideneaminooxy)-dimethylaminomorpholino)]uronium hexafluorophosphate (COMU)[32] as a coupling reagent[29] (Scheme 3). However, not unexpectedly,[33] treatment of the dipeptide derivative **32**[31] with piperidine to remove the Fmoc group



Scheme 2. Stereoselective synthesis of **5**. a) DDQ, 4Å M.S., DCE, RT, 98%; b) MeOTf, DTBMP, 4Å M.S., DCE, 40°C, 48 h; c) 10% TFA, $CHCl_3$, 0°C, 2 h; d) Ac_2O, py, 56% from **30**; e) 20% TFA, $CHCl_3$, RT, 2 h, 84%.

resulted in formation of a significant amount of the DKP **34** (MALDI-TOF MS: $[M+Na]^+$ calc for $C_{50}H_{66}N_2Na_1O_{30}$, 1197.35, found 1197.43). This problem was circumvented by Fmoc-SPPS using the chlorotrityl (ClTr) resin **30**.[34] Namely, after introducing **2**, the resin **31** was subjected to Fmoc removal and COMU coupling with **3** to give the resin-bound dipeptide **33**. Subsequent Fmoc removal was carried out with piperidine in the presence of 1-*O*-hydroxybenzotriazole



Scheme 3. Solid-phase glycopeptide synthesis toward **1**. a) DIPEA, RT, 16 h; b) pip, DMF, RT, 10 min; c) COMU, DIPEA, RT, 16 h; d) pip, DMF, RT, 84%; e) pip, HOBt, DMF, RT, 10 min; f) DIC, HOBt, RT, 16 h; g) TFA/(iPr)$_3$SiH/H_2O (190:5:5), 1 h; h) NaOH, MeOH, 0°C, 21% based on initial loading to ClTr resin. pip = piperidine.

(HOBt)[35] to trap free amino groups, and DIC–HOBt coupling[36] with **4** afforded the tripeptide **35** with suppression of the DKP formation. The entire sequence of **1** was successfully constructed by iterative chain elongation using **4** and **5** to give **37** via **36**. Finally, the resultant pentapeptide was cleaved from the resin and was deacetylated at 0 °C,[37] thus completing the synthesis of **1**. High-resolution ESI-TOF mass data ($[M+H]^+$ calcd for $C_{89}H_{142}N_5O_{64}$, 2304.8011, found 2304.7968) provided evidence to support the target structure.

For structural analysis, we opted to use the CID method for tandem MS-MS analysis, although glycosidic linkages would be cleaved to a larger extent than in ETD/ECD methods. The fragmentations of all sugar moieties one by one was observed to support the presence of all required sugar residues.[31] The structure of **1** was also analyzed by two-dimensional NMR techniques in detail. The corresponding 13 glycosidic linkages were identified in HSQC spectra (Figure 3), and full assignment of the peaks was made at 900 MHz.[31]

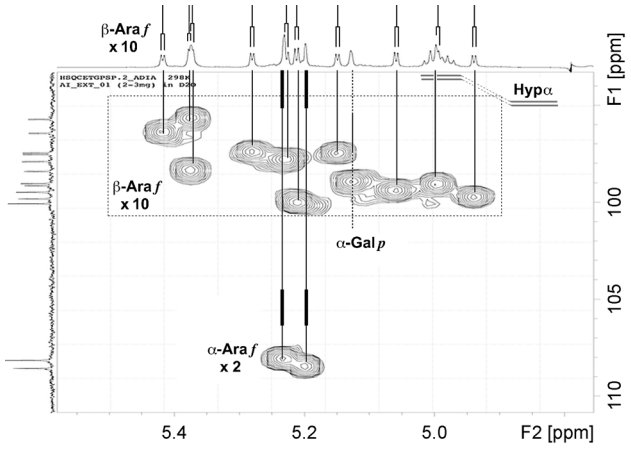


Figure 3. Anomeric reagion of HSQC spectra (900 MHz, D_2O) of **1**.

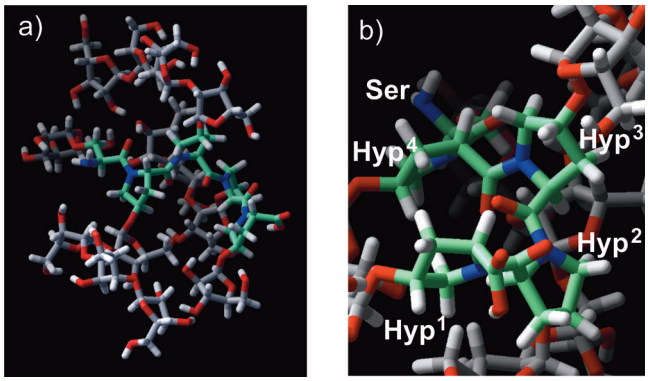


Figure 4. a) Global minimum structure of **1** in H_2O. Carbon atoms of amino acids are colored green. b) Focus on peptides.

The NOESY experiment revealed correlations between all pairs of neighboring α-H(Ser/Hyp) and δ-H_2(Hyp) units, even at 40 °C. Consequently, all Hyps in **1** were suggested to possess *s-trans*-configured amide bonds, thus making the glycopeptide into a left-handed polyPro II helixlike structure[38] (Figure 4), as has been found in nonglycosylated as well as in glycosylated Hyp derivatives.[39,31]

In summary, stereoselective synthesis of Ser(Galp_1)-Hyp-(Araf_4)-Hyp(Araf_4)-Hyp(Araf_3)-Hyp(Araf_1) (**1**) was realized by NAP-IAD for stereoselective constructions of all 1,2-*cis* glycosides containing the fragments Ser(Galp_1) and Hyp-(Araf_n) ($n = 1, 3, 4$). Assembly of the highly glycosylated pentapeptide unit was achieved by Fmoc solid-phase peptide synthesis.

Received: May 2, 2014
Published online: July 15, 2014

Keywords: glycopeptides · proteins · solid-phase synthesis · synthetic methods

[1] For recent reviews, see: M. J. Kieliszewski, D. T. A. Lamport, L. Tan, M. C. Cannon in *Ann. Plant Rev.: Plant Polysaccharides, Biosynthesis and Bioengineering*, *Vol. 41* (Ed.: P. Ulvskov), Wiley-Blackwell, Oxford, **2011**, pp. 321–342, and references therein.
[2] a) R. Hieta, J. Myllyharju, *J. Biol. Chem.* **2002**, *277*, 23965–23971; b) P. Tiainen, J. Myllyharju, P. Koivunen, *J. Biol. Chem.* **2005**, *280*, 1142–1148.
[3] B. L. Petersen, K. Faber, P. Ulvskov in *Ann. Plant Rev.: Plant Polysaccharides, Biosynthesis and Bioengineering*, *Vol. 41* (Ed.: P. Ulvskov), Wiley-Blackwell, Oxford, **2011**, pp. 305–320.
[4] S. M. Velasquez, et al, *Science* **2011**, *332*, 1401–1403.
[5] a) Y. Matsubayashi, *Plant Cell Physiol.* **2011**, *52*, 5–13; b) Y. Matsubayashi, *Genes Cells* **2012**, *17*, 1–10.
[6] K. Ohyama, H. Shinohara, M. Ogawa-Ohnishi, Y. Matsubayashi, *Nat. Chem. Biol.* **2009**, *5*, 578–580.
[7] D. T. A. Lamport, L. Katona, S. Roerig, *Biochem. J.* **1973**, *133*, 125–131.
[8] M. J. Kieliszewski, D. T. A. Lamport, *Plant J.* **1994**, *5*, 157–172.
[9] M. C. Cannon, K. Terneus, Q. Hall, L. Tan, Y. Wang, B. L. Wegenhart, L. Chen, D. T. A. Lamport, Y. Chen, M. J. Kieliszewski, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2226–2231.
[10] S. C. Fry, *Biochem. J.* **1982**, *204*, 449–455.
[11] L. Epstein, D. T. A. Lamport, *Phytochemistry* **1984**, *23*, 1241–1246.
[12] J. M. Estévez, M. J. Kieliszewski, N. Khitrov, C. Somerville, *Plant Physiol.* **2006**, *142*, 458–470.
[13] S. Gille, U. Hänsel, M. Ziemann, M. Pauly, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 14699–14704.
[14] a) Y. Akiyama, K. Kato, *Agric. Biol. Chem.* **1976**, *40*, 2343–2348; b) Y. Akiyama, K. Kato, *Agric. Biol. Chem.* **1977**, *41*, 79–81.
[15] a) F. M. Klis, H. Eeltink, *Planta* **1979**, *144*, 479–484; b) D. Ashford, N. N. Desai, A. K. Allen, A. Neuberger, M. A. O'Neil, R. R. Selvendran, *Biochem. J.* **1982**, *201*, 199–208.
[16] a) K. Fujita, Y. Takashi, E. Obuchi, K. Kitahara, T. Suganuma, *J. Biol. Chem.* **2014**, *289*, 6110–6119; b) K. Fujita, S. Sakamoto, Y. Ono, M. Wakao, Y. Suda, K. Kitahara, T. Suganuma, *J. Biol. Chem.* **2011**, *286*, 5143–5150.
[17] M. Ogawa-Ohnishi, W. Matsushita, Y. Matsubayashi, *Nat. Chem. Biol.* **2013**, 9, 726–730.
[18] T. Ito, K. Saikawa, S. Kim, K. Fujita, A. Ishiwata, S. Kaeothip, T. Arakawa, T. Wakagi, G. T. Beckham, Y. Ito, S. Fushinobu, *Biochen. Biophys. Res. Commun.* **2014**, *447*, 32–37.
[19] For recent reviews on 1,2-*cis* glycosylation, see: a) A. Ishiwata, Y. Ito, *Trends Glycosci. Glycotechnol.* **2009**, *21*, 266–289;

b) L. K. Mydock, A. V. Demchenko, *Org. Biomol. Chem.* **2010**, *8*, 497 – 510.
[20] For recent reviews, see: a) A. Imamura, T. L. Lowary, *Trends Glycosci. Glycotechnol.* **2011**, *23*, 134 – 152; For our studies on intermolecular approach, see: b) A. Ishiwata, H. Akao, Y. Ito, *Org. Lett.* **2006**, *8*, 5525 – 5528; c) A. Ishiwata, H. Akao, Y. Ito, M. Sunagawa, N. Kusunose, Y. Kashiwazaki, *Bioorg. Med. Chem.* **2006**, *14*, 3049 – 3061.
[21] For PMB-ether-mediated IAD, see: a) Y. Ito, T. Ogawa, *Angew. Chem.* **1994**, *106*, 1843 – 1845; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1765 – 1767; b) J. Désiré, J. Prandi, *Carbohydr. Res.* **1999**, *317*, 110 – 118; c) H. Shinohara, Y. Matsubayashi, *Plant Cell Physiol.* **2013**, *54*, 369 – 374; d) S. Okamoto, H. Shinohara, T. Mori, Y. Matsubayashi, M. Kawaguchi, *Nat. Commun.* **2013**, *4*, 2191.
[22] a) A. Ishiwata, Y. Munemura, Y. Ito, *Eur. J. Org. Chem.* **2008**, 4250 – 4263; b) A. Ishiwata, Y. Ito, *J. Am. Chem. Soc.* **2011**, *133*, 2275 – 2291.
[23] N. Xie, C. M. Taylor, *Org. Lett.* **2010**, *12*, 4968 – 4971.
[24] S. Kaeothip, G.-J. Boons, *Org. Biomol. Chem.* **2013**, *11*, 5136 – 5146.
[25] For recent examples, see: a) T. Tsujimoto, Y. Ito, *Tetrahedron Lett.* **2007**, *48*, 5513 – 5516; b) L.-D. Huang, H.-J. Lin, P.-H. Huang, W.-C. Hsiao, V. R. Reddy, S.-L. Fu, C.-C. Lin, *Org. Biomol. Chem.* **2011**, *9*, 2492 – 2504; c) Y. Kim, K. Oh, S. Song, D.-S. Lee, S. B. Park, *J. Med. Chem.* **2013**, *56*, 7100 – 7109.
[26] For recent reviews, see: a) I. Cumpstey, *Carbohydr. Res.* **2008**, *343*, 1553 – 1573; b) A. T. Carmona, A. J. Moreno-Vargas, I. Robina, *Curr. Org. Synth.* **2008**, *5*, 33 – 63; c) A. Ishiwata, Y. J. Lee, Y. Ito, *Org. Biomol. Chem.* **2010**, *8*, 3596 – 3608.
[27] A. Ishiwata, Y. Ito, *Yuki Gosei Kagaku Kyokaishi* **2012**, *70*, 382 – 394.
[28] For 1,2-*cis* pyranosylation, see: a) Y. J. Lee, A. Ishiwata, Y. Ito, *J. Am. Chem. Soc.* **2008**, *130*, 6330 – 6331; b) A. Ishiwata, A. Sakurai, Y. Nishimiya, S. Tsuda, Y. Ito, *J. Am. Chem. Soc.* **2011**, *133*, 19524 – 19535.
[29] S. Kaeothip, A. Ishiwata, Y. Ito, *Org. Biomol. Chem.* **2013**, *11*, 5892 – 5907.
[30] S. Kaeothip, A. Ishiwata, T. Ito, S. Fushinobu, K. Fujita, Y. Ito, *Carbohydr. Res.* **2013**, *382*, 95 – 100.
[31] See the Supporting Information.
[32] A. El-Faham, R. S. Funosas, R. Prohens F. Albericio, *Chem. Eur. J.* **2009**, *15*, 9404 – 9416.
[33] W. Lukenmaier, H. Zahn, *Justus Liebigs Ann. Chem.* **1970**, *740*, 1 – 17.
[34] P. Rovero, S. Viganò, S. Pegoraro, L. Quartara, *Lett. Pept. Sci.* **1995**, *2*, 319 – 323.
[35] J. L. Lauer, C. G. Fields, G. B. Fields, *Lett. Pept. Sci.* **1994**, *1*, 197 – 205.
[36] M. A. Fara, J. J. Diaz-Mochon, M. Bradley, *Tetrahedron Lett.* **2006**, *47*, 1011 – 1014.
[37] U. Westerlind, A. Hobel, N. Gaidzik, E. Schmitt, H. Kunz, *Angew. Chem.* **2008**, *120*, 7662 – 7667; *Angew. Chem. Int. Ed.* **2008**, *47*, 7551 – 7556.
[38] C. M. Deber, F. A. Bovey, J. P. Carver, E. R. Blout, *J. Am. Chem. Soc.* **1970**, *92*, 6191 – 6198.
[39] a) N. W. Owens, C. Braun, J. D. O'Neil, K. Marat, F. Schweizer, *J. Am. Chem. Soc.* **2007**, *129*, 11670 – 11671; b) N. W. Owens, A. Lee, K. Marat, F. Schweizer, *Chem. Eur. J.* **2009**, *15*, 10649 – 10657.

