Oligosaccharide Synthesis

DOI: 10.1002/ange.200604909

Highly Efficient and Versatile Synthesis of Proteoglycan Core Structures from 1,6-Anhydro-β-lactose as a Key Starting Material**

*Ken Shimawaki, Yoshinori Fujisawa, Fumihiro Sato, Naoki Fujitani, Masaki Kurogochi, Hiroko Hoshi, Hiroshi Hinou, and Shin-Ichiro Nishimura**

Cell-surface and extracellular-matrix glycosaminoglycans (GAGs) of proteoglycans (PGs) are known to play vital functional roles in various biological processes, such as cell proliferation, differentiation, adhesion, blood coagulation, and wound repair.[1] The major GAG chains found in PGs are divided into two categories, heparan sulfate/heparin and chondroitin sulfate/dermatan sulfate, which are high-molecular-weight polysaccharides that consist of repeating [GlcNAcα(1→4)GlcAβ(1→4)] and [GalNAcβ(1→4)GlcAβ-(1→3)] units, respectively. These GAG chains are further modified by some epimerases and sulfotransferases to provide considerable complexity and the functionality required for interaction with growth factors, cytokines, and many protein ligands. The biosynthetic pathway to the mature GAG chains is initiated from the core tetrasaccharide moiety, namely, GlcAβ(1→3)Galβ(1→3)Galβ(1→4)Xylβ1→, which is attached to the serine residue of core proteins. Modification of this GAG-protein linkage region by phosphorylation and/or sulfation seems to influence the molecular diversity in the biosynthetic pathway of the disaccharide repeating structures.[2,3] Furthermore, GAG precursors have been proposed as antiamyloid or antithrombotic agents in anticancer therapy.[4] Although extensive efforts have been made toward the synthesis of GAGs, with particular emphasis on the core structures,[3,5] we believed that an efficient and versatile strategy for assembling various proteoglycan initiators would greatly accelerate the process of gaining an understanding of the functional roles of the enzymes responsible for the construction of GAGs. We communicate herein a highly efficient and versatile synthesis of proteoglycan core structures via key intermediates derived from 1,6-anhydro-β-lactose,[6] which can be prepared readily from naturally occurring, abundant lactose (Figure 1). We demonstrate the versatility of the building blocks **1** and **9a** in the synthesis of model tetrapeptides H-Ser[1]-Gly[2]-Ser[3]-Gly[4]-OH with di-(Galβ(1→4)Xylβ1→), tri- (Galβ(1→3)Galβ(1→4)Xylβ1→), and tetrasaccharide (GlcAβ(1→3)Galβ(1→3)Galβ(1→4)Xylβ1→) branches at the hydroxy group of the Ser[3] residue.

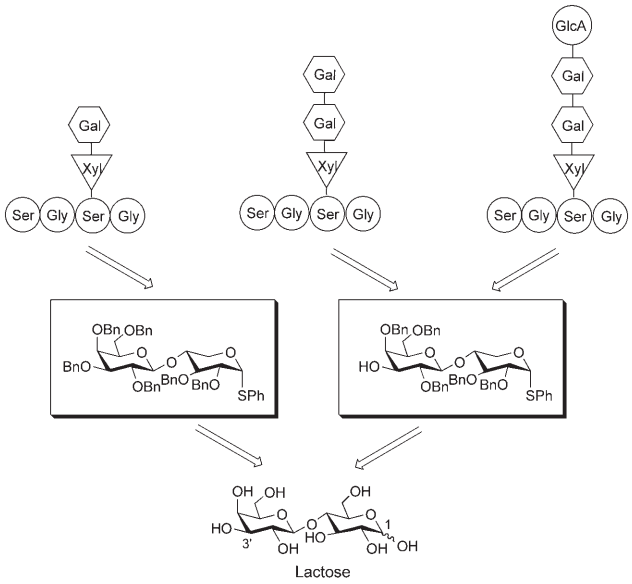


Figure 1. A synthetic strategy based on the conversion of 1,6-anhydro-β-lactose into the key intermediates **1** and **9a** for the construction of proteoglycan core linkage structures. Bn = benzyl.

As found in pioneering studies on the synthesis of GAGs,[5a–d] D-xylopyranose often makes selective chemical modification difficult. It seems likely that regioselective modifications and the stereoselective glycosidation of D-xylose derivatives would entail complicated synthetic protocols with tedious multistep protection/deprotection processes and the purification of anomeric mixtures from glycosidation reactions. We directed our attention toward reports that 1,6-anhydro-β-lactose can be used for the efficient synthesis of various important oligosaccharide derivatives.[6] We hypothesized that the use of 1,6-anhydro-β-lactose would permit the

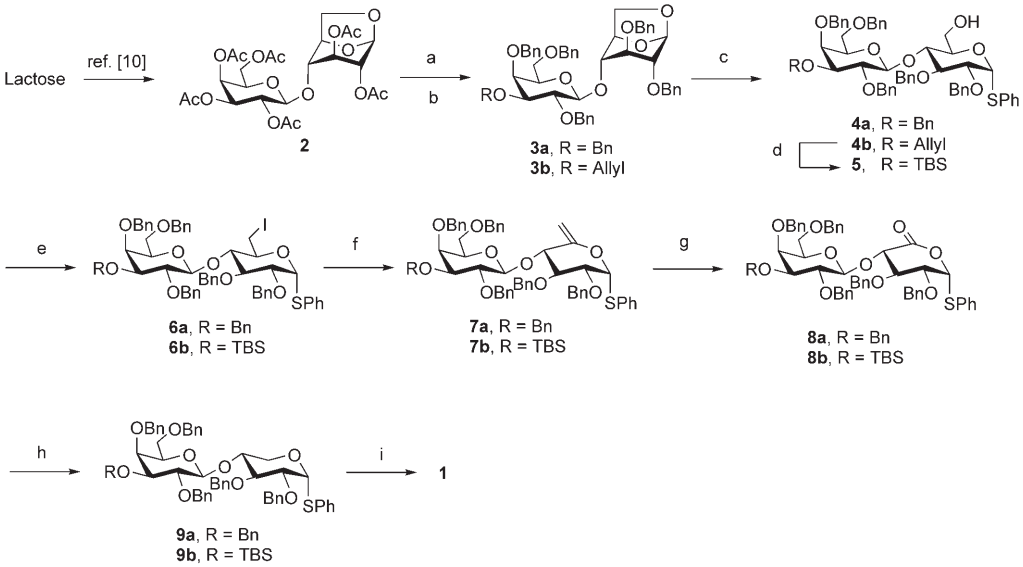
[*] K. Shimawaki, Y. Fujisawa, F. Sato, Dr. N. Fujitani, Dr. M. Kurogochi, Dr. H. Hoshi, Dr. H. Hinou, Prof. Dr. S.-I. Nishimura
Laboratory for Advanced Chemical Biology
Graduate School of Advanced Life Science
Frontier Research Center for Post-Genome Science and Technology
Hokkaido University, Sapporo 001-0021 (Japan)
Fax: (+81) 11-706-9042
E-mail: shin@glyco.sci.hokudai.ac.jp
Dr. H. Hinou, Prof. Dr. S.-I. Nishimura
Drug-Seeds Discovery Research Laboratory
National Institute of Advanced Industrial Science and Technology (AIST)
Sapporo 062-8517 (Japan)
Fax: (+81) 11-857-8451
E-mail: tiger.nishimura@aist.go.jp

[**] This research was partly supported by a grant for the "National Project on Functional Glycoconjugate Research Aimed at Developing New Industry" from the Ministry of Education, Science, and Culture of Japan. We thank S. Oka, M. Kiuchi, and T. Hirose of the Center of Instrumental Analysis, Hokkaido University, for high resolution mass spectrometric measurements and amino acid analysis.

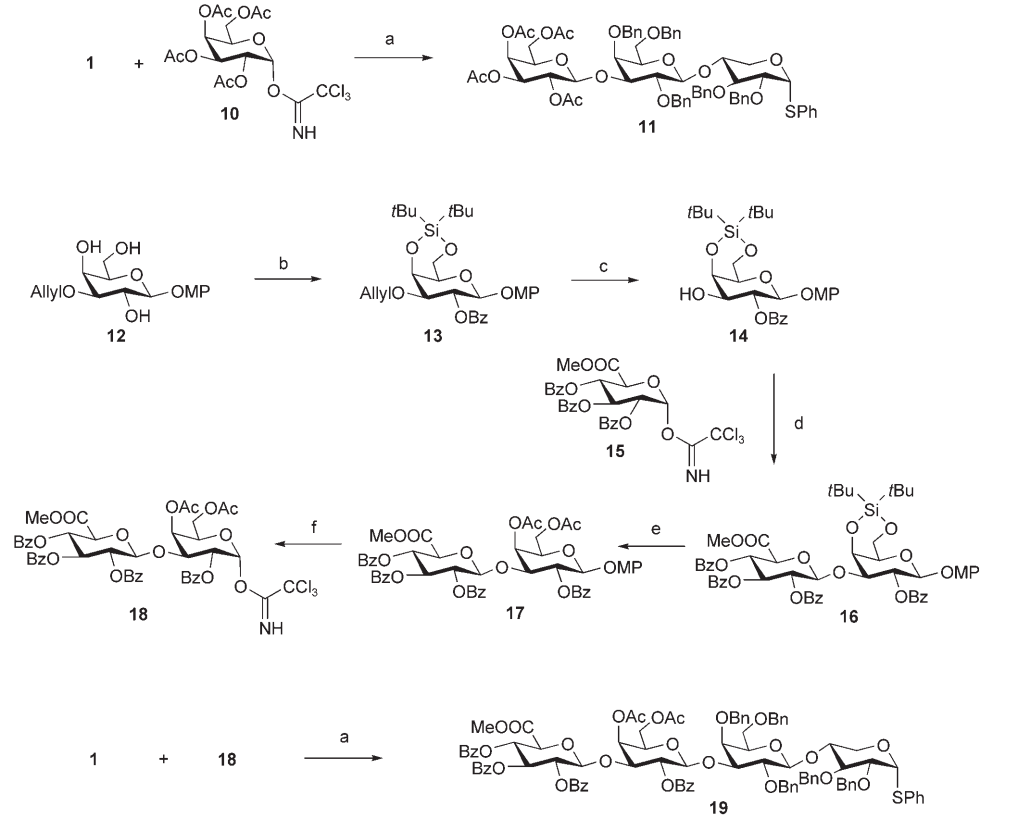
Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

rapid and efficient synthesis of a variety of Galβ(1→4)Xyl derivatives suitable for the construction of more complex and longer oligosaccharides as analogues of the linkage region of the proteoglycans.

Scheme 1 shows a synthetic route to the key intermediate **1**, which is suitably functionalized for both selective elongation with sugar residues at C3′ and glycosidation with core peptides at the anomeric (C1) position, and **9a**, which can also undergo glycosidation with core peptides at the anomeric position. The synthesis of compound **1** involves 1) regioselective modifications at the 3′-position of 1,6-anhydro-β-lactose, and 2) ring opening at the 1,6-anhydro bond and deletion of the hydroxymethyl group of the D-glucopyranose residue for replacement with a D-xylopyranose residue. First, the known compound **2**[6c] was subjected to de-*O*-acetylation and subsequent stannylation[7] with *n*-dibutyltin oxide (nBu$_2$SnO), followed by the introduction of an *O*-allyl group at the 3′-position and perbenzylation, to give compound **3b** in 87% overall yield. (Compound **3a** was prepared readily by direct O benzylation of free 1,6-anhydro-β-lactose in 93% yield.) The C6 hydroxy group was liberated by treatment with phenylthio(trimethyl)silane (TMSSPh) and trimethylsilyl trifluoromethanesulfonate (TMSOTf)[8] to afford alcohols **4a** (79%) and **4b** (70%). The allyl protecting group in **4b** was exchanged for a *tert*-butyldimethylsilyl group (**5**, 72% over four steps) for the planned ozonolysis. Next, the iodinated derivatives **6a** and **6b** were



Scheme 1. Synthetic route to the key intermediate **1**: a) NaOMe/MeOH; b) BnBr, NaH, DMF for **3a**; 1) nBu$_2$SnO, toluene, reflux; then allyl bromide, tetra-*n*-butyl ammonium bromide; 2) BnBr, NaH, DMF for **3b**; c) TMSSPh, TMSOTf, CH_2Cl_2; d) 1) Ac_2O, pyridine; 2) [Ir(cod)(PMePh$_2$)$_2$PF$_6$], H_2, EtOH; then 1 N HCl, acetone; 3) TBSOTf, 2,6-lutidine, CH_2Cl_2; 4) NaOMe/MeOH; e) 1) TsCl, pyridine; 2) NaI, 2-butanone; f) DBU, DMF, reflux; g) O_3, MeOH/THF; h) 1) DIBAL-H, CH_2Cl_2; 2) $(COCl)_2$, DMF, CH_2Cl_2; 3) $NaBH_3CN$, THF, reflux for **9a**, and $LiEt_3BH$, THF, reflux for **9b**; i) TBAF, THF. cod = cyclooctadiene, DMF = *N*,*N*-dimethylformamide, TBS = *tert*-butyldimethylsilyl, Tf = trifluoromethanesulfonyl, Ts = *p*-toluenesulfonyl.



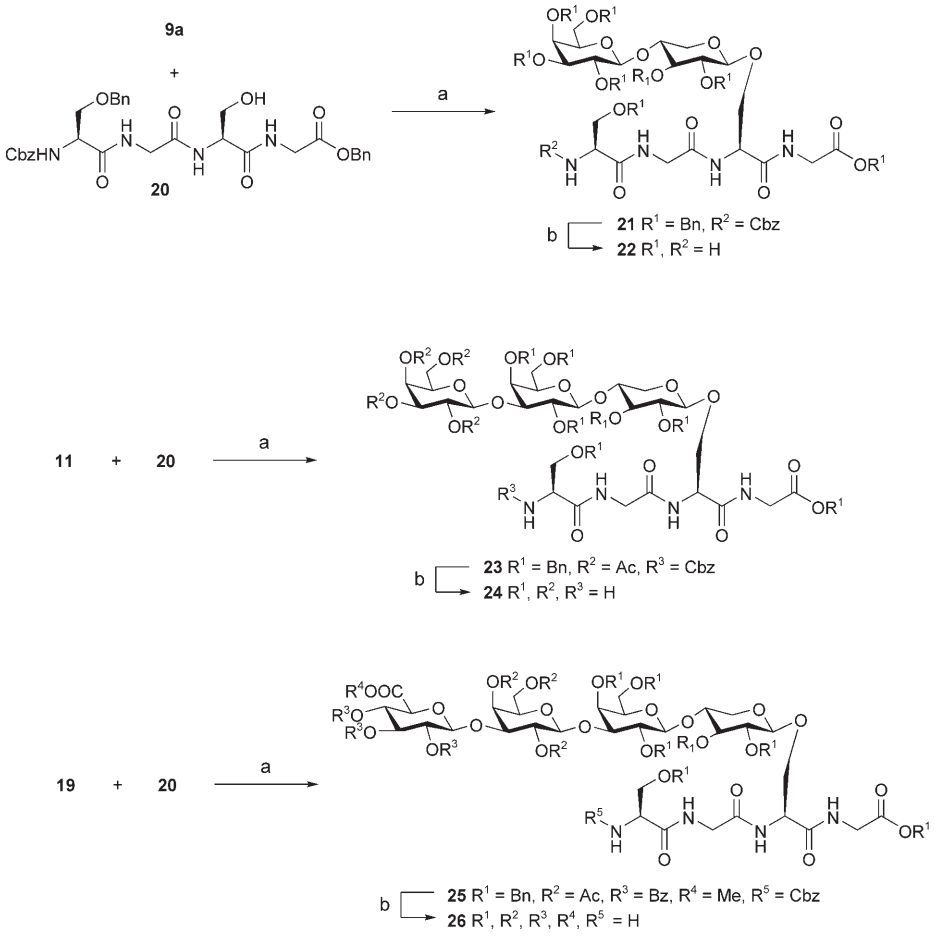
Scheme 2. Synthetic route to the thiophenyl glycosides **11** and **19**: a) TMSOTf, MS (AW300), CH_2Cl_2/CH_3CN; b) 1) DTBS(OTf)$_2$, 2,6-lutidine, CH_2Cl_2; 2) BzCl, pyridine, DMAP; c) [Ir(cod)(PMePh$_2$)$_2$PF$_6$], H_2, THF; then TsOH·H_2O; d) TMSOTf, MS (AW300), CH_2Cl_2/CH_3CN, **15**; e) 1) $(HF)_x$·pyridine, THF; 2) Ac_2O, pyridine; f) 1) CAN, CH_3CN/H_2O; 2) CCl_3CN, DBU. Bz = benzoyl, CAN = ceric ammonium nitrate, DTBS = di-*tert*-butylsilyl, DMAP = 4-dimethylaminopyridine, MP = 4-methoxyphenyl, MS = molecular sieves.

prepared by tosylation and a subsequent substitution reaction with sodium iodide in 87% and 89% yield, respectively. The β elimination of **6a** and **6b** with 1,8-diazabicyclo[5,4,0]undec-7-ene (DBU) gave the unsaturated *exo*-methylene derivatives **7a** (57%) and **7b** (75%), respectively, the treatment of which with O_3 afforded lactones **8a** (100%) and **8b** (90%). The hemiacetals generated by the reduction of **8a** and **8b** in the presence of diisobutylaluminum hydride (DIBAL-H) were subjected to chlorination with oxalyl chloride and subsequent reduction by treatment with sodium cyanoborohydride to afford **9a** in 47% and **9b** in 71% yield (over three steps). Finally, the removal of the silyl protecting group in **9b** by using tetra-*n*-butylammonium fluoride (TBAF) gave the key building block **1** in 98% yield.

Scheme 2 shows the synthetic procedures for the construction of the tri- and tetrasaccharide derivatives **11** and **19**, the glycosyl donors later conjugated with the tetrapeptide derivative **20**. Compound **11** was prepared in 58% yield by the glycosidation of **1** with the known imidate **10** in the presence of TMSOTf.[9] In the case of compound **19**, we had to establish an improved synthetic pathway for the construction of the imidate **18**, the designated glycosyl donor with a GlcAβ(1→3)Gal moiety. Owing to the extremely low reactivity of compound **15** as a result of the electron-withdrawing carboxy group, we chose to prepare the glycosyl acceptor **14** with a cyclic di-*tert*-butylsilylenediyl ether as a protecting group for the hydroxy groups at C4 and C6[10] as reported by Thollas and Jacquinet.[11] As anticipated, the enhanced nucleophilicity of the C3 hydroxy group of **14** due to the conformational stress induced by this protecting group facilitated its reaction with **15** to give the disaccharide derivative **16** in high yield (97%). After the removal of the silyl group and acetylation of the resulting diol (74% over two steps), the replacement of the methoxyphenyl group at the anomeric position with a trichloroacetimidate group afforded the desired glycosyl donor **18** in a satisfactory yield of 56% from **17**. The coupling reaction between **18** and **1** proceeded smoothly in the presence of TMSOTf to give the tetrasaccharide **19** in 66% yield.

The thioglycosides **9a**, **11**, and **19** underwent glycosidation with the tetrapeptide derivative **20** in the presence of trifluoromethanesulfonic acid (TfOH) and *N*-iodosuccimide (NIS)[12] to give the fully protected glycopeptides **21**, **23**, and **25** in 74, 34, and 48% yield, respectively (Scheme 3). Finally, the protecting groups were removed to afford the target glycopeptides **22** (100%), **24** (78%), and **26** (40%). Precise structural characterization of the proteoglycan linkage region was carried out for the first time by analysis by mass spectrometry (ESI and MALDI) and ^1H,^{13}C NMR spectroscopy. Fragmentation by MALDI-LIFT-TOF/TOF mass spectrometry[13] of the precursor ion derived from **26** (*m*/*z* 939.39) in the presence of 2,5-dihydroxybenzoic acid (DHB) occurred both at the glycoside linkages and at the peptide moiety to give meaningful ion peaks generated by the characteristic cleavage patterns represented in Figure 2. The present large-scale synthesis made a systematic study by high-resolution NMR spectroscopy possible. HSQC and TOCSY spectra[14] of **26** acquired at 600 MHz allowed complete assignment of ^1H and ^{13}C chemical shifts (Figure 3).

We tested the suitability of compounds **22** (disaccharide–tetrapeptide: 2s4p), **24** (3s4p), and **26** (4s4p) as proteoglycan initiators (substrates) for the elongation of GAG chains by using normal human articular chondrocyte (nHAC-Kn). The compounds were labeled with fluorescein isothiocyanate (FITC), then incubated with nHAC-Kn (1.0×10^5 cells) at 37 °C for 48 h. It was demonstrated that all synthetic initiators can be incorporated into cultured cells and modified by the



Scheme 3. Coupling of the thioglycosides with the tetrapeptide **20**: a) TfOH, NIS, MS (4 Å), CH_2Cl_2/CH_3CN; b) Pd/C, H_2 for **22**; 1) Pd/C, H_2; 2) NaOMe/MeOH for **24**; 1) Pd/C, H_2; 2) aqueous LiOH; 3) aqueous NaOH/MeOH for **26**.

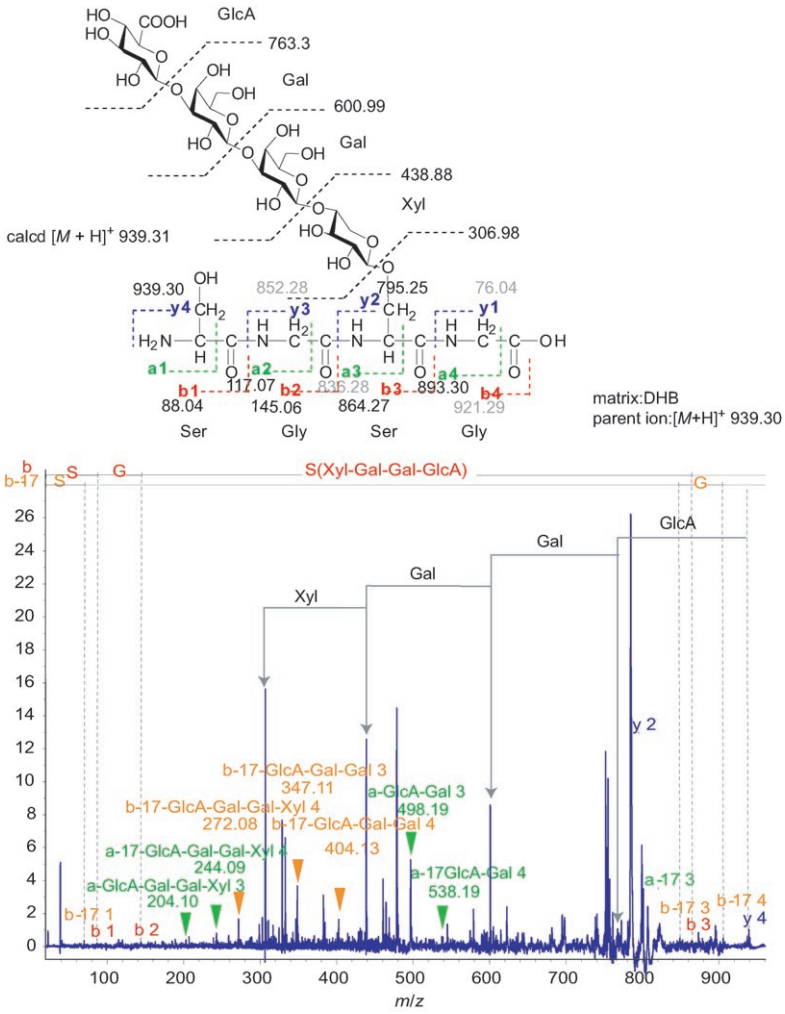


Figure 2. MALDI-LIFT-TOF/TOF MS analysis of the proteoglycan core glycopeptide **26**; b-17-GlcA-Gal-Gal 3 signifies the b3 fragment ion less 17 (NH_3), GlcA, Gal, and Gal.

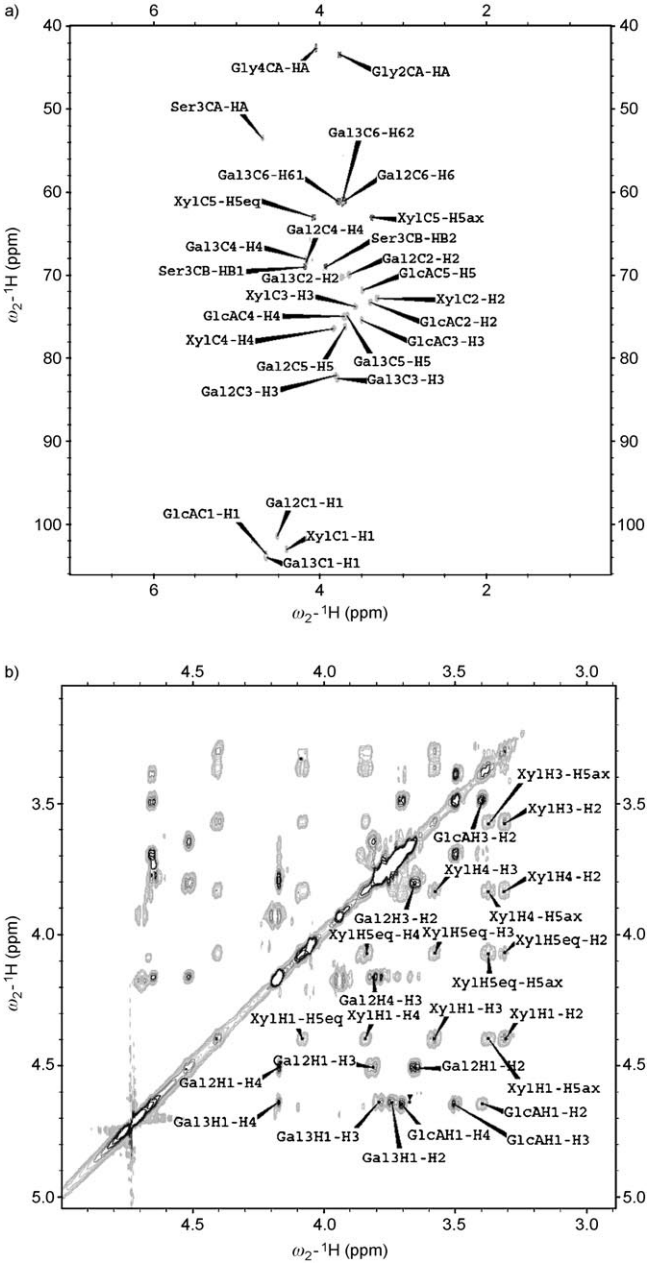


Figure 3. High-resolution NMR spectroscopic analysis of the proteoglycan core glycopeptide **26**: a) ^{13}C-edited HSQC spectrum; b) TOCSY spectrum.

biosynthetic pathways of GAG-chain-extension enzymes (Figure 4). Furthermore, it seems that **22** (2s4p) is a much better substrate than **24** (3s4p) and **26** (4s4p), thus indicating that the carbohydrate structure and/or molecular size of the initiator might strongly influence how well it is accepted by enzymes or the efficiency of its incorporation by nHAC-Kn. To verify this hypothesis, we synthesized tandem-repeating polypeptides from **22** as a macromonomer by direct polycondensation reactions according to the method reported previously (Figure 5).[15] We could conclude clearly that macromolecular initiators from the dimer (MW = 1186.61) to the heptamer (MW = 4145.60) can be used as good substrates for cellular GAG biosynthetic processes and afford a product fraction of high molecular weight estimated at 86 kDa (Figure 5b). These results indicate that the acceptability of synthetic glycopeptide initiators by human chondrocytes depends greatly on the carbohydrate structure in the linker moiety rather than on the molecular weight.

In conclusion, we have demonstrated the versatility of 1,6-anhydro-β-lactose in the rapid, efficient, and large-scale synthesis of important proteoglycan core structures. A key intermediate, phenyl *O*-(2,4,6-tri-*O*-benzyl-β-D-galactosyl)-(1→4)-2,3-di-*O*-benzyl-1-thio-α-D-xylopyranoside (**1**), was synthesized by selective chemical modifications at the 6-position of the reducing Glc residue and the 3′-position of the nonreducing Gal residue in high overall yield. The general synthetic route established herein facilitates the systematic and rapid synthesis of tetrapeptides (H-Ser^1-Gly^2-Ser^3-Gly^4-OH) with di-, tri-, and tetrasaccharide branches, such as Galβ(1→4)Xylβ1→ (in **22**), Galβ(1→3)Galβ(1→4)Xylβ1→ (in **24**), and GlcAβ(1→3)Galβ(1→3)Galβ(1→4)Xylβ1→ (in **26**), at the hydroxy group of the Ser^3 residue. These derivatives, which are related to proteoglycan core structures,

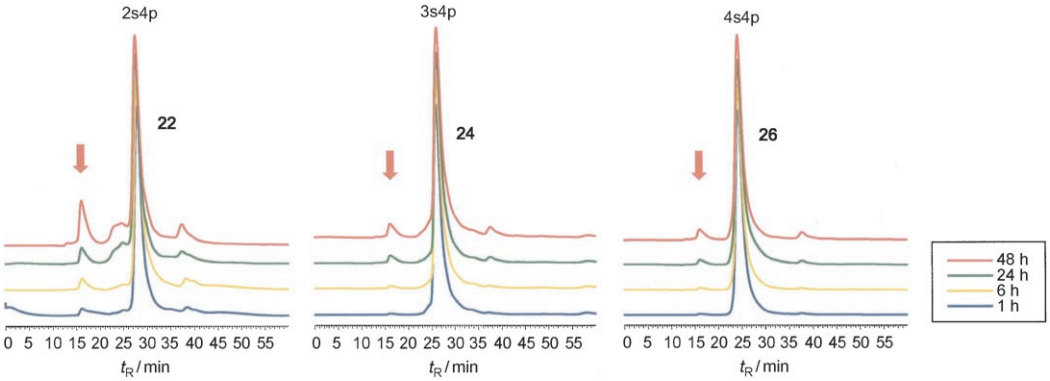


Figure 4. HPLC assay for sugar extension by human chondrocytes with the FITC-labeled initiators **22**, **24**, and **26**. t_R = retention time.

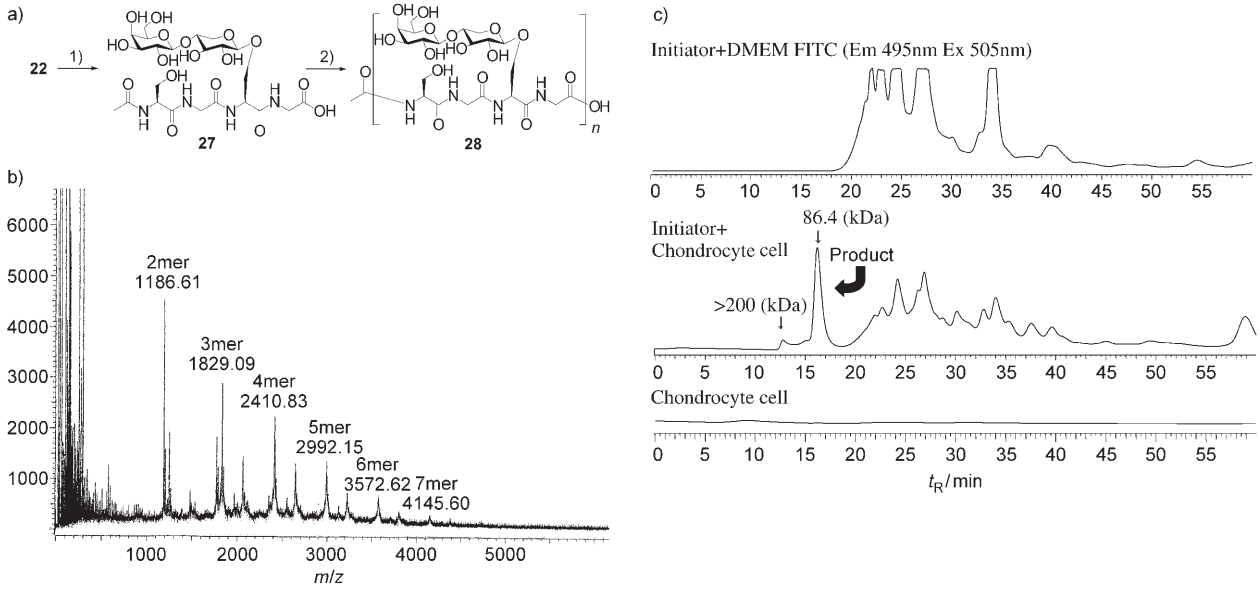


Figure 5. Tandem-repeating glycopeptide **28** derived from **22** and acceptability by human chondrocytes: a) Synthesis of tandem-repeating polypeptides by the diphenylphosphoryl azide (DPPA) method: 1) Ac_2O, DMF; 2) DPPA (1.2 equiv), Et_3N (2.2 equiv), DMF; b) MALDI-TOF mass spectrum showing the molecular weights of the polypeptides derived from **22**; c) HPLC assay of GAG extension by human chondrocytes; polymer **28** was used after conjugation with FITC at the C terminus.

can be used as potent synthetic primers in human cartilage cells to produce novel proteoglycans with mature glycosaminoglycan chains of high molecular weight. It is our belief that the present approach will greatly accelerate the acquisition of fundamental insight into the functional roles of the essential core linker moiety of the glycosaminoglycans, and will be able to be applied to the study of artificial proteoglycans in therapeutic and biomedical fields.

Received: December 5, 2006
Published online: March 15, 2007

Keywords: biosynthesis · carbohydrates · extracellular matrices · glycosaminoglycans · proteoglycans

[1] a) L. Kjellén, U. Lindahl, *Annu. Rev. Biochem.* **1991**, *60*, 443–475; b) A. Yayon, M. Klagsbrun, J. D. Esko, P. Leder, D. M. Ornitz, *Cell* **1991**, *64*, 841–848; c) A. C. Rapraeger, A. Krufka, B. B. Olwin, *Science* **1991**, *252*, 1705–1708; d) D. M. Ornitz, A. B. Herr, M. Nilsson, J. Westman, C. M. Svahn, G. Waksman, *Science* **1995**, *268*, 432–436; e) C. T. Esmon, *J. Biol. Chem.* **1989**, *264*, 4743–4746; f) S. Mizuguchi, T. Uyama, H. Kitagawa, K. H. Nomura, K. Dejima, K. Gengyo-Ando, S. Mitani, K. Sugahara, K. Nomura, *Nature* **2003**, *423*, 443–448.

[2] a) N. R. Krishna, P. K. Agrawal, *Adv. Carbohydr. Chem. Biochem.* **2000**, *56*, 201–234; b) R. M. Lauder, T. N. Huckerby, G. M. Brown, M. T. Bayliss, I. A. Nieduszynski, *Biochem. J.* **2001**, *358*, 523–528; c) S. Yamada, M. Oyama, Y. Yuki, K. Kato, K. Sugahara, *Eur. J. Biochem.* **1995**, *233*, 687–693.

[3] a) B.-T. Kim, H. Kitagawa, J. Tanaka, J.-I. Tamura, K. Sugahara, *J. Biol. Chem.* **2003**, *278*, 41618–41623; b) S. Gulberti, V. Lattard, M. Fondeur, J.-C. Jacquinet, G. Mulliert, P. Netter, J. Magdalou, M. Ouzzine, S. Fournel-Gigleux, *J. Biol. Chem.* **2004**, *280*, 1417–1425.

[4] a) R. Kisilevasky, W. A. Szarek, J. Ancsin, R. Vohra, Z. Li, S. Marone, *J. Mol. Neurosci.* **2004**, *4*, 167–172; b) N. B. Martin, P. Sepulchre, J. Theveniaux, J. Millet, F. Bellamy, *Semin. Thromb. Hemostasis* **1996**, *22*, 247–254; c) M. Belting, L. Borsig, M. M. Fuster, J. R. Brown, L. Persson, L. A. Fronsson, J. D. Esko, *Proc.*

Natl. Acad. Sci. USA **2002**, *99*, 371 – 376; d) S. E. Tully, R. Mabon, C. I. Gama, S. M. Tsai, X. Liu, L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* **2004**, *126*, 7736 – 7737.

[5] a) H. Kunz, H. Waldmann, *Angew. Chem.* **1984**, *96*, 49 – 50; *Angew. Chem. Int. Ed. Engl.* **1984**, *23*, 71 – 72; b) F. Goto, T. Ogawa, *Tetrahedron Lett.* **1992**, *33*, 5099 – 5102; c) M. Nilsson, J. Westman, C.-M. Svahn, *J. Carbohydr. Chem.* **1993**, *12*, 23 – 37; d) F. Goto, T. Ogawa, *Pure Appl. Chem.* **1993**, *65*, 793 – 801; e) S Rio, J.-M. Beau, J.-C. Jacquinet, *Carbohydr. Res.* **1994**, *255*, 103 – 124; f) J. I. Tamura, J. Nishihara, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1911 – 1914; g) J. I. Tamura, J. Nishihara, *J. Org. Chem.* **2001**, *66*, 3074 – 3083.

[6] a) T. Takamura, S. Tejima, *Chem. Pharm. Bull.* **1978**, *26*, 1117 – 1122; b) T. Takamura, T. Chiba, S. Tejima, *Chem. Pharm. Bull.* **1979**, *27*, 721 – 725; c) A. Morikawa, H. Kuzuhara, *J. Carbohydr. Chem.* **1990**, *9*, 167 – 179; d) S. I. Nishimura, S. Murayama, K. Kurita, H. Kuzuhara, *Chem. Lett.* **1992**, 1413 – 1416; e) S. I. Nishimura, K. Matsuoka, T. Furuike, N. Nishi, S. Tokura, K. Nagami, S. Murayama, K. Kurita, *Macromolecules* **1994**, *27*, 157 – 163; f) T. Tsuda, T. Furuike, S. I. Nishimura, *Bull. Chem. Soc. Jpn.* **1996**, *69*, 411 – 416; g) T. Furuike, K. Yamada, T. Ota, K. Monde, S. I. Nishimura, *Tetrahedron* **2003**, *59*, 5105 – 5113.

[7] a) T. Ogawa, T. Nukada, M. Matsui, *Carbohydr. Res.* **1982**, *101*, 263 – 270; b) J. Alais, A. Maranduba, A. Veyrieres, *Tetrahedron Lett.* **1983**, *24*, 2383 – 2386.

[8] L.-X. Wang, N. Sakairi, H. Kuzuhara, *J. Chem. Soc. Perkin Trans. 1* **1990**, 1677 – 1682.

[9] R. R. Schmidt, J. Michel, *Angew. Chem.* **1980**, *92*, 763 – 764; *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 731 – 732.

[10] D. Kumagai, M. Miyazaki, S. I. Nishimura, *Tetrahedron Lett.* **2001**, *42*, 1953 – 1956.

[11] B. Thollas, J. C. Jacquinet, *Org. Biomol. Chem.* **2004**, *2*, 434 – 442.

[12] a) G. H. Veeneman, S. H. van Leeuwen, J. H. van Boon, *Tetrahedron Lett.* **1990**, *31*, 1331 – 1334; b) P. Konradsson, U. E. Udodong, B. Fraser-Reid, *Tetrahedron Lett.* **1990**, *31*, 4313 – 4316.

[13] a) M. Kurogochi, T. Matsushita, S. I. Nishimura, *Angew. Chem.* **2004**, *116*, 4163 – 4167; *Angew. Chem. Int. Ed.* **2004**, *43*, 4071 – 4075; b) M. Kurogochi, S. I. Nishimura, *Anal. Chem.* **2004**, *76*, 6097 – 6101.

[14] a) T. Matsushita, H. Hinou, M. Fumoto, M. Kurogochi, N. Fujitani, H. Shimizu, S. I. Nishimura, *J. Org. Chem.* **2006**, *71*, 3051 – 3060; b) K. Naruchi, T. Hamamoto, M. Kurogochi, H. Hinou, H. Shimizu, T. Matsushita, N. Fujitani, H. Kondo, S. I. Nishimura, *J. Org. Chem.* **2006**, *71*, 9609 – 9621.

[15] a) T. Tsuda, S. I. Nishimura, *Chem. Commun.* **1996**, 2779 – 2780; b) Y. Tachibana, N. Matsubara, F. Nakajima, S. Tsuda, K. Monde, S.-I. Nishimura, *Tetrahedron* **2002**, *58*, 10213 – 10224; c) Y. Tachibana, G. L. Fletcher, N. Fujitani, S. Tsuda, K. Monde, S. I. Nishimura, *Angew. Chem.* **2004**, *116*, 874 – 880; *Angew. Chem. Int. Ed.* **2004**, *43*, 856 – 862.

