J|A|C|S
ARTICLES
Published on Web 11/01/2006

Sugar-Assisted Ligation of N-Linked Glycopeptides with Broad Sequence Tolerance at the Ligation Junction

Ashraf Brik,*,† Simon Ficht, Yu-Ying Yang, Clay S. Bennett, and Chi-Huey Wong*

Contribution from the Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received August 7, 2006; E-mail: abrik@bgu.ac.il; wong@scripps.edu

Abstract: A novel method for the synthesis of N-linked glycopeptides using the sugar-assisted ligation strategy from cysteine free peptides is presented. The ligation junction tolerates a variety of amino acids, favoring less hindered amino acids and those with side chains that could serve as a general base in the ligation pathway. Since our approach allows the ligation of difficult junctions, the method could be applied to the synthesis of large peptides by enzymatic removal of the sugar moiety. Alternatively, more complex glycopeptides can be synthesized using glycosyltransferases. Together, this sequence of reactions should be amenable to the synthesis of glycopeptides and glycoproteins and their deglycosylated products.

Introduction

The chemical synthesis of homogeneous glycopeptides and glycoproteins from readily available materials has gained considerable attention in organic synthesis.[1] This field of research is mainly driven by the fact that glycoproteins are naturally expressed as heterogeneous mixtures of glycoforms, rendering the study of the discrete effects of glycosylation in protein function a difficult task.[2] Chemical synthesis of glycopeptides and glycoproteins based on native chemical ligation (NCL) represents one of the useful methods for the synthesis of homogeneous glycopeptide structures.[3] However, the restriction imposed by the requirement for a cysteine residue at the ligation site poses a significant limitation to this methodology.[4] To overcome this limitation, several removable thiol-based auxiliaries were introduced to fulfill the function of the cysteine side chain in the ligation reaction. Kent and co-workers were the first to extend the applicability of NCL by using a temporary ethanethiol handle attached to the N-terminal peptide.[5] Subsequent to this initial study, peptides with N^{α}-linked auxiliaries based on 1-phenyl-2-mercaptoethyl and 2-mercaptobenzyl have also been investigated and successfully applied to the synthesis of large peptides.[6]

The use of auxiliaries to assist chemical ligation in the context of glycopeptide and glycoprotein synthesis has recently become a major goal for several research groups. Macmillan and co-workers investigated the applicability of 4,5,6-trimethoxy-2-mercaptobenzyl and 1-(2,4-dimethoxyphenyl)-2-mercaptoethyl auxiliaries for a cysteine-free NCL for glycopeptide synthesis.[7] The Danishefsky group successfully used the 4,5,6-trimethoxy-2-mercaptobenzyl auxiliary[6] for the construction of complex glycopeptides.[8] Recently this group reported a similar approach that takes advantage of an O→S followed by S→N migration with the attachment of both coupling partner to a benzylic auxiliary.[9] These elegant examples have demonstrated the power of auxiliaries in a cysteine-free glycopeptide ligation. In these studies, however, only ligation at Gly-Gly and Gly-AA (AA represents unhindered amino acids) junctions delivered the desired products, limiting their scope in the synthesis of proteins and glycoproteins. As a result, the development of auxiliaries that can assist the ligation of cysteine-free glycopeptides with tolerance to the ligation junction would be of great benefit to the field of glycopeptide and glycoprotein synthesis.

We recently described a new approach for the synthesis of β-O-linked glycopeptides lacking a cysteine residue.[10] This method that is based on sugar-assisted ligation (SAL) uses a peptide thioester and a glycopeptide in which *N*-acetylglucosamine (GlcNAc) β-linked to the serine side chain is modified

† Present address: Department of Chemistry, Ben Gurion University, Beer Sheva 84105, Israel.

(1) For selected reviews: (a) Liu, L.; Bennett, C. S.; Wong, C.-H. *Chem. Commun.* **2006**, *1*, 21−33. (b) Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58−68. (c) Davis, B. G. *Chem. Rev.* **2002**, *102*, 579−602. (d) Seitz, O. *ChemBioChem* **2000**, *1*, 214−246.
(2) Helenius, A.; Aebi M. *Science* **2001**, *291*, 2364−2369.
(3) (a) Hackenberger, C. P. R.; Friel, C. T.; Radford, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **2005**, *127*, 12882−12889. (b) T. J. Tolbert, D. Franke, C. H. Wong, *Bioorg. Med. Chem.* **2005**, *13*, 909−915. (c) Macmillan: D.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2004**, *43*, 1355−1359. (d) Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J. *J. Am. Chem. Soc.* 2004, *126*, 6576−6578. (e) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684−11689.
(4) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.
(5) Cane, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891−5896.
(6) (a) Offer, J.; Boddy, C.; Dawson, P. E. *J. Am. Chem. Soc.* **2002**, *124*, 4642−4646. (b) Low, D. W.; Hill, M. G.; Carrasco, M. R.; Kent, S. B. H. Botti, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6554−6559.
(7) Macmillan: D.; Anderson, D. W. *Org. Lett.* **2004**, *6*, 4659−4662.
(8) (a) Wu, B.; Chen, J.; Warren, J. D.; Chen, G.; Hua, Z.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 1−11. (b) Chen, J.; Warren, J. D.; Wu, B.; Chen, G.; Wan, Q.; Danishefsky, S. J. *Tetrahedron Lett.* **2006**, *47*, 1969−1972.
(9) Chen, G.; Warren, J. D.; Chen, J.; Wu, B.; Wan, Q.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2006**, *128*, 7460−7462.
(10) Brik, A.; Ying, Y. Y.; Ficht, S.; Wong, C.-H. *J. Am. Chem. Soc.* **2006**, *128*, 5626−5627.

10.1021/ja065601q CCC: $33.50

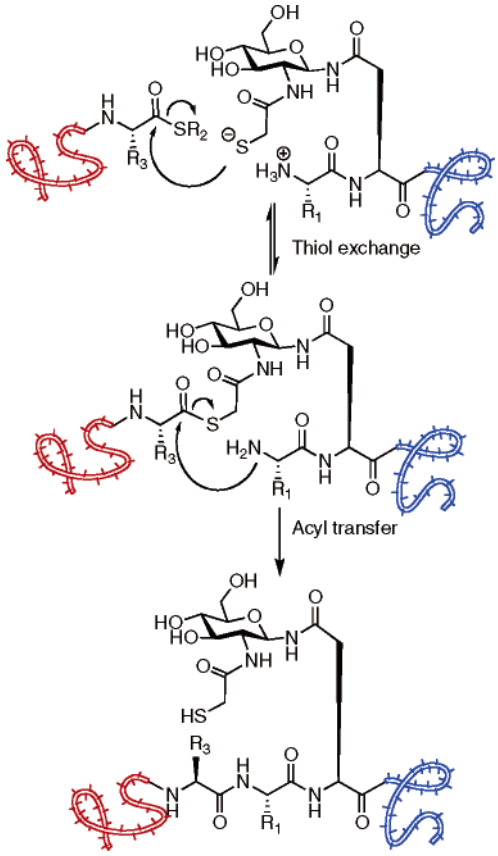
at the C-2 acetamide with a thiol handle to mimic the cysteine function. Upon completion of the ligation reaction, the thiol handle can be reduced with H_2/metal to the acetamide moiety, furnishing the unmodified glycopeptides.[11] Here we report that this new method can be applied to the efficient synthesis of N-linked glycopeptides. Despite the differences between the N- and O-linked glycopeptides, the ligation rates were similar. Importantly, the ligation chemistry tolerates a variety of amino acids at the N-terminal glycopeptide, favoring less hindered amino acids and those with side chains that can serve as a general base in the ligation pathway. Moreover, the modified glycopeptide can be transformed to the corresponding peptide by glycosidases or to a more complex glycopeptide using glycosyltransferases.

Results and Discussion

In addition to the naturally occurring O-linked glycopeptides, N-linked glycosylation, which is a more prevalent form of glycosylation, is found in a wide range of organisms ranging from archaea to mammals and other eukaryotes.[12] Due to the important role of glycoproteins in many biological systems, their preparation has been the goal of many laboratories for some time.[13] Encouraged by our previous results with SAL,[10] we sought to extend this strategy to the synthesis of N-linked glycopeptides, despite the structural variations compared to the O-linked case (Scheme 1). The differences could arise from the change in the ring size of the S→N acyl transfer intermediate, which is a 15-membered ring in the N-linked case versus a 14-membered ring in the O-linked case. More importantly, the reduced degree of rotational freedom around the glycan−amide bond in the N-linked case can affect the proximity of the nucleophilic amine to the thioester intermediate, thus hampering the S→N acyl transfer. On the other hand, N-linked glycopeptides bear the N-acetyl-glucosamine unit as the first glycan moiety attached to the peptide side chain. This ubiquitous connectivity allows us to modify the C-2 acetamide with the necessary thiol handle to assist the ligation.

Building Blocks and Glycopeptides Synthesis. In our previous report, we relied on solution-phase synthesis to prepare the thiol-modified O-linked glycopeptides. In order to make our approach applicable for the synthesis of large glycopeptides and ultimately for glycoproteins, we developed a solid-phase synthesis of the N-linked glycopeptide precursor. We first examined whether the necessary thiol-modified *β-N*-(GlcNAc)-Asn building block **6** can be synthesized efficiently and used in solid-phase peptide synthesis (SPPS). We started from the known glycosyl azide **1** (Scheme 2A),[14] which was reduced with propanedithiol[15] to give glycosyl amine **2** with a complete

Scheme 1. General Strategy of the Sugar-Assisted Chemical Synthesis of N-Linked Glycopeptides (R_1, R_3 = Amino Acid Side Chains, R_2 = Ph/$(CH_2)CONH_2$)

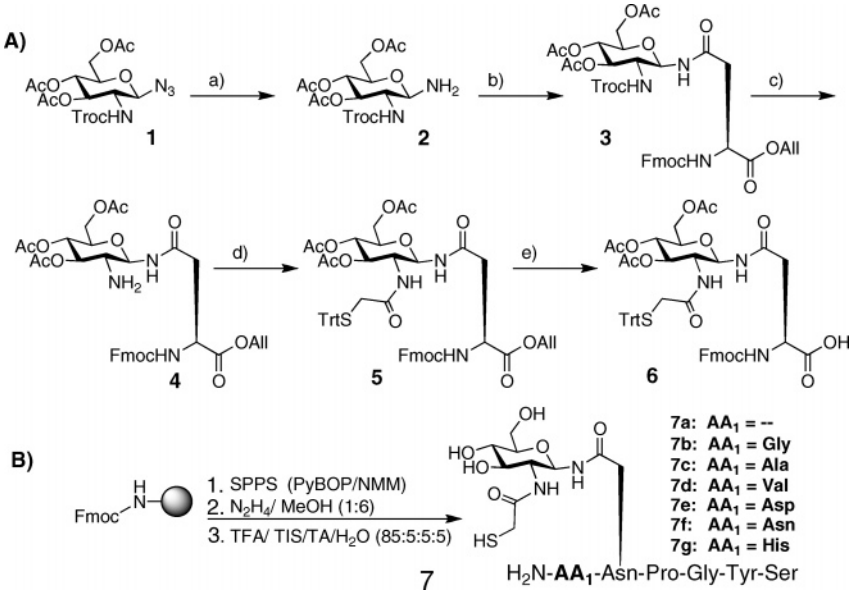


retention of the stereochemistry. The *β*-glycosylamine **2** was then coupled to the side chain of a protected aspartic acid by employing HBTU/DIEA coupling conditions to furnish compound **3**.[16] Following deprotection of the Troc protecting group,[17] the 2-amino functionality of compound **4** was coupled to S-trityl-2-mercaptoacetic acid[18] using HBTU/DIEA coupling conditions, giving the fully protected compound **5**. Finally, the acid functionality was unmasked by employing $Pd(PPh_3)_4$[19] to provide the desired building block **6** in gram quantities.

Glycopeptide synthesis bearing the thiol-modified *β-N*-(GlcNAc) was carried out on Fmoc-Rink amide resin employing routine PyBOP/NMM coupling conditions[16] (Scheme 2B). The coupling of the *β-N*-(GlcNAc)-Asn **6** to the N-terminal proline, which is known to be a difficult coupling, was quantitatively achieved as indicated by UV absorption at 302 nm of the Fmoc/piperidine adduct. Upon completion of coupling of the amino acid **6**, the resin was split into five portions in which four different amino acids (Gly, Ala, Val, Asp) were coupled to the free N-terminal glycopeptide (Scheme 2B). Glycopeptide **7a**, without further elongation, was also prepared to evaluate the effect of amino acid extension on the ligation reaction. Following hydrazine/methanol mixture (1:6) mediated acetate removal, the glycopeptide−resin was exposed to the TFA/TIS/TA/H_2O (85:5:5:5) mixture to release the deprotected glyco-

(11) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526−533.
(12) (a) Lechner, J.; Wieland, F. *Annu. Rev. Biochem.* **1989**, *58*, 173−194. (b) Kornfeld, R.; Kornfeld, S. *Annu. Rev. Biochem.* **1985**, *54*, 631−664.
(13) (a) Hackenberger, C. P. R.; Friel, C. T.; Radford, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **2005**, *127*, 12882−12889. (b) Miller, J. S.; Dudkin, V. Y.; Lyon, G. J.; Muir, T. W.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2003**, *42*, 431−434. (c) Pratt, M. R.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2003**, *125*, 6149−6159. (d) Hojo, H.; Watabe, J.; Nakahara, Y.; Ito, Y.; Nabeshima, K.; Toole, B. P. *Tetrahedron Lett.* **2001**, *42*, 3001−3004. (e) Tolbert, T. J.; Wong, C. H.; *J. Am. Chem. Soc.* **2000**, *122*, 5421−5428. (f) Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, I.; Kawakami, T.; Aimoto, S.; Yamamoto, K.; Inazu, T.; *J. Am. Chem. Soc.* **1999**, *121*, 284−290. (g) Guo, Z. W.; Nakahara, Y.; Ogawa, T. *Angew. Chem., Int. Ed.* **1997**, *36*, 1464−1466.
(14) Matsubara, K.; Mukaiama, T. *Chem. Lett.* **1994**, *15*, 247−250.
(15) Bayley, H.; Standring, D. N.; Knowles, J. R. *Tetrahedron Lett.* **1978**, *39*, 3633−3634.
(16) Albericio, F.; Carpino, L. A. *Methods Enzymol.* **1997**, *289*, 104−126.
(17) Woodward, R. B.; Heusler, K.; Gosteli, J.; Naegeli, P.; Oppolzer, W.; Ramage, R.; Rangamathan S.; Vorbrüggen, H. *J. Am. Chem. Soc.* **1966**, *88*, 852−853.
(18) Kupihar, Z.; Schmel, Z.; Kele, Z.; Penke, B.; Kovacs, L. *Bioorg. Med. Chem.* **2001**, *9*, 1241−1247.
(19) Ciommer, M.; Kunz, H. *Synlett* **1991**, *8*, 593−595.

Scheme 2. (A) Synthesis of the Building Block **6**[a] and (B) SPPS Synthesis of Glycopeptides on Fmoc-Rink Amide Polystyrene Resin



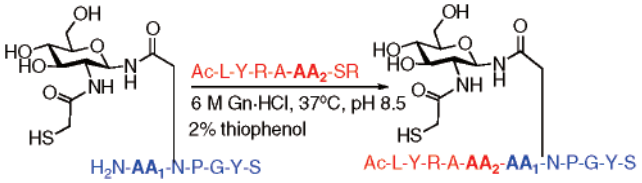
[a] Reagents and conditions: (a) propanedithiol, DIEA, MeOH, 4 h, rt, (b) Fmoc-(Asp)-OAll, HBTU, DIEA, DMF, 12 h, 90%; (c) Zn, AcOH, rt, 12 h, 85%; (d) $TrtSCH_2COOH$, HBTU, DIEA, DMF, 4 h, rt, 78%; (e) $Pd(PPh_3)_4$, NMA, THF, 1 h, rt, 90%.

peptide from the solid support. The crude mixture of each synthesis was purified by preparative HPLC to generate peptides **7a**−**7e** in 55−65% yield (See Supporting Information). To compare the ligation rates of the N- and the O-linked glycopeptides, we used the same thioester peptides bearing Gly, Ala, His, and Val at the C-terminus.[10]

Glycopeptide Ligation. Having the model peptides in hand, we turned our focus on the ligation study. The ligations of the unprotected synthetic peptides were performed by employing conditions similar to NCL (6 M guanidine•HCl, pH 7.5−8.5).[20] Thus, peptides were dissolved at a final concentration of 10−12 mM (1:1.2 molar ratio of peptide thioester to glycopeptide) followed by the addition of 2% thiophenol.[21] The ligation reactions were performed at 37 °C and were vortexed periodically to equilibrate the thiol additive. The progress of each ligation reaction was monitored using analytical HPLC and LCMS.

The rates of the various ligation reactions reported in Table 1 provided several insights into our strategy. As shown in Table 1, the observed ligation rate is dependent on the C-terminal amino acid of the peptide thioester. These results are in agreement with previous observation in the NCL.[20] The presence of a sterically hindered amino acid at the C-terminal thioester peptide leads to a slow reaction relative to the unhindered model. Indeed, in the case of the valine peptide thioester ligation with the glycopeptide **7b** (Table 1, entry 10), we observed the thioester intermediate as the major product by LCMS, in addition to a minor peak that corresponds to the desired product. The use of glycopeptide **7a** without the extension of an additional amino acid adjacent to the asparagine leads to a very slow and incomplete ligation (Table 1, entry 1), similar to the O-linked glycopeptide case. This observation may be to do with the decrease in the ring size of the S→N acyl transfer intermediate from a 15-membered ring to a 12-membered ring in the shorter case. Moreover, the steric hindrance induced by the modification of the side chain with the sugar moiety could contribute to the sluggish rate.

Table 1. Scope of the Sugar-Assisted Ligation



	ligand junction				mass (Da)	
entry	-AA_2-	-AA_1-	-AA_2-AA_1-	t (h)[a]	obsd[b]	calcd
1	Gly	−	Gly	>48	1373.5 ± 0.2	1373.6
2	Gly	Gly	Gly-Gly	∼12	1430.6 ± 0.2	1430.6
3	Gly	Ala	Gly-Ala	∼24	1444.5 ± 0.2	1444.6
4	Gly	Val	Gly-Val	∼48	1472.5 ± 0.2	1472.7
5	Gly	Asp	Gly-Asp	∼12	1488.5 ± 0.2	1488.7
6	Ala	Gly	Ala-Gly	∼40	1444.5 ± 0.2	1444.6
7	Ala	Ala	Ala-Ala	>48	1486.2 ± 0.2	1458.7
8	Ala	Val	Ala-Val	−	1486.2 ± 0.2	1486.3[c]
9	Ala	Asp	Ala-Asp	∼36	1504.2 ± 0.2	1503.6
10	Val	Gly	Val-Gly	−	1472.4 ± 0.2	1472.2[c]
11	His	Gly	His-Gly	∼12	1511.1 ± 0.2	1511.6
12	His	Asp	His-Asp	∼12	1569.2 ± 0.2	1569.7
13	Gly	Asn	Gly-Asn	∼36	1489.1 ± 0.2	1488.6
14	Gly	His	Gly-His	∼6	1511.2 ± 0.2	1511.6
15	His	His	His-His	∼8	1591.3 ± 0.2	1591.7

[a] t represents the time at which the thioester peptide was consumed, with the major peak being the desired product (Figure 1). [b] Characterized by LCMS. [c] After 48 h, the major product was the ligation thioester intermediate with less than 5% of product.

Interestingly, SAL tolerates several different amino acids at N-terminus of the glycopeptide. This can be seen from the examples where glycopeptide **7a** is extended with various amino acids. Despite the changes in the ligation rates going from glycine to alanine and to valine (Table 1, entries 2−4), these reactions proceeded well. In all of these cases the product was

(20) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068−10073.
(21) Dawson, P. E.; Churchill, M.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1997**, *119*, 4325−4329.

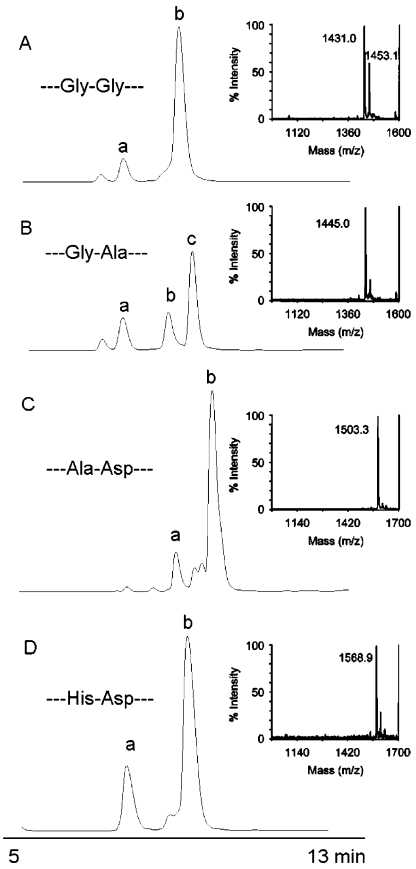


Figure 1. Representative analytical HPLC traces of ligation reactions and mass spectrometry analysis of the products (MALDI-TOF/MS). In all the examples, the thioester substrate was completely consumed. (A) Ligation reaction at 14 h (Table 1, entry 2): peak a, thioester substrate hydrolysis; peak b, ligation product with the expected mass of 1431.0 ± 1 Da. (B) Ligation reaction at 24 h (Table 1, entry 3): peak a, thioester substrate hydrolysis; peak b, ligation intermediate; peak c, ligation product with the expected mass of 1445.0 ± 1 Da, (C) Ligation reaction at 48 h (Table 1, entry 9): peak a, thioester substrate hydrolysis; peak b, ligation product with the expected mass of 1503.3 ± 1 Da. (D) Ligation reaction at 14 h (Table 1, entry 12): peak a, thioester substrate hydrolysis; peak b, ligation product with the expected mass of 1568.9 ± 1 Da.

observed by HPLC analysis as the major peak with a complete consumption of the starting material (Figure 1A,B). Preparative HPLC of the ligation mixture gave the desired glycopeptides in 70% yield (Table 1, entries 2, 3). Notably, ligation at the Ala-Ala junction was also observed and gave the product in 50% yield (Table 1, entry 7). It should be noticed that in our strategy the nucleophilic amine is primary, whereas in the other known auxiliaries it is secondary, rendering it less nucleophilic toward the thioester.[6,7] These differences could account for the increased sequence tolerance at the ligation junction and should be taken into consideration for the design of new ligation auxiliaries.

To our pleasant surprise, we found that in the case of glycopeptide **7e** containing N-terminal aspartic acid (Table 1, entry 5), despite being a hindered amino acid, the ligation rate was similar to that for the glycopeptide bearing a glycine residue (Table 1, entry 2). We hypothesized that the carboxylic acid side chain could serve as a general base in the ligation pathway, thus facilitating the S→N acyl transfer. To test this, we prepared a similar glycopeptide **7f**, in which the N-terminal aspartic acid was replaced with the sterically equivalent asparagine, which lacks the ability to serve as a general base. When we used this glycopeptide in the ligation with the same peptide thioester, the rate was 3-fold slower (Table 1, entry 13) and similar to the rate observed with glycopeptide **7d** bearing the valine extension (Table 1, entry 4). Furthermore, when glycopeptide **7a** was extended with a histidine residue, an amino acid known to serve as a general base in many biological systems, the rate was faster than the case of the glycine-extended glycopeptide ligation (Table 1, entries 14, 15), supporting our general base hypothesis.[22]

In all the ligation reactions presented in Table 1, the thioester intermediate eluted in the HPLC analysis slightly before the rearranged product. In our previous report we were able to identify the thioester intermediate for the O-linked glycopeptide by the mass spectrometry fragmentation pattern;[10] however, in the N-linked glycopeptide we did not observe similar mass fragmentation. In order to differentiate the intermediate from the product, we treated the suspected intermediate with a 10-fold excess of 2-mercaptoethansulfonate (MES) at pH 7.4 and the reaction was followed by LCSM. For example, when the intermediates from entries 8 and 10 (1486.2, 1472.4 Da, respectively) were treated with MES, they were rapidly converted to the corresponding α-MES thioester peptides (758.3 and 744.3 Da, respectively) and the glycopeptide starting material (841.3 and 871.3 Da, respectively).

In NCL, the unrearranged species has never been observed, and it is known that the S→N acyl transfer is extremely rapid and never becomes rate-limiting.[23] In our ligation strategy, in every instance we observed the intermediate in the reaction pathway. The level of its accumulation was clearly dependent on the ligation site. For the very difficult ligation junctions, the major product was the intermediate thioester (Table 1, entries 8, 10). For the faster ones (Table 1, entries 2, 12, 5, 11, 12, 14, 15), the intermediate was formed rapidly (within the first few minutes) and was eventually converted to the product in a rate that is dependent on the ligation junction. Our results agree with the studies on other auxiliaries,[5,6] where the rearrangement of the thioester intermediate is often rate-limiting.

Our study shows that the ligation reaction is more tolerant to the steric hindrance around the nucleophilic amine compared to the thioester elctrophile, as seen by comparing the ligation rates of entries 4 and 10 (Table 1). In both cases the ligation junction is Val-Gly; however, in entry 4 the thioester peptide contains Gly at the C-terminus, while the nucleophilic amine bears the Val residue ($t \sim 40$ h). On the other hand, in entry 10, where the thioester peptide bears Val at the C-terminus while the nucleophilic amine is part of a Gly residue, only the thioester intermediate was observed. Similar results were also obtained with the Gly-Ala junctions (Table 1, entries 3 and 6).

The peptides that are used in these ligations bear the side chains of Asp, Asn, His, Tyr, Ser, Thr, and Arg, in addition to the free hydroxyl groups presented on the glycan core. Our results show that SAL tolerates a variety of functional groups,

(22) It has been reported by Tam and coworkers (Zhang, L.; Tam, J. P. *Tetrahedron Lett.* **1997**, *38*, 3−6.) that the imidazole of an N-terminal histidine-containing peptide assisted in the ligation with a peptide thioester (obtained by activating the thiocaboxylic acid of the C-terminal peptide by Ellman's reagent). The authors proposed a covalently bonded N^{im}-acyl intermediate, without rolling out a mechanism that involves anchimeric assistance of the imidazole moiety.

(23) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640−6646.

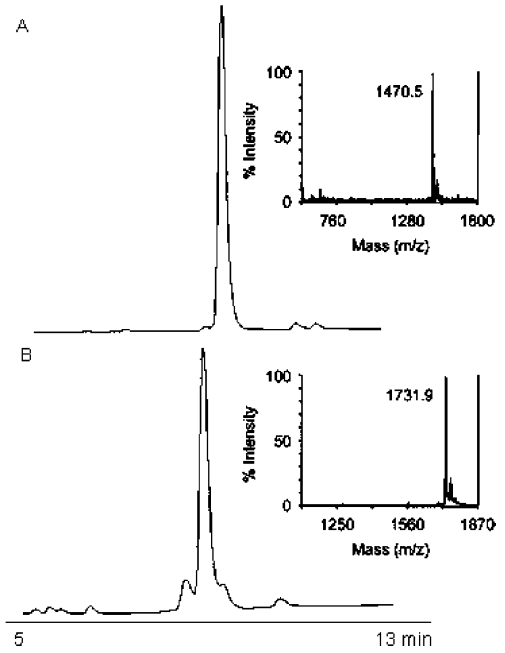


Figure 2. (A) HPLC and MALDI-TOF analysis of the *crude* desulfurization reaction (1 h). A full conversion of the starting material from entry 9 (Table 1) to the desired product with the mass of 1470.5 Da (calcd mass = 1471.5 Da) was observed. (B) HPLC and MALDI-TOF analysis of the *crude* enzymatic transfer reaction (40 h) of galactose to the glycopeptide ligation product (Table 1, entry 12). The desired product with the observed mass of 1731.9 Da (calcd mass = 1731.8 Da) was isolated in 80% yield.

suggesting that SAL is chemoselective reaction. To check for the compatibility of the lysine side chains under the ligation conditions, two thioester peptides bearing the lysine residue at different positions were prepared (Ac-LYRKG-SR, Ac-LYKAG-SR) and tested with glycopeptide **7e**. In both cases, similar results were obtained in terms of rate and yield to the non-lysine-containing peptides (see Supporting Information). It has been reported that a lysine residue at the C-terminal peptide could lead to lactam formation;[24] however, in our study the lactam formation was not observed.

Postligation Modification. Previously, we have shown that the thiol handle can be removed efficiently by using desulfurization conditions (H_2/Pd/Al_2O_3, 6.0 M guanidine containing 100 mM phosphate buffer, pH 5.8, 10 mM TCEP) to furnish the unmodified O-linked glycopeptide (Scheme 3).[10] The desulfurization reaction was first applied by Dawson and co-workers to extend NCL for cysteine free peptides[11] and was nicely used by Kent and co-workers for the synthesis of cysteine-free ubiquitin.[25] In this study, the desulfurization reaction was also successfully applied to the N-linked glycopeptide, and several thiol-modified glycopeptides were desulfurized in up to 90% isolated yield (Figure 2A). Although the removal of the thiol handle by applying the desulfurization reaction could restrict the use of our method to cysteine free peptides, several orthogonal protecting groups can in principle be used to protect the cysteine side chain during the reduction. Work toward performing the ligation reaction followed by desulfurization in the presence of cysteine residues is currently under investigation in our laboratory. Alternatively, the thiol group could also serve as a handle for modification with various alkylating agents such as thiol-reactive dyes for fluorescence labeling (Scheme 3).

Following SAL, the sugar moiety can be cleaved enzymatically or chemically from the asparagine side chain, affording the peptide structure (Scheme 3). In this case the sugar serves as an auxiliary to assist the ligation at difficult junctions that are not accessible with the existing auxiliaries (Table 1, entries 4, 7, 9, 12, 15).[5−8] Preliminary studies were conducted with the ligation product from Table 1, entry 9, and its desulfurized version (1503.6, 1471.3 Da). Both substrates were treated separately with PNGase A, an amidase that cleaves between the innermost GlcNAc and asparagine residues from N-linked glycopeptides and glycoproteins.[26] Our results show that in both reactions the product was formed after 2 h with the mass that corresponds to the nonglycosylated peptide product (1268.3 Da). However, the reaction was not completed after 24 h, providing the product in 30% yield, possibly due to the unnatural glycopeptide sequence that was used in our study.[26]

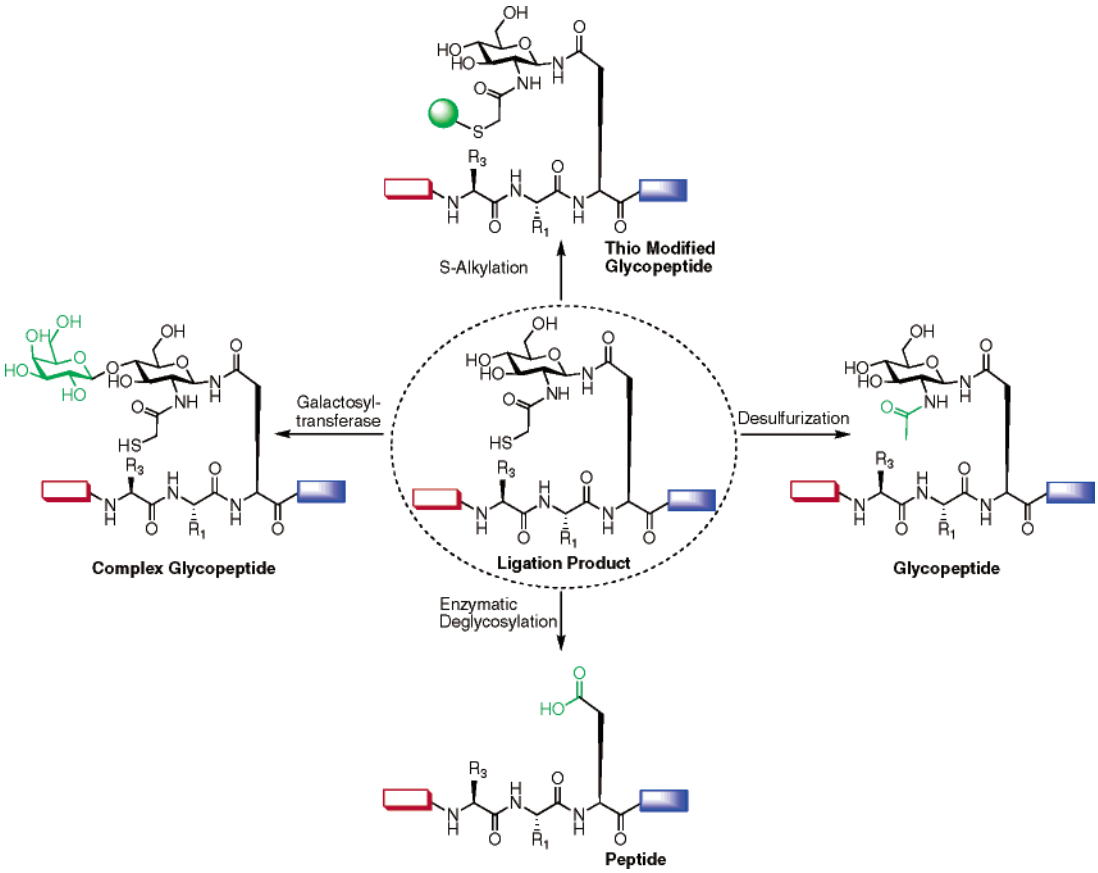
Alternatively, the thiol-containing GlcNAc moiety or the desulfurization product could serve as an acceptor for enzymatic elaboration to generate a more complex glycan (Scheme 3).[27] To this end, we tested the enzymatic transfer of galactose directly into the ligation product using β-1,4-galactosyltransferase (GalT).[28] Although the thiol-modified sugar is not a natural substrate for the transferase,[29] we observed a full conversion to the disaccharide glycopeptide when the ligation product (Table 1, entries 3, 12) was treated with GalT and UDP-galactose (Figure 2B). The reaction mixture was treated with tris(2-carboxyethyl)phosphine hydrochloride (TCEP) to reduce disulfide bonds formed during the enzymatic reaction followed by product isolation using preparative HPLC (81% yield). Transferring the sugar onto the ligation product before the thiol handle removal allows further elaboration of the glycan moiety through chemical dervitization. Further investigations into the use of SAL in the synthesis of glycoproteins and neoglycoconjugates are currently underway in our laboratory.

Conclusions

We have demonstrated that SAL can be applied to N-linked glycopeptide ligation. Despite the changes in the atom connectivity and the ring size of the S→N acyl transfer intermediate between N- and O-linked glycopeptides, the ligation reactions proceeded successfully and in comparable rates. Notably, the ligation junction tolerates a variety of sequences favoring less hindered amino acids and those with side chains that can serve as a general base in the ligation pathway. Moreover, since our approach afforded ligation at difficult junctions, this method can be applied to the synthesis of large peptides by subsequent enzymatic removal of the sugar moiety. Alternatively, by using GalT we were able to introduce galactose directly to the ligation product, furnishing the disaccharide glycopeptide. Together,

(24) de Koning, M. C.; Filippov, D. V.; Overhand, M; van der Marel G. A.; van Boom J. H. *Eur. J. Org. Chem.* **2004**, 850−857.
(25) Bang, D.; Makhatadze, G. I.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B. *Angew. Chem., Int. Ed.* **2005**, *44*, 3852 −3856.
(26) Fan, J. Q.; Lee, Y. C. *J. Biol. Chem.* **1997**, *272*, 27058−27064.
(27) For, recent reviews discussing the use of enzymes to elaborate glycoproteins, see: Gorgan, M. J.; Pratt, M. R.; Maracaurelle, L. A.; Bertozzi, C. A. *Annu. Rev. Biochem.* **2002**, *71*, 593−634.
(28) (a) Schanbacher, F. L.; Ebner, K. E. *J. Biol. Chem.* **1970**, *245*, 5057−5061. (b) Beyer, T. A.; Sadler, J. E.; Rearick, J. I.; Paulson, J. C.; Hill, R. L. *Adv. Enzymol.* **1981**, *52*, 23−175.
(29) For a previous study on the effect of substitution at C-2 of the GlcNAc acceptor on GalT activity, see: Berliner, L. J.; Davis, M. E.; Ebner, K. E.; Beyer, T. A.; Bell, J. E. *Mol. Cell. Biochem.* **1984**, *62*, 37−42.

Scheme 3. Possible Modifications of the Ligation Product



these sequences of reactions hold a great potential in the synthesis of glycopeptides and glycoproteins.

Experimental Procedures

General Methods. All chemical reagents were purchased from Aldrich or Acros and used without further purification. Amino acids and coupling reagent were purchased from Novabiochem. Tetrahydrofuran (THF) was distilled over sodium/benzophenone, and methylene chloride (CH_2Cl_2) was distilled over calcium chloride. Analytical thin-layer chromatography (TLC) was performed using silica gel 60 F_{254} glass plates, and compound spots were visualized by UV light (254 nm) and by staining with citric ammonium molybdate. Flash chromatography was performed on silica Sila-P flash silica gel (40−63 μm, Silicycle, Quebec, Canada). ^{1}H and ^{13}C NMR spectra were recorded on a Bruker AMX 500-MHz spectrometer. Coupling constants (J) are reported in hertz. All ^{1}H chemical shifts are reported in δ referenced to solvent. LCMS (Agilent 1100 LC coupled to an Agilent 1100 single quadrapole mass spectrometer, column is an Agilent SB C8 50 × 4.6 mm) and/or reverse-phase HPLC analysis were performed on a Vydac C18 column using a linear gradient of buffer A in buffer B over 30 min (buffer A, 0.1% trifuoroaetic acid (TFA) in water; buffer B, 0.1% TFA in acetonitrile).

Compound 2. Propanedithiol (24 mmol, 2.6 mL) was added to a degassed mixture of compound **1** (2.0 g, 4.0 mmol) and Et_3N (24.0 mmol, 3.4 mL) in methanol (20 mL). The reaction was allowed to stir at room temperature for 6 h, at which time it was concentrated, resulting in glycosylamine **2**, which was placed under high vacuum for 8 h and used directly for the next step. ^{1}H NMR ($CDCl_3$, 500 MHz): δ 5.11 (dd, J = 9.6, 9.6 Hz, 1H,), 5.06 (dd, J = 9.6, 9.6 Hz, 1H), 5.83 (d, J = 11.8 Hz, 1H), 4.63 (d, J = 12.1 Hz, 1H), 4.24 (dd, J = 12.5, 5.2 Hz, 1H), 4.17 (d, J = 9.2 Hz, 1H), 4.11 (dd, J = 12.1, 1.9 Hz, 1H), 3.64−3.68 (m, 2H), 2.10 (s, 3H), 2.04 (s, 3H), 2.03 (s, 3H). ^{13}C NMR ($CDCl_3$, 500 MHz): δ 171.11, 170.71, 169.42, 154.89, 95.36, 86.18, 74.55, 72.93, 72.84, 68.52, 62.35, 57.18, 20.77, 20.62, 20.60. HRMS (ESI-TOF) calcd for $C_{15}H_{21}Cl_3N_2O_9$ $[M + H]^+$: 479.0385. Found: 479.0375.

Compound 3. Compound **2** (2.0 g, 4.0 mmol) was added to a preactivated Fmoc-Asp-OAll (2.48 mg, 6.20 mmol) with HBTU (2.4 g, 6.2 mmol) and DIEA (2.0 mL, 12 mmol) in dry DMF (20 mL) at 0 °C. The reaction mixture was stirred at room temperature for 12 h and then diluted with 80 mL of ethyl acetate and washed with water and brine. The organic layer was dried over $MgSO_4$ and concentrated for flash column chromatography ($CHCl_3$/MeOH 10:1) to give 3.2 g of **3** (90%). ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.74 (d, J = 7.7 Hz, 2H), 7.60 (d, J = 7.4 Hz, 2H), 7.23−7.39 (m, 5H), 6.21 (dd, J = 8.8, 4.1 Hz, 1H), 5.87 (m, 1H), 5.19−5.31 (m, 4H), 5.08 (dd, J = 9.6, 9.6 Hz, 1H,), 4.76 (d, J = 12.1 Hz, 1H), 4.69 (d, J = 12.4 Hz, 1H), 4.62−4.66 (m, 2H), 4.42 (dd, J = 10.3, 7.0 Hz, 1H), 4.27−4.33 (m, 2H), 4.05 (d, J = 12.1 Hz, 1H), 3.80−3.90 (m, 2H), 2.97 (m, 1H), 2.81 (m, 1H), 2.01−2.03 (2s, 9H). ^{13}C NMR ($CDCl_3$, 500 MHz): δ 170.70, 170.66, 170.61, 170.44, 169.34, 165.55, 162.58, 156.03, 155.44, 143.68, 143.53, 141.02, 131.44, 127.51, 126.88, 125.05, 124.99, 119.78, 118.36, 95.24, 79.06, 74.35, 73.28, 72.55, 68.04, 67.11, 66.03, 54.93, 50.31, 46.86, 44.70, 37.47, 36.35, 31.32, 20.49, 20.43, 20.39. HRMS (ESI-TOF) calcd for $C_{37}H_{40}Cl_3N_3O_{14}$ $[M + H]^+$: 856.1649. Found: 856.1645.

Compound 4. Compound **3** (3.2 g) was dissolved in 20 mL of acetic acid followed by the addition of 3.0 g of zinc dust. The reaction mixture was stirred at room temperature for 8 h. After filtration the solvent was removed under reduced pressure. The residue was purified using flash column chromatography [ethyl acetate (EA)/hexanes (Hex), 3:1] to give 2.6 g of **4** (88%). ^{1}H NMR (DMSO-d_6, 500 MHz): δ 8.62 (d, J = 8.8 Hz, 1H), 7.88 (d, J = 7.7 Hz, 2H), 7.81 (d, J = 8.1 Hz, 1H), 7.69 (d, J = 7.4 Hz, 2H), 7.31−7.42 (m, 4H), 5.87 (m, 1H), 5.29 (d, J = 17.2 Hz, 1H), 5.17 (d, 1H, 10.7), 4.93−5.01 (m, 2H), 4.77 (dd, J = 9.9, 9.9 Hz, 1H), 4.49−4.60 (m, 3H), 4.21−4.35 (m, 3H), 4.15 (dd, J = 12.5, 4.0, 1H), 3.92 (d, J = 11.7 Hz, 1H), 3.82−3.84 (m, 1H),

2.68−2.75 (m, 1H), 2.61 (dd, J = 15.4, 7.7, 1H), 1.98−1.95 (3s, 9H). ^{13}C NMR (DMSO-d_6, 500 MHz): δ 171.26, 170.03, 169.87, 169.41, 169.35, 156.82, 143.75, 143.73, 140.71, 140.69, 132.30, 127.64, 127.10, 125.23, 125.19, 120.11, 117.54, 80.70, 75.36, 72.08, 68.64, 65.80, 64.99, 61.99, 54.93, 50.42, 46.55, 38.23, 36.93, 20.70, 20.52, 20.43. HRMS (ESI-TOF) calcd for $C_{34}H_{39}N_3O_{12}$ $[M + H]^+$: 682.2606. Found: 682.2583.

Compound 5. Compound **4** (2.0 g, 2.8 mmol) was added to preactivated *S*-trityl-2-mercaptoactic acid (1.86, 5.6 mmol) with HBTU (2.10 g, 5.6 mmol) and DIEA (2.0 mL, 11.2 mmoL) in dry DMF (20.0 mL). The reaction mixture was stirred at room temperature for 4 h, diluted with 80 mL of ethyl acetate, and washed with water and brine. The organic layer was dried over $MgSO_4$ and concentrated for flash column chromatography (EA/Hex 1:2) to give 2.0 g of **5** (89%). ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.74 (d, J = 7.4 Hz, 2H), 7.58 (d, J = 5.5 Hz, 2H), 7.18−7.39 (m, 22H), 6.03 (d, J = 8.4 Hz, 1H), 5.77−5.81 (m, 2H), 5.24 (d, J = 17.3 Hz, 1H), 5.14 (d, J = 10.3 Hz, 1H), 5.03 (dd, J = 9.5, 9.5 Hz, 1H), 4.93 (dd, J = 10.3, 10.3 Hz, 1H), 4.80 (dd, J = 8.8, 8.8 Hz, 1H), 4.54−4.64 (m, 2H), 4.38−4.51 (m, 2H), 4.20−4.35 (m, 3H), 4.04 (d, J = 12.1 Hz, 1H), 3.90−3.94 (m, 1H), 3.69 (d, J = 8.1 Hz, 1H), 2.84−2.89 (m, 1H), 2.64−2.69 (m, 1H), 2.04 (s, 3H), 2.01 (s, 3H), 1.95 (s, 3H). ^{13}C NMR ($CDCl_3$, 500 MHz): δ 171.51, 171.05, 170.75, 170.58, 169.21, 156.02, 143.82, 141.20, 141.17, 131.52, 129.34, 128.26, 127.62, 127.10, 126.99, 125.17, 125.10, 119.89, 118.33, 80.23, 73.39, 71.87, 68.00, 67.70, 6.17, 66.04, 61.54, 53.07, 50.34, 47.02, 37.47, 35.86, 20.65, 20.49. HRMS (ESI-TOF) calcd for $C_{55}H_{55}N_3O_{13}S$ $[M + Na]^+$: 1020.3348. Found: 1020.3351.

Compound 6. Compound **5** (2.0 g, 2.0 mmol) was suspended in THF (20 mL), and *N*-methylaniline (2.2 mL, 20 mmol) and $(PPh_3)_4Pd$ (231 mg, 0.2 mmol) were added subsequently. The reaction mixture was stirred at room temperature for 1 h. After removing the solvent under reduced pressure, the residue was subjected to column chromatography (MeOH/CH_3Cl 9:1) to give the product in 90% yield (1.8 g, 90%). ^{1}H NMR (MeOD + $CDCl_3$ (v/v = 1/4), 500 MHz): δ 7.72−7.77 (m, 2H), 7.55−7.64 (m, 3H), 7.50−7.55 (m, 1H), 7.34−7.41 (m, 3H), 7.24−7.31 (m, 7H), 7.18−7.23 (m, 5H), 7.11−7.15 (m, 2H), 5.15 (dd, J = 10.3, 9.6 Hz, 2H), 5.10 (d, J = 9.9 Hz, 1H), 4.96 (dd, J = 9.6, 9.6 Hz, 2H), 4.49 (m, 1H), 4.21−4.26 (m, 3H), 4.12 (dd, J = 7.4, 7.4 Hz, 1H), 4.00−4.07 (m, 3H), 3.78−3.81 (m, 1H), 1.99 (s, 3H), 1.98 (s, 3H), 1.95 (s, 3H). ^{13}C NMR (MeOD + $CDCl_3$ (v/v = 1/4), 500 MHz): δ 172.16, 171.65, 171.06, 145.10, 144.93, 144.77, 142.23, 142.21, 133.52, 133.49, 132.87, 132.79, 130.33, 129.76, 129.66, 129.57, 128.88, 128.56, 128.52, 127.94, 127.81, 127.79, 126.15, 126.08, 120.67, 79.28, 78.98, 74.32, 74.15, 69.59, 68.06, 67.96, 62.94, 53.61, 48.01, 38.89, 37.45, 20.75, 20.98. HRMS (ESI-TOF) calcd for $C_{52}H_{51}N_3O_{13}S$ $[M + Na]^+$: 980.3035. Found: 980.3032.

Glycopeptide Synthesis. Solid-phase chemistry was carried out in syringes equipped with Teflon filters, purchased from Torviq. If not differently described, all reactions were carried out at room temperature. Preparative HPLC was performed on a Hitachi (D-7000 HPLC system) instrument using a preparative column (Grace Vydac "Protein & Peptide" C18) and a flow rate of 9 mL/min. DMF was purchased from Aldrich in biotech grade. Commercial reagents were used without further purification. Resins, protected amino acids, and PyBOP were purchased from Novabiochem.

Preloading of the Rink Amide Resin. The resin (0.69 mmol/g) was washed (5× DCM, 5× DMF), followed by removal of the Fmoc group by treating it with 10% piperidine/DMF (2 × 5 min) and another washing step (5× DMF, 5× DCM, 5× DMF). For preactivation of the first protected amino acid, 1 equiv of PyBOP and 2 equiv of NMM were added to a solution of the building block (0.1 M) in DMF. After 5 min of preactivation, the mixture was added to the resin. After 2 h the resin was washed (5× DMF, 5× DCM, 5× DMF), capped with acetic anhydride/pyridine (1:9) (2 × 5 min), and washed (5× DMF, 5× DCM, 5× DMF).

Glycopeptide Solid-Phase Synthesis According to the Fmoc-Strategy: **Fmoc Cleavage.** After treatment with 10% piperidine/DMF (2 × 5 min) the resin was washed (5× DMF, 5× DCM, 5× DMF).

Coupling. After preactivation of 4 equiv of protected amino acid (final concentration 0.1 M in DMF) for 5 min using 4 equiv of PyBOP and 8 equiv of NMM, the solution was added to the resin. After 30 min, the resin was washed with DMF (5×), DCM (5×), and DMF (5×).

Capping. Acetic anhydride/pyridine (1:9) was added to the resin. After 5 min the resin was washed with DMF (5×), DCM (5×), and DMF (5×).

Coupling of the Sugar-Containing Monomer. After preactivation of 4 equiv of building block **6** (final concentration 0.1 M in DMF) for 5 min using 4 equiv of PyBOP and 8 equiv of NMM, the solution was added to the resin. After 12 h, the resin was washed with DMF (5×), DCM (5×), and DMF (5×).

Cleavage. 1. Removal of the Acetyl Protecting Groups of the Sugar. The resin was washed with MeOH (10×), treated with MeOH/hydrazine 6:1 (2 × 2.5 h), and washed (10× MeOH, 5× DMF, 10× DCM).

2. Cleavage from the Resin. A mixture of TFA, thioanisole, triisopropylsilane, and water (17:1:1:1) was added. After 2 h, the resin was washed with TFA (4 × 4 mL).

Workup. The combined solutions were concentrated in vacuo. The residue was dissolved in water, purified by preparative HPLC, and analyzed by MALDI-TOF/MS (matrix α-cyano-4-hydroxycinnamic acid).

General Procedure for Chemical Ligation. The ligation of unprotected glycopeptide segments was performed as follows: 0.1 M phosphate buffer (pH 8.5) containing 6 M guanidine was degassed with argon for 20 min before use. Peptide thioester and glycopeptide were dissolved at a final concentration of 10 mM followed by the addtion of 2% thiophenol. The ligation reaction was performed in a heating block at 37 °C and was vortexed periodically to equilibrate the thiol additive. Before analysis, TCEP (50 mM) was added to reduce any disulfide bonds. The reaction was monitored using LCMS and/or reverse-phase HPLC analysis performed on a Vydac C18 column using a linear gradient (95−60%) of buffer A in buffer B over 30 min.

General Procedure for Desulfurization. Desulfurization reactions were performed in a 0.1 M phosphate buffer containing 6.0 M guanidine at pH 5.8 and 10 mM TCEP at room temperature. The buffer was degassed by bubbling argon through for 20 min before each use. Pd/Al_2O_3 was added (10−20 times the weight of peptide) and the reaction was kept under hydrogen. The desulfurization reaction was monitored by analytical HPLC as described above.

Enzymatic Removal of the Glycan Moiety with *N*-Glycosidase A. A total of 0.6 nmol of substrate (Table 1, entry 9, and its desulfurized version, 1503.6 and 1471.3 Da) dissolved in 40 μL of pH 8.0, 50 mM sodium phosphate buffer was mixed with *N*-glycosidase A (10 U, purchased from New England Biolabs) and incubated at 37 °C. For LCMS analysis, 2.5 μL of reaction solution was taken from the Eppendorf tube and diluted into 35 μL of water.

Enzymatic Transfer. Glycopeptide product from entry 12, Table 1 (1.0 mg, 0.63 μmol) and UDP-Gal (0.6 mg, 1.0 μmol) were dissolved in 110 μL of HEPES buffer (130 mM, pH 7.4 + 0.25% Triton X-100) containing freshly prepared $MnCl_2$ solution (0.95 μmol). β-1,4-GalT (9.4 μL, 5 mU) and alkaline phosphatase (0.04 μL, 30 mU) were added, and the reaction was shaken gently at ambient temperature for 40 h. The reaction was then incubated with an excess of TCEP for 30 min and purified by HPLC employing a gradient of 100:0 A:B to 10:90 A:B over 30 min at a flow rate of 1.5 mL min^{-1}, where A = 0.1% TFA in H_2O and B = 0.1% TFA in MeCN. The retention time of the desired product was 9.16 min. Lyophilization of the desired fractions afforded pure disaccharide glycopeptide in 80% yield. Similar results were obtained with the ligation product from entry 3, Table 1 (calcd mass, 1605.7 Da; observed, 1606.4 Da).

Acknowledgment. This work was supported by the NIH and Skaggs Institute for Chemical Biology. S.F. is grateful for a postdoctoral fellowship provided by the Deutsche Akademische Austauschdienst (DAAD). C.S.B. is grateful for the Ruth L. Kirschstein NRSA postdoctoral fellowship provided by the NIH (GM073500-02).

Supporting Information Available: Experimental procedures and characterization of the products and other detailed results. This material is available free of charge via the Internet at http://pubs.acs.org.

JA065601Q