Glycoproteins

DOI: 10.1002/anie.200503900

Glycosylation of a Neoglycoprotein by Using Glycosynthase and Thioglycoligase Approaches: The Generation of a Thioglycoprotein**

*Johannes Müllegger, Hong Ming Chen, R. Antony J. Warren, and Stephen G. Withers**

Glycan structures, both intracellular and exposed on the surface of cells, play important roles in a large number of biological processes, such as cell–cell recognition, ligand–receptor binding, immunomodulation, and protein folding.[1,2] Elucidation of the biological function of glycoproteins is currently hampered mainly by the fact that they are not easily available as homogenous isoforms but exist as complex mixtures of different glycoforms.[3,4] As of yet, there are no generally useful strategies for the direct glycosylation of a protein to generate a natural glycoprotein: the strategies of Wong, Schultz, and co-workers are technically impressive but not yet applicable to the general production of glycoproteins on any scale.[5,6] Strategies therefore involve partial, specific deglycosylation of existing glycoproteins to create more homogenous populations of the glycoprotein, followed by chemoenzymatic remodeling.[7,8] These steps are required as the synthetic chemical assembly of larger glycostructures is essentially impossible on native proteins owing to the harsh conditions required. Two general strategies are to use glycosyl transferases[9] or the transglycosylation activity of glycosidases.[10] Only modest yields of shorter glycopeptides have generally been obtained by using the transglycosylation activity of *endo* glycosidases, although one very recent publication reports the high-yielding transglycosylation of a core N-glycan structure by using an activated oxazoline-donor sugar and a synthetic glycopeptide containing one N-linked β-*N*-acetylglycosaminyl residue as an acceptor.[11]

Glycosynthases and thioglycoligases are two new classes of mutant glycosidases recently developed for oligosaccharide synthesis, but they have not yet been applied to glycoprotein synthesis.[12,13] Indeed, the relatively high K_M values for their acceptors has caused concern that they would not be useful in this role. The goal of this study was to explore the applicability of these two classes of enzymes for the synthesis of glycoproteins by first using a neoglycoprotein acceptor as a model system to simplify the analysis. A particular goal was to explore the generation of glycoproteins in which the terminal sugar is linked through a glycosidase-resistant thioglycosidic bond and determine if such thioglycoproteins are recognized by other glycan-modifying enzymes. Through appropriate applications of such technology, it should be possible to generate therapeutic glycoproteins in which the terminal sugar is sulfur-linked. Such proteins should enjoy considerably longer serum half-lives.

The *endo* xylanase from *Bacillus circulans* (Bcx) was chosen as the protein base for a model neoglycoprotein for several reasons. First, it is a small, highly characterized protein of approximately the same size as cytokines; for some cytokines glycosylation is extremely important both for activity and serum half-life.[14] Second, as it is an enzyme, the effects of chemical and enzymatic manipulations on overall structural integrity can easily be followed by activity-based assays. Third, the three-dimensional structure of Bcx has been solved by X-ray crystallography and extensive NMR spectroscopy assignments.[15,16] Furthermore, the native protein contains no cysteines, thus unique thiol groups can be introduced at the surface by in vitro mutagenesis. This allows site-selective modification of Bcx with sugar entities through thiol-reactive reagents.[17] The sugar structures selected for conjugation were cellobiose and its 4′-thio analogue because these sugars function as good acceptors for *Agrobacterium* sp. (Abg) glycosynthase and Abg thioglycoligase, respectively. These are the enzymes that were used as glycosylation catalysts.

Thiol-reactive cellobioside **1** was synthesized by reduction and bromoacetylation of *p*-nitrophenyl cellobioside (see Supporting Information). After reaction with Bcx S22C for 1 h, analysis with MALDI-TOF MS showed complete labeling with a mass increase equivalent to one incorporated label (+474.2 Da) resulting in neoglycoprotein G2-Bcx (Figure 1A, B). This neoglycoprotein was then tested as an acceptor for Abg 2F6 glycosynthase[18] by using α-galactopyranosyl fluoride (α-GalF) as a donor and monitoring the reaction with MALDI-TOF MS. Complete conversion of the G2-Bcx into a glycoprotein bearing three sugar units (named LacG-Bcx) was observed after overnight incubation (Figure 1C, Figure 2G, Scheme 1A). No incorporation was observed in control reactions that used Bcx wild-type (wt) or the unlabeled mutant Bcx S22C as acceptors. Likewise, no reaction occurred upon incubation of G2-Bcx with only α-GalF over extended periods of time. The time course for this reaction showed conversion of $\geq 80\,\%$ in 200 minutes (see Supporting Information) when a molar ratio of 19:1 for acceptor/glycosynthase was used.

[*] Dr. J. Müllegger, Dr. H. M. Chen, Prof. Dr. S. G. Withers
Protein Engineering Network of Centres of Excellence
Department of Chemistry
University of British Columbia
Vancouver, B. C. V6T 1Z1 (Canada)
Fax: (+1) 604-822-8869
E-mail: withers@chem.ubc.ca
Dr. R. A. J. Warren
Protein Engineering Network of Centres of Excellence
Department of Microbiology and Immunology
University of British Columbia
Vancouver, B. C. V6T 1Z1 (Canada)

[**] The authors would like to thank Prof. L. P. McIntosh and Prof. W. W. Wakarchuk for Bcx clones, Dr. Y. W. Kim for Abg 2F6 glycosynthase, and L. Lairson for glycosyl transferases, as well as the proteomics core facility of UBC for providing assistance with, and access to, the mass spectrometer. The project was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Protein Engineering Network of Centres of Excellence of Canada (PENCE). J.M. was financed by a Schrödinger Fellowship from the Austrian FWF.

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

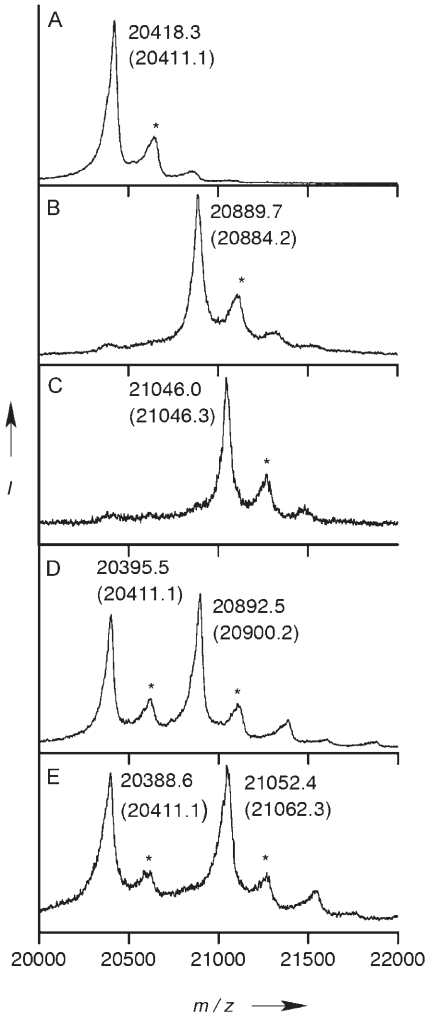


Figure 1. MALDI-TOF MS spectra of Bcx S22C before chemical glycosylation (A), G2-Bcx (B), LacG-Bcx (C), 4SG2-Bcx (D), and GalSG-Bcx (E). Average measured masses for the protein and calculated masses (in brackets) are shown. Asterisks (*) denote artifact peaks from sinapinic acid (SPA) matrix addition to the protein (≈206–224 Da). *I* = intensity.

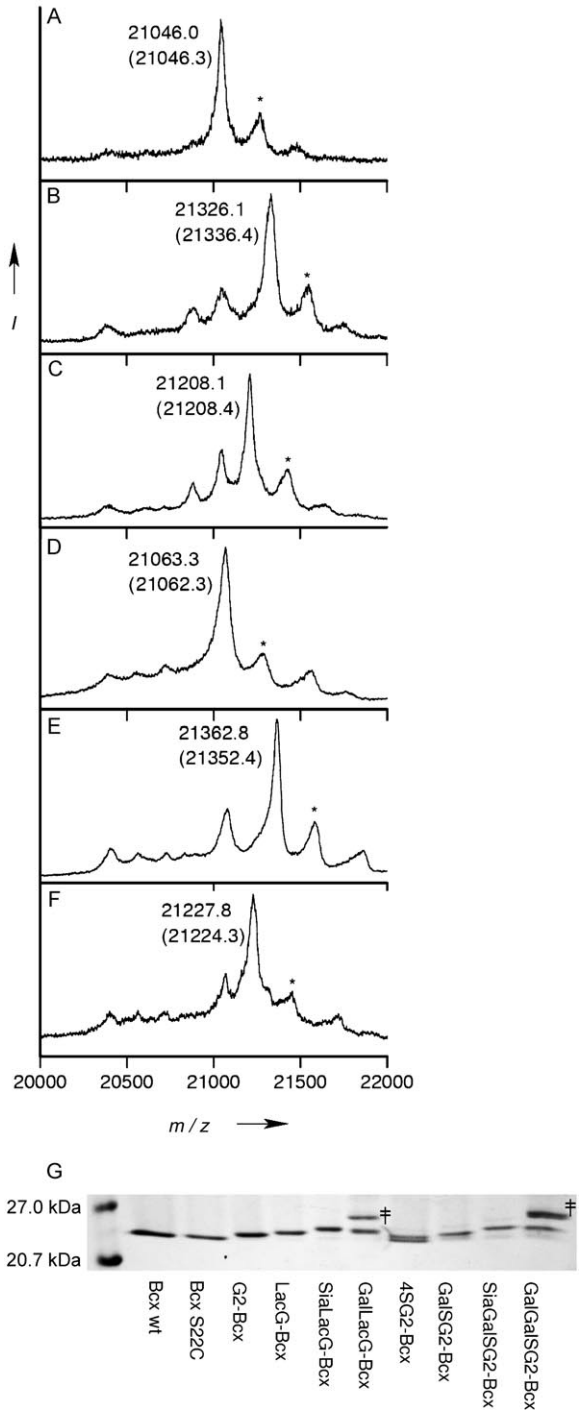
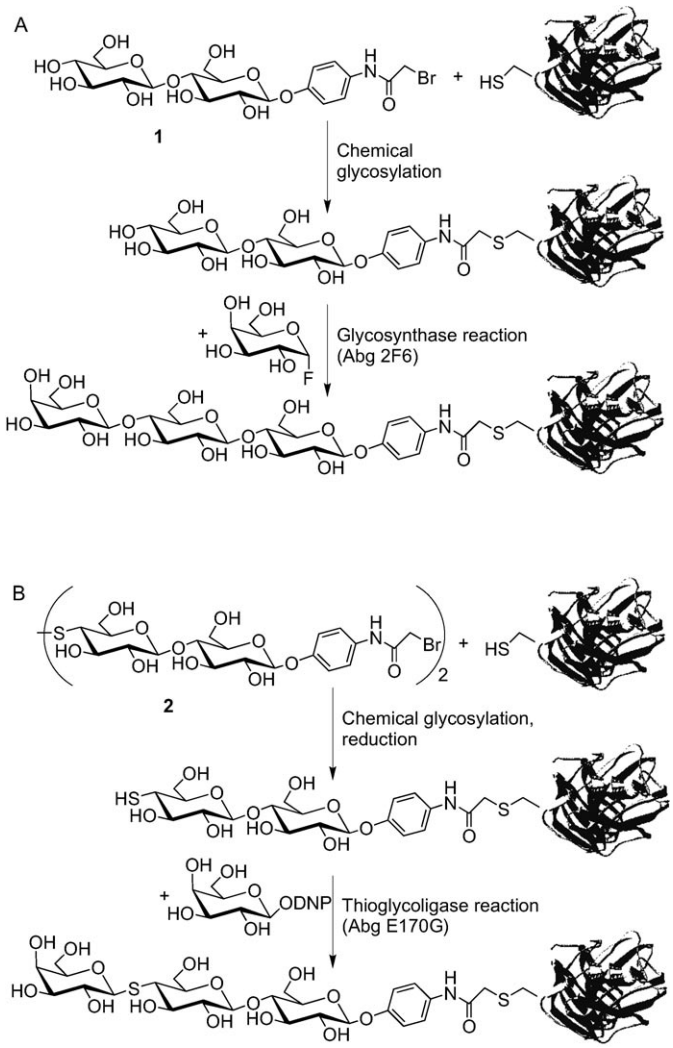


Figure 2. MALDI-TOF MS spectra of LacG-Bcx (A), and the products of glycosylation by Cst1 (B) and LgtC (C). Panels (D, E, F) show the respective spectra for GalSG2-Bcx. Asterisks (*) denote artifact peaks from SPA matrix addition to the protein (≈206–224 Da). Average masses measured for the protein and calculated masses (in brackets) are shown. (G) SDS-PAGE analysis of all Bcx neoglycoprotein variants. Small shifts in migration are observed, but more importantly the complete conversions achieved in the glycosynthase and thioglycoligase reactions are demonstrated. ‡ indicates protein contamination introduced with LgtC. *I* = intensity.

Two sets of experiments were undertaken to confirm the type of linkage formed by the glycosynthase on the surface of the neoglycoprotein. Cleavage of the terminal galactose by jack-bean β-galactosidase (β-Gal) and both the galactose and the glucose by β-glucosidase from Abg (bifunctional β-galactosidase/glucosidase), as shown by MALDI-TOF MS (see Supporting Information), clearly shows formation of the expected β linkage.[19] Confirmation of the terminal lactose structure on the neoglycoprotein was provided by the action of two glycosyl transferases with high specificities for lactose as an acceptor: LgtC, an α-1,4-galactosyl transferase from *Neisseria meningitides*,[20] and Cst1, an α-2,3-sialyltransferase from *Campylobacter jejuni*.[21] Both enzymes used LacG-Bcx as an acceptor, resulting in the formation of the Pk antigen (Gal α-1,4-lactose) and sialyllactose epitopes, respectively, on the surface of Bcx (Figure 2B, C). Finally, to confirm that all manipulations were carried out on a properly folded protein and that the glycosylation did not affect this conformation, the xylanase activities of Bcx wt, Bcx S22C, and LacG-Bcx were measured with 2,5-dinitrophenyl xylobioside as the



Scheme 1. Reaction scheme for the chemical glycosylation and subsequent transfer of a galactose by Abg glycosynthase (A) or Abg thioglycoligase (B).

substrate. No significant differences in k_{cat} or K_M from literature values were observed,[22] confirming the integrity of the protein (see Supporting Information).

Parallel studies to test the thioglycoligase approach required incorporation of a 4′-deoxy-4′-thiocellobiosyl moiety onto the neoglycoprotein. This was achieved by incubating Bcx S22C with an excess of **2** (used in its disulfide form to minimize self-polymerization) overnight, followed by reduction with dithiothreitol (DTT). MALDI-TOF MS analysis revealed about 50% incorporation to create thioneoglycoprotein Bcx-4SG2 (Figure 1D). This mixture of labeled and unlabeled protein was directly used as an acceptor for the Abg thioglycoligases, with the unreacted Bcx S22C acting as a useful internal control. Three different thioglycoligase mutants, namely Abg E170A, Abg E170G, and Abg E170Q, were tested by using 2,4-dinitrophenyl β-D-glucopyranoside (DNPGlc) and 2,5-dinitrophenyl β-D-galactopyranoside (DNPGal) as donors. Glycosylation of Bcx-4SG2 was observed by MALDI-TOF MS when each of the thioglycoligase mutants were used with DNPGlc (creating GSG2-Bcx), whereas only Abg E170G was active when using DNPGal as a donor (creating GalSG2-Bcx; Figure 1E and Figure 2G). This specificity is in full accordance with earlier observations.[23] Control reactions carried out with Bcx wt and the "internal" control Bcx S22C showed no signs of glycosylation, suggesting specific attachment of the hexose unit to the thioglycoside part of the thio-neoglycoprotein. The presence of a thioglycosidic linkage was established in three ways. First, the mixture of GalSG2-Bcx and Bcx S22C was treated with the thiol-reactive compound (+)-biotinyl-iodoacetamidyl-3,6-dioxaoctanediamine (Pierce), and the reaction mixture analyzed by MALDI-TOF MS. The mass of Bcx S22C, but not that of GalSG2-Bcx, increased by the mass of the biotin label (+415.5 Da), confirming the absence of a free thiol in the latter case and thus selective glycosylation of the 4-thio residue on the thio-neoglycoprotein. Further, GalSG2-Bcx was then easily separated from the contaminating Bcx S22C by capturing the biotin-labeled protein on streptavidin beads (Pierce). Second, GalSG2-Bcx was glycosylated by LgtC and Cst1, with the thio-neoglycoprotein proving to be a good acceptor for both transferases, as shown by MALDI-TOF MS (Figure 2D,E,F). This confirms that a thiolactose structure is present and that the presence of sulfur does not affect the action of the transferase. Third, the formation of a glycosidase-resistant thioglycoside linkage by the thioglycoligase was confirmed by subjecting the samples of GalSG2-Bcx to digestion with jack-bean β-Gal and Abg.[24,25] In both cases cleavage was not observed by MALDI-TOF MS. Thus, not only is a thioglycoprotein produced for the first time, but it has also proven to be resistant to glycosidase digestion.

In summary, we demonstrate unambiguously the utility of the glycosynthase and thioglycoligase approaches for the remodeling of glycoproteins, and describe the first generation of a thioglycosylated glycoprotein. This work thereby complements the recent demonstration of the synthesis of thioglycoside-containing glycolipids by using the very low inherent activities of a few glycosyltransferases to catalyze thioglycoside formation.[26,27] The resistance of this thioglycoprotein, but not of its oxygen analogue, to glycosidase digestion and the further glycosylation of both structures by glycosyl transferases directly demonstrates the potential of this approach for generating metabolically stable glycoprotein structures. Work is currently underway to extend this strategy with other glycosynthases and thioglycoligases to other glycosylated structures, particularly those which terminate in a thiosialoside, and to install these in natural glycoproteins.

Received: November 4, 2005
Revised: January 21, 2006
Published online: March 10, 2006

Keywords: enzyme models · glycoproteins · glycosylation · thioglycoligases · thioglycoproteins

[1] P. Van den Steen, P. M. Rudd, R. A. Dwek, G. Opdenakker, *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 151–208.

[2] A. Helenius, M. Aebi, *Science* **2001**, *291*, 2364 – 2369.
[3] J. F. Nemeth, G. P. Hochgesang, Jr., L. J. Marnett, R. M. Caprioli, *Biochemistry* **2001**, *40*, 3109 – 3116.
[4] C. Neususs, U. Demelbauer, M. Pelzing, *Electrophoresis* **2005**, *26*, 1442 – 1450.
[5] R. Xu, S. R. Hanson, Z. Zhang, Y. Y. Yang, P. G. Schultz, C. H. Wong, *J. Am. Chem. Soc.* **2004**, *126*, 15654 – 15655.
[6] Z. Zhang, J. Gildersleeve, Y. Y. Yang, R. Xu, J. A. Loo, S. Uryu, C. H. Wong, P. G. Schultz, *Science* **2004**, *303*, 371 – 373.
[7] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579 – 602.
[8] M. J. Grogan, M. R. Pratt, L. A. Marcaurelle, C. R. Bertozzi, *Annu. Rev. Biochem.* **2002**, *71*, 593 – 634.
[9] K. Witte, P. Sears, R. Martin, C. H. Wong, *J. Am. Chem. Soc.* **1997**, *119*, 2114 – 2118.
[10] K. Yamamoto, S. Kadowaki, J. Watanabe, H. Kumagai, *Biochem. Biophys. Res. Commun.* **1994**, *203*, 244 – 252.
[11] B. Li, Y. Zeng, S. Hauser, H. Song, L. X. Wang, *J. Am. Chem. Soc.* **2005**, *127*, 9692 – 9693.
[12] M. Jahn, S. G. Withers, *Biocatal. Biotransform.* **2003**, *21*, 159 – 166.
[13] G. Perugino, A. Trincone, M. Rossi, M. Moracci, *Trends Biotechnol.* **2004**, *22*, 31 – 37.
[14] L. C. Wasley, G. Timony, P. Murtha, J. Stoudemire, A. J. Dorner, J. Caro, M. Krieger, R. Kaufman, *Blood* **1991**, *77*, 2624 – 2632.
[15] W. W. Wakarchuk, R. L. Campbell, W. L. Sung, J. Davoodi, M. Yaguchi, *Protein Sci.* **1994**, *3*, 467 – 475.
[16] L. A. Plesniak, W. W. Wakarchuk, L. P. McIntosh, *Protein Sci.* **1996**, *5*, 1118 – 1135.
[17] D. Macmillan, R. M. Bill, K. A. Sage, D. Fern, S. L. Flitsch, *Chem. Biol.* **2001**, *8*, 133 – 145.
[18] Y. W. Kim, S. S. Lee, R. A. Warren, S. G. Withers, *J. Biol. Chem.* **2004**, *279*, 42787 – 42793.
[19] L. F. Mackenzie, Q. P. Wang, R. A. J. Warren, S. G. Withers, *J. Am. Chem. Soc.* **1998**, *120*, 5583 – 5584.
[20] L. L. Lairson, C. P. Chiu, H. D. Ly, S. He, W. W. Wakarchuk, N. C. Strynadka, S. G. Withers, *J. Biol. Chem.* **2004**, *279*, 28339 – 28344.
[21] M. Gilbert, J. R. Brisson, M. F. Karwaski, J. Michniewicz, A. M. Cunningham, Y. Wu, N. Young, W. Wakarchuk, *J. Biol. Chem.* **2000**, *275*, 3896 – 3906.
[22] L. Ziser, I. Setyawati, S. G. Withers, *Carbohydr. Res.* **1995**, *274*, 137 – 153.
[23] J. Müllegger, M. Jahn, H. M. Chen, R. A. J. Warren, S. G. Withers, *Protein Eng. Des. Sel.* **2005**, *18*, 33 – 40.
[24] M. Jahn, J. Marles, R. A. J. Warren, S. G. Withers, *Angew. Chem.* **2003**, *115*, 366 – 368; *Angew. Chem. Int. Ed.* **2003**, *42*, 352 – 354; .
[25] C. Orgeret, E. Seillier, C. Gautier, J. Defaye, H. Driguez, *Carbohydr. Res.* **1992**, *224*, 29 – 40.
[26] J. Rich, W. Wakarchuk, D. Bundle, *Chemistry* **2005**, *12*, 845 – 848.
[27] J. Rich, A. Szpacenko, M. Palcic, D. Bundle, *Angew. Chem.* **2004**, *116*, 623 – 625; *Angew. Chem. Int. Ed.* **2004**, *43*, 613 – 615.

