ORGANIC LETTERS

2002
Vol. 4, No. 9
1467−1470

Solid-Phase Synthesis of Thioether-Linked Glycopeptide Mimics for Application to Glycoprotein Semisynthesis

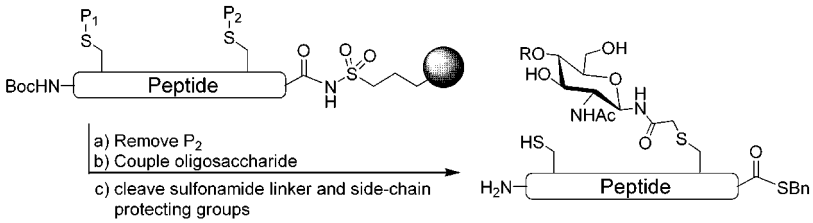
Derek Macmillan,* Alison M. Daines, Monika Bayrhuber, and Sabine L. Flitsch*

Department of Chemistry, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh EH9 3JJ, Scotland

derek.macmillan@ed.ac.uk

Received January 26, 2002

ABSTRACT



Glycoproteins are particularly suited to protein semisynthesis since homogeneous samples for biological analyses are not readily available using traditional recombinant techniques. Here we apply glycosyl iodoacetamides, normally used for the modification of bacterially derived proteins, to solid-phase glycopeptide synthesis. This provides access to glycopeptide α-thioesters, which may lend themselves to the semisynthesis of homogeneous *N*-linked glycoprotein mimics and novel glycopeptide libraries.

It is well established that protein glycosylation can be a vital co- and post-translational modification for the normal growth and development of organisms. This is compounded by the knowledge that the glycan moiety of glycoproteins is implicated in an increasing number of important biological processes from correct protein folding and secretion to cell−cell recognition.[1] Understanding how and why these highly specific interactions come about has always been hindered by the difficulties associated with the "microheterogeneity" of glycoproteins.[2]

As a result of the problems associated with the synthesis of glycopeptides and glycoproteins, chemists have sought novel methods for the synthesis of simpler structural analogues (glycopeptide mimics) that have been employed successfully in various glycobiological studies.[3] The main aim of such an approach is to simplify the synthesis relative to the natural products. However, some additional advantageous properties such as resistance to hydrolysis may also be conferred while remaining functionally silent.

The first report on the use of glycosyl iodoacetamides such as **1** (Figure 1) for the synthesis of *N*-linked glycoprotein mimics[4] was soon followed by a report discussing modification of short unprotected synthetic peptides sequences containing a single cysteine residue in solution.[5] More

* Email for S.L.F.: s.flitsch@ed.ac.uk.

(1) Parodi, A. J. *Annu. Rev. Biochem.* **2000**, *69*, 69. Varki, A. *Glycobiology* **1993**, *3*, 97. Rudd, P. M.; Wormald, M. R.; Stanfield, R. L.; Huang, M. D.; Mattsson, N.; Speir, J. A.; DiGennaro, J. A.; Fetrow, J. S.; Dwek, R. A.; Wilson, I. A. *J. Mol. Biol.* **1999**, *293*, 351. Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364. Frigerio, L.; Lord, J. M. *Curr. Biol.* **2000**, *10*, R674. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.

(2) Rush, R. S.; Derby, P. L.; Smith, D. M.; Merry, C.; Rogers, G.; Rohde, M. F.; Katta, V. *Anal. Chem.* **1995**, *67*, 1442.

(3) Reviewed in: Marcaurelle, L. A.; Bertozzi, C. R. *Chem. Eur. J.* **1999**, *5*, 1384. Davis, B. G.; Jones, J. B. *Synlett* **1999**, 1495. See also: Allen, J. R.; Harris, C. R.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2001**, *123*, 1890. Marcaurelle, L. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2001**, *123*, 1587. Schafer, A.; Klich, G.; Schreiber, M.; Paulsen, H.; Thiem, J. *Carbohydr. Res.* **1998**, *313*, 107. Wang, L. X.; Tang, M.; Suzuki, T.; Kitajima, K.; Inoue, Y.; Inoue, S.; Fan, J. Q.; Lee, Y. C. *J. Am. Chem. Soc.* **1997**, *119*, 11137.

(4) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32*, 6793.

10.1021/ol025627w CCC: $22.00
Published on Web 04/09/2002

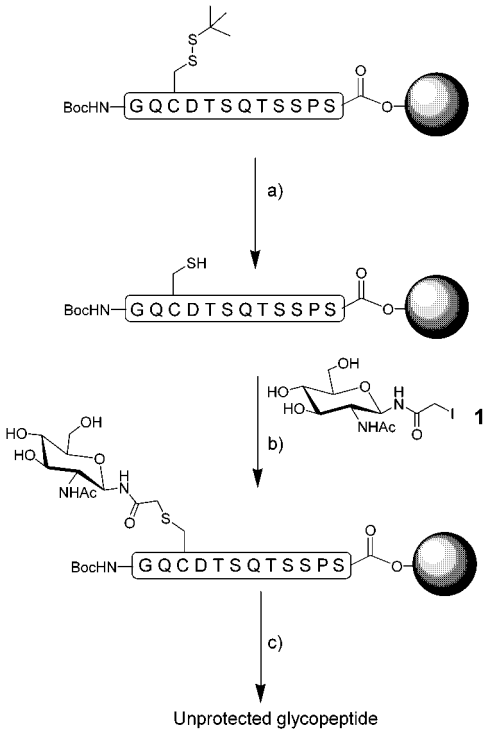
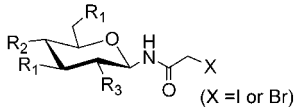


Figure 1. Model CD52 study. (a) 10% DTT sat. NH_4HCO_3/DMF, 24 h, quant. (b) Glycosyl iodoacetamide (3 equiv), DMF, 2.5% v/v pyridine, 3 h, 95−100%. (c) 95% TFA, 3 h.

recently this and closely related methodologies have provided the first examples of homogeneously glycosylated, bacterially derived proteins. In such cases, the site of modification has been specifically targeted using a mutagenesis/modification approach. In this way specific protein glycoforms in the 20−25 KDa molecular weight range have been prepared.[6] The iodoacetamide methodology may, however, be limited by obstacles, including incomplete modification of free thiols on recombinant proteins, which then require further separation. Furthermore, the introduction of extra cysteine residues into proteins may not be wholly tolerated, for example, where the introduction of more than one cysteine residue into a recombinantly derived protein is desirable, yet the protein already contains many other cysteine residues normally involved in disulfide bond formation. These potential limitations have driven the search for a means of expanding the present methodology.

Methods for protein semisynthesis, particularly native chemical ligation[7] and expressed protein ligation[8] are particularly suited to glycoprotein synthesis since they enable

Table 1. Examples of Glycopeptide Mimics Formed Using Solid-Phase Chemistry Depicted in Figure 1 and "thio"-CD52 as the Peptide Scaffold[a]

saccharide haloacetamide	glyco-peptide	calc. m/z	obs. m/z
			
1 R_1= R_2 = OH, R_3 = NHAc	**6**	1457.9	1457.9
UDP-Galactose, β1,4 –Gal T			
"R_1=OH,R_2=β-Gal,R_3= NHAc"	**7**	1619.9	1619.8
2 R_1= R_2 =R_3 = OH	**8**	1416.7	1416.6
3 R_1= R_2= OAc, R_3= NHAc	**9**	1583.6	1583.6
4 R_1= OH, R_2= β-GlcNAc, R_3= NHAc	**10**	1660.9	1661.1
5 R_1=OH, R_2= β-Gal, R_3= OH	**11**	1578.9	1578.7

[a] Enzymatic transformations can also be carried out on the glycopeptide products.

the synthesis of peptides containing post-translational modifications such as phosphorylation[9] and glycosylation[10] that are not under direct genetic control. We therefore aimed to widen the scope of the original iodoacetamide method to include protein semisynthesis. Through chemical synthesis, it should be trivial to differentiate between cysteines involved in disulfide bonds and those to be modified by glycosylation using orthogonally protected cysteine residues. We also aimed to conduct all reactions and glycopeptide synthesis on the solid phase using 4-sulfamylbutyryl-AM resin, allowing glycopeptide mimetic α-thioesters to be prepared using standard Fmoc chemistry.[11] This facilitated application

(5) Wong, S. Y. C.; Guile, G. R.; Dwek, R. A.; Arsequell, G. *Biochem. J.* **1994**, *300*, 843.
(6) Bill, R. M.; Winter, P. C.; McHale, C. M.; Hodges, V. M.; Elder, G. E.; Caley, J.; Flitsch, S. L.; Bicknell, R.; Lappin, T. R. J. *Biochim. Biophys. Acta-Gene Struct. Express.* **1995**, *1261*, 35. Davis, B. G. *J. Chem. Soc., Perkin Trans. 1* **1999**, 3215. Davis, B. G.; Lloyd, R. C.; Jones, J. B. *Bioorg. Med. Chem.* **2000**, *8*, 1527. Davis, B. G. *Chem. Commun.* **2001**, 351. Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. *Chem. Biol.* **2001**, *8*, 133.
(7) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923.
(8) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705. Noren, C. J.; Wang, J. M.; Perler, F. B. *Angew. Chem., Int. Ed.* **2000**, *39*, 451.
(9) Huse, M.; Holford, M. N.; Kuriyan, J.; Muir, T. W. *J. Am. Chem. Soc.* **2000**, *122*, 8337. Flavell, R. R.; Huse, M.; Goger, M.; Trester-Zedlitz, M.; Kuriyan, J.; Muir, T. W. *Org. Lett.* **2002**, *4*, 165.
(10) 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. Winans, K. A.; King, D. S.; Rao, V. R.; Bertozzi, C. R. *Biochemistry* **1999**, *38*, 11700. Marcaurelle, L. A.; Mizoue, L. S.; Wilken, J.; Oldham, L.; Kent, S. B. H.; Handel, T. M.; Bertozzi, C. R. *Chem. Eur. J.* **2001**, *7*, 1129. Tolbert, T. J.; Wong, C.-H. *J. Am. Chem. Soc.* **2000**, *122*, 5421. Macmillan, D.; Bertozzi, C. R. *Tetrahedron* **2000**, *56*, 9515.
(11) Backes, B. J.; Ellman, J. A. *J. Org. Chem.* **1999**, *64*, 2322. 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.
(12) Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495.
(13) Diekman, A. B.; Norton, E. J.; Klotz, K. L.; Westbrook, V. A.; Herr, J. C. *Immunol. Rev.* **1999**, *171*, 203.
(14) For peptide synthesis, see: Macmillan, D.; Bertozzi, C. R. *Tetrahedron* **2000**, *56*, 9515.
(15) St Hilaire, P. M.; Meldal, M. *Angew. Chem., Int. Ed.* **2000**, *39*, 1163.

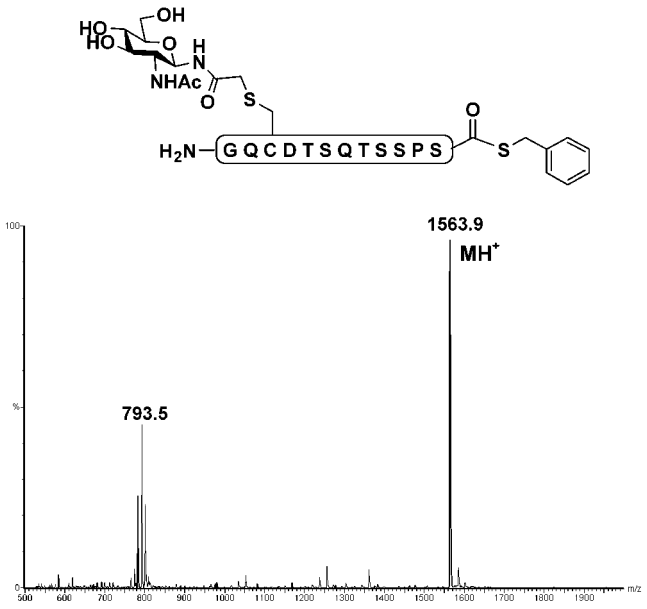


Figure 2. Extension to sulfonamide resin affords the "thio"-CD52 glycopeptide mimic α-benzylthioester after cleavage. Calculated mass = 1563.9 Da, observed mass = 1563.9 Da.

to native chemical ligation or further C-terminal modification. This also overcomes various problems associated with forming native linkages between the carbohydrate and peptide moieties such as aspartamide formation and anomerization of glycosylamines (which results in an α/β mixture of glycopeptide products).[12]

We began with a model system on which to optimize the chemistry. CD52, a cell surface glycopeptide, was chosen as it contains *N*-linked glycosylation and is modified at its C-terminus with a GPI anchor.[13] The peptide moiety was first synthesized on acid-labile NovaSyn-TGT resin, allowing facile monitoring of each process. The peptide was prepared using standard Fmoc chemistry, replacing the native asparagine residue with the commercially available S-*t*Buthio protected cysteine. Following peptide synthesis,[14] the *tert*-butylthio protecting group was removed by subjecting the resin to 10% DTT in DMF (stirred previously with solid $(NH_4)_2CO_3$). The free thiol was then derivatized using synthetic saccharide bromoacetamides or iodoacetamides (Figure 1).

Generally, quantitative sulfhydryl modification was achieved within 1.5 h using 3 equiv of glycosyl haloacetamide. The resin-bound products were then cleaved with 95% TFA, purified by semipreparative HPLC, and characterized by mass spectrometry (Table 1).

The ability for different saccharides to be efficiently ligated to a peptide scaffold may also facilitate glycopeptide library synthesis,[15] potentially useful for the display of carbohydrate-dependent T-cell antigens, where the peptide moiety is present only to enable binding within the major histocompatibility complex (MHC) peptide binding cleft.[16]

Next, the potential for this chemistry to be applied to native chemical ligation and expressed protein ligation was ex-

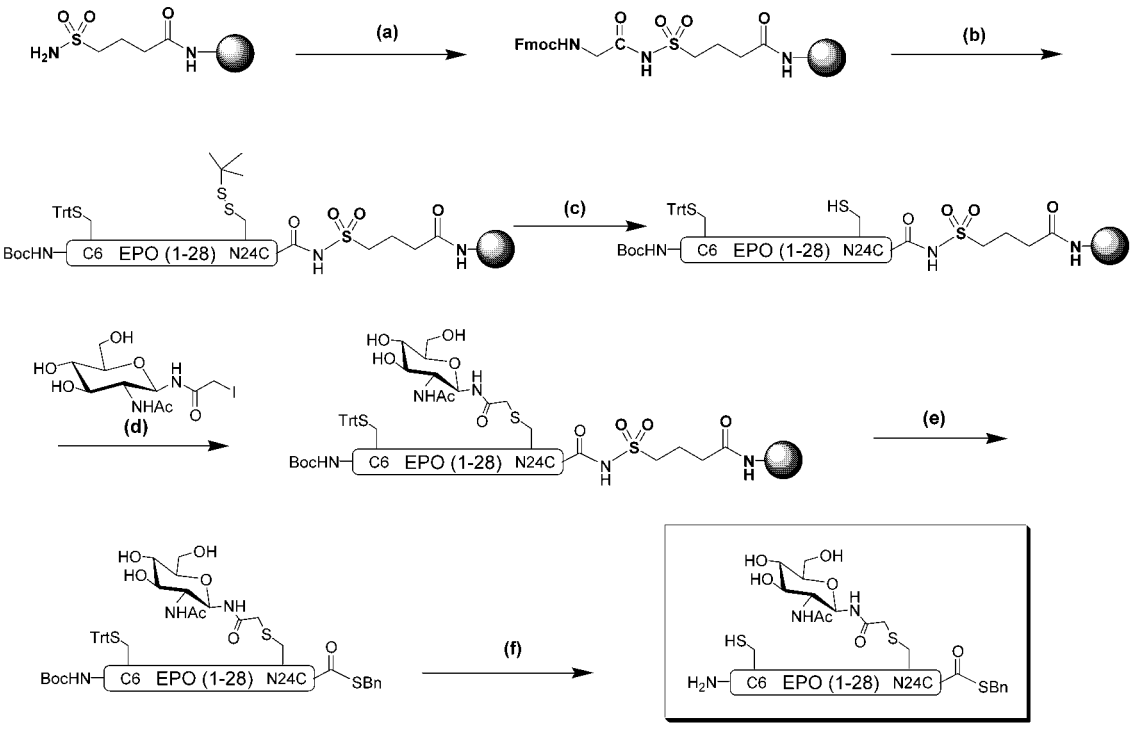


Figure 3. (a) Fmoc-Gly-OH, PyBOP, DIPEA, DMF, rt. (b) Fmoc SPPS. (c) 10% DTT, 2.5% v/v DIPEA, DMF, 24 h. (d) 2.5% v/v pyridine, DMF, 3 h (e) (1) ICH_2CN, DIPEA, DMF, 16 h; (2) BnSH, NaSPh, THF, 16 h. (f) 85% TFA, 5% H_2O 5% EDT, 5% thioanisole.

plored. The same CD52 glycopeptide mimic was thus prepared on 4-sulfamylbutyryl-AM (safety-catch) resin and cleaved as described.[17] The glycopeptide was released from an aliquot (50 mg) of resin with benzyl mercaptan in the presence of catalytic NaSPh after activation with iodoacetonitrile. The backbone protecting groups were then cleaved (85% TFA, 5% ethanedithiol, 5% thioanisole, 5% water, 3 h), and the pure glycopeptide α-thioester product (2.5 mg, 32%) was obtained after ether precipitation and semipreparative HPLC (Figure 2).

To test the methodology further and to test specifically for the ability to differentiate between cysteine residues within the same peptide sequence, we applied this method to the synthesis of the first 28 amino acids of the glycoprotein hormone erythropoietin (EPO). Within this fragment is contained the first glycosylation site (Asn24) and a cysteine residue (Cys-6) that is involved in disulfide bond formation. We aimed to differentiate between the cysteine residues using the commercially available *tert*-butylthio-protected cysteine as described above (Figure 1) to release the free thiol for modification (Asn24Cys) and *S*-trityl modification for Cys-6. The peptide was prepared using standard Fmoc solid-phase synthesis on 4-sufamylbutyryl resin. This would ultimately afford a C-terminal thioester for native chemical ligation. Figure 3 demonstrates how this peptide was prepared.

Again, the glycopeptide product was characterized by electrospray mass spectrometry after HPLC purification (Figure 4) and will be used for the synthesis of semisynthetic erythropoietin in future studies.

In summary, the use of glycosyl haloacetamides for the incorporation of oligosaccharides into bacterially derived proteins has been expanded to encompass solid-phase glycopeptide synthesis, providing access to homogeneously glycosylated *N*-linked glycopeptide mimic α-thioesters. The products can be applied to native chemical ligation and expressed protein ligation. Furthermore, we can achieve quantitative modification of specific, introduced cysteine residues using as little as 3 equiv of glycosyl haloacetamides. In doing so, some potential problems associated with the original approach such as incomplete sulfhydryl modification, lengthy purification, and ambiguity in the site of modification can be abolished. It should be noted that it is the stability of the thioether-linked saccharides that allows this chemistry to be applied to native chemical ligation. Other common approaches that rely on linking oligosaccharides through a disulfide bond to cysteine would likely be unsuitable for such methods.

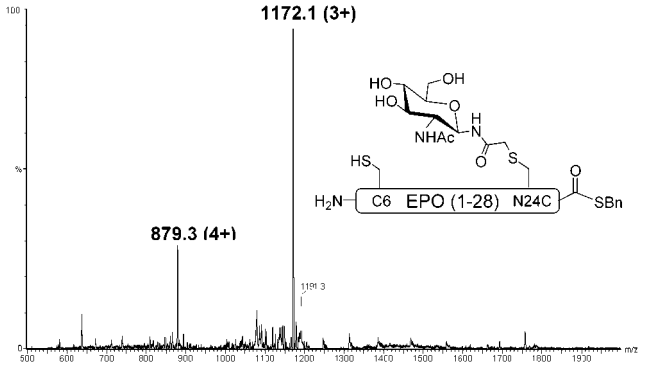


Figure 4. Electrospray mass spectrum of the glycosylated EPO fragment (residues 1−28), prepared as a C-terminal thioester.

Acknowledgment. We acknowledge financial support from the BBSRC.

Supporting Information Available: Full experimental procedures, MS and ^{1}H and ^{13}C NMR data for glycosyl haloacetamide, and mass spectra for selected glycopeptides (**6**−**8**, **10**, **11**). This material is available free of charge via the Internet at http://pubs.acs.org.

OL025627W

(16) Komba, S.; Meldal, M.; Werdelin, O.; Jensen, T.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1999**, 415. Gagne, R.; Werdelin, O.; Gad, M.; Komba, S.; Meldal, M. *Glycobiology* **2000**, *10*, 202. Keil, S.; Claus, C.; Dippold, W.; Kunz, H. *Angew. Chem., Int. Ed.* **2001**, *40*, 366. Hilaire, P. M. S.; Cipolla, L.; Franco, A.; Tedebark, U.; Tilly, D. A.; Meldal, M. *J. Chem. Soc., Perkin Trans. 1* **1999**, 3559.

(17) The resin-bound, protected glycopeptide mimetic (50−100 mg) was taken up in DMF (2.0 mL), treated with DIPEA (100 μL) and then iodoacetonitrile (filtered through basic alumina, 100 μL), and stirred at room temperature with the exclusion of light. After 16 h the resin was washed (DMF, then DCM, then THF) and taken up in anhydrous THF (2.0 mL). Benzyl mercaptan (100 μL) and solid NaSPh (2 mg) were added, and stirring was continued. After 16 h the solvent was removed by evaporation, and the crude peptide was obtained after precipitation with ether. See also: 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. Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. *J. Am. Chem. Soc.* **1999**, *121*, 11369.