Tetrahedron Letters, Vol.32, No.46, pp 6793-6796, 1991
Printed in Great Britain

0040-4039/91 $3.00 + .00
Pergamon Press plc

# A NOVEL METHOD FOR THE SPECIFIC GLYCOSYLATION OF PROTEINS

**Nicola J. Davis and Sabine L. Flitsch***

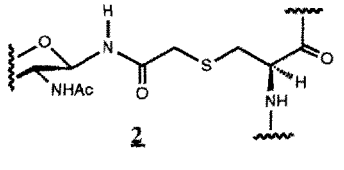
The Dyson Perrins Laboratory, South Parks Road, Oxford OX1 3QY, Great Britain

*Key Words*: Glycosylation; Cysteine Sidechains; Iodoacetamide; Glutathione; Bovine Serum Albumin

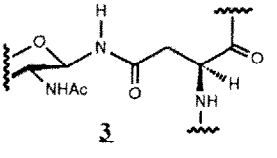
***Abstract***: The N-acetyl-glucosamine derivative **1** was synthesised and reacted with peptides and proteins containing cysteine residues. Analysis of products by nmr and mass spectroscopy showed that the thiol groups of cysteine were selectively glycosylated *via* a new sugar-protein linkage (as in **2**).

The importance of oligosaccharide sidechains for the biological function, biodistribution, and stability of glycoproteins has recently become recognised.[1-3] Oligosaccharides on cell surface glycoproteins have been shown to be responsible for cell-cell recognition and adhesion in biological processes, such as development[4], lymphocyte migration[5] and oncogenesis.[6] In order to systematically study these biological processes and possibly develop novel therapeutic agents, single glycoforms of glycoproteins, which have the same oligosaccharides attached to specific protein glycosylation sites, need to be available. We are interested in developing a chemical synthesis of such glycoforms by attaching suitably derivatised oligosaccharides selectively to protein sidechains.

The most important sugar protein linkages in glycoproteins are either to serine and threonine ("O-linked") or to asparagine residues ("N-linked").[7] The specific chemical synthesis of these linkages has been achieved in small peptides[7-11], but these methods are not applicable to whole proteins. We here describe a novel procedure which allows the selective glycosylation of a protein at one or several sites by reacting the iodoacetamide **1** with cysteine sidechains. The cysteine-sugar linkage thus generated (as in **2**) is structurally similar to the natural asparagine-sugar linkage (as in **3**) except that it is extended by a $-CH_2-S-$ group.




**2**

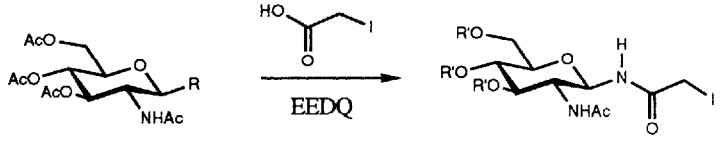



**3**

Iodoacetamide **1** was synthesised from protected azido-N-acetylglucosamine **4**[8] in three steps (Scheme 1). Reduction of **4** with $PtO_2/H_2$ led to the glycosylamine **5**, which was immediately coupled to iodoacetic acid

using EEDQ as a condensing agent. Deprotection of **6** with catalytic amounts of sodium methoxide gave **1** in 89% yield.[12]

*Scheme 1:*

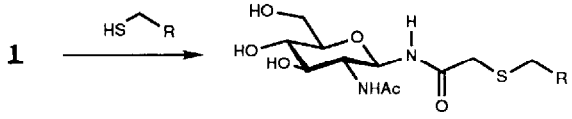



**4** R = $N_3$
**5** R = $NH_2$

**6** R' = Ac
**1** R' = H

The iodoacetamide **1** was chosen because the selectivity of iodoacetamide derivatives with protein thiol groups in aqueous buffer at pH 8 is well documented even in the presence of other nucleophiles such as amino- and hydroxyl-groups. Scheme 2 shows the reaction of **1** with several thiols. N-Acetyl-cysteine methylester and glutathione were chosen first since the products **7** and **8** respectively could be fully characterised spectroscopically. They both were treated with equimolar amounts of **1** and the products were purified by High Performance Liquid Chromatography (**7**; Spherisorb ODS2; water/acetonitrile/formic acid = 99/2/0.05) or gel permeation chromatography (**8**; P4 Biogel column; water) and characterised by high field nmr and mass spectrometry.[13] The analytical data confirmed that in compound **8** the sugar residue was only attached to the cysteinyl thiol group and not to the γ-glutamyl-α-amino group of glutathione.

*Scheme 2:*

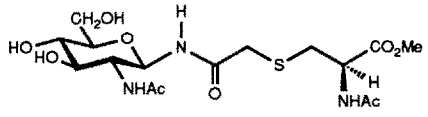



| $HSCH_2R$ | Product |
|---|---|
| $HSCH_2$-CH(NHAc)$CO_2CH_3$ | **7** |
| glutathione (γ-L-Glu-L-Cys-Gly) | **8** |
| bovine serum albumin | **9** |

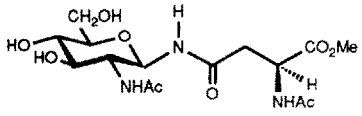
As an easily accessible protein for glycosylation reactions we chose Bovine Serum Albumin (BSA). BSA is known to contain one free cysteinyl group in its protein sequence. The derivatisation of this sulphydryl sidechain was monitored indirectly by using Ellman's Reagent.[14] This reagent reacts with free thiol groups with release of 5-thio-2-nitrobenzoic acid, which shows a strong absorption at 412 nm. Derivatisation of the free BSA thiol groups was achieved by treatment with sevenfold excess of **1** in ammonium carbonate (50 mM) for 40 min.[15] Although further investigations into the specificity of **1** for cysteines are necessary, the results shown

here demonstrate that compound **1** can be used to derivatise cysteine residues in proteins.

It is very likely that the biological function of an oligosaccharide is correlated to its conformation.[7] We were therefore concerned to establish that the new thioether linkage which is designed to mimic the natural N-linkage, would not change the conformation of the sugar. The previously described compound **10**[16] was therefore synthesised. This allowed us to compare the proton nmr spectra of **7** and **10**, two glycoconjugates which only differ in their glycopeptide linkage.




**7**




**10**

The high resolution $^1$H-nmr spectra (500 MHz) of **7** and **10** were found to be very similar. In particular, the coupling constants between the sugar ring protons, which are expected to be very sensitive to any conformational changes, were found to be nearly identical (Table 1). These preliminary results suggest that inserting the -$CH_2$-S- group into the natural N-linkage does not disturb the conformation of the sugar.

Table 1: Comparison of Coupling Constants (in Hz) of the Sugar Ring Protons in **7** and **10**

| Compound | $J_{1\text{-}2}$ | $J_{2\text{-}3}$ | $J_{3\text{-}4}$ | $J_{4\text{-}5}$ | $J_{5\text{-}6}$ | $J_{5\text{-}6'}$ | $J_{6\text{-}6'}$ |
|---|---|---|---|---|---|---|---|
| **7** | 9.7 | 10.1 | 8.5 | 9.8 | 5.0 | 2.1 | 12.4 |
| **10** | 9.7 | 10.2 | 8.5 | 9.8 | 5.1 | 2.1 | 12.4 |

In summary, we believe that the thioether linkage described here will be a useful mimic of naturally occuring sugar-protein linkages. It should allow us to synthesise defined glycoforms of glycoproteins since the derivatisation of cysteines is selective and there are generally very few cysteines that can act as possible attachment sites in proteins. We have so far only described the linkage of a monosaccharide (N-acetylglucosamine) to a protein. We are currently investigating the extention of this method to oligosaccharides, either by starting with the appropriate oligomeric glycosylamines or by extending the monosaccharide after attachment to the protein with glycosyltransferases.[16,17]

***Acknowledgements***: We thank Mrs. E. McGuinness and Dr. Robin Aplin for NMR and MS measurements respectively.

### *References and Notes*


1. Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.*, **1988**, *57*, 785.
2. Feizi, T. *Nature*, **1985**, *314*, 53.
3. Dube, S.; Fisher, J. W.; Powell, J. S. *J. Biol. Chem.*, **1988**, *263*, 17516.
4. Eggens, I.; Fenderson, B.; Toyokuni, T.; Dean, B.; Stroud, M.; Hakomori, S. *J. Biol. Chem.*, **1989**, *264*, 9476.
5. Springer, T. A. *Nature*, **1990**, *346*, 425.
6. Kojima, N. and Hakomori, S. *J. Biol. Chem.*, **1989**, *264*, 20159.
7. For a recent review see: Paulsen, H. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 823.
8. Spinola, M.; Jeanloz, R. *J. Biol. Chem.* **1970**, *245*, 4158.
9. Kunz, H. *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 294.
10. Anisfeld, S. T.; Lansbury, P. T. *J. Org. Chem.* **1990**, *55*, 5560.
11. Otvos, L.; Urge, L.; Hollosi, M.; Wroblewski, K.; Graczyk, G.; Fasman, G. D.; Thurin, J. *Tetrahedron Lett.* **1990**, *31*, 5889.
12. **1**: $\delta_H$(200MHz; $D_2O$) 1.84 (3H, s, $NHCOCH_3$), 3.23 (2H, s, $CH_2I$), 3.36 (1H, dd, *J* 8.6 and 9.8Hz, C(4)H), 3.39 (1H, m, C(5)H), 3.49 (1H, d, *J* 8.6 and 10.2Hz, C(3)H), 3.59 (1H, m, C(6)H), 3.71 (1H, m, C(2)H), 3.80 (1H, m, C(6)H), 4.93 (1H, d, *J* 9.7Hz, C(1)H). $\delta_C$(125.78MHz; $D_2O$) -3.29 ($CH_2I$), 22.24 ($COCH_3$), 54.24, 60.53, 69.56, 73.94, 77.60, 78.60 (C(1)-C(6)), 172.53, 174.54 (CO). *m/z* (FAB$^+$, DTTDTE/NaOAc) 411 ($MNa^+$), 389 ($MH^+$).
13. **7**: $\upsilon_{max}$ (KBr)/cm$^{-1}$ 3400-3100 (br, NH, OH), 1740 (C=O, ester), 1650 (C=O, amide). $\delta_H$(500MHz; $D_2O$) 1.91 (3H, s, $NHCOCH_3$), 1.96 (3H, s, $NHCOCH_3$), 2.85 (1H, dd, *J* 8.2 and 14Hz, $CH_2S$), 2.99 (1H, dd, *J* 4.9 and 14Hz, $CH_2S$), 3.23 (2H, m, $COCH_2S$), 3.40 (1H, dd, *J* 8.5 and 9.8Hz, C(4)H), 3.43 (1H, ddd, *J* 2.1, 5.0 and 9.8Hz, C(5)H), 3.53 (1H, dd, *J* 8.5 and 10.1Hz, C(3)H), 3.65 (1H, dd, *J* 5.0 and 12.4Hz, C(6)H), 3.68 (3H, s, OMe), 3.74 (1H, m, C(2)H), 3.78 (1H, dd, *J* 2.1 and 12.4Hz, C(6)H), 4.54 (1H, m, CH), 4.97 (1H, d, *J* 9.7Hz, C(1)H). $\delta_C$(125.78MHz; $D_2O$) 21.41, 21.80 ($COCH_3$), 32.92, 35.11 ($COCH_2S$, $CH_2S$), 52.20, 52.90 ($OCH_3$, CHNH), 54.07, 60.45, 69.48, 73.86, 77.44, 78.54 (C(1)-C(6)), 172.10, 172.47, 173.88, 174.42 (CO). *m/z* (FAB+, DTTDTE/Oxalic acid) 460 ($MNa^+$), 438 ($MH^+$).
**8**: $\delta_H$(500MHz; $D_2O$) 1.89 (3H, s, $COCH_3$), 2.04 (2H, m, $COCH_2CH_2$), 2.43 (2H, m, $COCH_2CH_2$), 2.80 (1H, dd , J 8.8 and 14Hz, $CH_2S$), 2.97 (1H, dd, *J* 5.1 and 14Hz, $CH_2S$), 3.22 (2H, m, $COCH_2S$), 3.38 (1H, dd, *J* 8.7 and 9.8Hz, C(4)H), 3.42 (1H, ddd, *J* 2.1, 4.9 and 9.8Hz, C(5)H), 3.52 (1H, dd, *J* 8.7 and 10.1Hz, C(3)H), 3.64 (1H, m, C(6)H), 3.65 (3H, m, $HO_2CCH_2$ and $CH_2CH_2CH$). 3.71 (1H, m, C(2)H), 3.77 (1H, d, *J* 2.1 and 12.4Hz, C(6)H), 4.53 (1H, dd, *J* 5.0 and 8.7Hz, $CHCH_2S$), 4.97 (1H, d, *J* 9.7Hz, C(1)H). *m/z* (Electrospray) 568.612 ($MH^+$), 590.620 ($MNa^+$), 606.627 ($MK^+$).
14. Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70.
15. Typically BSA (52mg) and iodoacetamide **1** (2mg) in 50 mM ammonium carbonate (2 ml) were incubated in the dark at room temperature. Thiol content was measured by taking 0.05 ml aliquots, diluting with 1 ml of buffer and adding 0.01 ml of Ellman's reagent (0.01M). The absorbance at 412 nm was used as a measure of thiol content. A control reaction, to which no **1** was added, did not show any decrease of thiol groups during the incubation.
16. Thiem, J.; Wiemann, T. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 80.
17. Unverzagt, C.; Kunz, H.; Paulson, J. C. *J. Am. Chem. Soc.* **1990**, *112*, 9308.


(Received in UK 9 September 1991)