OBC
www.rsc.org/obc

Synthesis of homogenous site-selectively glycosylated proteins

Richard S. Swanwick,[a] Alison M. Daines,[b] Sabine L. Flitsch[b] and Rudolf K. Allemann*[a]

[a] *School of Chemistry, University of Birmingham, Edgbaston, Birmingham, UK B15 2TT. E-mail: r.k.allemann@bam.ac.uk; Fax: 41 121 4147871; Tel: 41 121 4144359*

[b] *School of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, EH9 3JJ, UK. E-mail: s.flitsch@ed.ac.uk; Fax: 44 131 6504737; Tel: 44 131 6504737*

Received 10th November 2004, Accepted 6th December 2004
First published as an Advance Article on the web 24th December 2004

Apparently homogenous glycoproteins can be synthesised in good yield by a combination of site directed mutagenesis, a highly flexible but selective chemical derivatisation and efficient purification through the use of glycosyl thiosulfonates such as 2-((biotinoyl)-amino)-ethyl methanethiosulfonate.

The presence of carbohydrates has dramatic effects on the physical, chemical and biological properties of glycoproteins. Surface glycoproteins act as markers in the communication between cells during the immune response,[1] inflammation[2] and microbial virulence.[3] The correct glycosylation of proteins is important for their proper folding *in vivo*.[4] The presence of sugars on proteins has been shown to modulate their kinetic properties[5,6] and their stability.[2,7–13] The study of many of these events is complicated by the fact that natural glycoproteins normally occur as mixtures of glycoforms that carry different sugars in different positions on the same protein backbone. The biological importance of glycoconjugates and their increasing importance as therapeutic agents make the synthesis of homogenous glyoproteins an important task.

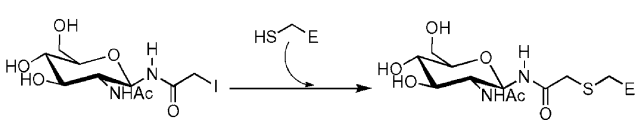
The chemical synthesis of homogenous glycoproteins is complicated mainly by two factors. Despite significant advances in convergent synthesis using approaches such as native peptide ligation and ligation of a peptide with expressed protein fragments,[14] the chemical synthesis of proteins is still a formidable task. In addition, while synthetic methods for the synthesis of large structurally defined oligosaccharides have recently been developed significantly,[15] oligosaccharide synthesis is still not routine.

Several methods have been described for the generation of neoglycoproteins through site selective glycosylation of proteins by chemical modification of biotechnologically produced proteins. In one such approach the free thiol group of cysteine residues is reacted with a chemically synthesised glycosyl iodoacetamide to produce a stable linkage between the protein and the carbohydrate that resembles that found in native glycosylation of asparagine residues (Scheme 1).[16–18] The usefulness of this approaches has been proven by the synthesis of homogeneous glycoforms of the human glycoprotein hormone erythropoietin.[18]

Glycosyl thiosulfonates have also been used to produce glycoproteins. In this case, the sugar moiety is attached to the cysteine thiol through a disulfide bond and such glycosyl thiosulfonates therefore suffer from instability towards reducing agents like those generally found in cellular environments.[19–21]

For many applications, the purity of the synthesised glycoproteins is paramount. However, all of the above protocols can suffer from reduced selectivity. At increased pH values where the reactivity of the cysteine nucleophile is maximal, reaction with the amino groups of lysines is an important side reaction. At lower pH values, where this side reaction is largely suppressed, the reduced reactivity of the protonated nucleophile leads to mixtures of the glycosylated and the sugar-free parent proteins. The small difference in size and overall structure between the glycosylated and unglycosylated proteins often precludes efficient purification and the generation of homogenous glycoproteins. Previously, affinity chromatography on lectin containing solid supports has been used for the purification of synthetic glycoproteins, but this method is limited to the purification of proteins containing sugars recognised by available lectins and by the often only modest affinity of lectins for glycoproteins. We report here a general method for the purification of *in vitro* glycosylated proteins that yields apparently homogenous glycoproteins.

To study the effects of site-specific glycosylation on the kinetics and stability of the naturally unglycosylated *E. coli*



Scheme 1 Generation of neoglycoproteins through glycosylation of cysteine containing proteins with glycosyl iodoacetamides.[22]

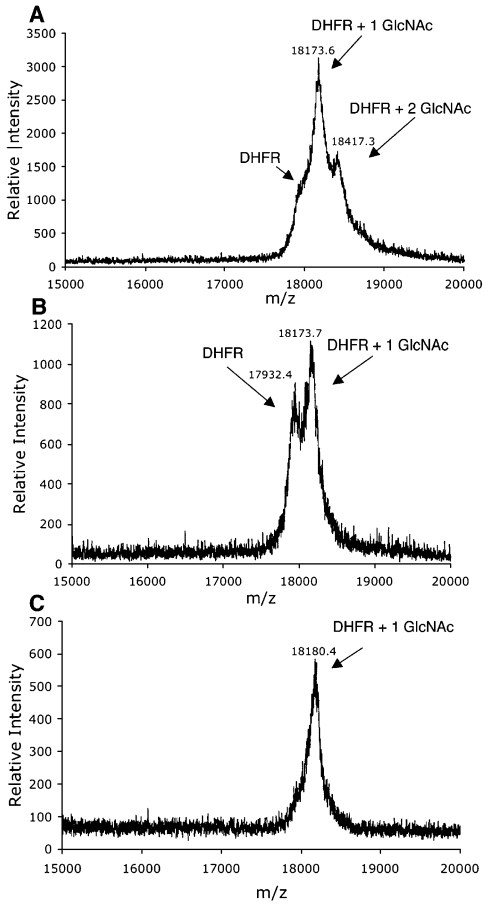
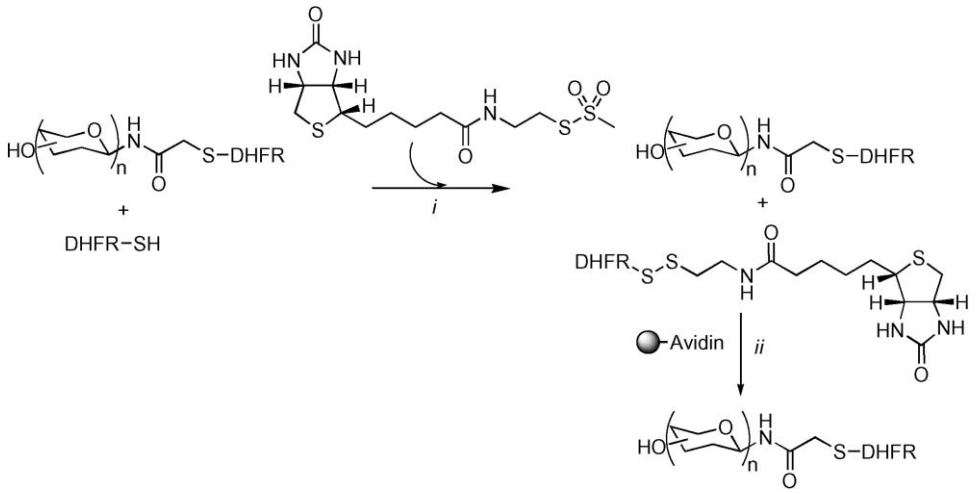


Fig. 1 MALDI-TOF mass spectra of reaction products of DHFR-C85AC152SE120C treated with *N*-acetyl-glucosyliodoacetamide at pH 9 (A) and pH 7 (B), respectively. (C) MALDI-TOF mass spectrum of purified DHFR-C85AC152SE120C carrying one GlcNAc on C120.

DOI: 10.1039/b417133k



Scheme 2 Derivatisation of sugar free proteins with MTSEA and purification on avidin containing resin. *Reagents: i*, 5 equivalents of MTSEA in DMSO was added to protein mixture (1 mg ml^{-1}) in 25 mM MES pH 7.5, 100 mM NaCl at 4 °C; after incubation for 12 h, reaction products were dialysed against 10 mM K_iPO_4, pH 7.5, 100 mM NaCl. *ii*, The mixture was applied to avidin immobilized on agarose (Promega, ~30 nmol avidin per ml of resin) and unbiotinylated proteins were eluted with 10 mM K_iPO_4, pH 7.5, 100 mM NaCl. Overall yields were 65, 55, 55, and 25% for GlcNAc, lactose, maltose, and maltotriose, respectively.

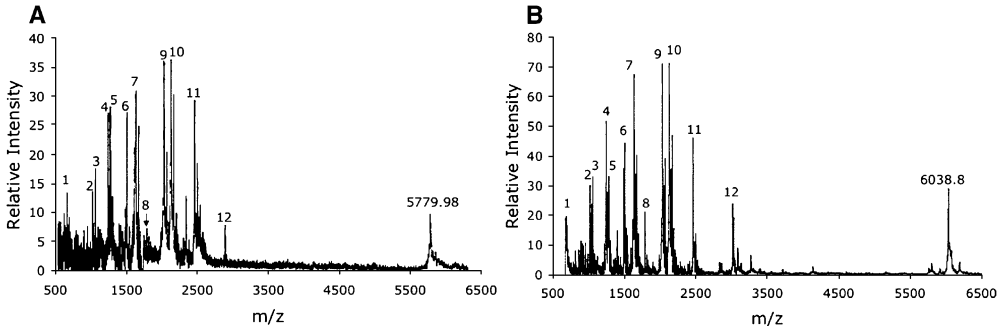


Fig. 2 Tryptic digests of DHFR-C85AC152S (A) and DHFR-C85AC152SE120C carrying a GlcNAc on residue 120 (B). 5779.98 and 6038.80 are the singly charged ions for peptide THIDAEV**C**GDTHFPDYEPDDWESVFSEFHDADAQNSHSYSFEILER, in its unglycosylated and *N*-acetylglucosamine carrying forms, respectively. Peaks 1–12 correspond to the other tryptic peptides.

protein dihydrofolate reductase (DHFR), selectively modified and highly purified proteins need to be produced. However, attempts to react at pH 9 glycosyliodoacetamides with the thiolate of a cysteine residue, which had been introduced by site directed mutagenesis to replace the solvent exposed Glu 120 of a cysteine-free variant of DHFR,† resulted in glycosylation of lysine residues in addition to Cys 120 (Fig. 1A). At pH 7, the reaction yielded mono-glycosylated proteins in relatively modest yields in addition to significant amounts of unreacted proteins (Fig. 1B).

In order to efficiently separate glycosylated proteins from sugar free protein the increased reactivity of thiosulfonates was taken advantage of (Scheme 2). Mixtures of glycosylated and unglycosylated proteins were treated at pH 7 with 2-((biotinoyl)-amino)-ethyl methanethiosulfonate (MTSEA) to produce mixtures of proteins that were either glycoslyated or biotinylated on Cys 120. Purification by affinity chromatography on an avidin resin resulted in the retention of biotinylated DHFR and elution of essentially pure mono-glycosylated proteins (Fig. 1C).

The glycosylated proteins were further analysed by proteolytic cleavage with trypsin. MALDI-TOF MS of the resulting peptides indicated that glycosylation had occurred only on the desired fragment confirming that Cys 120 had been modified (Fig. 2). Using this method, homogenous samples of DHFR carrying *N*-acetylglucosamine, glucose, mannose, lactose and maltotriose on residue 120 were successfully prepared in good overall yields (Scheme 2).

In summary we have reported here a route for the purification to homogeneity of substantial quantities of neoglycoproteins. Site-directed mutagenesis to introduce cysteine residues in specific positions of a target protein, derivatisation with chemically synthesised glycosyl iodoacetamides and purification through scavenging of the unreacted sugar free proteins by biotin containing thiosulfonates followed by affinity chromatography yielded essentially pure proteins. An efficient route for the synthesis of specifically glycosylated proteins will facilitate the study of the physical, chemical and biological effects of protein glycosylation. In addition, it should be noted that the present methodology is very general and should also be applicable to the derivatisation of cysteine sidechains with electrophilic thiol specific reagents other than glycosides, such as spectroscopic reporter groups, to produce homogeneously labelled material.[23–25]

Acknowledgements

We would like to thank the BBSRC and the Wellcome Trust for funding.

Notes and references

† *General procedure*. *Protein production:* DHFR-C85AC152SE120C was generated in DHFR-C85AC152S, a cysteine free mutant with properties identical to those of the wild type protein [26] (Swanwick & Allemann, unpublished)-using the QuikChange™ site directed mutagenesis kit. DHFR-C85AC152SE120C was produced and purified as previously described.[27] *Protein alkylation:* Proteins were dialysed against 25 mM *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES), pH 7.5, 100 mM NaCl. After reduction with 0.5 mM tris(2-carboxethyl) phosphine, proteins (2 mg ml^{-1}) were incubated with a 50-fold molar excess of glycolsyliodoacetamides at 5 °C in the dark for 72 h. The reaction products were dialysed against 25 mM 2-morpholinoethanesulfonic acid (MES) pH 7.5, 100 mM NaCl. *Trypsin digestion:* Proteins were dialysed

against 10 mM ammonium bicarbonate, pH 7.5 and the proteolytic cleavage initiated by the addition of 1μl of a 1μM solution of trypsin. After incubation at 37 °C for 12 hours, the resulting peptides were lyophilised and analysed by MALDI-TOF-MS.

1 J. B. Lowe, *Cell*, 2001, **104**, 809.
2 L. A. Lasky, *Annu. Rev. Biochem.*, 1995, **64**, 113.
3 N. Sharon and H. Lis, Lectins — Proteins with a sweet tooth: Functions in cell recognition, in *Essays in Biochemistry*, Vol. 30, 1995, p 59.
4 A. Helenius and M. Aebi, *Science*, 2001, **291**, 2364.
5 A. Kohen, T. Jonsson and J. P. Klinman, *Biochemistry*, 1997, **36**, 6854.
6 P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker and R. A. Dwek, *Biochemistry*, 1994, **33**, 17.
7 C. Q. Wang, M. Eufemi, C. Turano and A. Giartosio, *Biochemistry*, 1996, **35**, 7299.
8 K. S. Kwon and M. H. Yu, *Biochim. Biophys. Acta*, 1997, **1335**, 265.
9 W. O. Baek and M. A. Vijayalakshmi, *Biochim. Biophys. Acta*, 1997, **1336**, 394.
10 A. P. Gimenez-Roqueplo, J. Celerier, G. Lucarelli, P. Corvol and X. Jeunemaitre, *J. Biol. Chem.*, 1998, **273**, 21232.
11 J. Masarova, D. Mislovicova, P. Gemeiner and E. Michalkova, *Biotechnol. Appl. Biochem.*, 2001, **34**, 127.
12 K. Broersen, A. G. J. Voragen, R. J. Hamer and H. H. J. de Jongh, *Biotechnol. Bioeng.*, 2004, **86**, 78.
13 S. Vijayalekshmi, S. K. George, L. K. Andersson, J. Kihlberg and L. Baltzer, *Org. Biomol. Chem.*, 2003, **1**, 2455.
14 T. W. Muir, *Annu. Rev. Biochem.*, 2003, **72**, 249.
15 O. J. Plante, E. R. Palmacci and P. H. Seeberger, *Science*, 2001, **291**, 1523.
16 N. J. Davis and S. L. Flitsch, *Tetrahedron Lett.*, 1991, **32**, 6793.
17 R. M. Bill, P. C. Winter, C. M. McHale, V. M. Hodges, G. E. Elder, J. Caley, S. L. Flitsch, R. Bicknell and T. R. J. Lappin, *Biochim. Biophys. Acta*, 1995, **1261**, 35.
18 D. Macmillan, R. M. Bill, K. A. Sage, D. Fern and S. L. Flitsch, *Chem. Biol.*, 2001, **8**, 133.
19 G. M. Watt and G. J. Boons, *Carbohydr. Res.*, 2004, **339**, 181.
20 J. B. Watney, P. K. Agarwal and S. Hammes-Schiffer, *J. Am. Chem. Soc.*, 2003, **125**, 3745.
21 D. P. Gamblin, P. Garnier, S. J. Ward, N. J. Oldham, A. J. Fairbanks and B. G. Davis, *Org. Biomol. Chem.*, 2003, **1**, 3642.
22 S. L. Flitsch, *Curr. Opin. Chem. Biol.*, 2000, **4**, 619.
23 S. L. Flitsch and H. G. Khorana, *Biochemistry*, 1989, **28**, 7800.
24 C. Altenbach, S. L. Flitsch, H. G. Khorana and W. L. Hubbell, *Biochemistry*, 1989, **28**, 7806.
25 W. L. Hubbell, D. S. Cafiso and C. Altenbach, *Nat. Struct. Biol.*, 2000, **7**, 735.
26 M. Iwakura, B. E. Jones, J. B. Luo and C. R. Matthews, *J. Biochem. (Tokyo)*, 1995, **117**, 480.
27 R. S. Swanwick, P. J. Shrimpton and R. K. Allemann, *Biochemistry*, 2004, **43**, 4119.