ChemComm

The controlled glycosylation of a protein with a bivalent glycan: towards a new class of glycoconjugates, glycodendriproteins†‡

Benjamin G. Davis

Department of Chemistry, University of Durham, Science Laboratories, South Road, Durham, UK DH1 3LE. E-mail: ben.davis@durham.ac.uk

Received (in Liverpool, UK) 10th November 2000, Accepted 12th January 2001
First published as an Advance Article on the web 7th February 2001

The use of a novel bivalent carbohydrate modification reagent, based on a flexible, branched divalent core in a combined site-directed mutagenesis and chemical modification strategy has allowed the first controlled synthesis of a pure protein bearing a branched glycan or a first generation glycodendriprotein.

The glycosylation of proteins plays a key role in determining their expression, folding,[1] thermal and proteolytic stability[2] and in the case of enzymes, catalytic activity.[3] Furthermore, the role of glycoproteins as cell surface markers in communication events such as microbial invasion,[4] inflammation,[5] immune responses and tissue development[6] depends crucially on the correct glycosylation pattern.[7] There is evidence that even very slight alterations in the sugars that decorate the exterior of a protein can cause remarkable changes in these properties. For example, we have recently shown that correct and controlled glycosylation of a model enzyme system with a single saccharide unit allows fine tuning of activity to levels that are up to 8.4-fold greater than the native unglycosylated enzyme.[8] Models for glycoprotein interactions, including activity and binding to corresponding receptor proteins, have been suggested[9,10] and the elucidation of the mechanism of this binding and its consequences is a dominant primary goal in glycoscience and continues to drive the synthesis of glycoconjugates.[11] Access to well-defined scaffolds to probe the nature of these models is essential.[12]

Despite their very shallow binding sites, sugar-binding proteins, lectins, show a remarkable specificity in their binding of multivalent complex carbohydrate structures.[13] The monosaccharide–lectin interaction stands out as an unusually weak and relatively undiscriminating one ($K_d \sim 10^{-3}$ M).[14] However, when more than one saccharide of the right type and in the right orientation are clustered together there is a rapid increase in both affinity and specificity by the corresponding lectin.[15] This increase appears in some cases to be greater than would be expected due to the increase in local concentration (statistical effect) alone and has been termed the 'cluster' or 'multivalent effect', although its origin is yet to be rigorously determined.[10]

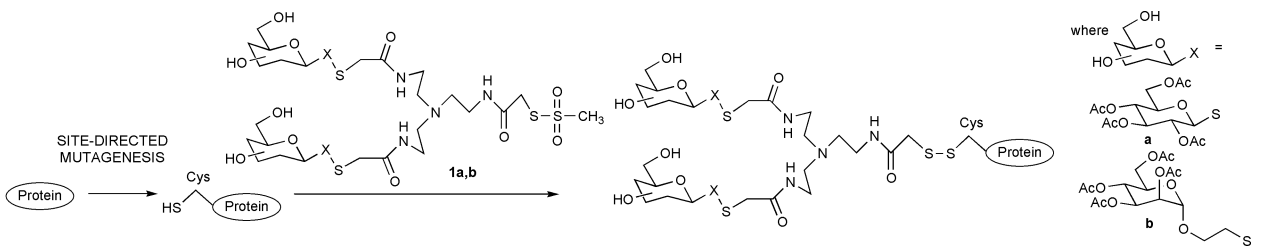
† Some of this work was presented at a joint meeting of the RSC Carbohydrate and Bioorganic groups, Warwick, July 6, 2000 and at the 20th International Carbohydrate Symposium, Hamburg, August 30, 2000.
‡ Electronic supplementary information (ESI) available: selected NMR data. See http://www.rsc.org/suppdata/cc/b0/b009184g/

Exploitation of lectin binding does not necessarily rely on mimicking natural multiantennary structures as long as an energetically efficient method for their presentation to binding sites may be found.[15] We have recently demonstrated that glycoproteins bearing a single monosaccharide ligand bind to lectins with a low affinity that is dependent both on the nature of the carbohydrate and the site of glycosylation.[16] We have set ourselves the goal of constructing glycoproteins bearing branched multivalent glycans at predetermined sites to probe the nature of this binding interaction further. Furthermore, by choosing an enzyme as our protein model we aim to investigate the effect of multivalent site-selective glycosylation, by displaying multiple copies of a glycan that previously allowed dramatic enhancement of enzyme actvities.[8] This communication describes the successful construction of the first generation of this new class of glycoconjugates: glycodendriproteins.

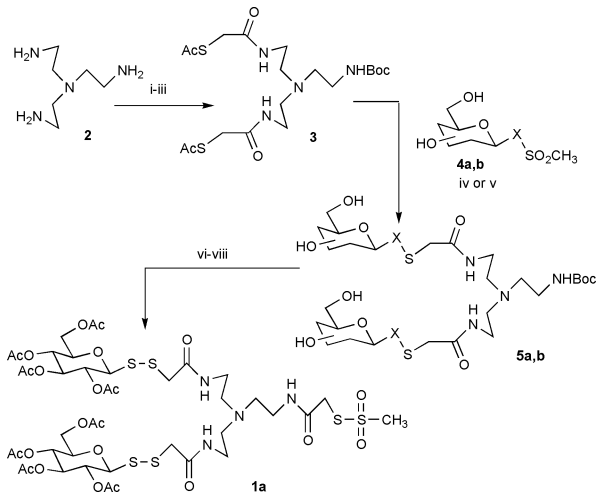
To achieve our goals we employed a combined site-directed mutagenesis (SDM) and chemical modification approach (Scheme 1).[17] This strategy involves the introduction of cysteine as a chemoselective tag at preselected positions within a given protein and then reaction of its thiol residue with (polyglyco)methanethiosulfonate reagents, such as **1**. Methanethiosulfonate (MTS) reagents react specifically and quantitatively with thiols[18] and allow the controlled formation of neutral disulfide linkages. We chose as our model protein the serine protease enzyme, subtilisin *Bacillus lentus* (SBL, EC 3.4.21.14).[19] SBL is an ideal model since wildtype (WT)-SBL contains no natural cysteines and methanethiosulfonate reagents therefore react only with the cysteine residue that has been introduced by SDM. In addition, SBL has been well characterized,[20] has been over-expressed and purified,[21] and its crystal structure is known.[22]

Two representative bivalent branched glycan MTS reagents **1a**,**b** based on a trivalent tris(2-aminoethyl)amine (TREN) core were chosen as targets (Scheme 2). **1a** bears at the end of its two glycan branches the same untethered peracetylglucose unit that had previously allowed dramatic enhancement of enzyme activity.[8] **1b** would bear ethyl-tethered mannose moieties that had been used in the construction of previous glycoproteins that had shown low levels of lectin binding.[16] These two reagents would therefore allow the introduction of multivalent, tethered or untethered, glycans with α or β anomeric stereochemistry from different parent carbohydrates systems.

After differentiation of one of the amine termini of TREN **2** through selective protection as its mono-Boc derivative,[23] the two remaining free amine termini were reacted with chloro-



Scheme 1



Scheme 2 *Reagents and conditions*: i, 1 equiv. Boc_2O, CH_2Cl_2, −78 °C, 68%; ii, $(ClCH_2CO)_2O$, pyridine, CH_2Cl_2, 97%; iii, KSAc, DMF, 50 °C, 88%; iv, NaOH (aq.) then **4a**, 73%; v, NaOH (aq,) then **4b**, 62%; vi, CF_3COOH, CH_2Cl_2, 88%; vii, $(ClCH_2CO)_2O$, pyridine, CH_2Cl_2, 87%; viii, $NaSSO_2CH_3$, DMF, 50 °C, 68%.

acetic anhydride to give the corresponding bis-α-chloroamide. Treatment of this branched dichloride with the potassium salt of thioacetic acid gave the bis-thioester **3** in a good overall yield (58% over 3 steps from **2**). One-pot selective deprotection and glycosylations of **3** were achieved by treatment with dilute aqueous NaOH solution to hydrolyse the labile thioacetates and then modification of the free thiol groups produced with the appropriate untethered β-gluco **4a** or tethered α-manno **4b** methanethiosulfonate reagents to yield the corresponding bivalent branched glycans **5a**,**b** in 73 and 62% yield, respectively. It should be noted that the use of a basic TREN-core as a scaffold allowed the scavenging of **5a**,**b** from reaction mixtures using acidic ion exchange resin and therefore greatly simplified their purification. With the ability to introduce two distinct glycan endgroups **a** or **b** thus suitably demonstrated, **5a** was selected and the remaining Boc-protected amine terminus elaborated to introduce a methanethiosulfonate group and to create **1a** as the first representative reagent that would allow protein modification. Thus, **5a** was deprotected through treatment with CF_3COOH and the free amine produced converted to the corresponding α-chloroamide. Displacement of α-chloro group through treatment with $NaSSO_2CH_3$ in DMF at 50 °C proceeded smoothly and yielded the target bis-glycan MTS **1a** in good yield (52% over 3 steps from **5a**).

We used **1a** as the first example of a bivalent protein glycosylating reagent to treat mutant protein SBL-S156C in aqueous buffer (Scheme 1). S156C was obtained through site-directed mutagenesis of SBL[8] and bears an outwardly directed reactive thiol on its surface close to its substrate-binding region. Gratifyingly, the reaction was rapid and quantitative, as judged by monitoring changes in specific activity and by titration of residual free thiols with Ellman's reagent.[24] The first generation glycodendriprotein S156C-(S-**a**)$_2$ formed was purified by size-exclusion chromatography and dialysis, and its structures confirmed by rigorous ES-MS analysis (Found 27811, Expected 27805 Da, Fig. 1b). The protein also appeared as a single band on non-denaturing gradient PAGE (Fig. 1(a)), thereby establishing its formation as a single glycoform. This is the first example of a homogeneous protein bearing branched multivalent glycans, or first-generation glycodendriprotein, in which both the site of glycosylation and the structure of the glycan introduced has been predetermined.

Work towards the synthesis of higher generations based on this novel-type of disulfide constructed dendrimeric system is now in progress and the details of kinetics and binding analysis of glycodendriprotein S156C-(S-**a**)$_2$ will be published in due course.

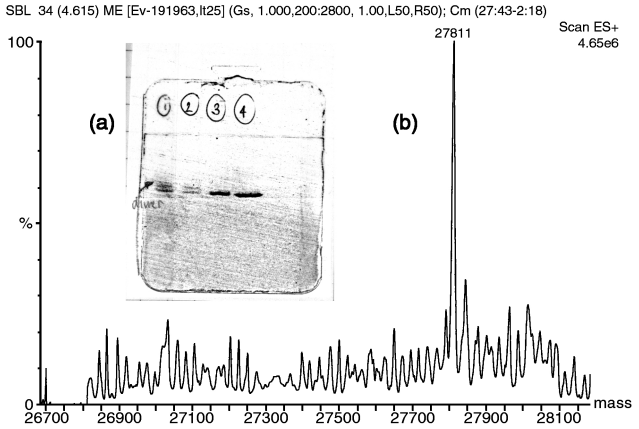


Fig. 1(a) PAGE analysis: Lanes 1,2 crude S156C, Lane 3 pure S156C, Lane 4 pure glycodendriprotein S156C-(S-**a**)$_2$ (b) MaxEnt™ deconvoluted ES^+-MS spectrum of S156C-(S-**a**)$_2$ indicating found mass 27811, expected 27805.

We gratefully acknowledge the generous financial support of Genencor International Inc. and the University of Durham. We also thank Dr R. R. Bott and Professor J. B. Jones for helpful discussions and J.P Marston and Dr M. Jones for technical assistance.

Notes and references

1 A. Helenius, *Mol. Biol. Cell*, 1994, **5**, 253.
2 G. Opdenakker, P. M. Rudd, C. P. Ponting and R. A. Dwek, *FASEB J.*, 1993, **7**, 1330.
3 P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker and R. A. Dwek, *Biochemistry*, 1994, **33**, 17.
4 N. Sharon and H. Lis, *Essays Biochem.*, 1995, **30**, 59.
5 L. A. Lasky, *Annu. Rev. Biochem.*, 1995, **64**, 113; W. I. Weis and K. Drickamer, *Annu. Rev. Biochem.*, 1996, **65**, 441.
6 K. Brueckner, L. Perez, H. Clausen and S. Cohen, *Nature*, 2000, **406**, 411.
7 A. Varki, *Glycobiol.*, 1993, **3**, 97; R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
8 R. C. Lloyd, B. G. Davis and J. B. Jones, *Bioorg. Med. Chem.*, 2000, **8**, 1537.
9 R. B. Parekh and C. J. Edge, *TIBTECH*, 1994, 339.
10 S. M. Dimick, S. C. Powell, S. A. McMahon, D. N. Moothoo, J. H. Naismith and E. J. Toone, *J. Am. Chem. Soc.*, 1999, **121**, 10286; J. B. Corbell, J. J Lundquist and E. J. Toone, *Tetrahedron: Asymmetry*, 2000, **11**, 95.
11 B. G. Davis, *J. Chem. Soc., Perkin Trans. 1*, 1999, 3215.
12 B. G. Davis, *Chem. Ind.*, 2000, 134.
13 N. Sharon and H. Lis, *Essays Biochem.*, 1995, **30**, 59.
14 Y. C. Lee and R. T. Lee, *Acc. Chem. Res.*, 1995, **26**, 323.
15 Y. C. Lee, R. R. Towsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson and H. Lönn, *J. Biol. Chem.*, 1983, **258**, 199.
16 B. G. Davis, D. Hodgson, A. Ullman, K. Khumtaveeporn, R. Sala, R. R. Bott and J. B. Jones, unpublished work. Lectin binding led to enhanced selectivity in the degradation of a mannose specific lectin by SBL glycosylated with a single mannose residue.
17 B. G. Davis, R. C. Lloyd and J. B. Jones, *J. Org. Chem.*, 1998, **63**, 9614; B. G. Davis and J. Bryan Jones, *Synlett*, 1999, 1495; B. G. Davis, M. A. T. Maughan, M. P. Green, A. Ullman and J. B. Jones, *Tetrahedron: Asymmetry*, 2000, **11**, 245.
18 R. Wynn and F. M. Richards, *Methods Enzymol.*, 1995, **201**, 351.
19 C. Betzel, S. Klupsch, G. Papendorf, S. Hastrup, S. Branner and K. S. Wilson, *J. Mol. Biol.*, 1992, **223**, 427.
20 H. Gron, M. Meldal and K. Breddam, *Biochemistry*, 1992, **31**, 6011.
21 M. R. Stabile, W. G. Lai, G. DeSantis, M. Gold, J. B. Jones, C. Mitchinson, R. R. Bott, T. P. Graycar and C.-C. Liu, *Bioorg. Med. Chem. Lett.*, 1996, **6**, 2501.
22 P. Kuhn, M. Knapp, S. M. Soltis, G. Ganshaw, M. Thoene and R. Bott, *Biochemistry*, 1998, **37**, 13446.
23 P. Tecilla, U. Tonellato, A. Veronese, F. Felluga and P. Scrimin, *J. Org. Chem.*, 1997, **62**, 7261.
24 G. L. Ellman, K. D. Courtney, V. Andres and R. M. Featherstone, *Biochem. Pharmacol.*, 1961, **7**, 88.