COMMUNICATION

OBC
www.rsc.org/obc

Glycosyl phenylthiosulfonates (Glyco-PTS): novel reagents for glycoprotein synthesis†‡

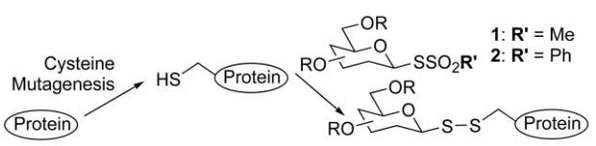
David P. Gamblin, Philippe Garnier, Sarah J. Ward, Neil J. Oldham, Antony J. Fairbanks* and Benjamin G. Davis*

Dyson Perrins Laboratory, University of Oxford, South Parks Road, Oxford, UK OX1 3QY. E-mail: Ben.Davis@chem.ox.ac.uk; Antony.Fairbanks@chem.ox.ac.uk; Fax: +44 1865 275674; Tel: +44 1865 275652

Received 19th June 2003, Accepted 28th July 2003
First published as an Advance Article on the web 11th August 2003

Controlled site-selective glycosylation can be achieved by combining site-directed cysteine mutagenesis with chemical modification of the introduced thiol; a new class of more efficient chemoselective reagents, glycosyl phenylthiosulfonates, allow rapid glycosylations of representative simple thiols, peptides and proteins.

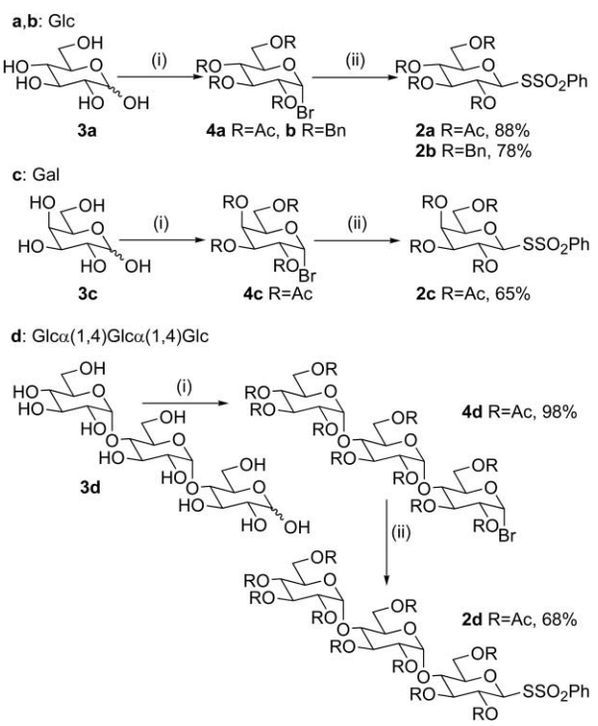
The glycosylation§ of proteins plays a vital role in their biological behaviour, destination and stability.[1] The chemical synthesis of glycoproteins offers certain key advantages,[2] not least of which is more ready access to pure glycoprotein glycoforms.[3] Several alternative and complementary strategies that provide access to pure glycoforms have been described.[2,4] We have previously described the first examples of a site-selective glycosylation process that combines site-directed mutagenesis with chemoselective modification.[4b] In this two-step strategy a cysteine is introduced through mutagenesis at a preselected position in a given protein to create a protein with a single free thiol, which is then chemoselectively modified with a thiol-selective carbohydrate reagent (Scheme 1) thereby allowing full control of both site and sugar. This early system exploited glycosyl methanethiosulfonate (glyco-MTS) reagents[5] **1** due to the strong history of use of methanethiosulfonates (MTS) as potent protein modifying reagents.[6–10] However, during the course of our work on glyco-[4b,5] and other[10] MTS reagents we have considered methods for increasing their utility yet further and describe here our synthesis of an improved class of carbohydrate thiosulfonates, the glycosyl phenylthiosulfonates (glyco-PTS) **2**.



Scheme 1

Leaving group alteration is a tried-and-tested method for tuning the reactivity of electrophilic structures.[11] We therefore focussed on the aglycon group of our glyco-MTS reagents to improve their use. Shortcomings of MTS reagents include occasionally moderate yields and difficulties in their preparation and occasional instability under the basic conditions in which they are often used. Additionally, any benefits in the form of enhanced rate and efficiency of reaction and *in situ* modification monitoring would also be welcome. We considered that simple alteration of the methyl group in the methanethiosulfonate (MTS) aglycon to a phenyl group to form the novel class of phenylthiosulfonates (PTS) might provide solutions or advantages in all of these respects. In particular, the removal of the potentially acidic methyl group site α to the sulfone unit (pK_a ~ 14)[12] might enhance stability both during preparation and reaction. Moreover, the introduction of this UV-active chromophore into the aglycon would also advantageously allow direct monitoring of reagent formation and use in, for example, protein glycosylation.

Representative mono- and oligo-saccharide glyco-PTS reagents **2a**–**d** were readily synthesized by displacement of halide from the appropriate glycosyl halides[13,14] using sodium phenylthiosulfonate (NaPTS),[15] in acetonitrile at 70 °C in the presence of 0.1 equiv. of tetrabutylammonium halide (Scheme 2).



Scheme 2 *Reagents and conditions:* (i) ref. 13 for **4a**,**c**; ref. 14 for **4b**; Ac_2O, NaOAc then HBr, AcOH for **4d**; (ii) $NaSSO_2Ph$, CH_3CN, 70 °C, 0.1 equiv. Bu_4NBr for **2a**,**c**, Bu_4NI for **2d**.

NaPTS is more conveniently prepared than sodium methanethiosulfonate.[15] In all cases, syntheses of the glyco-PTS reagents **2** were achieved in superior yields to corresponding syntheses of glyco-MTS reagents **1** (Table 1). Moreover, the costs of the starting materials of these reagents are lower than

† This is one of a number of contributions from the current members of the Dyson Perrins Laboratory to mark the end of almost 90 years of organic chemistry research in that building, as all its current academic staff move across South Parks Road to a new purpose-built laboratory.
‡ Electronic supplementary information (ESI) available: experimental procedures, characterization, protein ESI-MS spectra and crystal data. See http://www.rsc.org/suppdata/ob/b3/b306990g/

DOI: 10.1039/b306990g

Table 1 Comparison of the synthesis and glycosylation reactions of glyco-MTS **1** and glyco-PTS **2** reagents

Glycosylating reagent	Preparation [a]		EtSH (**5**) glycosylation [b]		Peptide (**6**) glycosylation [c]		Protein (**7**, SBL-Cys156) glycosylation [d]		Protein (**8**, BSA-Cys58) glycosylation [d]	
	Total yield (%)	Steps	Yield (%)	Time/ h	Yield (%)	Time/ h	Conversion (%)	Time/ min	Conversion (%)	Time/ min
Glc(Ac)$_4$β-MTS **1a**	46 [e]	3	96 [f]	3	62 [f]	5	100 [e]	50 [e]	—[g]	—
Glc(Ac)$_4$β-PTS **2a**	64	3	82	1	99	5	100	30	100	30
Glc(Bn)$_4$β-MTS **1b**	43 [f]	5	78 [f]	15	65	4	—	—	—	—
Glc(Bn)$_4$β-PTS **2b**	67	5	95	1.5	82	5	—	—	—	—
Gal(Ac)$_4$β-MTS **1c**	47	3	83	1	—	—	—	—	—	—
Gal(Ac)$_4$β-PTS **2c**	65	3	91	1	95	2	100	30	100	30
Glc(Ac)$_4$α(1,4)Glc(Ac)$_3$α(1,4)-Glc(Ac)$_3$β-PTS **2d**	60	3	93	1	74	3	100	30	—	—

[a] From the corresponding parent carbohydrate D-glucose (Glc), D-galactose (Gal) or Glcα(1,4)Glcα(1,4)Glc according to Scheme 2; [b] Et_3N, DCM, RT, 1 equiv. of thiosulfonate; [c] Et_3N, DCM/MeOH (20 : 1), RT, 1 equiv. of thiosulfonate; Peptide [P]-Cys-Ser-OMe, [P] = Ac except for reaction with **2d** where [P] = Boc; [d] 70 mM CHES, 5 mM MES, 2 mM $CaCl_2$ pH 9.5 or 50 mM Tris·HCl, pH 7.7, RT, ~30 equiv. for **1**, ~10 equiv. for **2a**,**c** + **7**, ~20 equiv. for **2a**,**c** + **8**, ~40 equiv. for **2d** + **7**. Conversion was determined by HPLC, A280 and ESI-MS; [e] Taken from refs. 4*b* and 5; [f] Taken from ref. 21; [g] — indicates reaction not studied rather than an absence of reaction.

those for **1** by almost 10-fold.[16] Excellent (>99% β for **2a**,**c**,**d** due to anchimeric assistance of *O*-2 acetate)[17] or workable (3 : 1 α : β for **2b** due to solvent participation)[18] β stereoselectivity was observed. The structure, and importantly the anomeric configuration, of **2a** was further confirmed by X-ray crystallography (Fig. 1).¶

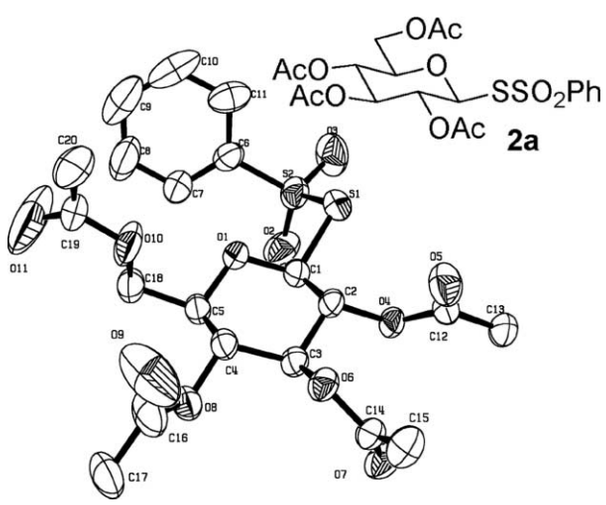
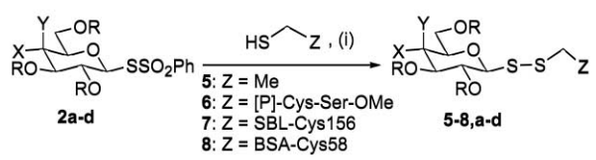


Fig. 1 Crystal structure of **2a**.¶

Glyco-PTS reagents **2a**–**c** were investigated in the glycosylation of representative thiols, peptides and proteins **5**–**8** under basic conditions. We were delighted that in most cases glyco-PTS reagents proved superior to glyco-MTS reagents and yielded glycosylated products **5a**–**c**–**6a**–**c** in 82–99% yields (Scheme 3, Table 1), and in all cases with absolute control of anomeric stereochemistry.



Scheme 3 *Reagents and conditions:* (i) for **5** DCM, Et_3N; for **6** MeOH/DCM, Et_3N; for **7** 70 mM CHES, 5 mM MES, 2 mM $CaCl_2$, pH 9.5; for **8** 50 mM Tris·HCl, pH 7.7.

Notably, these useful glyco-PTS reagents allowed quantitative glycosylation in aqueous buffer solution (HEPES, CHES, MES or Tris pH 7–9.5) of the model single thiol-containing proteins, subtilisin *Bacillus lentus* mutant S156C (**7**, SBL-Cys156) and bovine serum albumin (**8**, BSA-Cys58). The high purity of the glycoprotein products was confirmed by ESI-MS, Ellman's titration[19] and PAGE.

Pleasingly, the glyco-PTS reagents **2a**,**c** exhibited enhanced stability and efficiency under reaction conditions and this allowed the use of reduced ratio of reagent to protein (~10–20 equivalents‖) as compared with the corresponding glyco-MTS **1** reagents (~30 equivalents)[5] and the recovery of unused reagent. Post-glycosylation deprotection of these peracetylated glycoproteins, which has been demonstrated by us previously,[5] also allows access to deacetylated variants.

As a further, more stringent test of the utility of the glyco-PTS method, we investigated the use of the Glcα(1,4)-Glcα(1,4)Glcβ-trisaccharide PTS **2d** (Scheme 2). This allowed the successful glycosylation not only of thiols **5** and **6** but also of protein **7** to create **7d** (Fig. 2). To the best of our knowledge, this is one of the largest carbohydrates to have been used in a convergent and site-selective protein glycosylation to date.[20]

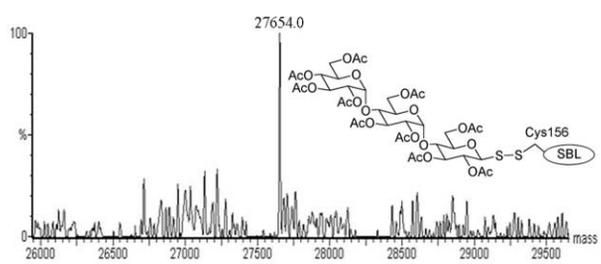


Fig. 2 ESI-MS spectrum of **7d** (Found 27654, Theoretical 27649).

Finally, we explored the potential of the $PhSO_2$– group in glyco-PTS reagents **2a**–**d** as an *in situ* indicator of protein glycosylation. Advantageously the $PhSO_2$– chromophore displays a maximum in the UV spectrum at ~265 nm that is not obscured by either protein or buffer associated chromophores (Fig. 3). The $PhSO_2$– moiety is present in both glyco-PTS reagent and in the $PhSO_2^-$ that is the by-product of glycosylation but the associated extinction coefficients differ sufficiently (*e.g.* $\varepsilon = 2308$ M^{-1} cm^{-1} for **2c**, $\varepsilon = 986$ M^{-1} cm^{-1} for $PhSO_2^-$) for the rate of protein glycosylation to be monitored. Initial results indicate that glycosylations with **2a**–**d** are typically complete within 20 min. Moreover, rates of glycosylation appear to vary with the size (*e.g.* **2d** *vs.* **2a**) of the carbohydrate introduced during protein glycosylation. This interestingly raises the possibility that the steric bulk of the carbohydrate could be used to control the kinetics of protein glycosylation. Further kinetic analyses of these glycosylation reactions are under investigation and will be published in due course.**

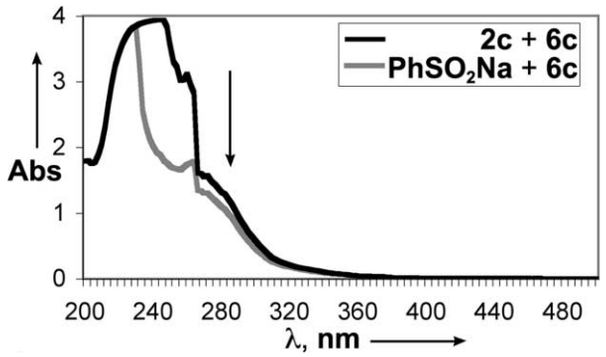


Fig. 3 UV Monitoring of the PTS group: the presence of the 265 nm $PhSO_2$– chromophore in the PTS group allows the monitoring of its presence and hence glycosylation. This is illustrated by UV absorbance spectra of solutions of **2c** (black) and NaPTS (grey) with **6c** (1.85 mM of each component in 1 : 1 pH 7 HEPES : CH_3CN) which show reduced absorbance at 265 nm for the $PhSO_2^-$ anion as compared with glyco-PTS **2c**.

In summary, this communication describes the synthesis of glyco-PTS reagents **2a**–**d** of double utility not only as protein, peptide and thiol glycosylating reagents but also as reagents that may be used for the preparation of glycosyl disulfide glycosyl donors such as **5a**–**d**.[21] Moreover, the successful use of phenylthiosulfonates as protein modifying reagents, demonstrated here for carbohydrate modifications, highlights the potentially broader utility of other PTS reagents as improved variants of their useful MTS counterparts.[6,7]

Acknowledgements

We thank the EPSRC (D. P. G., S. J. W.) and Glycoform for funding (D. P. G., P. G.); Mario Polywka for useful discussions; Barbara O'Dell, Tim D. Claridge, Andrew R. Cowley, Joanna Kirkpatrick for invaluable technical support and the EPSRC for access to the Mass Spectrometry Service at Swansea, the Chemical Database Service at Daresbury and for the award of quota (S. J. W.) and DTA (D. P. G.) studentships.

Notes and references

§ The Gene Ontology Consortium has defined biological process term number GO:0006486 *protein amino acid glycosylation* as "The addition of a sugar unit to a protein amino acid, *e.g.* the addition of glycan chains to proteins." [see *Genome Res.*, 2001, **11**, 1425.]. In this communication we similarly use the term glycosylation to refer to the general process of addition of a glycosyl unit to another moiety *via* a covalent linkage. Glycation has also been suggested by the IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) [see *Eur. J. Biochem.*, 1986, **159**, 1; 1989, **185**, 485; *Glycoconjugate J.*, 1986, **3**, 123; *J. Biol. Chem.*, 1987, **262**, 13; *Pure Appl. Chem.*, 1988, **60**, 1389; *Amino Acids Pept.*, 1990, **21**, 329; and in *Biochemical Nomenclature and Related Documents*, 2nd edition, Portland Press, London, 1992, pp. 84–89] as a general term for the product of all reactions that covalently link a sugar starting molecule to a protein or peptide.

¶ *Crystal data* for **2a**: $C_{20}H_{24}O_{11}S_2$, $M = 504.52$, monoclinic, space group $P2_1$, $a = 10.7470(3)$, $b = 8.3113(2)$, $c = 13.5613(4)$ Å, $\beta = 100.4321(12)°$, $V = 1191.3$ Å^3, $Z = 2$, $T = 150$ K, $\mu = 0.280$ mm^{-1}, reflections measured = 10203, unique reflections = 4759, $R_{int} = 0.024$, $R = 0.0417$, $wR = 0.0465$. CCDC reference number 213324. See http://www.rsc.org/suppdata/ob/b3/b306990g/ for crystallographic data in CIF or other electronic format.

‖ It should be noted that 10–20 equivalents of reagent represents a stoichiometry that is well below that typically used in protein modifications, which are often of the order of 1000 equivalents. See B. G. Davis, *Curr. Opin. Biotechnol.*, 2003, **14**, in press.

** These investigations include the potential use of other thiosulfonates such as *para*-nitrophenylthiosulfonates (*p*NPTS).

1 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
2 B. G. Davis, *Chem. Rev.*, 2002, **102**, 579.
3 T. W. Rademacher, R. B. Parekh and R. A. Dwek, *Annu. Rev. Biochem.*, 1988, **57**, 785.
4 For leading and recent additional examples see: (*a*) K. Witte, P. Sears, R. Martin and C.-H. Wong, *J. Am. Chem. Soc.*, 1997, **119**, 2114; (*b*) B. G. Davis, R. C. Lloyd and J. B. Jones, *J. Org. Chem.*, 1998, **63**, 9614; (*c*) Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman and C. R. Bertozzi, *J. Am. Chem. Soc.*, 1999, **121**, 11684; (*d*) D. Macmillan, R. M. Bill, K. A. Sage, D. Fern and S. L. Flitsch, *Chem. Biol.*, 2001, **8**, 133; (*e*) H. Liu, L. Wang, A. Brock, C.-H. Wong and P. G. Schultz, *J. Am. Chem. Soc.*, 2003, **125**, 1702.
5 B. G. Davis, M. A. T. Maughan, M. P. Green, A. Ullman and J. B. Jones, *Tetrahedron: Asymmetry*, 2000, **11**, 245.
6 G. L. Kenyon and T. W. Bruice, *Methods Enzymol.*, 1977, **47**, 407.
7 R. Wynn and F. M. Richards, *Methods Enzymol.*, 1995, **251**, 351.
8 P. Berglund, G. DeSantis, M. R. Stabile, X. Shang, M. Gold, R. R. Bott, T. P. Graycar, T. H. Lau, C. Mitchinson and J. B. Jones, *J. Am. Chem. Soc.*, 1997, **119**, 5265.
9 G. DeSantis and J. B. Jones, *Curr. Opin. Biotechnol.*, 1999, **10**, 324.
10 K. Matsumoto, B. G. Davis and J. B. Jones, *Chem. Eur. J.*, 2002, **8**, 4129.
11 For select examples see: J. I. Brauman, W. N. Olmstead and C. A. Lieder, *J. Am. Chem. Soc.*, 1974, **96**, 4030; P. J. Stang and A. G. Anderson, *J. Org. Chem.*, 1976, **41**, 781; M. J. Gresser and W. P. Jencks, *J. Am. Chem. Soc.*, 1977, **99**, 6970; D. N. Kevill and G. M. L. Lin, *Tetrahedron Lett.*, 1978, 949; M. J. Pellerite and J. I. Brauman, *J. Am. Chem. Soc.*, 1980, **102**, 5993; A. Arcoria, F. P. Ballistreri, G. Musumarra and G. A. Tomaselli, *J. Chem. Soc., Perkin Trans. 2*, 1981, 221; K. B. Sloan and S. A. M. Koch, *J. Org. Chem.*, 1983, **48**, 3777.
12 R. G. Pearson and R. L. Dillon, *J. Am. Chem. Soc.*, 1953, **75**, 2439.
13 P. G. Scheurer and F. J. Smith, *J. Am. Chem. Soc.*, 1954, **76**, 3224.
14 H. P. Wessel and D. Bundle, *J. Chem. Soc., Perkin Trans. 1*, 1985, 2251.
15 T. G. R. Sato, Y. Takakawa and S. Takizawa, *Synthesis*, 1980, 615.
16 2003 preparation costs: 1 g of NaMTS £11.02, 1 g of NaPTS £1.21.
17 G. Wulff and G. Röhle, *Angew. Chem., Int. Ed. Engl.*, 1974, **13**, 157.
18 A. Marra, J. Esnault, A. Veyrières and P. Sinaÿ, *J. Am. Chem. Soc.*, 1992, **114**, 6354.
19 G. L. Ellman, K. D. Courtney, V. Andres and R. M. Featherstone, *Biochem. Pharmacol.*, 1961, **7**, 88.
20 For one other convergent trisaccharide-protein glycosylation see ref. 4*e*.
21 B. G. Davis, S. J. Ward and P. M. Rendle, *Chem. Commun.*, 2001, 189.