DOI: 10.1002/cbic.201100125

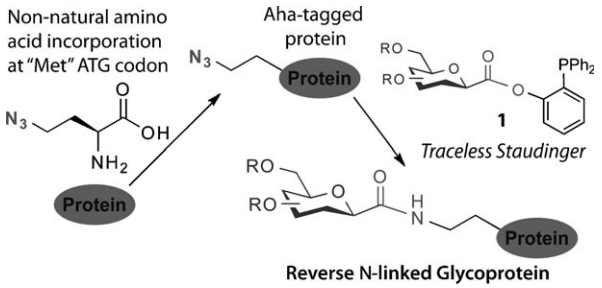
Site-Selective Traceless Staudinger Ligation for Glycoprotein Synthesis Reveals Scope and Limitations

Gonçalo J. L. Bernardes, Lars Linderoth, Katie J. Doores, Omar Boutureira, and Benjamin G. Davis*[a]

Efficient, chemoselective reactions enable access to well-defined post-translationally modified proteins.[1–4] One approach involves the positioning of a reactive moiety within a side chain that may be chemoselectively addressed (a tag-and-modify approach).[5] Among functionalities, the azide group has received attention since it can be installed in biopolymers (e.g., by using auxotrophic or non-natural expression as well as metabolic and enzymatic processes[6,7]) and has been used in installation of post-translational modifications and mimics[8,9] and/or protein labelling[2–4,10] in combination with Huisgen cycloaddition reactions that are Cu^{I} catalysed[11,12] (CCHC) or strain promoted.[13–16]

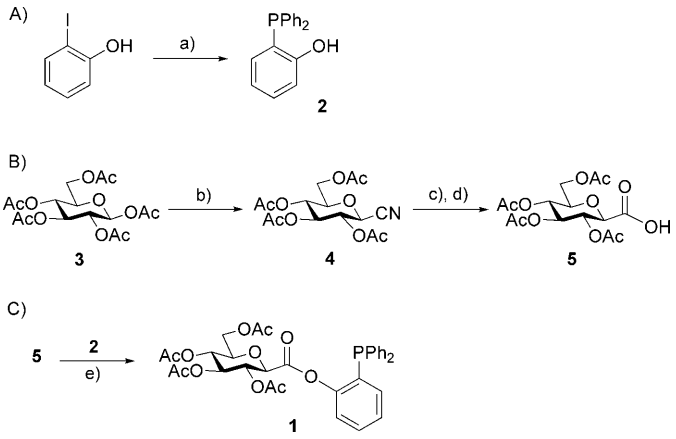
The Staudinger reaction[17] is another azide-selective reaction that has been used, for example, to modify cell surface carbohydrates,[18,19] and to label proteins that contain azidohomoalanine (Aha).[20] The reaction proceeds by attack of phosphine on azide to form an aza-ylid intermediate (Scheme S1 A in the Supporting Information). Intramolecular trapping of the aza-ylid by ester gives amide-linked phosphine oxide after hydrolysis. While the reaction has found widespread application,[18,19] ideally it would produce an amide without an intervening triarylphosphine oxide group. To circumvent this problem, traceless variants[21–25] displace an ester with attached phosphonium group. Hydrolysis of the rearranged adduct produces an amide bond and liberates phosphine oxide (Scheme S1 B in the Supporting Information). It has been used for ligation of protected glyco-amino acid and peptide fragments,[26,27] peptide cyclisation[28] or conjugation of antigens to carriers.[29] Its chemo- and site-selectivity in proteins has been less studied and a single example of selective labelling of proteins containing *p*-azidophenylalanine by using this reaction has been reported.[30] Here, we explore its use in the construction of well-defined, "reverse amide" variants of *N*-linked glycoproteins, an important class of post-translationally modified proteins, involved in many cellular recognition processes.[31]

Our convergent chemical approach produces glycoproteins with a reversed amide linkage between sugar and protein when compared to natural *N*-linked glycans, which are attached in a β-*N*-glycosidic linkage to the carboxamido nitrogen of asparagines (Asn). Such reverse amides can confer additional stability towards hydrolases as previously demonstrated for other natural substrates.[32,33] The construction of glycoproteins bearing this unnatural motif may also circumvent problems encountered in chemoenzymatic syntheses of glycoproteins by using endoglycosidases[34] and provide a linkage more resistant to hydrolysis than natural Asn-linked glycoconjugates.[35] Our strategy is based on a "tag-and-modify" approach[5] and exploits the introduction into proteins of Aha[20] (tag) that may be chemoselectively addressed (modified) using traceless Staudinger conditions mild enough to retain protein activity throughout (Scheme 1).



Scheme 1. "Tag-and-modify" strategy by using the traceless Staudinger reaction for the construction of reverse amide *N*-linked glycoproteins.

Sugar phosphine **1** was designed bearing three key features (Scheme 2). First, the carbohydrate is mounted at the acyl moiety, which is destined for amide bond formation. Second, the acyl component is attached to an aryl group through a cleavable ester linkage. In particular, phenol esters were previously shown to be optimal substrates for the traceless Stau-



Scheme 2. Synthesis of carbohydrate functionalised phosphine **1**. a) Ph_2PH, KOAc, $Pd(OAc)_2$, DMA, 110 °C; b) $HgBr_2$, nitromethane, TMSCN, room temperature, 48 h, 61%; c) NaOMe, MeOH then aq. NaOH (12.5%), reflux; d) Ac_2O, $ZnCl_2$, 60%; e) DCC, DMAP, 110 °C, 70%.

[a] *Dr. G. J. L. Bernardes, Dr. L. Linderoth, Dr. K. J. Doores, Dr. O. Boutureira, Prof. B. G. Davis*
Department of Chemistry, University of Oxford
Chemistry Research Laboratory, 12 Mansfield Road
OX1 3TA Oxford (UK)
Fax: (+44) 1865-285002
E-mail: ben.davis@chem.ox.ac.uk

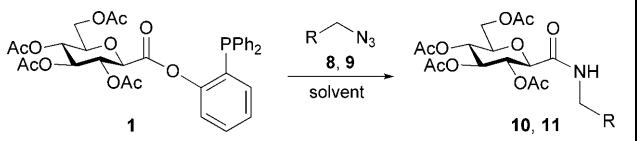
Supporting information for this article including full experimental details, ^{1}H and ^{13}C NMR spectra for all new compounds and ESI-MS for all protein samples is available on the WWW under http://dx.doi.org/10.1002/cbic.201100125.

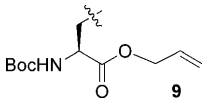
dinger reaction;[21] the rest of the carbohydrate hydroxyls were differentially protected as simpler acetate esters known to allow enhanced uptake of sugar reagents into eukaryotic cells prior to intracellular revelation.[36,37] Finally, two aromatic phosphine substituents are present to impart better stability towards oxidation.

Assembly involved the attachment of phosphine moiety **2** to the carbohydrate by using carbodiimide-mediated esterification. Moiety **2** was prepared through Pd-mediated coupling of 2-iodophenol with diphenylphosphine (Scheme 2A).[38] Next, partially protected guloheptonic acid (**5**) was prepared starting from β-glucosyl cyanide (**4**), which was, in turn, synthesised in 61% yield by using a mild Lewis acid, mercury(II) bromide[39] from **3**. Hydrolysis of nitrile **4** followed by acetylation by using $ZnCl_2$ in Ac_2O[40] afforded **5** in 60% yield (Scheme 2B). Finally, coupling of **5** and **2** by using DCC and DMAP produced the desired carbohydrate-functionalised phosphine **1** in 70% yield (Scheme 2C).

Initial investigations of **1** used model azide coupling partners benzyl azide (**8**) and protected L-azido homoalanine (**9**; see the Supporting Information for synthesis details). Reactions were optimised for solvent system, temperature and duration (Table 1). Conversions were high but yields were more moderate (45% for **8**, entry 4; and 47% for **9**, entry 7) and highly dependent on solvent. Prolonged reaction and DMF proved important. Addition of water was critical for rearrangement of the rather stable, even isolable aza-ylid intermediate,[26,27] to product (Scheme S2 in the Supporting Information). Improvements obtained by using aqueous conditions were clearly of direct beneficial relevance to biomolecules.

Table 1. Model traceless Staudinger glycoconjugation reactions.



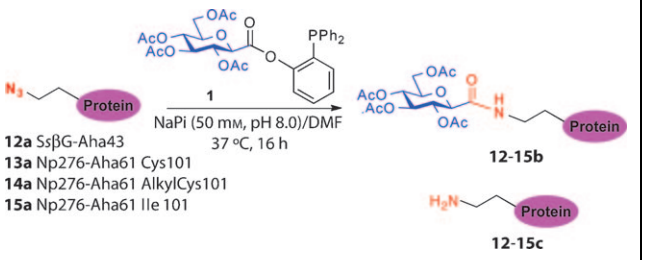
Entry[a]	R	Solvent	T [°C]	t [h]	Yield [%][b]
1	 **8**	DCM	40	30	11
2		$CHCl_3$	70	20	17
3		DMF	70	16	18
4		DMF/H_2O[c]	70	16	45
5	 **9**	$CHCl_3$	70	16	8
6		$CHCl_3$/H_2O[c]	70	20	27
7		DMF/H_2O[c]	70	20	47

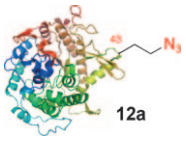
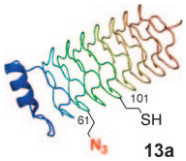
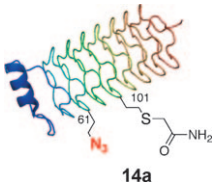
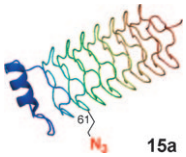
[a] General conditions: typically, sugar phosphine (0.05 mmol) and the azide (0.04 mmol) were dissolved in dry solvent (1 mL) and the reaction was stirred and heated under argon. The reaction was concentrated in vacuo and purified by flash chromatography (EtOAc/petrol); see the Supporting Information for full experimental details; [b] isolated yield; [c] H_2O (0.5 mL) was added to the reaction and the mixture was stirred at 70 °C for 24 h.

tiforme.[42] Non-natural amino acid Aha was site-specifically introduced into these protein systems as an azide "tag" through expression of corresponding gene sequences in an auxotrophic strain of *E. coli* (B834(DE3)).[20] Gene sequences were designed to create proteins displaying azide at sites for which the position could simply be controlled by the "Met" triplet codon ATG. Replacement of wild-type (WT) methionine residues, with the near-isosteric amino acid isoleucine allows reassignment of the codons in the gene sequence and enables incorporation of Aha as a "tag". SsβG is from a generic fold, the TIM barrel,[41] which we have used to create synthetic protein probes for in vitro and in vivo applications.[9]

Use of aqueous media and physiological pH at or near room temperature are important prerequisites for protein stability. Model reactions had required heating and use of DMF as a co-solvent. Since proteins can be stable in up to 50% DMF (*v*/*v*), we started by optimising solvent ratio (buffer/DMF) for solubilising **1** and with reaction at 37 °C (Table 2). Use of SsβG–Aha43

Table 2. Traceless Staudinger reactions on azide-tagged proteins.



Entry[a]	Protein–Aha	1 [equiv]	DMF [%]	Product conv. [%][b]
1	 **12a**	100	10	SM[c]
2		100	20	SM[c]
3		150	20	SM[c]
4		150	40	**12b** (55) **12c** (45)
5		250	40	–[d]
6	 **13a**	150	20	SM[c]
7		150	40	SM[c]
8	 **14a**	150	20	SM[c]
9		150	40	SM[c]
10	 **15a**	150	20	SM[c]
11		150	40	**15b** (45) **15c** (50)
12		250	40	–[d]

[a] See the Experimental Section and the Supporting Information for full details; [b] determined by LC-MS analysis; [c] only the starting azide–protein was detected by LC-MS; [d] protein precipitation.

(**12a**) as a substrate allowed optimization (entries 1–5) of protein ligations; phosphate buffer pH 7.4 with 40% (*v/v*) DMF for 16 h at 37 °C led to complete consumption of **12a** giving the desired product **12b** along with competing hydrolysis of aza-ylid to give aminohomoalanine containing protein **12c** in an about 1:1 ratio (entry 4).[43] DMF (40%, *v/v*) and 150 equiv of **1** were necessary for successful ligation; at these protein concentrations further excess of phosphine resulted in protein precipitation (entry 5). Relatively slow ligation rate, possibly in the intramolecular cyclization step,[21] and hydrolysis of phenol ester can result in competition between product formation and hydrolysis of aza-ylid to amine. While for certain fragile protein substrates the optimised conditions may result in some degree of protein denaturation, the protein used here, SsβG–Aha43 (**12a**), remained enzymatically active (see the Supporting Information).

Having established traceless Staudinger conditions for the successful site-selective formation of reverse amide *N*-linked glycoproteins in SsβG–Aha43, we explored other azide-tagged proteins. Cuboid protein Np276–Aha61 (**13a**), which contained a site-selectively positioned azide tag, was treated by using the optimised procedure and also proved to be stable. However, with this substrate no conversion of the azide was observed (entry 6). This mutant of Np276 has a single Cys at position 101 while our first successful substrate SsβG–Aha43 (**12a**) contained no Cys residues. Given the thiophilicity of phosphines,[44,45] we speculated that the sulfur in the side chain of Cys might interact unfavourably with **1**. To test this hypothesis, and determine whether this "poisoning" might be dependent on thiol or thioether, we alkylated the Cys of **13a** to generate **14a**. Product **14a** was also unreactive under these conditions (entry 8). However, by replacing Cys101 with an isoleucine residue containing no sulfur in its side chain, complete conversion of protein **15a** to products **15b** and **15c** (~1:1) was observed. The possibility that these reactivity differences were due to modulation of reaction site 61 accessibility through changes in the side chain of site 101 from CH_2SH(**13a**)→$CH_2SCH_2CONH_2$-(**14a**)→$CH(CH_3)CH_2CH_3$(**15a**)—whilst unlikely, given their relative remoteness—was also tested. In another azide-modifying reaction, the glyco-CCHC reaction,[8] the azides in **14a** and **15a** were fully reactive (see the Supporting Information). Together, these comparisons suggest that the S found in Cys residues, even as a thioether, is incompatible with conditions developed for phosphine **1**.[46]

In conclusion, we have synthesised and used a specifically modified sugar phosphine for the construction of reverse *N*-linked glycoproteins using the traceless Staudinger reaction. The reaction was applied to different azidoprotein substrates. The azide protein is fully consumed to give equimolar amounts of ligated amide product and reduced amine product. The formation of these novel reverse amide glycoproteins exploits the reactivity and selectivity between azides and phosphines. However, these reactions have also revealed the first examples of an apparent incompatibility with certain other residues: proteins differing only in the presence of free and alkylated Cys residues did not react under identical conditions. The mechanistic origins of this incompatibility are unclear and might be limited to the examples listed here. In initial studies with addition of exogeneous small-molecule thioethers and thiols we have also observed modulation of reactivity.

However, no intermediates or possible side-products (e.g., thioesters from transthioesterification) were identified. Given the increasingly widespread use of the Staudinger reaction in both in vitro[18] and even in vivo[19] biology these findings further delineate both the potential (in forming valuable new bioconjugates, such as the reverse amide glycoprotein isostere made here) and the possible limitations (competing reduction and sulfur interaction). This suggests the need for detailed study in the development of novel functionalised phosphines for this purpose; other phosphines, such as **2**, are capable of reducing azide to amine. Various factors could influence the rate of cyclization to favour amide over amine formation;[25] these include the use of differently substituted aryl substituents[47] or those with altered solubility.[48] The potential of the novel sugar phosphines we have described here, for example, as cell-permeable reagents for intracellular glycoprotein synthesis, is currently under investigation in our laboratory.

Experimental Section

Typical "traceless" Staudinger procedure for protein glycosylation: Sugar phosphine **1** (150 equiv) was dissolved in DMF (80 μL, 40% of total reaction volume) and added to the protein solution (ca. 0.5 mg mL^{-1}; 120 μL) in sodium phosphate buffer (pH 7.4, 50 mM). The mixture was vortexed and shaken at 37 °C for 16 h. The reaction was analyzed directly by LC-MS. Control experiments to determine protein stability in 40% DMF were performed by incubating the protein sample (120 μL) with DMF (80 μL) at 37 °C for 16 h.

Acknowledgements

We thank FCT Portugal, the Royal Society of Chemistry J. W. T. Jones fellowship (G.J.L.B.), and the EC (Marie Curie IEF, O.B.) for financial support. We also thank Prof. John S. Blanchard for providing pET28a/NP276 plasmid and Matthew Young, Dr. James Errey for technical assistance. B.G.D. is a Royal Society Wolfson Merit Award recipient and is supported by an EPSRC LSI Platform Grant.

Keywords: glycoproteins · protein modification · Staudinger ligation · sugar phosphines

[1] B. G. Davis, *Science* **2004**, *303*, 480.
[2] G. J. L. Bernardes, J. M. Chalker, B. G. Davis in *Ideas in Chemistry and Molecular Sciences: Where Chemistry Meets Life* (Ed.: B. Pignataro), Wiley-VCH, Weinheim, **2010**, p. 59.
[3] E. M. Sletten, C. R. Bertozzi, *Angew. Chem.* **2009**, *121*, 7108; *Angew. Chem. Int. Ed.* **2009**, *48*, 6974.
[4] C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, *120*, 10182.
[5] B. G. Davis, *Pure Appl. Chem.* **2009**, *81*, 285.
[6] K. L. Kiick, D. A. Tirrell, *Tetrahedron* **2000**, *56*, 9487.
[7] C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413.
[8] O. Boutureira, F. D'Hooge, M. Fernandez-Gonzalez, G. J. L. Bernardes, M. Sanchez-Navarro, J. R. Koeppe, B. G. Davis, *Chem. Commun.* **2010**, *46*, 8142.

[9] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105.
[10] J. Kalia, R. T. Raines, *Curr. Org. Chem.* **2010**, *14*, 138.
[11] C. W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.
[12] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596.
[13] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046.
[14] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664.
[15] P. V. Chang, J. A. Prescher, E. M. Sletten, J. M. Baskin, I. A. Miller, N. J. Agard, A. Lo, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2010**, *107* 1821.
[16] X. Ning, J. Guo, M. A. Wolfert, G.-J. Boons, *Angew. Chem.* **2008**, *120*, 2285; *Angew. Chem. Int. Ed.* **2008**, *47*, 2253.
[17] J. M. H. Staudinger, *Helv. Chim. Acta* **1919**, *2*, 635.
[18] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007.
[19] J. A. Prescher, D. H. Dube, C. R. Bertozzi, *Nature* **2004**, *430*, 873.
[20] K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 19.
[21] E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141.
[22] B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, *2*, 1939.
[23] B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2001**, *3*, 9.
[24] Y. He, R. J. Hinklin, J. Chang, L. L. Kiessling, *Org. Lett.* **2004**, *6*, 4479.
[25] M. B. Soellner, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2006**, *128*, 8820.
[26] A. Bianchi, A. Russo, A. Bernardi, *Tetrahedron: Asymmetry* **2005**, *16*, 381.
[27] A. Bianchi, A. Bernardi, *J. Org. Chem.* **2006**, *71*, 4565.
[28] R. Kleineweischede, C. P. R. Hackenberger, *Angew. Chem.* **2008**, *120*, 6073; *Angew. Chem. Int. Ed.* **2008**, *47*, 5984.
[29] C. Grandjean, A. Boutonnier, C. Guerreiro, J.-M. Fournier, L. A. Mulard, *J. Org. Chem.* **2005**, *70*, 7123.
[30] M. L. Tsao, F. Tian, P. G. Schultz, *ChemBioChem* **2005**, *6*, 2147.
[31] K. Ohtsubo, J. D. Marth, *Cell* **2006**, *126*, 855.
[32] J. Coleman, L. Kirby, R. Klein, *J. Lipid Res.* **1995**, *36*, 901.
[33] D. Fegley, S. Kathuria, R. Mercier, C. Li, A. Goutopoulos, A. Makriyannis, D. Piomelli, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8756.
[34] J. R. Rich, S. G. Withers, *Nat. Chem. Biol.* **2009**, *5*, 206.
[35] M. Fernandez-Gonzalez, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey, B. G. Davis, *Chem. Sci.* **2010**, 709.
[36] A. K. Sarkar, T. A. Fritz, W. H. Taylor, J. D. Esko, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3323.
[37] A. K. Sarkar, K. S. Rostand, R. K. Jain, K. L. Matta, J. D. Esko, *J. Biol. Chem.* **1997**, *272*, 25608.
[38] O. Herd, A. Heβler, M. Hingst, P. Machnitzki, M. Tepper, O. Stelzer, *Catal. Today* **1998**, *42*, 413.
[39] P. Phiasivongsa, J. Gallagher, C.-N. Chen, P. R. Jones, V. V. Samoshin, P. H. Gross, *Org. Lett.* **2002**, *4*, 4587.
[40] W. Weiser, J. Lehmann, C. F. Brewer, E. J. Hehre, *Carbohydr. Res.* **1988**, *183*, 287.
[41] C. F. Aguilar, I. Sanderson, M. Moracci, M. Ciaramella, R. Nucci, M. Rossi, L. H. Pearl, *J. Mol. Biol.* **1997**, *271*, 789.
[42] M. W. Vetting, S. S. Hegde, K. Z. Hazleton, J. S. Blanchard, *Protein Sci.* **2007**, *16*, 755.
[43] A similar ratio of reduction versus amide formation has been also observed by Schultz and co-workers when labelling a protein containing *p*-azidophenylalanine with a fluorophore; see ref. [30].
[44] G. J. L. Bernardes, E. J. Grayson, S. Thompson, J. M. Chalker, J. C. Errey, F. E. Oualid, T. D. W. Claridge, B. G. Davis, *Angew. Chem.* **2008**, *120*, 2276; *Angew. Chem. Int. Ed.* **2008**, *47*, 2244.
[45] D. Crich, V. Subramanian, M. Karatholuvhu, *J. Org. Chem.* **2009**, *74*, 9422.
[46] Notably, the single report of selective functionalization of a protein containing an azide by using traceless Staudinger conditions was also performed in a Cys-free protein; see ref. [30].
[47] M. B. Soellner, A. Tam, R. T. Raines, *J. Org. Chem.* **2006**, *71*, 9824.
[48] A. Tam, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2007**, *129*, 11421.

Received: February 22, 2011
Published online on May 19, 2011

