# Direct deprotected glycosyl–asparagine ligation†

**Katie J. Doores,[a] Yusuke Mimura,[bc] Raymond A. Dwek,[b] Pauline M. Rudd,[b] Tim Elliott[c] and Benjamin G. Davis*[a]**

***Received (in Cambridge, UK) 1st November 2005, Accepted 31st January 2006***
***First published as an Advance Article on the web 6th March 2006***
**DOI: 10.1039/b515472c**

**A simple and efficient synthesis of *N*-linked glycoamino acids and glycopeptides from deprotected sugars using the Staudinger reaction.**

*N*-Linked and *O*-linked glycoproteins are key to many biological processes including cell–cell communication[1] and the immune response.[2] These glycoproteins exist as glycoforms[3] each with different carbohydrate structures which are key to their biological role.[4] It is important to be able to synthesise single glycoforms of glycoproteins in order to understand these roles more clearly.[5]

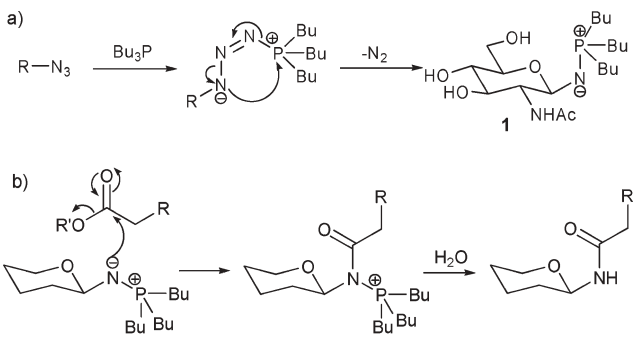
Strategies for the synthesis of *N*-linked glycoproteins include stepwise solid phase peptide synthesis[6] where the *N*-linked asparagine moiety is introduced as a glycoamino acid. Resulting glycopeptide chains can later be utilised in native chemical ligation[7] to afford larger glycoproteins. This strategy requires a method for the synthesis of *N*-linked glycoamino acids with orthogonal *N*- and *C*-terminal protection in a linear strategy. Alternatively, direct coupling of glycosylamines to protected peptides as demonstrated in Lansbury's prescient paper for a single glycosyl amine,[8] allows a flexible convergent approach. We describe here a simple one-pot method which utilises stable, deprotected sugars which can be applied to amino acids and larger peptide fragments.

Current methods for the synthesis of the amide link between glycosyl and asparagine moieties typically utilise glycosylamines with peptide coupling reagents to yield the *N*-linked glycoamino acid.[9] Glycosylamines are accessible by the Kochetkov reaction where a fully protected reducing sugar or deprotected sugar is treated with 50 times excess of ammonium bicarbonate for 6 days.[10] Bejugam *et al.*[11] have recently shown that microwave irradiation can accelerate the Kochetkov reaction which forms the glycosyl amine. Wen *et al.*[12] used a two step synthesis where reduction of a deprotected sugar anomeric azide catalysed by palladium on carbon to access the amine was followed by reaction with an activated ester of aspartic acid. Lansbury has been able to extend this methodology in the convergent synthesis of glycopeptides.[8] However, glycosyl amines are relatively unstable[13] and the hemiaminal is prone to mutarotation to give a mixture of anomers thus making stereocontrol difficult.[14] Diglycosylamines are also observed as by-products.[11,15] Ishiwata *et al.* have been able to react deprotected Alloc glycosyl amines to form the *N*-linkage.[16] However, this requires the use of unnatural acid fluoride functionalised aspartic acid residues.

Methods that do not involve an anomeric amine are therefore advantageous and have been investigated although many of these are over-complicated and/or use only protected sugars and therefore require further manipulation of the sugar which can lead to racemisation of the amino acid.[17] The simplest is that developed by Inazu *et al.*[18] which used a three-component Staudinger reaction of a carboxylic acid and an azide in the presence of a tertiary phosphine to afford an *N*-linked glycoamino acid from protected sugar moieties in yields ranging from 23–77% depending significantly on the solvent. He *et al.*[19] used a two component Staudinger ligation where an activated ester and phosphine exist in the same molecule. However, this requires multiple synthetic steps to synthesise the phosphinothioesters required for the ligation as well as only being useful for protected sugars. Similar to this is the work of Bianchi *et al.*,[20] which requires a multiple step synthesis of a functionalised phosphine from *o*-diphenylphosphinophenol. Bertozzi *et al.* have also used a 'traceless' Staudinger Ligation for the chemoselective synthesis of amide bonds where a methoxycarbonyl group on one of the aryl rings of a phosphine traps the aza-ylid intermediate and the phosphine oxide is removed by subsequent hydrolysis.[21]

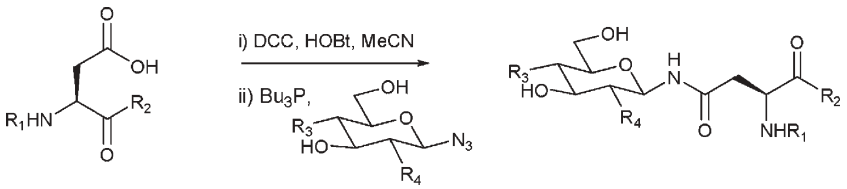
To simplify and enhance efficient access to glycosyl asparagines we desired a novel reaction i) for deprotected substrates, ii) compatible with linear and convergent approaches, iii) that is stereocontrolled and iv) free from the need for complex auxilliaries.

Having surveyed the literature, it was felt that the three component Staudinger (Scheme 1) was a promising route for use with deprotected sugars as it requires no specialist activation or protection. A range of non-polar solvents were surveyed. However, several such as DCM were incompatible with




**Scheme 1** a) Intermediate observed, b) suggested mechanism for Staudinger Ligation.

[a]*Department of Chemistry, Oxford University, Chemistry Research Laboratory, Mansfield Road, Oxford, UK OX1 3TA. E-mail: ben.davis@chem.ox.ac.uk; Fax: +44 (0)1865 275674; Tel: +44 (0)1865 275654*
[b]*Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK OX1 3QU*
[c]*Cancer Sciences Division, Southampton General Hospital, University of Southampton, Tremona Road, Southampton, UK SO16 6YD*
† Electronic supplementary information (ESI) available: Experimental section. See DOI: 10.1039/b515472c




**Scheme 2** Synthesis of *N*-linked glycoamino acids using Staudinger methodology.

**Table 1** Examples of the Staudinger Ligation

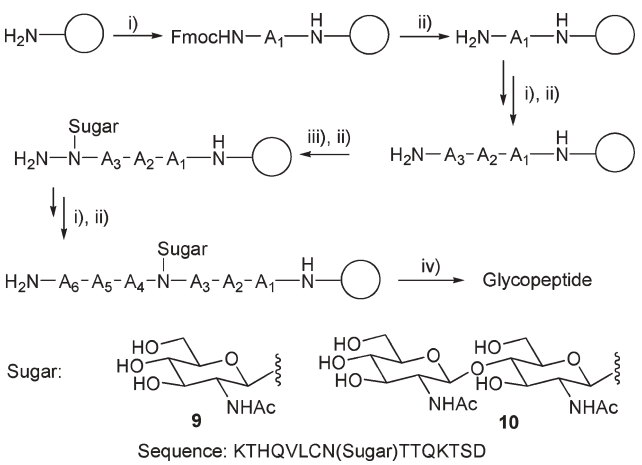
| Entry | $R_1$ | $R_2$ | $R_3$ | $R_4$ | Solvent | Time (h) | Product | Yield % |
|---|---|---|---|---|---|---|---|---|
| a | Fmoc | OtBu | OH | NHAc | MeCN | 16 | **2** | 87 |
| b | Fmoc | OtBu | OH | NHAc | DMF | 16 | **2** | 30 |
| c | Fmoc | OtBu | OH | NHAc | DMF | 72 | **2** | 60 |
| d | Fmoc | OtBu | OH | NHAc | Dioxane | 16 | **2** | 25 |
| c | Fmoc | OAllyl | OH | NHAc | MeCN | 16 | **3** | 62 |
| d | Fmoc | OtBu | OH | OH | MeCN | 16 | **4** | 50 |
| e | Boc | OBn | OH | NHAc | MeCN | 16 | **5** | 62 |
| f | Boc | NHPh | OH | NHAc | MeCN | 16 | **6** | 53 |
| g | Boc | NHPh | OH | OH | MeCN | 16 | **7** | 47 |
| h | Fmoc | OtBu | GlcNAc(β1- | NHAc | MeCN | 16 | **8** | 61 |
| i | Fmoc-S- | -L-T-$NH_2$ | OH | NHAc | DMF | 72 | **12** | 77 |

deprotected sugars due to solubility problems; reactions were therefore attempted in acetonitrile under the same conditions using tributylphosphine but no product was observed. Careful analysis of the crude reaction mixture revealed the presence of intermediate **1** (Scheme 1a). This aza-ylid **1** appears to be stable and yet unreactive in acetonitrile perhaps due to intramolecular hydrogen bonding. Addition of a base to encourage reaction did not lead to the desired product.

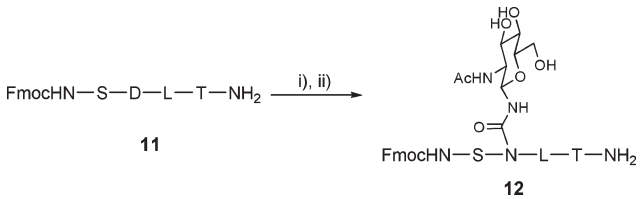
The intended Staudinger reaction proceeds through a putative nitrogen nucleophile; we reasoned that carboxyl preactivation would enhance electrophilicity (Scheme 1b). *N*-α-Fmoc-protected L-aspartic acid α-*tert*-butyl ester was first stirred with DCC and HOBt and then the monosaccharide GlcNAc azide was added followed by tributyl phosphine. This novel direct one pot three component approach gave the desired product **2** in 87% yield (Scheme 2, Table 1) allowing direct access to anomerically pure β-*N*-linked glycoamino acids from simple, stable starting materials. Changing the solvent slowed the reaction rate giving the product in only 30% (DMF), 25% (dioxane) after 16 h but 60% yield after 72 h.

Once the conditions for the Staudinger Ligation had been optimised for the GlcNAc *N*-linked glycoamino acid **2**, we applied this method to other sugar azides, 1-azido glucose and the disaccharide chitobiose, to afford **4** and **8** in 50% and 61% yield respectively (Scheme 2). Interestingly, generally lower yields were observed for glucose (Glc) and it cannot be discounted that the C-2 NHAc may play a role in enhancing the reactivity of the aza-ylid **1**. Alternative *C*- and *N*-protected amino acids were also used to explore compatibility with ester, amide and carbamate functionality. These successfully gave glycoaminoacids **3**, **5**, **6** and **7** in 62%, 62%, 53% and 47% yields respectively.

To access biologically relevant glycopeptides we adopted two approaches: linear (insertion of a glycoamino acid into a growing peptide chain) and convergent (direct attachment of sugar to polypeptide).[5] Candidate epitopes of the tumour antigen gp90, a murine leukaemia virus envelope protein produced by colorectal tumour cells,[22] were chosen because they were known to bind to MHC class II. Glycoamino acids **2** and **8** were used in the stepwise solid phase synthesis of 15 mer glycopeptide epitopes of gp90 (Scheme 3). The *t*-butyl ester groups were removed in quantitative yield using TFA to afford the necessary deprotected building blocks. The first amino acid was loaded onto a rink amide resin using PyBOP, HOBt and DIPEA. Deprotection was achieved using piperidine in DMF followed by sequential coupling and deprotection of the appropriately protected amino acids again using PyBOP, HOBt and DIPEA (Scheme 3). For coupling of the glycoamino acid and the two adjacent amino acids, it was found that double coupling rounds and use of the coupling agent HATU ensured completion. The product 15 mer glycopeptides, **9** and **10**, were finally cleaved from the resin using TFA/TIS/$H_2O$ and obtained in 59% and 56% overall yields respectively. These GlcNAc and chitobiose containing glycopeptides **9** and **10** were




**Scheme 3** i) Fmoc-A-OH, PyBOP, HOBt, DIEA, DMF, ii) 20% piperidine, DMF, iii) glycoamino acid, HATU, HOBt, DIEA, DMF, iv) TFA/TIS/$H_2O$.




**Scheme 4** i) HBTU, HOBt, DMF, ii) $GlcNAcN_3$, $Bu_3P$, 77% over 2 steps.

used to probe the recognition site in gp90 by gp90-specific T cells; interestingly initial results suggest that T cell recognition is independent of glycosylation.[23]

As well as being compatible with such linear strategies, the straightforward method outlined here is also compatible with the convergent synthesis of *N*-linked glycopeptides. A 4 mer glycopeptide epitope of gp90 was targeted (Scheme 4). Although acetonitrile gave higher yields in the case of single amino acids, DMF was the preferred solvent due to the higher solubility of the peptides used. Tetrapeptide **11** was synthesised on the solid phase using standard conditions as described above and cleaved from the resin using TFA/TIS/$H_2O$. The carboxylate side chain of the aspartic acid was activated first with HBTU and HOBt followed by addition of GlcNAc azide and $Bu_3P$ (Scheme 4) giving the desired glycopeptide **12** in 77% yield.[24]

In conclusion, we have developed a novel straightforward, one-pot method for the synthesis of *N*-linked glycoamino acids from deprotected sugars with complete stereocontrol and applied this to a range of sugar and peptide functionality allowing the linear and convergent synthesis of biologically relevant *N*-linked glycopeptides.

We would like to thank the Leverhulme Trust for funding and Dr Marie Squire for technical assistance.

## Notes and references

1 T. Okajima and K. D. Irvine, *Cell*, 2002, **111**, 893.
2 J. B. Lowe, *Cell*, 2001, **104**, 804.
3 T. W. Rademacher, R. B. Parekh and R. A. Dwek, *Annu. Rev. Biochem.*, 1988, **57**, 785; P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker and R. A. Dwek, *Biochemistry*, 1994, **33**, 17.
4 A. Varki, *Glycobiology*, 1993, **3**, 97; R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
5 B. G. Davis, *Chem. Rev.*, 2002, **102**, 579.
6 E. Meinjohanns, M. Meldal, H. Paulsen, R. A. Dwek and K. Bock, *J. Chem. Soc., Perkin Trans. 1*, 1998, 549.
7 P. E. Dawson, T. W. Muir and I. Clark-Lewis, *Science*, 1994, **266**, 776.
8 S. T. Cohen-Anisfeld and P. T. Lansbury, *J. Am. Chem. Soc.*, 1993, **115**, 10531.
9 For examples see: I. L. Deras, K. Takegawa, A. Kondo, I. Kato and Y. C. Lee, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 1763; I. Matsuo, Y. Nakahara, Y. Ito, T. Nukada, Y. Nakahara and T. Ogawa, *Bioorg. Med. Chem.*, 1995, **3**, 1455.
10 L. Likhosherstov, O. Novikova, V. A. Derveitskaja and N. K. Kochetkov, *Carbohydr. Res.*, 1986, **146**, C1–C5.
11 M. Bejugam and S. L. Flitsch, *Org. Lett.*, 2004, **6**, 4001.
12 S. Wen and Z. Guo, *Org. Lett.*, 2001, **3**, 3773.
13 S. J. Danishefsky, S. Hu, P. F. Cirillo, M. Eckhardt and P. H. Seeberger, *Chem.–Eur. J.*, 1997, **3**, 1617.
14 M. Amadori, *Atti. Accad. Nazl. Lincei.*, 1925, **2**, 337.
15 D. Vetter and M. A. Gallop, *Bioconjugate Chem.*, 1995, **6**, 316.
16 A. Ishiwata, M. Takatani, Y. Nakahara and Y. Ito, *Synlett*, 2002, **4**, 634.
17 P. Sjolin, M. Elofsson and J. Kihlberg, *J. Org. Chem.*, 1996, **61**, 560; H. Kunz, *Angew. Chem., Int. Ed. Engl.*, 1987, **26**, 294.
18 T. Inazu and K. Kobayashi, *Synlett*, 1993, 869.
19 Y. He, R. J. Hinklin, J. Chang and L. L. Kiessling, *Org. Lett.*, 2004, **6**, 4479.
20 A. Bianchi, A. Russo and A. Bernardi, *Tetrahedron: Asymmetry*, 2005, **16**, 381.
21 E. Saxon, J. I. Armstrong and C. R. Bertozzi, *Org. Lett.*, 2000, **2**, 2141; E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007.
22 D. Golgher, F. Korangy, B. Gao, K. Gorski, E. Jaffee, M. Edidin, D. M. Pardoll and T. Elliott, *J. Immunol.*, 2001, **167**, 147.
23 Y. Mimura, Y. Mimura-Kimura, K. J. Doores, B. G. Davis, R. A. Dwek, P. M. Rudd and T. Elliot, paper submitted.
24 Based on HPLC analysis.