Synthesis of *N*-linked glycopeptides *via* solid-phase aspartylation†

Trent Conroy, Katrina A. Jolliffe and Richard J. Payne*

Received 26th February 2010, Accepted 24th May 2010
First published as an Advance Article on the web 21st June 2010
DOI: 10.1039/c003673k

An efficient strategy for the preparation of *N*-linked glycopeptides is described. The method relies on the use of side chain protecting groups on aspartic acid residues, namely the allyl and Dmab esters, which are orthogonal to those utilised in Fmoc-strategy SPPS. After peptide assembly these protecting groups were selectively removed and the resulting free side chains derivatised with a glycosylamine to afford a resin bound glycopeptide bearing a native *N*-linkage. Initially, *N*-linked glycopeptides were successfully synthesised according to this strategy, however, yields varied substantially depending on the nature of the amino acid residue situated adjacent (*C*-terminal) to the putative glycosylation site. This was due to generation of substantial quantities of aspartimide by-products. Aspartimide formation was overcome by incorporation of a 2,4-dimethoxybenzyl (Dmb) backbone amide protecting group on the residue adjacent to an allyl- or Dmab-protected aspartic acid residue. *N*-linked glycopeptides were prepared in excellent yield after the solid-phase aspartylation reactions. The utility and orthogonality of the allyl and Dmab ester solid-phase approaches were exploited in the preparation of an *N*-linked glycodecapeptide bearing two different carbohydrate moieties. This exemplified the efficiency of the solid-phase methodology for the preparation of glycopeptides bearing various combinations of *N*-linked glycans.

Introduction

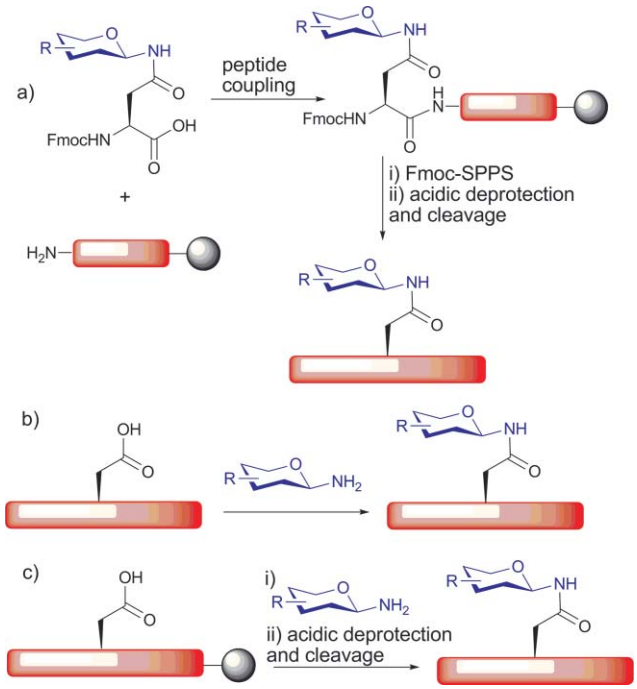
The majority of proteins in the human body are known to be post-translationally modified. The most ubiquitous modification is glycosylation, which introduces an enormous level of structural diversity to proteins.[1] It has been reported that over 50% of all human proteins are glycosylated, a modification known to be essential for a variety of key biological processes including protein folding, transport, cell adhesion, cell differentiation and cell growth.[2-4] In addition, glycosylation plays a critical role in the regulation of biological half life and the activity of many hormones by maintaining structural integrity and by protecting the peptide from proteolytic degradation.[5,6] Aberrant glycosylation of proteins is known to perturb intracellular recognition and, as such, has been implicated in a number of serious illnesses including cancer and autoimmune diseases. For example, in tumour progression there are remarkable changes in cell surface carbohydrate profiles which appears to be associated with the state of metastasis.[7,8] As they frequently serve as cell differentiation markers, glycoproteins are involved in binding of pathogens to cells. Indeed, many pathogens have evolved to use heavily glycosylated, membrane bound proteins as a point of entry into host cells.[9,10]

Broadly speaking, almost all native protein glycosylation can be classified into two types: *O*-glycosides, whereby a glycan is α- or β-linked to the hydroxyl of serine, threonine or tyrosine, or *N*-glycosides, in which *N*-acetylglucosamine is β-linked to the amide side chain of asparagine present within an Asn-Xaa-Ser/Thr consensus sequence. Despite the importance of glycoproteins for a plethora of important biological processes, their detailed study has been hampered by the fact that they are difficult to access in pure form. This is a result of the glycosylation process which, unlike protein synthesis, is not under templated control. Instead, glycoproteins are generated as mixtures of glycoforms, with the glycosylation patterns dependent on the relative activities of a number of cellular glycosyltransferases. These glycoforms are inseparable in most cases and, as such, recombinant methods can not normally be used for the production of homogeneous glycoproteins. It is currently accepted that synthetic chemistry can be used to access homogeneous *O*- and *N*-linked glycopeptides and glycoproteins.[11–23]

We were interested in the preparation of *N*-linked glycopeptides for use in the fragment-based assembly of homogeneous glycoproteins. The most common method for the preparation of *N*-linked glycopeptides is *via* the so-called cassette approach, whereby preformed per-*O*-protected glycosylamino acid building blocks are first synthesised and incorporated into the growing peptide backbone during solid-phase peptide synthesis (SPPS) (Scheme 1a).[24–29] Synthesis of the glycopeptide is generally achieved *via* the Fmoc-strategy due to the vulnerability of glycosidic linkages in common oligosaccharides to the strongly acidic conditions employed under the Boc-strategy. A major drawback of the cassette approach is the greatly reduced efficiency of peptide coupling reactions, both during glycosylamino acid insertion and subsequent peptide elongation. This is due primarily to the sterically hindered nature of amino acids bearing large glycans.[30] The lack of convergence of the cassette-based approach can also result in the loss of the precious building block over the iterative coupling and deprotection steps. An alternative, more convergent approach was developed by Lansbury and Cohen-Anisfeld.[31,32] This method, dubbed the Lansbury Aspartylation, has been used

School of Chemistry, The University of Sydney, NSW 2006, AUSTRALIA. E-mail: richard.payne@sydney.edu.au; Fax: +61 2 9351 3329; Tel: +61 2 9351 5877

† Electronic supplementary information (ESI) available: Additional experimental procedures and analytical data for all novel compounds and *N*-linked glycopeptides. See DOI: 10.1039/c003673k



Scheme 1 Synthesis of *N*-linked glycopeptides *via*: a) coupling of preformed glycosylamino acid cassettes; b) solution phase coupling of glycosylamines ("Lansbury Aspartylation"); c) solid-phase coupling of glycosylamines (solid-phase aspartylation).

for the synthesis of a number of complex glycopeptides.[33–37] The reaction involves coupling of a glycosylamine with a free aspartic acid residue on a peptide backbone and is generally achieved using standard peptide coupling reagents with a suitable base (Scheme 1b). The major disadvantage of this method is the need for extensive protection of the peptide backbone in order to prevent unwanted side reactions such as the formation of aspartimides during aspartic acid side chain activation.[31] Additional protecting groups orthogonal to Fmoc-strategy SPPS are also necessary to allow for the selective modification of aspartic acid residues in the presence of other aspartic acid and glutamic acid residues within the peptide or protein sequence.[21,33,34,37]

The solid-phase preparation of *N*-linked glycopeptides utilising a late stage aspartylation step represents a less common approach (Scheme 1c).[38–42] Although on-resin aspartylation represents a more convergent approach toward the synthesis of glycopeptide fragments, significant problems have been encountered with regards to unwanted aspartimide formation during both peptide elongation and aspartic acid side chain activation during the glycan coupling step. Although aspartimide suppression has been achieved with glycine *C*-terminal to the glycosylation site through the use of the 2-hydroxy-4-methoxybenzyl (Hmb) backbone amide protecting group,[39,43,44] limitations arise from its use. While this manuscript was in preparation, Tolbert and Chen reported an elegant strategy for the solid-phase preparation of glycopeptides bearing a high mannose *N*-linked glycan which could be achieved with reduced aspartimide formation compared with previous approaches.[45] Specifically, this method relied on the use of a 2-phenylisopropyl-protected aspartic acid residue which could be liberated selectively with weak acid and aspartylated on the solid phase with a glycosylamine. In the present work, we were interested in the development of a convergent solid-phase aspartylation strategy for the high yielding preparation of *N*-linked glycopeptides which would completely avoid the formation of aspartimide by-products.

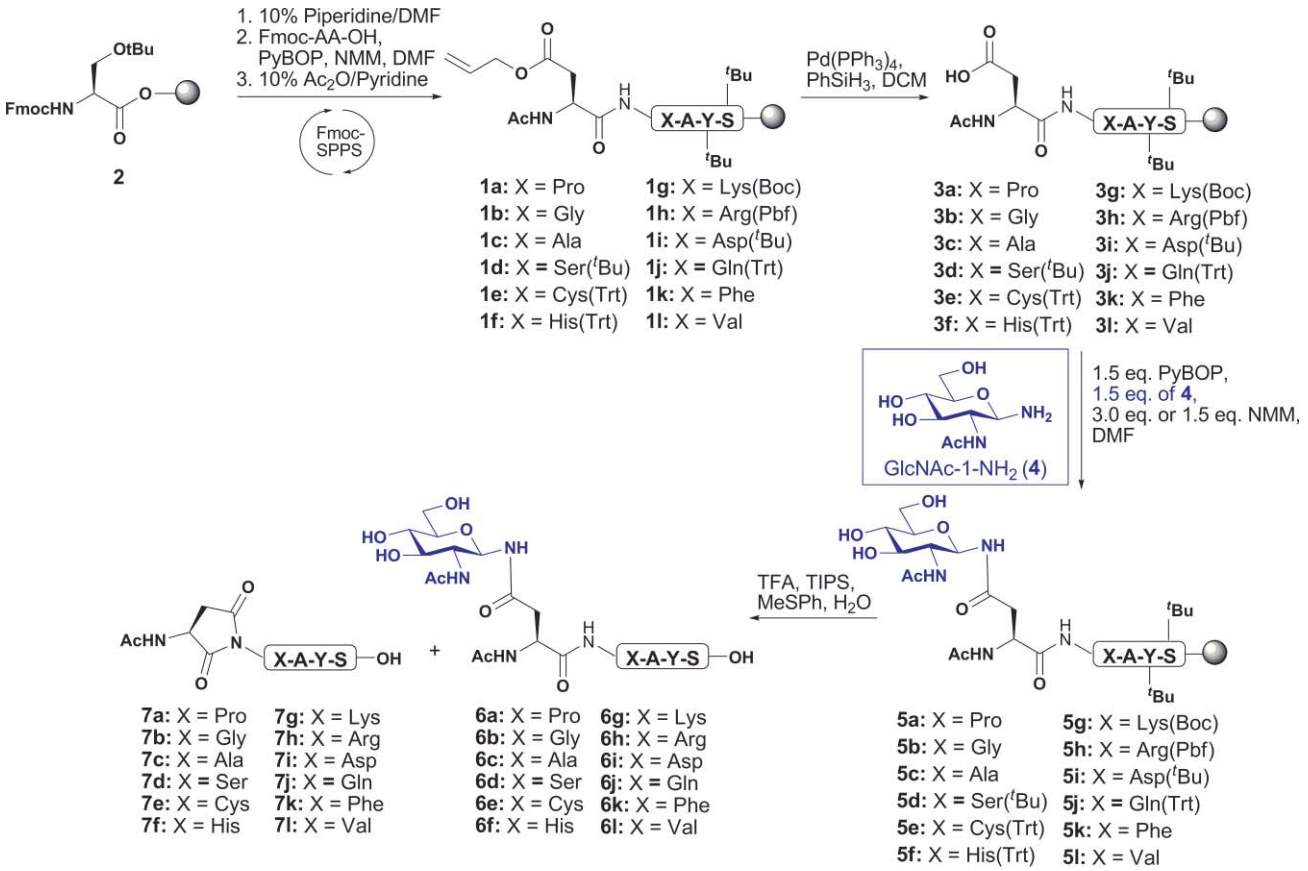
Results and Discussion

Results of previous solid-phase aspartylation studies suggest that neighbouring amino acids play a critical role in the outcome of the reactions and the extent of unwanted by-products such as aspartimides.[39,40] To date there have been no detailed investigations into the effect of amino acid substitution *C*-terminal to the proposed aspartylation site (to be modified with a carbohydrate) on the solid-phase. The initial goal of this study therefore entailed the first detailed investigation of the role of neighbouring residues on solid-phase aspartylation reactions by substituting with a range of amino acids *C*-terminal to the putative aspartylation sites in a model peptide system.

To this end, the protected pentapeptides **1a–1l** were synthesised from preloaded Wang resin **2** using Fmoc-strategy SPPS (Scheme 2). After assembly of the protected resin bound tripeptide (AYS) from **2**, the resin was split and a variety of amino acids coupled as the penultimate residues. Eleven amino acids were chosen as a representative selection of the 20 proteinogenic building blocks. The final amino acid incorporated into the growing peptides was commercially available Fmoc-Asp(OAll)–OH. This residue represented the putative glycosylation site and, as such, was incorporated with the β-carboxy side chain derivatised with an orthogonal protecting group. It was proposed that chemoselective removal of the allyl ester using Pd(0) would selectively liberate the carboxylate side chain for the ensuing aspartylation reaction on the solid phase.[46,47] Following Fmoc deprotection of the fully assembled pentapeptide, the *N*-terminus was acetylated with acetic anhydride in pyridine to afford the fully protected, resin bound pentapeptides **1a–1l**.

Deprotection of the allyl ester from the fully assembled resin bound pentapeptides **1a–1l** was achieved by treating the resin bound peptide with tetrakis(triphenylphosphine) palladium(0) in the presence of phenylsilane as a scavenger to afford **3a–3l** (Scheme 2).[48] Following deallylation, the resin was treated with Py-BOP, 2-acetamido-2-deoxy-β-D-glucopyranosylamine (GlcNAc-1-NH_2) **4** (prepared from *N*-acetylglucosamine *via* Kotchetkov amination[49]) and *N*-methylmorpholine (NMM) in DMF to generate the desired resin bound *N*-linked glycopeptides **5a–5l**. Following this, deprotection and cleavage of the glycopeptides from resin was achieved using an acidic cocktail. Purification was accomplished by preparative reverse-phase HPLC to afford the desired glycopeptides **6a–6l** and the corresponding aspartimides **7a–7l** which, as anticipated, were produced in highly variable yields depending on the nature of the penultimate amino acid residue (Table 1).

The results from this study can be rationalised by assessing the steric and electronic characteristics of the amino acid preceding the aspartic acid residue. For example, under standard conditions (3 equivalents of NMM relative to the peptide) glycine, the amino acid with lowest steric bulk in its side chain, gave rise to the greatest levels of aspartimide formation (98%, entry 2, Table 1). In contrast, incorporation of a penultimate valine residue, bearing a sterically hindered side chain, led to much lower



Scheme 2 Synthesis of *N*-linked glycopeptides **6a–6l** *via* solid-phase aspartylation.

Table 1 Yields of *N*-linked glycopeptides **6a–l** and aspartimides **7a–l**

		3 eq. NMM		1.5 eq. NMM	
Entry	X	Product (**6**) yield (%)	Aspartimide (**7**) yield (%)	Product (**6**) yield (%)	Aspartimide (**7**) yield (%)
1	Pro	91	0	ND	ND
2	Gly	2	98	48	52
3	Ala	51	45	82	10
4	Ser	83	14	77	4
5	Cys	65	28	45	4
6	His	77	20	83	8
7	Lys	72	16	84	6
8	Arg	80	15	77	11
9	Asp	42	53	87	11
10	Gln	81	17	65	20
11	Phe	81	14	92	4
12	Val	87	7	90	8

levels of aspartimide formation (7%) and therefore the desired glycopeptide **6l** was provided in high yield (87%, entry 12, Table 1). Resin bound peptides containing penultimate amino acid residues including serine (**1d**), histidine (**1f**), lysine (**1g**), arginine (**1h**), glutamine (**1j**) and phenylalanine (**1k**) gave good yields of *N*-linked glycopeptides after solid-phase aspartylation and purification (77–83%), however, unacceptable levels of aspartimides (14–20%) were generated in all of these cases (Table 1). High levels of aspartimide formation was observed when aspartic acid was found adjacent to the glycosylation site (**1i**), which can perhaps be attributed to increased acidity of the backbone amide.[50] As such, the corresponding glycopeptide **6i** was only isolated in 42% yield with 53% of the aspartimide byproduct **7i** generated in the process (entry 9, Table 1). Glycopeptide **6a** was provided in high yield (91%) from resin bound peptide **1a** with no aspartimide related byproducts observed (entry 1, Table 1). This result can be explained by the presence of a tertiary amide adjacent to the glycosylation site which renders the backbone amide non-nucleophilic. Using the results of this study it is possible to extrapolate likely outcomes for other amino acids based on their steric and electronic properties. However, given the unacceptable levels of aspartimides formed during the solid-phase aspartylation reactions of most peptides studied here it was necessary to modify the conditions in order to suppress aspartimide formation.

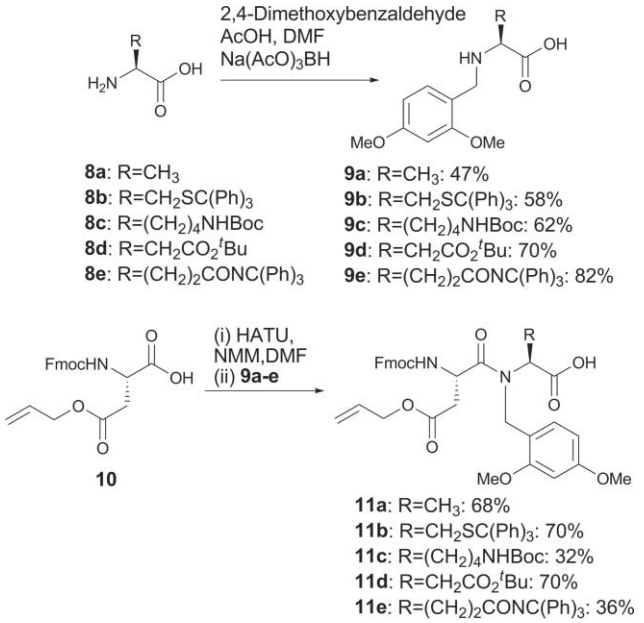
Aspartimide formation is known to be intimately linked with the quantity of base used in the reaction.[31] As such, it was proposed that lowering the number of equivalents of base used during the aspartylation reactions would reduce the propensity of the peptides to form aspartimides, thereby improving the efficiency of the solid-phase aspartylation reactions. To this end, reactions were repeated for peptides **1a–1l**, this time reducing the amount of NMM to 1.5 equivalents relative to the resin bound peptide (Table 1). Following solid-phase aspartylation with glycosylamine **4**, the desired glycopeptides were deprotected and cleaved from the resin as previously described. Subsequent purification by preparative reverse-phase HPLC then afforded the desired glycopeptides. Gratifyingly, the reduction in base during solid-phase aspartylation diminished aspartimide formation for all amino acid substitutions. The effect was particularly profound

in the case of glycine, alanine, and aspartic acid where almost 50% reduction in aspartimide formation was observed (entries 2, 3 and 9, Table 1). Resin bound peptides **1d–1g** bearing penultimate serine, cysteine, histidine, lysine, and **1k** bearing a phenylalanine residue exhibited less than 10% aspartimide formation during the solid-phase aspartylation and, as such, these conditions represent significant improvements to the methodology. However, although these modified conditions minimised aspartimide formation, the efficiency of the aspartylation reaction to install the glycan proved to be less efficient. As a result, the yields of glycopeptides obtained from these reactions were lower than those achieved previously *e.g.* glycopeptides bearing a penultimate serine (**6d**) cysteine (**6e**) or glutamine residue (**6j**) (45–77% yields *cf.* 65–83% using 3 eq. of NMM). Finding an alternative method for the prevention of aspartimides, whilst providing high solid-phase aspartylation yields, was therefore essential for the efficient preparation of *N*-linked glycopeptides as was the goal of this research.

The introduction of a backbone amide protecting group on the amino acid directly preceding the putative glycosylation site was therefore proposed to prevent aspartimide formation and improve the efficiency of the solid-phase aspartylation methodology for the preparation of *N*-linked glycopeptides. We chose the 2,4-dimethoxybenzyl (Dmb) protecting group as an alternative to the commonly employed Hmb group for the protection of the backbone amides. The Dmb group was chosen by virtue of its acid lability allowing concomitant cleavage of the protecting group after generation of the resin bound glycopeptides without further modification (*cf.* Hmb). With the exception of Fmoc-(Dmb)Gly-OH, Dmb-protected amino acids were not commercially available and therefore needed to be synthesised. The N^{α}-protection of amino acids was accomplished *via* reductive amination of 2,4-dimethoxybenzaldehyde with the appropriately side chain protected amino acids. To this end, treatment of amino acids **8a–8e** with 2,4-dimethoxybenzaldehyde and sodium triacetoxyborohydride in the presence of acetic acid afforded the desired Dmb-protected amino acids **9a–9e** in 47–82% yield (Scheme 3).

In order to increase the convergency of our approach, we chose to incorporate **9a–9e** into the resin bound peptides as preformed dipeptides. This would allow for a faster, more efficient solution-phase coupling to the Dmb-secondary amine. In addition, it would eliminate the need for Fmoc protection and deprotection of the Dmb-protected derivatives in subsequent coupling steps on the solid support. To this end, careful pre-activation of Fmoc-Asp(OAll)-OH (**10**) with HATU and NMM followed by addition of the Dmb-protected amino acids **9a–9e** gave the desired dipeptides **11a–11e.** Unfortunately, in our hands the dipeptides could not be easily purified by silica chromatography, however reverse-phase HPLC afforded the desired building blocks in moderate to high yields in most cases.

Due to the commercial availability of Fmoc-(Dmb)Gly-OH (**12**) this was not incorporated as a dipeptide, but rather was coupled as an individual amino acid followed by Fmoc-Asp(OAll)-OH (**10**) to resin bound tripeptide **13** to give **14a** (Scheme 4). Dipeptides **11a–11e** were coupled to **13** using HATU, to afford the desired resin bound pentapeptides **14b–14f**, possessing a Dmb backbone amide protecting group on the penultimate residue (Scheme 4). Gratifyingly the dipeptides could be coupled to **13** in quantitative yield in 2 h (as determined by measuring the piperidine–fulvene adduct at 301 nm upon Fmoc deprotection).



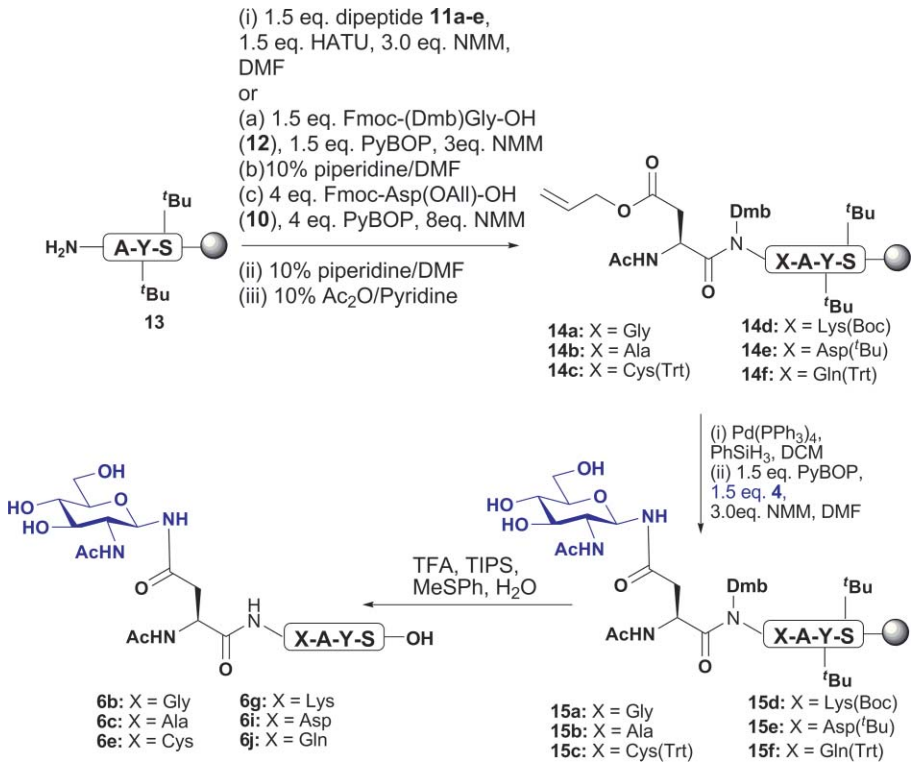
Scheme 3 Synthesis of Dmb-amino acids **9a–9e** and Fmoc-Asp(OAll)-(Dmb)-AA-OH dipeptides **11a–11e**.

Following the complete assembly of **14a–14f**, deallylation was achieved *via* the palladium-catalysed conditions previously described. The peptides were subsequently submitted to the solid-phase aspartylation conditions (1.5 equiv. PyBOP, 3.0 equiv. NMM) with glycosylamine **4** (1.5 equiv.) to afford **15a–15f**. The resin bound glycopeptides were then cleaved from the resin and deprotected *via* acidolysis. Following purification by preparative reverse-phase HPLC, the desired glycopeptides **6b, 6c, 6e, 6g, 6i** and **6j** were obtained in high yield in all cases (Table 2). Indeed, analysis of these reactions by analytical HPLC indicated that preparation of the desired *N*-linked glycopeptides was highly efficient with the backbone amide (Dmb) protection completely eliminating aspartimide formation. Following the 14 linear steps involved in the solid-phase synthesis, HPLC analysis indicated a >90% yield of the desired *N*-linked glycopeptides in all cases. The isolated yields were slightly lower due to loss during purification of the products by HPLC. A representative analytical HPLC trace for the crude alanine containing glycopeptide **6c** clearly depicts the efficiency of the solid-phase aspartylation of peptides bearing Dmb protection on the amino acid residue *C*-terminal to the glycosylation site (Fig. 1).

Having successfully established an efficient means for the aspartimide-free solid-phase preparation of *N*-linked glycopeptides, our next goal was to extend the methodology to allow

Table 2 Preparation of *N*-linked glycopeptides *via* solid-phase aspartylation of backbone protected peptides

Entry	Amino Acid X	HPLC yield (%)	Isolated yield (%)	Aspartimide yield (%)
1	Gly	91	64	0
2	Ala	90	58	0
3	Cys	93	67	0
4	Lys	91	69	0
5	Asp	91	59	0
6	Gln	94	73	0



Scheme 4 Synthesis of *N*-linked glycopeptides *via* solid-phase aspartylation from backbone amide protected resin bound peptides **14a–14f**.

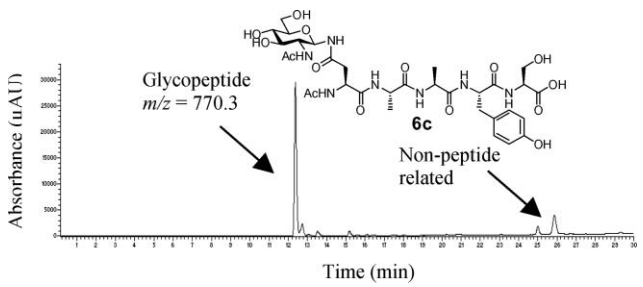
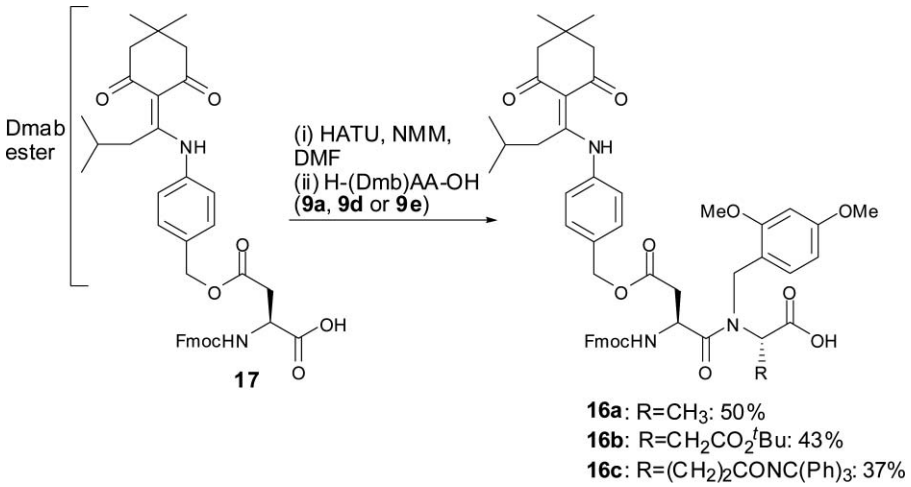


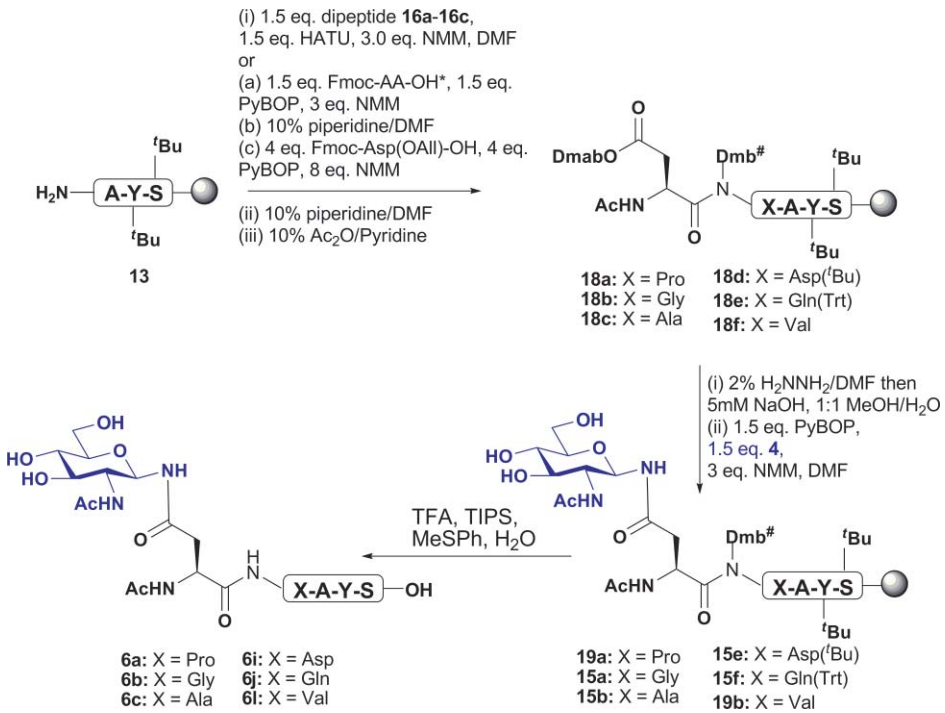
Fig. 1 Crude analytical HPLC of a representative *N*-linked glycopeptide **6c** after cleavage from the resin (λ = 280 nm, gradient: 0–50%B over 30 min).

for the incorporation of two different sugars onto the same resin bound peptide fragment. To this end, an alternative side chain protecting group was sought which would be orthogonal to the allyl ester and the acid labile protecting groups utilised in Fmoc-strategy SPPS. In this regard, the 4-{*N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino} benzyl (Dmab) ester has recently shown promise as an orthogonal protecting group for the modification of peptides on resin.[51-53] We recently showed that the Dmab ester could be deprotected in quantitative yield on the solid phase by treatment with 2% hydrazine to remove the 1-(4,4-dimethyl-2,6-dioxo-cyclohexylidene)-3-methylbutyl (ivDde) moiety[52] followed by treatment with dilute aqueous hydroxide to facilitate elimination of the resultant *p*-aminobenzyl ester.[54] It was anticipated that this early work could be extended in our solid phase aspartylation methodology for the facile preparation of a range of *N*-linked glycopeptides.

Dmab protected dipeptides **16a–16c** possessing Dmb protection of the backbone amide were synthesised in solution by pre-activation of Fmoc-Asp(ODmab)-OH **17** with HATU and NMM followed by treatment with Dmb-amino acids **9a, 9d** and **9e**. Subsequent purification by preparative reverse-phase HPLC afforded the desired dipeptides **16a–16c** in 37–50% yield (Scheme 5). Dmb protected dipeptides **16a–16c** were next coupled to tripeptide **13** using HATU to afford the desired resin bound pentapeptides **18c–18e** in quantitative yield (as determined by measuring the UV absorbance of the piperidine–fulvene adduct at 301 nm upon Fmoc removal, Scheme 6). Fmoc-Pro-OH, Fmoc-(Dmb)Gly-OH and Fmoc-Val-OH were incorporated as individual amino acids followed by coupling of Fmoc-Asp(ODmab)-OH to afford **18a, 18b** and **18f** also in quantitative yield. Valine was coupled without a backbone amide protecting group, as previous observations indicated that the steric bulk of the side chain would preclude significant aspartimide formation (Table 1). Following assembly, the Dmab groups of pentapeptides **18a–18f** were removed by treatment with 2% hydrazine, followed by 5mM NaOH in 1 : 1 v/v MeOH–H_2O to promote aminobenzyl ester hydrolysis.[54] The resulting peptides were then subjected to the solid-phase aspartylation conditions (**4**, PyBOP and NMM in DMF) to afford the resin bound *N*-linked glycopeptides. Following acidolytic deprotection and cleavage from the resin, the glycopeptides were purified by preparative reverse-phase HPLC to afford **6a–6c, 6i, 6j** and **6l** in high yields in all cases (80–96% yield, entries 1–6, Table 3). As expected, aspartimides were not generated during the solid-phase aspartylation reactions, with the exception of **18f** bearing a penultimate valine residue which, after aspartylation, provided glycopeptide **6l** in 96% yield, along with a 3% yield of the corresponding aspartimide which was generated due to the lack of a backbone amide protecting group.



Scheme 5 Synthesis of Fmoc-Asp(ODmab)(Dmb)-AA-OH dipeptides **16a–16c**.



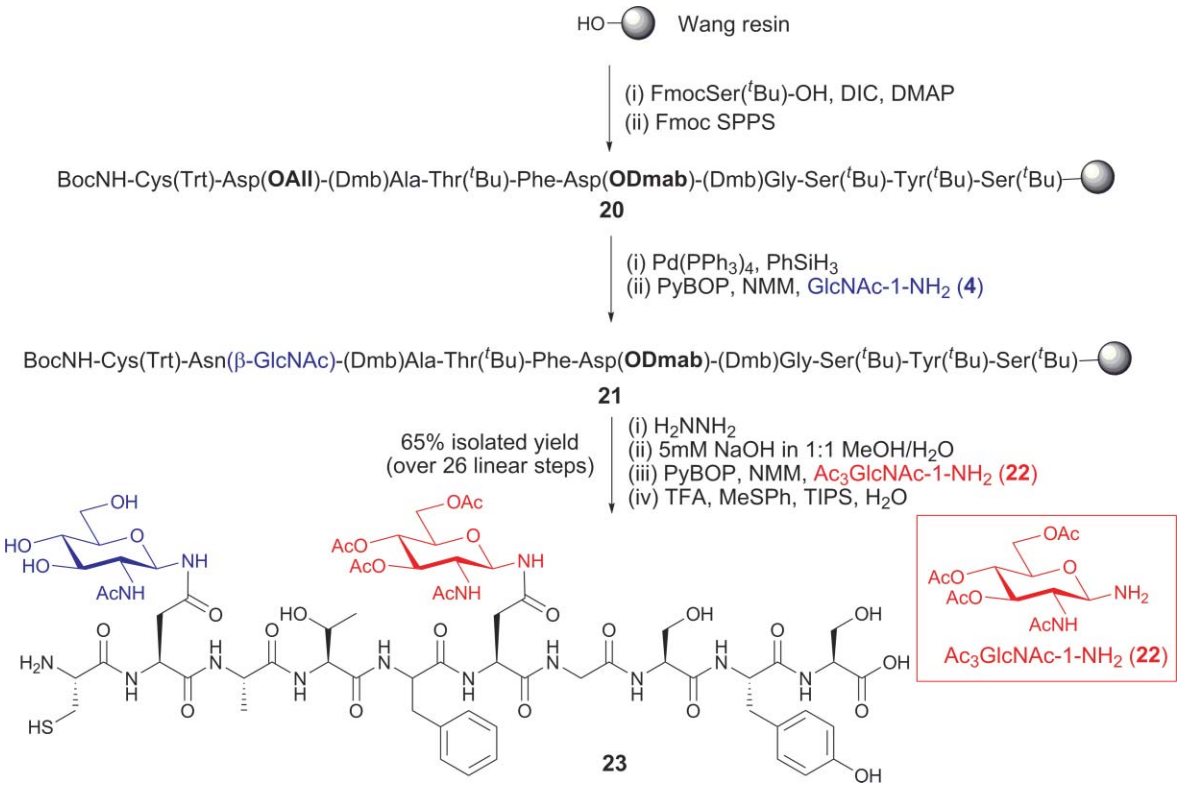
Scheme 6 Synthesis of *N*-linked glycopeptides *via* solid-phase aspartylation of Dmab ester protected resin bound peptides **18a–18f**. * Fmoc-AA-OH includes Fmoc-Pro-OH, Fmoc-(Dmb)Gly-OH (**12**) and Fmoc-Val-OH. # **18a**, **18f**, **19a** and **19b** were not Dmb-protected.

Table 3 Yields of *N*-linked glycopeptides from resin bound Dmab-protected peptides **18a–18f**

Entry	Amino Acid **X**	HPLC yield (%)	Isolated yield (%)	Aspartimide yield (%)
1	Pro[a]	95	86	0
2	Gly	92	50	0
3	Ala	88	71	0
4	Asp	91	58	0
5	Gln	80	65	0
6	Val[a]	96	56	3

[a] No Dmb protection.

Having successfully demonstrated that the Allyl and Dmab esters could be used interchangeably for the preparation of *N*-linked glycopeptides *via* solid-phase aspartylation, the next goal was to combine these two strategies to selectively aspartylate a resin-bound peptide with two different carbohydrate moieties. To demonstrate the feasibility of this approach, the model decapeptide **20** was assembled on Wang resin using Fmoc-strategy SPPS (Scheme 7). Dmb-protected glycine and alanine residues were incorporated on the *C*-terminal side of two aspartic acid residues which were side chain protected with Dmab and Allyl esters respectively in **20**. Peptide couplings until the final aspartic acid residue proceeded with greater than 95% efficiency



Scheme 7 Synthesis of diglycosylated peptide **23** *via* a double solid-phase aspartylation reaction.

as determined by measuring the absorbance of the piperidine–fulvene adduct after Fmoc deprotection of this residue. The final cysteine residue was introduced as the *N*-Boc protected derivative thereby allowing the final acidolytic deprotection and cleavage step to afford a free *N*-terminus. Having successfully assembled the fully protected peptide on resin, it was possible to commence solid phase aspartylation. Accordingly, deallylation of the corresponding aspartic acid residue was achieved with palladium(0) in the presence of phenylsilane. Solid-phase aspartylation with glycosylamine **4** then gave resin bound glycopeptide **21**. The next step involved deprotection of the Dmab ester which was achieved in two steps using 2% hydrazine in DMF followed by a 5 mM aqueous solution of NaOH. A second solid phase aspartylation was then performed using the peracetylated glycosylamine **22** (prepared in three steps from *N*-acetylglucosamine[55]). The final acidolytic side chain deprotection and resin cleavage step followed by purification by preparative reverse phase HPLC afforded the desired diglycopeptide **23** in an excellent 65% isolated yield over 26 linear steps (starting from preloaded Wang resin). Although this first example was conducted using two monosaccharides, it is anticipated that the double solid-phase aspartylation will provide a useful means for the incorporation of large oligosaccharides. Specifically, the selective incorporation of bridgehead monosaccharides as described here should prove useful for enzymatic elaboration *via* the use of endo-β-*N*-acetylglucosaminidases (Endo-A and Endo-M) to afford more complex *N*-linked glycans on the peptide backbone.[56-65] Alternatively, larger oligosaccharides bearing anomeric amine moieties could be incorporated to prepare glycopeptides bearing more complex glycans.[45]

Conclusions

In summary, we have developed a rapid and efficient method for the synthesis of *N*-linked glycopeptides *via* solid-phase aspartylation. An extensive study was first conducted in order to ascertain the extent to which amino acids adjacent (*C*-terminal) to the putative glycosylation site affect aspartimide formation. Peptides bearing allyl ester protected aspartic acid side chains were deprotected with palladium(0) before aspartylation with an glycosylamine. With the exception of peptides containing a penultimate glycine residue, the use of less molar equivalents of the base NMM (1.5 equiv.) provided glycopeptides in satisfactory yields (45–92%) with minimum aspartimide formation (4–20%). The incorporation of a 2,4-dimethoxybenzyl (Dmb) backbone amide protecting group adjacent to the putative glycosylation site completely prevented aspartimide formation during the solid-phase aspartylation reaction, thus providing *N*-linked glycopeptides in high yields in all cases. A second orthogonal protecting group, the Dmab ester was also utilised as an aspartic acid side chain protecting group within resin bound peptides. Deprotection using hydrazine and dilute aqueous base followed by aspartylation with a glycosylamine again provided the desired *N*-linked glycopeptides in high yields. The presence of backbone amide protecting groups in these peptides also prevented aspartimide by-products. To demonstrate the scope of the methodology and the compatibility of the allyl and Dmab esters, a resin bound peptide bearing both of these protecting groups was prepared. Selective deprotection and aspartylation at each putative glycosylation site was achieved with two different glycans to afford a glycopeptide containing two different *N*-glycans in high yield.

Given the simplicity and efficiency of this methodology it is anticipated that it will have far reaching applications in the preparation of *N*-linked glycopeptides and *N*-linked glycopeptide libraries. Future directions in our laboratory are focussed on the use of the solid-phase aspartylation methodology developed here for the synthesis of *N*-linked glycopeptides bearing native, complex *N*-linked glycans. This will then be extended to facilitate the total synthesis of a homogeneous glycoprotein by the use of convergent ligation chemistry.

Experimental

General procedure A: N^{α}-Dmb protection of amino acids 9a–9e

To the appropriately side chain protected L-amino acid (8 mmol) in DMF (20 mL) containing acetic acid (0.2 mL) was added 2,4-dimethoxybenzaldehyde (1.40 g, 8.4 mmol) and sodium triacetoxyborohydride (1.87 g, 8.8 mmol) and the reaction stirred at 22 °C for 14 h. Water (5 mL) and acetic acid (2 mL) were added and the reaction warmed under reduced pressure until effervescence ceased. The reaction was then concentrated *in vacuo* to afford a viscous oil which was purified by column chromatography.

H-(Dmb)-Ala-OH (9a). L-Alanine **8a** (0.71 g, 8.0 mmol) was Dmb protected using general procedure A. The product was purified by column chromatography (eluent: 3 : 20 v/v MeOH–DCM) to afford **9a**, as a pale yellow solid (0.90 g, 47%). R_f [3 : 20 MeOH–DCM containing 1% AcOH] = 0.24; m.p. = 216–218 °C; $[\alpha]_D^{20} = -0.41$ ($c = 2.2$, MeOH); IR (thin film) $\nu_{max} = 2993, 2939, 2831, 2360, 2337, 1612$; ^{1}H NMR (400 MHz, CD_3OD) δ 7.28 (d, 1H, $J = 8.2$ Hz, Ar H), 6.61 (d, 1H, $J = 2.0$ Hz, Ar H), 6.57 (dd, 1H, $J = 8.2$, 2.0 Hz, Ar H), 4.15 (d, 1H, $J = 16.0$ Hz, C*H*HN), 4.13 (d, 1H, $J = 16.0$ Hz, CH*H*N), 3.89 (s, 3H, OCH_3), 3.81 (s, 3H, OCH_3), 3.52 (q, 1H, $J = 7.2$ Hz, α-H), 1.49 (d, 3H, $J = 7.2$ Hz, CH_3); ^{13}C NMR (75 MHz, CD_3OD) δ 174.7, 164.4, 161.0, 133.9, 113.3, 106.6, 99.9, 58.9, 56.6, 56.4, 47.0, 16.5; MS (ESI) m/z 239.9 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{12}H_{17}NO_4Na$: MNa$^+$, 262.1050 found MNa$^+$, 262.1052.

H-(Dmb)-Cys(Trt)-OH (9b). *S*-Trityl-L-Cysteine **8b** (2.90 g, 8.0 mmol) was Dmb protected using general procedure A. The product was purified by column chromatography (eluent: 1 : 9 v/v MeOH–DCM) to afford **9b** as a pale yellow solid (2.34 g, 57%). R_f [1 : 9 v/v MeOH–DCM containing 1% AcOH] = 0.40; m.p. = 108–110 °C; $[\alpha]_D^{20} = -0.3$ ($c = 1.5$, $CHCl_3$); IR $\nu_{max} = 3726, 3679, 3610, 3031, 3016, 3001, 2970, 2893, 2399, 2360, 1751, 1666, 1620$; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (d, 6H, $J = 7.6$ Hz, Ar H), 7.23–7.14 (m, 9H, Ar H), 7.03 (d, 1H, $J = 8.0$ Hz, Ar H), 6.38 (dd, 1H, $J = 1.5$, 8.0 Hz, Ar H), 6.37 (br. s, 1H, Ar H), 3.95 (d, 1H, $J = 13.2$ Hz, C*H*HN), 3.80 (d, 1H, $J = 13.2$ Hz, CH*H*N), 3.78 (s, 3H, OCH_3), 3.65 (s, 3H, OCH_3), 2.98 (m, 1H, C*H*HS), 2.77 (m, 2H, α-H + CH*H*S); ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.8, 162.5, 159.2, 144.6, 133.0, 129.9, 128.5, 127.3, 112.5, 105.0, 99.0, 67.9, 59.4, 55.9, 55.7, 47.6, 32.9; MS (ESI) m/z 514.1 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{31}H_{31}NO_4SNa$: MNa$^+$, 536.1866 found MNa$^+$, 536.1875.

H-(Dmb)-Lys(Boc)-OH (9c). *N*-ε-Boc-L-Lysine **8c** (1.97 g, 8.0 mmol) was Dmb protected using general procedure A. The product was purified by column chromatography (eluent: 1 : 9 v/v MeOH–DCM) to afford **9c** as a pale yellow solid (1.96 g, 62%). R_f [1 : 9 v/v MeOH–DCM containing 1% AcOH] = 0.18; m.p. = 146–148 °C; $[\alpha]_D^{20} = -5.3$ ($c = 1.6$, $CHCl_3$); IR $\nu_{max} = 3448, 3001, 2970, 2939, 2399, 1697, 1620$; ^{1}H NMR (400 MHz, $CDCl_3$) 7.29 (d, 1H, $J = 8.2$ Hz, Ar H), 6.44–6.39 (m, 2H, Ar H), 4.13 (d, 1H, $J = 13.4$ Hz, C*H*HN), 4.02 (d, 1H, $J = 13.4$ Hz, CH*H*N), 3.82 (s, 3H, OCH_3), 3.74 (s, 3H, OCH_3), 3.31 (m, 1H, α-H), 3.02 (m, 2H, CH_2), 1.82 (m, 2H, CH_2), 1.43–1.38 (m, 13 H, 2 x CH_2, 3 x CH_3); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.7, 162.4, 159.4, 156.6, 133.1, 112.2, 105.1, 98.9, 79.1, 61.0, 55.9, 55.8, 45.6, 40.3, 30.1, 29.9, 28.9, 23.3; MS (ESI) m/z 397.1 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{20}H_{32}N_2O_6Na$: MNa$^+$, 419.2153 found MNa$^+$, 419.2153.

H-(Dmb)-Asp(O^{t}Bu)-OH (9d). L-Aspartic acid-β-*tert*-butyl ester **8d** (1.51 g, 8.0 mmol) was Dmb protected using general procedure A. The product was purified by column chromatography (eluent: 1 : 9 MeOH–DCM) to afford **9d** as a pale yellow solid (1.90 g, 70%). R_f [1 : 9 v/v MeOH–DCM containing 1% AcOH] = 0.37; m.p. = 49–51 °C; $[\alpha]_D^{20} = -6.65$ ($c = 1.5$, $CHCl_3$); IR $\nu_{max} = 3679, 3016, 3001, 2970, 2399, 1712, 1620$; ^{1}H NMR (400 MHz, CD_3OD) δ 7.29 (d, 1H, $J = 8.4$ Hz, Ar H), 6.65 (d, 1H, $J = 2.0$ Hz, Ar H), 6.58 (dd, 1H, $J = 8.4$, 2.0 Hz, Ar H), 4.29 (d, 1H, $J = 10.8$ Hz, C*H*HN), 4.26 (d, 1H, $J = 10.8$ Hz, CH*H*N), 3.94 (s, 3H, OCH_3), 3.87 (dd, 1H, $J = 7.2$, 4.4 Hz, α-H), 3.84 (s, 3H, OCH_3), 2.98 (dd, 1H, $J = 18.0$, 4.4 Hz, C*H*H), 2.87 (dd, 1H, $J = 18.0$, 7.2 Hz, CH*H*), 1.49 (s, 9H, CH_3); ^{13}C NMR (75 MHz, CD_3OD) δ 172.1, 171.7, 164.6, 161.0, 134.0, 112.9, 106.7, 100.0, 84.0, 58.7, 56.6, 56.4, 48.5, 36.2, 28.7; MS (ESI) m/z 340.0 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{17}H_{25}NO_6Na$: MNa$^+$, 362.1574 found MNa$^+$, 362.1573.

H-(Dmb)-Gln(Trt)-OH (9e). *N*-γ-Trityl-L-glutamine **8e** (3.10 g, 8.0 mmol) was Dmb protected using general procedure A. The product was purified by column chromatography (eluent: 1 : 9 v/v MeOH–DCM containing 1% AcOH) to afford **9e** as a pale yellow solid (3.53 g, 82%). R_f [1 : 9 v/v MeOH–DCM containing 1% AcOH] = 0.36; m.p. = 119–121 °C; $[\alpha]_D^{20} = -2.26$ ($c = 1.5$, $CHCl_3$); IR $\nu_{max} = 3687, 3425, 3085, 3031, 3008, 2977, 2399, 1650, 1620$; ^{1}H NMR (400 MHz, CD_3OD) δ 7.25-7.12 (m, 15H, Trt Ar H), 6.99 (d, 1H, $J = 8.8$ Hz, Dmb Ar H), 6.36-6.32 (m, 2H, 2 x Ar H), 3.88 (d, 1H, $J = 12.8$ Hz, C*H*HN), 3.80 (d, 1H, $J = 12.8$ Hz, CH*H*N), 3.76 (s, 3H, OCH_3), 3.58 (s, 3H, OCH_3), 3.33 (dd, 1H, $J = 5.2$, 5.2 Hz, α-H), 2.52-2.33 (m, 2H, CH_2), 2.16-2.06 (m, 1H, C*H*H), 1.99-1.89 (m, 1H, CH*H*); ^{13}C NMR (75 MHz, CD_3OD) δ 173.8, 170.7, 162.5, 159.3, 144.8, 132.8, 129.2, 128.2, 127.2, 111.8, 104.9, 98.9, 71.1, 61.1, 55.8, 51.1, 48.2, 31.3, 25.4; MS (ESI) m/z 539.1 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{33}H_{34}N_2O_5Na$: MNa$^+$, 561.2360 found MNa$^+$, 561.2364.

General procedure B: Synthesis of Fmoc-Asp(OAll)-(Dmb)AA-OH dipeptides 11a–11e. A solution of Fmoc-Asp(OAll)-OH **10** (0.20 g, 0.34 mmol), HATU (0.13 g, 0.34 mmol) and NMM (110 μL, 1 mmol) in DMF (5 mL) was stirred for 30 min before the addition of Dmb amino acid **9a–9e** (0.50 mmol). The reaction was stirred for 3 h before dilution with EtOAc (20 mL) and saturated aqueous NH_4Cl solution (20 mL). The aqueous layer was separated and re-extracted with EtOAc (20 mL). The combined organic layers were washed with brine (20 mL), dried over Na_2SO_4, concentrated *in vacuo* and the resulting residue purified by preparative HPLC, gradient: 30–100%B over 45 min (solvent A: water + 0.1% TFA, solvent B: MeCN + 0.1%TFA).

Fmoc-Asp(OAll)-(Dmb)Ala-OH (11a). H-(Dmb)-Ala-OH **9a** (0.12 g, 0.50 mmol) was coupled to Fmoc-Asp(OAll)-OH **10** using general procedure B to afford dipeptide **11a** as a white solid (0.14 g, 68%). m.p. = 45–47 °C; $[\alpha]_D^{20} = -28.97$ ($c = 0.9$, $CHCl_3$); IR ν_{max} = 3687, 3610, 3047, 3001, 2970, 2900, 2437, 2408, 2355, 1735, 1720, 1650, 1612; ^{1}H NMR (400 MHz, CD_3OD, major rotamer) δ 7.53 (d, 2H, $J = 7.2$ Hz, 2x Ar H), 7.38 (m, 2H, 2 x Ar H), 7.17 (t, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.07 (t, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.02 (d, 1H, $J = 8.2$ Hz, Ar H), 6.22 (br. s, 1H, Ar H), 6.16 (d, 1H, $J = 8.2$ Hz, Ar H), 5.71-5.61 (m, 1H, C*H*=CH_2), 5.06 (d, 1H, $J = 17.2$ Hz, CH=C*H*H), 4.97 (m, 2H, CH=CH*H*, CH), 4.44 (d, 1H, $J = 16.4$ Hz, C*H*HN), 4.34 (m, 3H, CH_2, OC*H*H), 4.29 (d, 1H, $J = 16.4$ Hz, CH*H*N), 4.06 (d, 1H, $J = 7.2$ Hz, OCH*H*), 3.96 (m, 1H, CH), 3.84 (q, 1H, $J = 7.2$ Hz, CH), 3.55 (s, 3H, OCH_3), 3.48 (s, 3H, OCH_3), 2.66 (dd, 1H, $J = 16.2$, 7.2 Hz, C*H*H), 2.46 (dd, 1H, $J = 16.2$, 6.8 Hz, CH*H*), 1.02 (d, 3H, $J = 7.2$ Hz, CH_3); ^{13}C NMR (75 MHz, CD_3OD, major rotamer) δ 173.8, 171.1, 170.8, 161.3, 161.2, 158.8, 144.1, 141.6, 132.2, 130.2, 128.0, 127.4, 125.4, 120.2, 118.5, 116.7, 104.5, 98.8, 67.4, 65.9, 56.1, 55.9, 55.5, 55.4, 49.1, 47.4, 37.7, 14.3; MS (ESI) *m/z* 639.3 [(M+Na)$^+$, 100%] HRMS Calcd for $C_{34}H_{36}N_2O_9Na$: MNa^+, 639.2313 found MNa^+, 639.2301.

Fmoc-Asp(OAll)-(Dmb)Cys(Trt)-OH (11b). H-(Dmb)Cys-(Trt)-OH **9b** (0.26 g, 0.50 mmol) was coupled to Fmoc-Asp(OAll)-OH **10** using general procedure B to afford dipeptide **11b** as a white solid (0.21 g, 70%). m.p. = 79–81 °C $[\alpha]_D^{20} = -4.61$ (c = 0.7, $CHCl_3$); IR ν_{max} = 3687, 3417, 3294, 3055, 3036, 3024, 3008, 2947, 2839, 1720, 1650, 1612; ^{1}H NMR (400 MHz, 1 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 7.76 (d, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.60 (d, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.38 (t, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.35-7.07 (m, 18H, 18 x Ar H), 6.39 (m, 1H, Ar H), 6.37 (br. s, 1H, Ar H), 5.90-5.75 (m, 1H, C*H*=CH_2), 5.34-5.15 (m, 3H, CH=CH_2, CH), 4.58-4.50 (m, 3H, CH_2, CH), 4.37 (m, 1H, CH), 4.26 (d, 1H, $J = 10.4$ Hz C*H*HN), 4.22 (d, 1H, J = 10.4 Hz, CH*H*N), 4.09 (m, 1H, CH), 3.77 (s, 3H, OCH_3), 3.71 (s, 3H, OCH_3), 3.57 (m, 1H, CH), 3.01-2.72 (m, 4H, 2 x CH_2); ^{13}C NMR (75 MHz, 1 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 172.6, 172.2, 170.4, 161.3, 160.7, 155.8, 144.4, 143.9, 141.3, 132.0, 131.9, 129.6, 127.9, 127.8, 127.2, 126.8, 125.3, 120.0, 118.6, 115.4, 103.9, 98.7, 67.4, 67.3, 65.8, 60.9, 55.4, 55.2, 49.5, 47.2, 46.7, 38.0, 30.3; MS (ESI) *m/z* 913.7 [(M+Na)$^+$, 100%]; HRMS Calcd for $C_{53}H_{50}N_2O_9SNa$: MNa^+, 913.3129 found MNa^+, 913.3126.

Fmoc-Asp(OAll)-(Dmb)Lys(Boc)-OH (11c). H-(Dmb)Lys-(Boc)-OH **9c** (0.20 g, 0.50 mmol) was coupled to Fmoc-Asp(OAll)-OH **10** using general procedure B to afford dipeptide **11c** as a white solid (0.84 g, 32%). m.p. = 54–56 °C; $[\alpha]_D^{20} = -11.17$ ($c = 0.7$, $CHCl_3$); IR ν_{max} = 3687, 3101, 3024, 3016, 2931, 2399, 1720, 1650, 1612; ^{1}H NMR (400 MHz, 2 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 7.74 (d, 2H, $J = 6.0$ Hz, 2 x Ar H), 7.60 (m, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.38 (m, 2H, 2 x Ar H), 7.29 (m, 2H, 2 x Ar H), 7.19 (m, 1H, Ar H), 6.43 (br. s, 1H, Ar H), 6.41 (m, 1H, Ar H), 5.94-5.81 (m, 1H, C*H*=CH_2), 5.50-5.16 (m, 3H, CH=CH_2, CH), 4.69 (m, 1H, C*H*HN), 4.60-4.55 (m, 3H, CH*H*N, CH_2), 4.36 (m, 1H, OC*H*H), 4.28 (m, 1H, OCH*H*), 4.21 (t, 1H, $J = 6.4$ Hz, C*H*CH_2), 3.83-3.67 (m, 7H, 2 x OCH_3, CH), 3.10-2.62 (m, 4H, 2 x CH_2), 2.10-1.60 (m, 2H, CH_2), 1.40 (s, 9H, 3 x CH_3), 1.31-1.02 (m, 4H, 2 x CH_2); ^{13}C NMR (75 MHz, 2 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 173.9, 172.3, 170.8, 161.7, 159.3, 156.1, 156.0 144.2, 141.7, 132.3, 131.5, 128.1, 127.5, 125.7, 120.3, 118.9, 115.9, 104.4, 99.1, 79.5, 67.7, 66.0, 61.1, 60.6, 55.8, 55.7, 49.3, 47.5, 40.6, 37.9, 29.7, 29.0, 28.8, 23.7; MS (ESI) *m/z* 796.3 [(M+Na)$^+$, 100%]; HRMS Calcd for $C_{42}H_{51}N_3O_{11}Na$: MNa^+, 796.3416 found MNa^+, 796.3407.

Fmoc-Asp(OAll)-(Dmb)Asp(O′Bu)-OH (11d). H-(Dmb)Asp-(O′Bu)-OH **9d** (0.17 g, 0.50 mmol) was coupled to Fmoc-Asp(OAll)-OH **10** using general procedure B to afford dipeptide **11d** as a white solid (0.17 g, 70%). m.p. = 50–52 °C; $[\alpha]_D^{20}$ = –26.97 ($c = 0.7$, $CHCl_3$); IR ν_{max} = 3621, 3448, 3024, 3001, 2893, 2399, 2360, 1720, 1650, 1612; ^{1}H NMR (400 MHz, $CDCl_3$, major rotamer) δ 7.77 (m, 2H, 2 x Ar H), 7.62 (m, 2H, 2 x Ar H), 7.41 (m, 2H, 2 x Ar H), 7.33 (m, 3H, 3 x Ar H), 6.47 (br. s, 1H, Ar H), 6.42 (m, 1H, Ar H), 5.95-5.83 (m, 1H, C*H*=CH_2), 5.43 (m, 1H, CH), 5.34-5.18 (m, 2H, CH=CH_2), 4.74 (m, 1H, OC*H*H), 4.66-4.51 (m, 3H, OCH*H*, CH_2), 4.42-4.18 (m, 4H, CH_2N, 2 x CH), 3.80 (s, 3H, OCH_3), 3.78 (s, 3H, OCH_3), 3.22-2.72 (m, 3H, CH_2, C*H*H), 2.35-2.10 (m, 1H, CH*H*), 1.39 (s, 9H, CH_3); ^{13}C NMR (75 MHz, $CDCl_3$, major rotamer) δ 173.1, 171.8, 170.5, 170.2, 161.6, 158.9, 156.0, 144.0, 141.3, 131.9, 131.1, 127.8, 127.2, 125.4, 120.0, 118.7, 104.3, 104.0, 98.8, 81.2, 67.5, 65.1, 65.0, 58.0, 55.4, 49.0, 47.1, 37.7, 35.3, 31.0, 28.0; MS (ESI) *m/z* 716.8 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{39}H_{44}N_2O_{11}Na$: MNa^+, 739.2837 found MNa^+, 739.2836.

Fmoc-Asp(OAll)-(Dmb)Gln(Trt)-OH (11e). H-(Dmb)Gln-(Trt)-OH **9e** (0.27 g, 0.50 mmol) was coupled to Fmoc-Asp(OAll)-OH **10** using general procedure B to afford dipeptide **11e** as a white solid (0.11 g, 36%). m.p. = 59–61 °C; $[\alpha]_D^{20} = -23.40$ (c = 0.5, $CHCl_3$); IR ν_{max} = 3687, 3425, 3047, 2977, 2360, 2337, 1720, 1689, 1650, 1612; ^{1}H NMR (400 MHz, 2 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 7.77 (d, 2H, $J = 7.6$ Hz, 2 x Ar H), 7.60 (t, 2H, $J = 6.8$ Hz, 2 x Ar H), 7.41 (t, 2H, $J = 7.2$ Hz, 2 x Ar H), 7.36-7.15 (m, 17H, 17 x Ar H), 7.06 (d, 1H, $J = 8.0$ Hz, Ar H), 6.40 (br. s, 1H, Ar H), 6.32 (d, 1H, $J = 8.4$ Hz, Ar H), 5.88-5.74 (m, 1H, C*H*=CH_2), 5.37 (m, 1H, CH), 5.22 (m, 2H, CH=CH_2), 4.67 (d, 1H, $J = 14.8$ Hz, C*H*HN), 4.45-4.22 (m, 5H, CH*H*N, OCH_2, CH_2), 4.20 (t, 1H, $J = 7.2$ Hz, C*H*CH_2) 3.92 (m, 1H, CH), 3.71 (s, 6H, OCH_3), 3.00 (dd, 1H, $J = 16.0$, 7.6 Hz, C*H*H), 2.77 (dd, 1H, $J = 16.0$, 4.0 Hz, CH*H*), 2.50-2.23 (m, 4H, 2 x CH_2); ^{13}C NMR (75 MHz, $CDCl_3$, major rotamer) δ 173.5, 173.0, 172.9, 170.8, 161.7, 159.3, 155.9, 144.7, 144.5, 141.7, 132.1, 129.0, 128.4, 128.3, 128.1, 127.7, 127.5, 125.6, 120.4, 118.9, 115.4, 104.5, 99.1, 81.0, 71.1, 67.7, 65.9, 59.8, 55.7, 49.2, 48.8, 47.4, 38.2, 34.1, 25.1; MS (ESI) *m/z* 938.4 [(M+Na)$^+$, 100%]; HRMS Calcd for $C_{55}H_{53}N_3O_{10}Na$: MNa^+, 938.3623 found MNa^+, 938.3620.

General procedure C: Synthesis of Fmoc-Asp(ODmab)-(Dmb)AA-OH dipeptides 16a–16c. A solution of Fmoc-Asp(ODmab)-OH **17** (0.10 g, 0.15 mmol), HATU (57 mg, 0.15 mmol) and NMM (33 µL, 0.30 mmol) in DMF (5 mL) were stirred for 30 min before the addition of Dmb amino acid **9a**, **9d** or **9e** (0.23 mmol). The reaction was stirred for 3 h before diluting with EtOAc (20 mL) and saturated aqueous NH_4Cl solution (20 mL). The aqueous layer was separated and re-extracted with EtOAc (20 mL). The combined organic layers were washed with brine (20 mL), dried over $NaSO_4$, and concentrated *in vacuo* to afford a viscous oil which was purified by preparative reverse-phase HPLC,

gradient: 30–100%B over 45 min (solvent A: water + 0.1% TFA, solvent B: MeCN + 0.1% TFA).

Fmoc-Asp(ODmab)-(Dmb)-Ala-OH (16a). H-(Dmb)Ala-OH **9a** (0.05 g, 0.23 mmol) was coupled to Fmoc-Asp(ODmab)-OH **17** (0.10 g, 0.15 mmol) using general procedure C to afford dipeptide **16a** as a yellow oil (0.07 g, 50%). $[\alpha]_D^{20} = -5.32$ ($c = 0.5$, $CHCl_3$); IR $\nu_{max} = 3692, 3448, 3016, 3001, 2939, 2893, 2677, 2553, 2229, 2067, 1789, 1720, 1643$; ^{1}H NMR (400 MHz, 1 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 7.76 (d, 2H, $J = 7.6$ Hz, 2 x Ar H), 7.61 (d, 2H, $J = 6.8$ Hz, 2 x Ar H), 7.44-7.35 (m, 4H, 4 x Ar H), 7.31 (m, 2H, 2 x Ar H), 7.21 (m, 1H, Ar H), 7.10 (m, 2H, 2 x Ar H), 6.43 (br. s, 1H, Ar H), 6.38 (m, 1H, Ar H), 5.27 (m, 1H, CH), 5.17 (s, 2H, OCH_2), 4.68 (m, 1H, NC*H*H), 4.56 (m, NCH*H*), 4.37-4.14 (m, 3H, CH_2, CH), 4.08 (t, 1H, $J = 7.2$ Hz, C*H*CH_2), 3.77 (s, 3H, OCH_3), 3.71 (s, 3H, OCH_3), 2.98-2.72 (m, 4H, 2 x CH_2), 2.46 (s, 4H, 2 x CH_2), 1.86-1.77 (m, 1H, CH), 1.26 (d, 3H, $J = 7.2$ Hz, CH_3), 1.09 (s, 6H, 2 x CH_3), 0.77 (d, 6H, $J = 6.8$ Hz, 2 x CH_3); ^{13}C NMR (75 MHz, 1 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 198.6, 176.4, 173.1, 172.9, 170.2, 160.6, 160.5, 158.1, 155.6, 143.2, 140.8, 135.9, 135.4, 129.5, 128.6, 127.3, 126.8, 126.2, 126.1 124.7, 119.6, 116.2, 115.9, 107.1, 103.6, 98.0, 66.7, 65.4, 55.4, 55.2, 54.8, 52.3, 46.8, 38.1, 37.2, 37.1, 29.6, 29.2, 27.6, 25.2, 22.0, 13.6; MS (ESI) m/z 910.1 [(M+Na)$^+$, 100%]; HRMS Calcd for $C_{51}H_{57}N_3O_{11}Na$: MNa^+, 910.3885 found MNa^+, 910.3884.

Fmoc-Asp(ODmab)-(Dmb)Asp(O′Bu)-OH (16b). H-(Dmb)-Asp(O′Bu)-OH **9d** (0.08 g, 0.23 mmol) was coupled to Fmoc-Asp(ODmab)-OH **17** (0.10 g, 0.15 mmol) using general procedure C to afford dipeptide **16b** as a yellow oil (0.06 g, 43%). $[\alpha]_D^{20} = -16.56$ ($c = 0.5$, $CHCl_3$); IR $\nu_{max} = 3680, 3634, 3433, 3008, 2978, 2677, 2399, 2252, 2068, 1720, 1604$; ^{1}H-NMR (400 MHz 3 : 1 v/v $CD_3OD/CDCl_3$ major rotamer) δ 7.77 (d, 2H, $J = 7.6$ Hz, 2 x Ar H), 7.63 (m, 2H, 2 x Ar H), 7.46-7.05 (m, 9H, 9 x Ar H), 6.61-6.39 (m, 2H, 2 x Ar H), 5.40-5.15 (m, 1H, CH), 4.92-4.16 (m, 7H, 3 x CH_2, CH), 3.17-2.80 (m, 4H, 2 x CH_2), 2.75-2.63 (m, 6H, 3 x CH_2), 2.49-2.32 (m, 3H, CH_2, C*H*H), 2.18-2.06 (m, 1H, CH*H*), 1.82-1.70 (m, 1H, CH), 1.38 (s, 9H, 3 x CH_3), 1.08 (s, 6H, 2 x CH_3), 0.96 (d, 6H, $J = 6.4$ Hz, 2 x CH_3); ^{13}C NMR (100 MHz, $CDCl_3$, major rotamer) δ 199.3, 177.6,172.5, 171.7, 170.5, 170,3, 161.5, 161.1, 158.8, 155.7, 143.7, 141.2, 136.3, 135.6, 131.3, 130.9, 129.1, 127.7, 127.1, 126.5, 126.4, 125.2, 120.0, 115.1, 114.8, 107.3, 104.1, 103.9, 98.7, 81.0, 67.4, 65.9, 57.0, 56.9, 55.3, 52.3, 49.0, 47.0, 38.8, 37.6, 35.2, 30.1, 29.7, 28.1, 27.9, 22.5. MS (ESI) m/z 988 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{56}H_{66}N_3O_{13}$: MH^+, 988.4590 found MH^+, 988.4589.

Fmoc-Asp(ODmab)-(Dmb)-Gln(Trt)-OH (16c). H-(Dmb)-Gln(Trt)-OH **9e** (0.12 g, 0.23 mmol) was coupled to Fmoc-Asp(ODmab)-OH **17** (0.10 g, 0.15 mmol) using general procedure C to afford dipeptide **16c** as a yellow oil (0.07 g, 37%). $[\alpha]_D^{20} = -14.23$ ($c = 0.5$, $CHCl_3$); IR $\nu_{max} = 3618, 3448, 3031, 3001, 2977, 2893, 2461, 1720, 1650$; ^{1}H NMR (400 MHz, 1 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 7.74 (m, 2H, 2 x Ar H), 7.60 (m, 2H, 2 x Ar H), 7.39-7.04 (m, 24 H, 24 x Ar H, Dmb Ar H), 6.39-6.36 (m, 2H, 2 x Ar H), 5.38 (m, 1 H, CH), 5.09-4.40 (m, 4H, 2 x CH_2), 4.30 (d, 2H, $J = 7.2$ Hz, CH_2), 4.17 (t, 1H, $J = 7.2$ Hz, CH), 3.76-3.60 (m, 7H, 2 x OCH_3, CH), 3.02-2.74 (m, 6 H, 3 x CH_2), 2.54 (s, 4H, 2 x CH_2), 2.35-2.13 (m, 2H, CH_2), 1.87-1.80 (m, 1H, CH), 1.11 (s, 6H, 2 x CH_3), 0.76 (m 6H, 2 x CH_3); ^{13}C NMR (75 MHz, 1 : 1 v/v $CDCl_3/CD_3OD$, major rotamer) δ 198.8, 176.6, 173.0, 172.9, 170.7, 162.3, 160.9, 158.7, 156.2, 144.1, 143.7, 143.4, 136.1, 129.6, 129.2, 128.4, 127.6, 126.9, 126.7, 126.4, 126.4, 124.9, 119.7, 115.7, 112.8, 107.4, 103.8, 98.2, 78.0, 70.1, 66.9, 65.3, 55.5, 54.9, 53.2, 52.5, 46.8, 38.2, 36.1, 33.6, 33.4, 29.8, 29.4, 27.8, 24.7, 22.2; MS (ESI) m/z 1188.4 [(M+H)$^+$, 100%]; HRMS Calcd for $C_{72}H_{74}N_4O_{12}Na$: MNa^+, 1209.5201 found MNa^+, 1209.5193.

Solid-phase peptide assembly (25 μmol scale). *Fmoc Deprotection*: A solution of piperidine/DMF (5 mL, 1 : 9 v/v) was added to pre-loaded Wang resin, shaken for 3 min and the procedure repeated. The resin was subsequently washed with DMF (5 × 3 mL), DCM (5 × 3 mL), and DMF (5 × 3 mL). *Amino Acid Coupling*: A solution of protected amino acid (100 μmol), PyBOP (52 mg, 100 μmol) and NMM (22 μL, 200 μmol) in DMF (1 mL) was added to the resin and shaken. After 1 h the resin was washed with DMF (5 × 3 mL), DCM (5 × 3 mL), and DMF (5 × 3 mL). *Capping*: Acetic anhydride/pyridine (1 : 9 v/v) was added to the resin and shaken. After 3 min the resin was washed DMF (5 × 3 mL), DCM (5 × 3 mL), and DMF (5 × 3 mL). *Final N-terminal deprotection*: When the peptide had been fully assembled the resin was treated with piperidine/DMF (5 mL, 1 : 9 v/v) and shaken for 3 min (x 2). The *N*-terminus was then acetylated by treatment with Ac_2O/pyridine (5 mL, 1 : 9 v/v) for 5 min and subsequently washed with DMF (10 × 3 mL) then DCM (10 × 3 mL).

Coupling of dipeptides. Following assembly of tripeptide **13** on Wang resin using the solid-phase assembly procedure described above, a solution of dipeptide **11a–11e** or **16a–16c** (37.5 μmol), HATU (14.3 mg, 37.5 μmol) and NMM (8.25 μL, 75 μmol) in DMF (0.5 mL) were added to the resin and placed on a shaker for 2 h. The resin was subsequently washed with DMF (5 × 3 mL), DCM (5 × 3 mL), and DMF (5 × 3 mL).

Allyl ester deprotection. A solution of tetrakis(triphenylphosphine)palladium(0) (20 mg, 18 μmol) and phenylsilane (123 μL, 1 mmol) in dry DCM (1 mL) was added to the fully assembled resin bound peptide (25 μmol). The resin was placed on a shaker for 1 h and the procedure repeated. The resin was subsequently washed with DMF (10 × 3 mL), DCM (10 × 3 mL), and DMF (5 × 3 mL).

Dmab ester deprotection. A solution of hydrazine monohydrate in DMF (2% v/v, 3 mL) was added to the resin and placed on a shaker for 3 min. Removal of the ivDde moiety was monitored by measuring the UV absorbance at 290 nm and the above procedure repeated until no further indolone adduct could be detected. Following ivDde removal, the resin was washed with H_2O/Methanol (3 mL, 1 : 1 v/v) containing 5 mM NaOH for 3 h.

Solid-phase aspartylation. A solution of 2-acetamido-2-deoxy-β-D-glucopyranosylamine **4** (8.3 mg, 37.5 μmol), PyBOP (19.5 mg, 37.5 μmol) and NMM (8.3 μL, 75 μmol) in DMF (0.5 mL), was added to the deprotected resin bound peptide and placed on a shaker for 16 h. The resin was subsequently washed with DMF (10 × 3 mL) then DCM (10 × 3 mL).

Deprotection and cleavage. A mixture of TFA/thioanisole/triisopropylsilane/water (17 : 1 : 1 : 1 v/v/v/v) was added to the resin and placed on a shaker for 1.5 h. The resin was then washed with TFA (3 × 3 mL), and the combined cleavage and washing solutions were concentrated *in vacuo*. The resulting residue was

dissolved in DMSO (1.5 mL), purified by preparative HPLC and lyophilised to afford the desired glycopeptides.

Acknowledgements

We would like to thank the Australian Research Council (ARC) for funding this project (DP0986632) and an Australian Postgraduate Award for PhD funding (TC). We would also like to thank Dr Ian Luck and Dr Kelvin Picker for technical support.

Notes and references

1 R. G. Spiro, *Glycobiology*, 2002, **12**, 43R–56R.
2 R. Apweiler, H. Hermjakob and N. Sharon, *Biochim. Biophys. Acta, Gen. Subj.*, 1999, **1473**, 4–8.
3 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683–720.
4 A. Varki, *Glycobiology*, 1993, **3**, 97–130.
5 P. M. Rudd, R. J. Woods, M. R. Wormald, G. Opdenakker, A. K. Downing, I. D. Campbell and R. A. Dwek, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1995, **1248**, 1–10.
6 O. Seitz, *ChemBioChem*, 2000, **1**, 214–246.
7 D. H. Dube and C. R. Bertozzi, *Nat. Rev. Drug Discovery*, 2005, **4**, 477–488.
8 T. Buskas, S. Ingale and G.-J. Boons, *Glycobiology*, 2006, **16**, 113R–136R.
9 K. A. Karlsson, *Trends Pharmacol. Sci.*, 1991, **12**, 265–272.
10 L. H. Miller, M. F. Good and G. Milon, *Science*, 1994, **264**, 1878–1883.
11 B. G. Davis, *Chem. Rev.*, 2002, **102**, 579–601.
12 M. R. Pratt and C. R. Bertozzi, *Chem. Soc. Rev.*, 2005, **34**, 58–68.
13 L. Liu, C. S. Bennett and C.-H. Wong, *Chem. Commun.*, 2006, 21–33.
14 A. Brik, S. Ficht and C.-H. Wong, *Curr. Opin. Chem. Biol.*, 2006, **10**, 638–644.
15 H. Hojo and Y. Nakahara, *Biopolymers*, 2007, **88**, 308–324.
16 D. P. Gamblin, E. M. Scanlan and B. G. Davis, *Chem. Rev.*, 2009, **109**, 131–163.
17 R. J. Payne and C. H. Wong, *Chem. Commun.*, 2010, **46**, 21–43.
18 N. Yamamoto, Y. Tanabe, R. Okamoto, P. E. Dawson and Y. Kajihara, *J. Am. Chem. Soc.*, 2008, **130**, 501–510.
19 C. Piontek, P. Ring, O. Harjes, C. Heinlein, S. Mezzato, N. Lombana, C. Pöhner, M. Püttner, D. V. Silva, A. Martin, F. X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1936–1940.
20 C. Piontek, D. V. Silva, C. Heinlein, C. Pöhner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1941–1945.
21 Z. P. Tan, S. Y. Shang, T. Halkina, Y. Yuan and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 5424–5431.
22 Y. Yuan, J. Chen, Q. Wan, Z. P. Tan, G. Chen, C. Kan and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 5432–5437.
23 C. Kan, J. D. Trzupek, B. Wu, G. Chen, Z. P. Tan, Y. Yuan and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 5438–5443.
24 H. Kunz, *Angew. Chem., Int. Ed. Engl.*, 1987, **26**, 294–308.
25 H. Kunz and B. Dombo, *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 711–713.
26 L. Otvos Jr, L. Urge, M. Hollosi, K. Wroblewski, G. Graczyk, G. D. Fasman and J. Thurin, *Tetrahedron Lett.*, 1990, **31**, 5889–5892.
27 L. Urge, E. Kollat, M. Hollosi, I. Laczko, K. Wroblewski, J. Thurin and L. Otvos Jr, *Tetrahedron Lett.*, 1991, **32**, 3445–3448.
28 L. Urge, L. Otvos Jr, E. Lang, K. Wroblewski, I. Laczko and M. Hollosi, *Carbohydr. Res.*, 1992, **235**, 83–93.
29 M. Meldal and K. Bock, *Tetrahedron Lett.*, 1990, **31**, 6987–6990.
30 C. S. Bennett and C.-H. Wong, *Chem. Soc. Rev.*, 2007, **36**, 1227–1238.
31 S. T. Anisfeld and P. T. Lansbury, *J. Org. Chem.*, 1990, **55**, 5560–5562.
32 S. T. Cohen-Anisfeld and P. T. Lansbury, *J. Am. Chem. Soc.*, 1993, **115**, 10531–10537.
33 B. Wu, Z. Tan, G. Chen, J. Chen, Z. Hua, Q. Wan, K. Ranganathan and S. J. Danishefsky, *Tetrahedron Lett.*, 2006, **47**, 8009–8011.
34 R. M. Wilson and S. J. Danishefsky, *Pure Appl. Chem.*, 2007, **79**, 2189–2216.
35 I. J. Krauss, J. G. Joyce, A. C. Finnefrock, H. C. Song, V. Y. Dudkin, X. Geng, J. D. Warren, M. Chastain, J. W. Shiver and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2007, **129**, 11042–11044.
36 V. Y. Dudkin, J. S. Miller and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2004, **126**, 736–738.
37 B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2006, **45**, 4116–4125.
38 S. A. Kates, B. G. Delatorre, R. Eritja and F. Albericio, *Tetrahedron Lett.*, 1994, **35**, 1033–1034.
39 J. Offer, M. Quibell and T. Johnson, *J. Chem. Soc., Perkin Trans. 1*, 1995, 175–182.
40 D. Vetter, D. Tumelty, S. K. Singh and M. A. Gallop, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 60–63.
41 J. Y. Roberge, X. Beebe and S. J. Danishefsky, *Science*, 1995, **269**, 202–204.
42 C. M. Kaneshiro and K. Michael, *Angew. Chem., Int. Ed.*, 2006, **45**, 1077–1081.
43 L. C. Packman, *Tetrahedron Lett.*, 1995, **36**, 7523–7526.
44 E. Nicolás, M. Pujades, J. Bacardit, E. Giralt and F. Albericio, *Tetrahedron Lett.*, 1997, **38**, 2317–2320.
45 R. Chen and T. J. Tolbert, *J. Am. Chem. Soc.*, 2010, **132**, 3211–3216.
46 H. Kunz and H. Waldmann, *Angew. Chem., Int. Ed. Engl.*, 1984, **23**, 71–72.
47 F. Guibe, *Tetrahedron*, 1998, **54**, 2967–3042.
48 S. Ficht, R. J. Payne, R. T. Guy and C.-H. Wong, *Chem.–Eur. J.*, 2008, **14**, 3620–3629.
49 L. M. Likhosherstov, O. S. Novikova, V. A. Derevitskaja and N. K. Kochetkov, *Carbohydr.Res.*, 1986, **146**, C1–C5.
50 J. L. Radkiewicz, H. Zipse, S. Clarke and K. N. Houk, *J. Am. Chem. Soc.*, 2001, **123**, 3499–3506.
51 B. W. Bycroft, W. C. Chan, S. R. Chhabra and N. D. Hone, *J. Chem. Soc., Chem. Commun.*, 1993, 778–779.
52 W. C. Chan, B. W. Bycroft, D. J. Evans and P. D. White, *J. Chem. Soc., Chem. Commun.*, 1995, 2209–2210.
53 T. Johnson, M. Liley, T. J. Cheeseright and F. Begum, *J. Chem. Soc., Perkin Trans. 1*, 2000, 2811–2820.
54 T. Conroy, K. A. Jolliffe and R. J. Payne, *Org. Biomol. Chem.*, 2009, **7**, 2255–2258.
55 D. Macmillan, A. M. Daines, M. Bayrhuber and S. L. Flitsch, *Org. Lett.*, 2002, **4**, 1467–1470.
56 B. Li, Y. Zeng, S. Hauser, H. J. Song and L. X. Wang, *J. Am. Chem. Soc.*, 2005, **127**, 9692–9693.
57 Y. Zeng, J. S. Wang, B. Li, S. Hauser, H. G. Li and L. X. Wang, *Chem.–Eur. J.*, 2006, **12**, 3355–3364.
58 B. Li, H. J. Song, S. Hauser and L. X. Wang, *Org. Lett.*, 2006, **8**, 3081–3084.
59 T. W. D. F. Rising, T. D. W. Claridge, N. Davies, D. P. Gamblin, J. W. B. Moir and A. J. Fairbanks, *Carbohydr. Res.*, 2006, **341**, 1574–1596.
60 T. W. D. F. Rising, T. D. W. Claridge, J. W. B. Moir and A. J. Fairbanks, *ChemBioChem*, 2006, **7**, 1177–1180.
61 C. D. Heidecke, Z. L. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons and A. J. Fairbanks, *ChemBioChem*, 2008, **9**, 2045–2051.
62 T. W. D. F. Rising, C. D. Heidecke, J. W. B. Moir, Z. Ling and A. J. Fairbanks, *Chem.–Eur. J.*, 2008, **14**, 6444–6464.
63 H. Ochiai, W. Huang and L. X. Wang, *J. Am. Chem. Soc.*, 2008, **130**, 13790–13803.
64 W. Huang, C. Li, B. Li, M. Umekawa, K. Yamamoto, X. Zhang and L. X. Wang, *J. Am. Chem. Soc.*, 2009, **131**, 2214–2223.
65 L. X. Wang, *Carbohydr. Res.*, 2008, **343**, 1509–1522.