Supporting Information

for

Angew. Chem. Int. Ed. Z52673

© Wiley-VCH 2003

69451 Weinheim, Germany
Modular Assembly of Glycoproteins: Towards the Synthesis of GlyCAM-1 Using Expressed Protein Ligation.

D. Macmillan\textsuperscript{1*} and C. R. Bertozzi \textsuperscript{2}

\textsuperscript{1} School of Chemistry, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh, EH9 3JJ, Scotland.

\textsuperscript{2} Center for New Directions in Organic Synthesis,\textsuperscript{[*]} Departments of Chemistry and Molecular and Cell Biology and Howard Hughes Medical Institute, University of California, Berkeley, CA 94720

*To whom correspondence should be addressed

TEL UK (131) 650 4825

e-mail: derek.macmillan@ed.ac.uk
Materials and methods

Automated peptide synthesis was carried out on an Applied Biosystems model 431A peptide synthesizer using 4-sulfamylbutyryl-AM resin, pre-loaded NovaSyn®TGT resin and Fmoc amino acids from Novabiochem. The IMPACT CN system and all other molecular biology reagents were obtained from New England Biolabs. Oligonucleotide primers were obtained fully desalted and deprotected from Genosys. PCR was carried out on a Perkin Elmer Cetus thermal cycler. Mass spectra were obtained on a Hewlett Packard LC-MSD1100 series electrospray LC-MS and a micromass Platform II mass spectrometer connected to a Waters Alliance HPLC system. LC-MS was performed using a Zorbax C_{18} LC-MS column (2.1×150 mm) and a gradient of 5-95% acetonitrile containing 0.1% TFA over 25 minutes (flow rate of 0.2 mL/min). Semi-preparative HPLC was performed using a Rainin DYNAMAX C18 column and a gradient of 10-90% acetonitrile containing 0.1% TFA over 50 minutes (flow rate of 3.0 mL/min). All other chemical reagents were obtained from Aldrich.

Solid-phase glycopeptide synthesis

Automated solid-phase peptide synthesis was carried out on a 0.1 mmol scale using 1.0 mmol of each Fmoc amino acid per coupling reaction and DCC/HOBt as coupling reagents. Coupling time was 3 h. Manual solid-phase peptide synthesis was conducted using 0.5 mmol of each Fmoc amino acid (including Fmoc-Ser((AcO)_{3}GalNAc)-OH or Fmoc-Thr((AcO)_{3}GalNAc)-OH) and HBTU/HOBt as coupling reagents. On average the coupling time was 7 h and the reaction progress was monitored using LCMS and the Kaiser ninhydrin test.
Isolated Yields of synthetic peptides and peptide thioesters after HPLC purification.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Mass (observed) / Da</th>
<th>Mass (av) (calculated)/ Da</th>
<th>Yield /mg (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5, N-terminal Glycopeptide (residues L1-S40)-S-benzyl thioester</td>
<td>6643.6</td>
<td>6642.7</td>
<td>40(10)</td>
</tr>
<tr>
<td>6, (Acm)-C-terminal Glycopeptide (residues C78-S101)-S-benzyl thioester</td>
<td>4713.2</td>
<td>4710.8</td>
<td>23 (5)</td>
</tr>
<tr>
<td>7, unglycosylated peptide (C102-S132)</td>
<td>3253.9</td>
<td>3253.7</td>
<td>180 (55)</td>
</tr>
</tbody>
</table>

Isolation of peptide 4a

The recombinant peptide-intein-chitin binding domain (CBD) fusion was purified on chitin beads from 0.5 L bacterial culture as previously described[8a] and incubated at room temperature for 48 h in cleavage buffer (100 mM sodium phosphate, pH 8.0, 100 mM NaCl, 50 mM DTT). The first 4.5 mL were eluted from the chitin column and loaded directly onto a semi-prep reversed-phase HPLC column. Fractions containing the cleaved peptide were lyophilized to afford the purified peptide, 3.0 mg, as a white solid.
General procedure for one-pot Factor Xa cleavage, ligation (synthesis of glycoforms 1 and 3)*:

The peptide or glycopeptide bearing a factor Xa epitope (IEGR) protected N-terminal cysteine residue was dissolved in ligation buffer (100 mM sodium phosphate buffer, pH 8.0, 100 mM NaCl, 2 % MESNA, 0.5-1.0mL) to a final concentration of 4 mgmL$^{-1}$. This solution was then transferred to an eppendorf tube containing lyophilized peptide or glycopeptide S-benzyl thioester. CaCl$_2$ was then added to a final concentration of 2.5 mM, followed by Factor Xa (Novagen, 5 µL, 4.5 U). The reaction mixture was then allowed to stand at room temperature with occasional mixing (by inversion) and the reaction was monitored using LC-ESI-MS. Typically, within one hour the benzyl ester was quantitatively transformed to the MESNA thioester. After 3 h the cleaved peptide MIEGR m/z = 605.3 was observed as well as the ligation product and the reaction was complete within 24 h.

*NOTE: We conducted the reactions in 2 % MESNA with no evidence of non-specific proteolysis at room temperature (23 °C). The promiscuity of Factor Xa for cleavage sites similar to IEGR is well known and our one-pot procedure is likely to be protein-dependant to some degree.

Synthesis of glycopeptide 8

The C-terminal peptide 7 (15.0 mg) was dissolved in ligation buffer containing 4% MESNA (2.5 mL) and the glycopeptide $\alpha$-thioester 6 (10 mg, 0.5 equiv) was added as a lyophilized solid. The reaction mixture was allowed to stand at room temperature with occasional mixing (by inversion). After 24 h LC-MS indicated that the formation of ligated glycopeptide was complete. The reaction mixture was then treated directly with hydrazine hydrate (to a final concentration of 5 %) and DTT (to a final concentration of 50 mM). After a further 8 h, the product glycopeptide 8 was purified by loading the reaction mixture directly onto a semi-preparative C18 HPLC column. Fractions containing the product were lyophilized (weight approx. = 10 mg, 67 %)
Iodoacetamide capping Cys(SAcm) deprotection: synthesis of glycopeptide 9

Lyophilised fractions of 8 (see above) were re-dissolved in 50 mM sodium phosphate buffer, pH 8.0 to a concentration of approximately 5 mgmL⁻¹ and treated with 10 mM iodoacetamide. The reaction was determined to be quantitative by LC-MS as only the product could be observed after 1 h. The reaction mixture was again purified by semi-prep HPLC to afford the product as a white solid after lyophilization. The iodoacetamide capped and acetamidomethyl protected peptide (8) was dissolved in 10% acetic acid (1.0 mL) at a concentration of 5 mgmL⁻¹ and mercury(II)acetate (3.0 mg) was added followed shortly afterwards by DTT, to a final concentration of 50mM. The mixture was allowed to stand at room temperature under argon with occasional agitation. After 16 h LC-MS indicated that the reaction was complete and the thick white precipitate was removed by centrifugation at 13000 rpm for 5 min. The remaining solution was purified directly by semi-preparative HPLC and fractions lyophilized to afford pure 9 as a fluffy white solid.

General procedure for protein ligation with recombinant thioesters:

The recombinant peptide-intein-chitin binding domain (CBD) fusion was purified on chitin beads from 1 L bacterial culture as previously described. The immobilized fusion protein was then washed with ligation buffer (100mM sodium phosphate, pH 8.0, 100 mM NaCl, 2 % MESNA, 5.0 mL) and allowed to drain. The column was then plugged and incubated at room temperature. After 24 h the column was allowed to drain into 15.0 mL sterile tube. The resin beads were washed with ligation buffer and a total of 4.0 mL were collected. This thioester solution (10 or 11) was first analysed by LC-ESI-MS and then added to lyophilised glycopeptide 9. The reaction mixture was overlain with argon and then shaken at 200 rpm for 48 h. If ligation yields were unsatisfactory (as judged by LC-MS) the product and unreacted glycopeptide were co-purified from the ligation mixture and were resubjected to freshly prepared thioester.
Selected LC-MS DATA

Synthesis of the SAcm Protected glycopeptide thioester 8 (prior to carbohydrate deacetylation): Calculated Mwt. = 7841.0 Da Observed molecular weight = 7844.9 Da).
LC-MS analysis of C-terminal mucin domain 8-9:

**HPLC**
(deacetylated GalNAcs)

**TIC**
After iodoacetamide capping

Electrospray charge Spectrum of glycopeptide 9 (iodoacetamide capped, prior to Acm deprotection)
Deconvoluted Data: calculated mwt. = 7139.4 observed mwt. = 7140.56

Cys(SAcm) deprotection: glycopeptide 9

Calculated Mol. Wt. = 7067.4, observed Mol. Wt. = 7067.7
Glycoform 2.

Deconvoluted Data (calculated molecular weight 15509 Da, observed molecular weight = 11510.0 Da),
Synthesis of glycoform 3: (containing C and N-terminal mucin domains)

Production of the central IEGR-Cys(41-S77)C(O)SCH2CH2SO3- fragment: Calculated molecular weight (MESNA thioester) = 4804 Da, observed molecular weight = 4805.3 Da.

Ligation between thioester 11 and glycopeptide 9:

LC-MS analysis of the ligation reaction (inject 25μL of 4.0 ml reaction):
Analysis of species under each peak:

The m/z species corresponding to the ligated product 12 and associated charge states are shown above.

**Deconvoluted Data:**

The m/z species corresponding to the ligated product 12 and associated charge states are shown above.
After 72 h the remaining bacterially derived thioester was > 50% hydrolysed so the mixed fraction containing 9 and 12 was purified by semi-prep HPLC, lyophilised and re-subjected to freshly prepared bacterial thioester 11 to drive the reaction to completion.

Deconvoluted data after resubjection to bacterial thioester 11:

Once a satisfactory conversion was accomplished the ligated product 12 was purified and treated with synthetic thioester 5 as for the synthesis of glycoform 1. To simplify MS analysis the ligation was reaction was treated directly with hydrazine hydrate to remove acetyl ester protecting groups which are partially cleaved over prolonged reaction times under the reaction conditions.

Deacetylated product calc. Mwt = 16777 Da, observed Mwt = 16776 Da
LC-MS analysis of Glycoform 3