

Post-Synthetic Modifications | Hot Paper |

Chemoselective Glycosylation of Peptides through S-Alkylation Reaction

Enrica Calce,^[a] Giuseppe Digilio,^[b] Valeria Menchise,^[c] Michele Saviano,^[d] and Stefania De Luca^{*[a]}

Abstract: An efficient and rapid procedure for synthesizing S-linked glycopeptides is reported. The approach uses activated molecular sieves as a base to promote the selective S-alkylation of readily prepared cysteine-containing peptides, upon reaction of appropriate glycosyl halides. Considering the very mild conditions employed, the chemoselective link-

age of the electrophilic sugar with a peptide sulfhydryl group occurred in satisfactory yield, allowing the incorporation of mono and disaccharide moieties. The sugar-peptide conjugates obtained from α -D-glycosyl derivatives adopt a β -S-configuration, indicating the high stereoselectivity of the substitution reaction.

Introduction

Protein glycosylation is an important post-translational modification that regulates a wide range of biological processes, including folding and stability of biomolecules, inflammation, immune response, and cell adhesion. Two main types of glycosylated proteins are known in the secretory pathway. They are classified as either N-glycosylated, where the glycan is linked to the side-chain of an asparagine residue, or as O-glycosylated, where the linkage occurs at the hydroxyl side-chain of a serine or threonine residue.^[1] In addition, S-glycosylation on cysteine residues was recently found as an additional post-translational glycosylation that is often linked with antimicrobial activity.^[2-6] S-Glycosylated peptides are attracting growing interest because they are well tolerated by most biological systems and are characterized by an enhanced chemical stability and resistance to proteolytic enzymes.

A number of approaches have been proposed for the chemical synthesis of S-linked glycopeptides, mostly exploiting the exceptional nucleophilicity of the thiol group.^[7-10] Some of these methods rely on site-directed glycoconjugation on pepti-

dic substrates bearing chemical functions that promptly react with appropriate nucleophilic thiol-sugar derivatives.

Among them, the convergent strategies involve a chemoselective linkage of the carbohydrate derivative on a peptide bromide, on a dehydro-peptide, or on an aziridine-2-carboxylic acid-containing peptide.^[11-13]

However, the most commonly employed methodologies for the synthesis of glycopeptides require the preparation of glycosylated amino acid as building blocks, which are incorporated into stepwise solid-phase peptide synthesis (SPPS).^[8,14] This approach allows for complete control over the site-specificity of carbohydrate linkage. Some protocols for the synthesis of SPPS building blocks consist of reacting the nucleophilic thiol-carbohydrate with a suitable amino acid derivative. In a recently published work aimed at incorporating a carbohydrate moiety in Tn antigen peptides, the S-alkylation of a thio- α -D-galactopyranose (GalNAc- α -SH) with a bromoalanine derivative was carried out.^[15] The reaction was promoted by activated molecular sieves (MS). The glycosylated amino acid building block approach is perhaps the most convenient method for the synthesis of glycopeptides, even though issues might arise from merging of the carbohydrate chemistry and the peptide chemistry. However, the synthesis of the building block, in particular the formation of the glycosidic bond, may be challenging, especially when polysaccharides are involved. Therefore, new and more efficient S-glycosylation methods are needed.

In this context, the versatility shown by the cysteine thiol as the nucleophilic substrate for introducing chemical modifications by activation with molecular sieves could represent an alternative route to link a carbohydrate moiety on peptide sequences.^[16-20] Here, we present a novel synthetic strategy to prepare cysteine S-linked peptides. It relies on a cysteine residue, inserted into a preformed peptide sequence, as a sensitive substrate to bind carbohydrate derivatives through post-synthetic S-alkylation. Such a method uses only activated molecular sieves (MS) as the base and the appropriate glycosyl

[a] Dr. E. Calce, Dr. S. De Luca
Institute of Biostructures and Bioimaging, National Research Council
80134 Naples (Italy)
E-mail: stefania.deluca@cnr.it

[b] Dr. G. Digilio
Department of Science and Technologic Innovation
Università del Piemonte Orientale "A. Avogadro", 15121 Alessandria (Italy)

[c] Dr. V. Menchise
Institute of Biostructures and Bioimaging, National Research Council
c/o Molecular Biotechnology Center, 10126 Turin (Italy)

[d] Dr. M. Saviano
Institute of Crystallography, National Research Council, 70126 Bari (Italy)

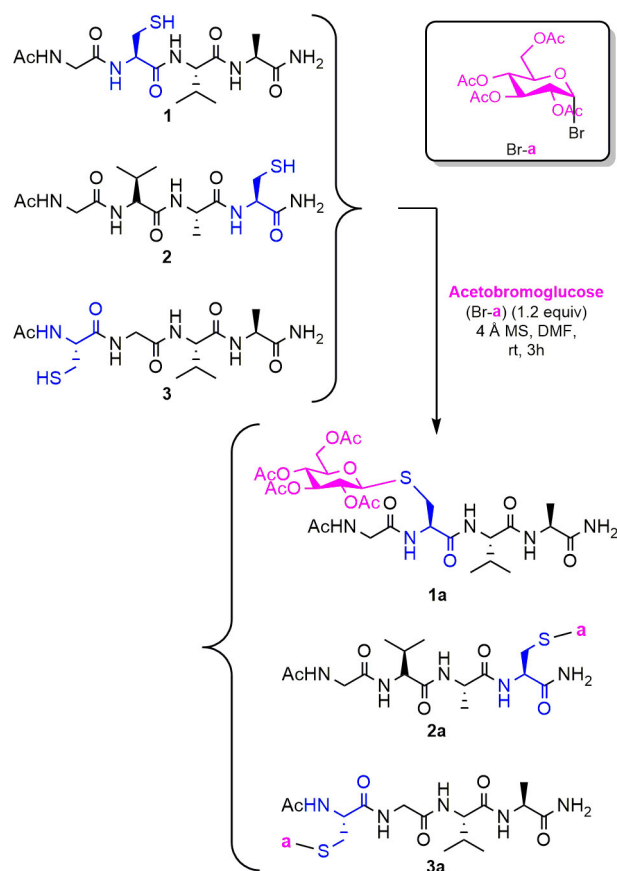
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/chem.201800265>.

halide. Moreover, it allows a good selectivity for the functionalization of the cysteine thiol group, even in the presence of other sensitive nucleophilic groups.

Results and Discussion

A number of model peptides were synthesized to investigate the efficiency of the proposed procedure for *S*-glycosylation. All model peptides are protected at the N-terminus by an acetyl group, and amidated at the C-terminus. Peptides **1–3a** were synthesized to assess the reactivity of the cysteine thiol group upon treatment with acetobromo- α -D-glucose under mild reaction conditions (Scheme 1). Namely, the thiol-containing peptides were dissolved in *N,N*-dimethylformamide (DMF) under an argon atmosphere and in the presence of activated 4 Å molecular sieves ($T=280^{\circ}\text{C}$ for 4 h under vacuum). Then, 1.2 equivalents of the same glycosyl bromide were added and the obtained mixture was kept under stirring for three hours at room temperature (Scheme 1).

As shown in Scheme 1, the cysteine position in the peptide sequence was varied. The aim was to tune the *S*-glycosylation reaction parameters and to assess whether the percentage of the peptide functionalization with the carbohydrate moiety was related to the specific position of the reactive site. In **1a**, the cysteine was in the middle of the sequence, in **2a** it was at



Scheme 1. *S*-Glycosylation of peptide sequences with acetobromo- α -D-glucose.

the C-terminus and in **3a** it was at the N-terminal. The yield of the performed reaction, as estimated by integration of the HPLC peak of the glycosylated peptide compared with that of the starting peptide (and other by-products, if any), was about 60% for all three peptides (Table 1). On peptide **1a**, the removal of the acetyl groups was performed by using a solution of sodium methoxide in methanol, to yield compound **1a'**. Mass spectrometry analysis confirmed full de-acetylation of the glycosylated peptide, without substantial loss of compound.

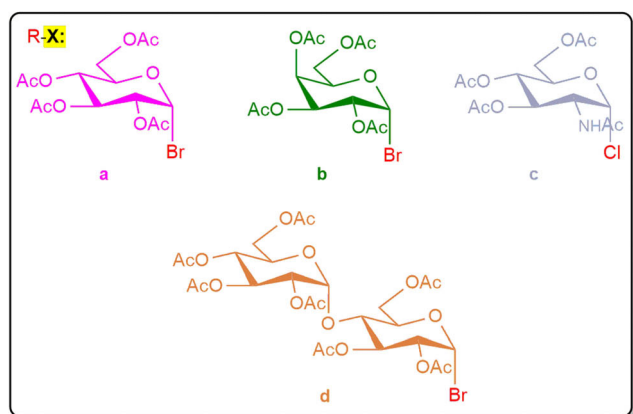
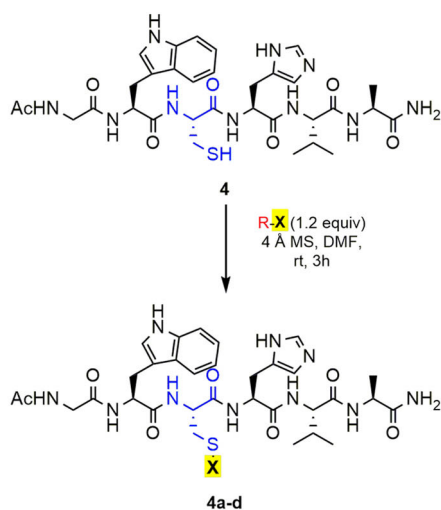
Table 1. Efficiency of the *S*-glycosylation reaction performed on peptide models with different carbohydrate halides.

Peptide	Yield [%]
1a AcGlyCys(Aceto-Glucose)ValAlaNH ₂	60
2a AcGlyValAlaCys(Aceto-Glucose)NH ₂	60
3a AcCys(Aceto-Glucose)GlyValAlaNH ₂	60
4a AcGlyTrpCys(Aceto-Glucose)HisValAlaNH ₂	60
5a AcGlyTrpHisValAlaCys(Aceto-Glucose)NH ₂	60
6a AcCys(Aceto-Glucose)GlyTrpHisValAlaNH ₂	60
4b AcGlyTrpCys(Aceto-Galactose)HisValAlaNH ₂	45
5b AcGlyTrpHisValAlaCys(Aceto-Galactose)NH ₂	40
6b AcCys(Aceto-Galactose)GlyTrpHisValAlaNH ₂	45
4c AcGlyTrpCys(<i>N</i> -acetyl-glucosamine triacetate)HisValAlaNH ₂	35
7d Fmoc-Cys(Aceto-Maltose)NH ₂	45
4d AcGlyTrpCys(Aceto-Maltose)HisValAlaNH ₂	45
8d AcCys(Aceto-Maltose)GlyThrValAlaNH ₂	45
9d AcCys(Aceto-Maltose)GlyArgValAlaNH ₂	40
10d AcCys(Aceto-Maltose)GlyGluValAlaNH ₂	40
11a Ac-Cys(Acm)AlaCys(Aceto-Glucose)SerProCys(Acm)NH ₂	45

Next, the reactivity of the cysteine thiol group upon reaction with several glycosyl halides was explored. The AcGlyTrpCysHisValAlaNH₂ peptide sequence (**4**) was reacted with acetobromo- α -D-glucose (leading to **4a**), or acetobromo- α -D-galactose (leading to **4b**), or 1-chloro-*N*-acetyl-glucosamine triacetate (leading to **4c**) (Scheme 2). The cysteine thiol reacted with all glycosyl halides, although with different yields (Table 1).

It is worth noting that model peptide **4** was chosen because it contains potentially competing nucleophilic groups (namely, Trp and His ring nitrogen atoms) to assess the chemoselectivity for cysteine of the glycosylation reaction. 2D-NOESY NMR spectroscopy was used to establish unambiguously the glycosylation site of the reaction product having the mass expected for **4a**. NOE correlations between the anomeric Glc H₁ and Cys H_β/H_α and the lack of NOEs between the Glc and the His/Trp moieties confirmed the chemoselectivity of the reaction (Figure 1; see the Supporting Information for the complete ¹H NMR assignment).

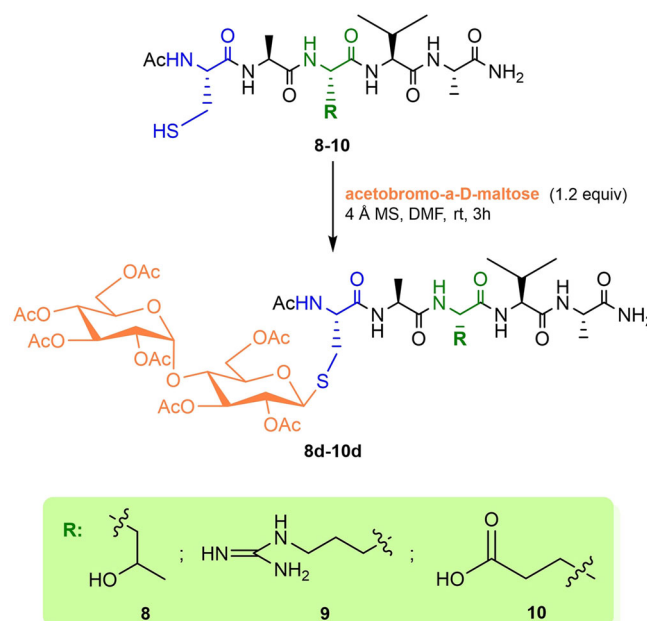
The study was extended to compounds AcGlyTrpHisValAlaCys(Aceto-Glucose)NH₂ (**5a**) and AcCys(Aceto-Glucose)GlyTrpHisValAlaNH₂ (**6a**), to assess whether the position of the cysteine within the polypeptide sequence could affect the thiol nucleophilicity, and hence reaction yields and chemoselectivity. Reaction yields for **5a** and **6a** were similar to that of **4a** (Table 1). As found for **4a**, NOE contacts between Glc H₁ and Cys H_β/H_α (see the Supporting Information) confirmed an excellent chemoselectivity for the Cys thiol over potentially com-



Scheme 2. S-Glycosylation of model peptide **4** with different carbohydrate moieties.

peting His and Trp sites. The ^1H NMR chemical shift of the anomeric Glc H₁ was approximately 4.9 ppm (and $^3J_{\text{H}_1, \text{H}_2} = 9.8$ Hz) for all three compounds, indicating the β -configuration of the glucopyranose ring.^[21] This is consistent with the expected $\text{S}_{\text{N}}2$ reaction mechanism.^[8]

The versatility of the developed protocol was also assessed by carrying out S-glycosylation with a disaccharide unit, which is known to be more challenging because of steric hindrance effects.^[8] Fmoc-Cys-OH (compound **7**) and model peptides **4**, **8**, **9** and **10** were derivatized with acetobromo- α -D-maltose under the S-alkylation conditions described above, to give compounds **7d**, **4d**, **8d**, **9d**, and **10d**, respectively (Scheme 2 and Scheme 3). HPLC analysis of the crude reaction products



Scheme 3. S-Glycosylation of different peptide sequences with acetobromo- α -D-maltose.

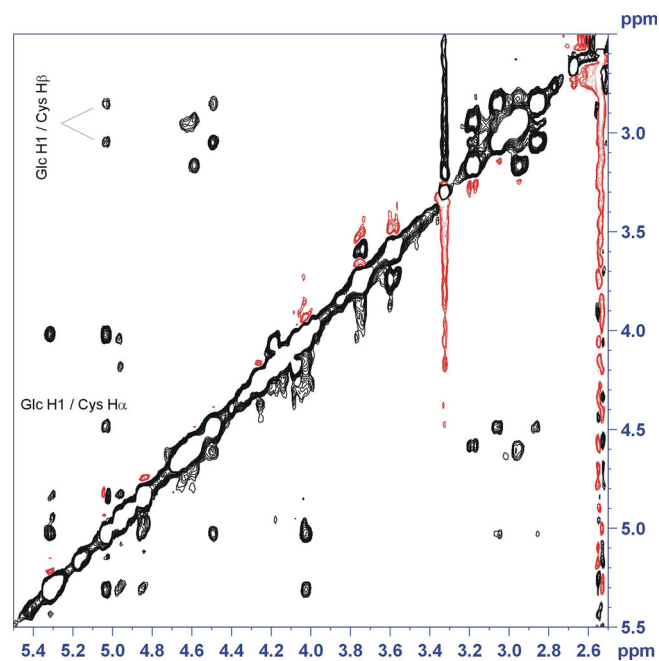


Figure 1. Expansion of the 2D-NOESY NMR spectrum of **4a** (300 K, $[\text{D}_6]\text{DMSO}$) showing NOE contacts between Glc H₁ and Cys H_β/H_α.

showed two main peaks (accounting for >90% of the total AUC), corresponding to the glycosylated species and the unreacted peptide as verified by MS analysis of purified fractions. The reaction yields for S-glycosylation with maltose were invariably in the range 40–45%, slightly less than that of 60% typically obtained for analogous reaction with glucose. Such a decrease in yield is in line with increasing steric hindrance on going from mono- to disaccharide units.

The peptide sequences used to test S-alkylation contained a number of residues potentially competing with cysteine for alkylation: namely, histidine and tryptophan (**4d**), threonine (**8d**), arginine (**9d**), and glutamic acid (**10d**) (Scheme 2 and Scheme 3).

Therefore, these compounds were subjected to NMR spectroscopic analysis to ascertain the glycosylation site (sample ^1H NMR spectra are given in Figure 2). Full assignment of all proton resonance was achieved by analysis of 2D-TOCSY, 2D-COSY, and 2D-NOESY spectra (see the Supporting Information). All compounds showed the expected NOE contacts between

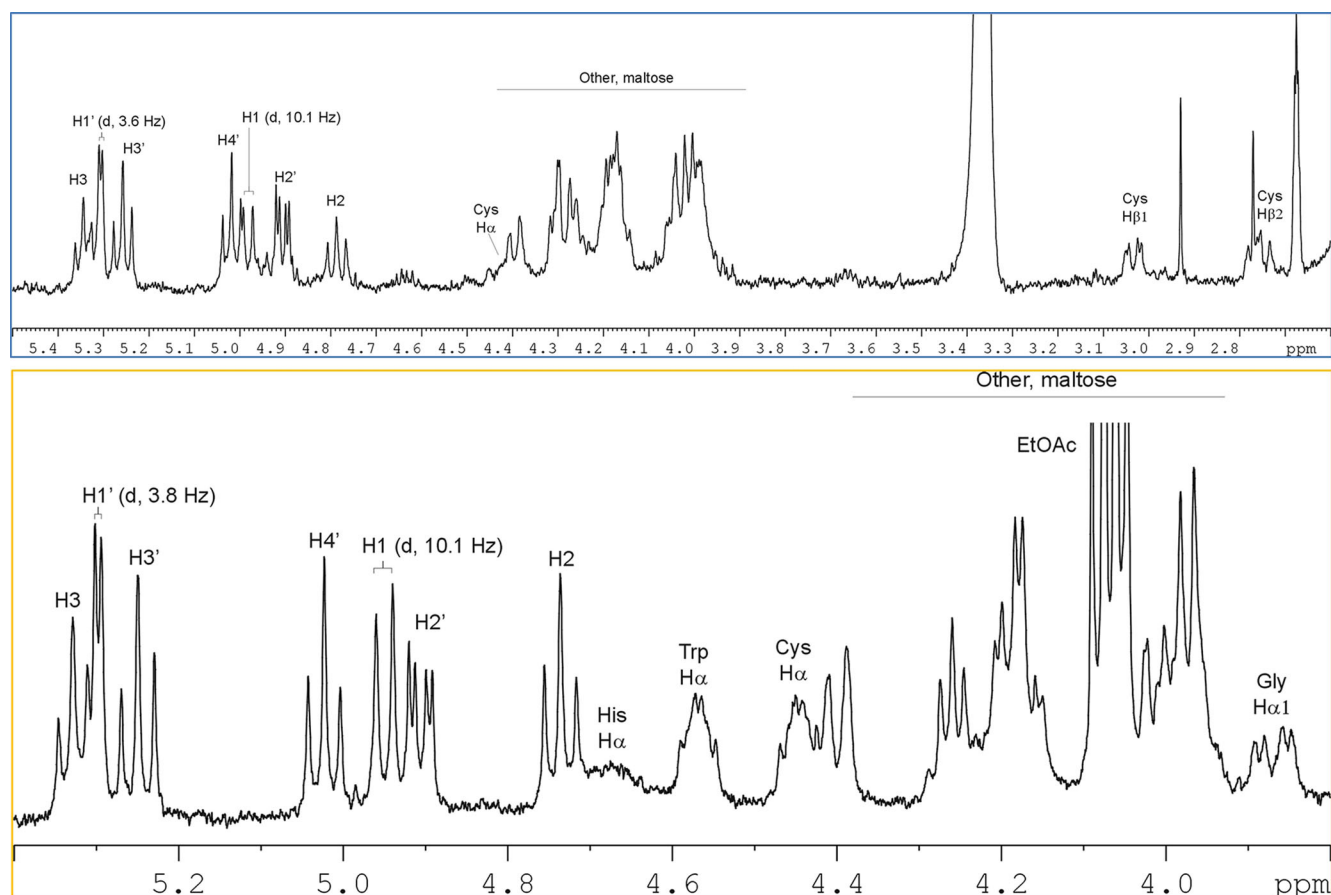


Figure 2. Expansion of the ^1H NMR spectra ($[\text{D}_6]\text{DMSO}$, 300 K) of **7d** (top) and **4d** (bottom), showing the assignment of the β -heptacetyl-maltosyl moiety. See the Supporting Information for atom numbering and full assignment.

Cys $\text{H}_\beta/\text{H}_\alpha$ and the anomeric H_1 proton of the maltose moiety, and showed very similar chemical shifts of the Cys H_β protons and maltose protons (see also the Supporting Information). Finally, the $^3J_{\text{H}_1, \text{H}_2}$ of the S-linked glycopyranose ring had a value of ca. 10 Hz for all compounds considered. Taken together, these data clearly indicated chemoselectivity for the cysteine thiol, with the S-linked glycopyranose ring adopting a β -configuration as expected (Figure 2).

Free lysine as a competing alkylation site was not included in the chemoselectivity study because it is well known that it does compete for alkylation and it is expected to lower to a certain extent the yield of the mono-S-glycosylated peptide.^[16] However, the formation of poly-glycosylated products can be easily avoided by employing protected lysines, such as Fmoc-Lys(ivDde)-OH and Fmoc-Lys(Alloc)-OH, which are routinely used for solid-phase peptide synthesis.

Next, microwave radiation was applied to the developed glycosylation post-synthetic protocol to shorten the reaction time. In particular, the microwave heating procedure was employed to synthesize compounds **4a**, **4d**, and **7d**. The obtained mixture was irradiated and stirred for 5 min at 40 °C in a microwave source apparatus (Initiator, Biotage Sweden AB, Uppsala, Sweden), while all other reaction parameters were kept unchanged. The desired products were obtained with a lower yield (ca. 10%), although a by-product, which was not detect-

ed under the standard reaction conditions, was generated and could not be isolated by HPLC from the glycosylated peptide. Mass spectrometry analysis revealed that this by-product results from E2 elimination from the sugar moiety, as competing reaction to the sulfhydryl nucleophilic substitution.^[22–23] This side reaction was found to be particularly relevant where the glycosylation involved the disaccharide moiety, in both cases, upon reaction with the peptide **4** and with the Fmoc-Cys-OH (**7**). In Figure 3 the mass spectra of **7d** and **4d** obtained under MW irradiation are reported; both clearly showed the presence of the glycal product derived from sugar elimination reaction ($[M + \text{H}]^+ = 619.3$). The results obtained discourage the application of microwave radiation as a procedure for functionalizing fully unprotected peptides with glycan moieties, in the presence of activated MS.

As an example of a practical application of our synthetic strategy, glycopeptide **11a** was synthesized. It is derived from Cys-Ala-Ser(β -O-Glc)-Ser-Pro-Cys, which is a common consensus sequence present in several naturally occurring glycoproteins. The design, the synthesis, and the characterization of this peptide segment have been previously reported.^[24] Namely, Ser3 was substituted with a cysteine residue and an S-glycosylation was performed under our standard protocol, while both Cys1 and Cys6 were protected with an Ac group (**11a**) (Figure 4). The desired compound **11a** was obtained

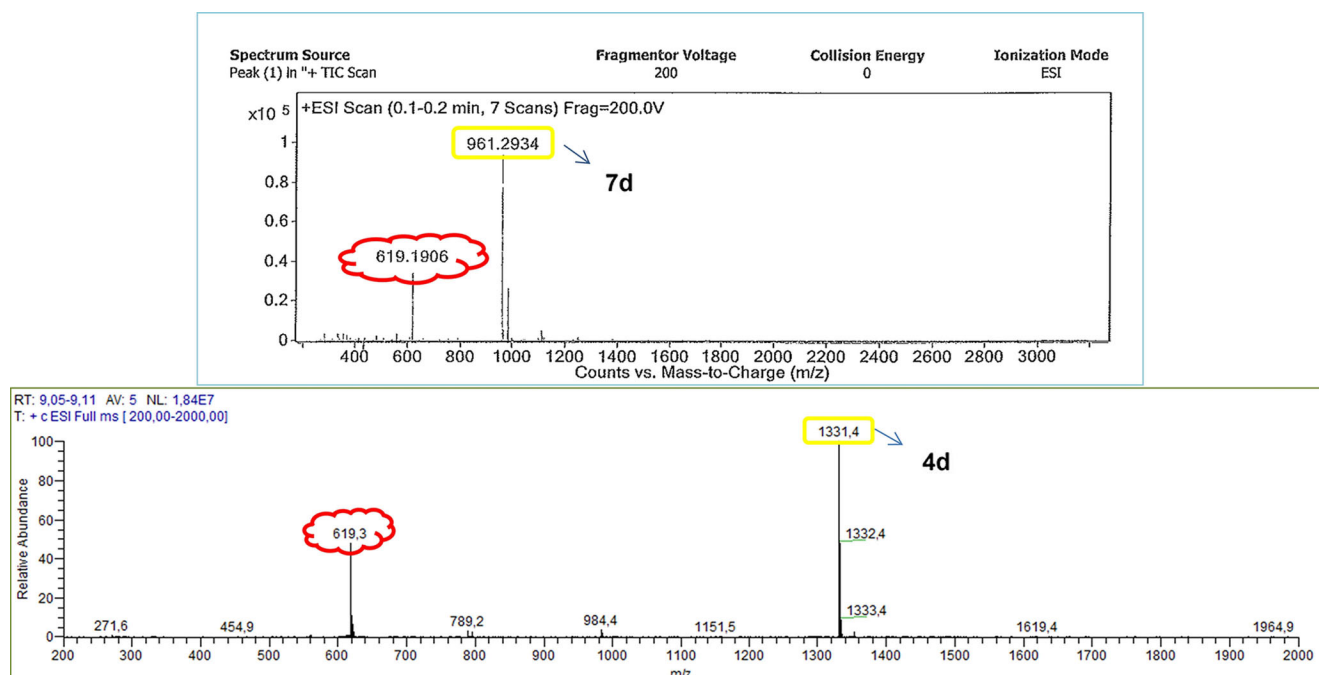


Figure 3. MS profile of compounds 7d and 4d.

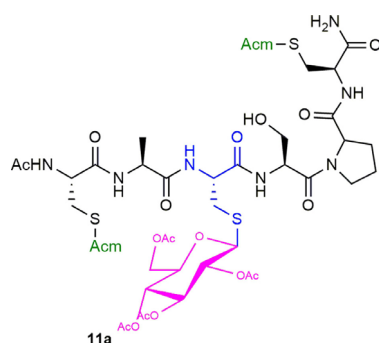


Figure 4. Modified consensus peptide sequence of natural glycoprotein.

with 45% yield. The final removal of the Acm protecting group, performed by the standard procedure described in the Experimental Section, was characterized by a quantitative yield of the desired deprotected compound.

Considering all the performed glycosylation reactions, the percentage of functionalization spans from 40 to 60%, which is reduced by a further 20% after purification by RP-chromatography. Although these yields are moderate, they are still interesting for a post-synthetic reaction, especially when considering that an excess of only 1.2 equivalent of the sugar derivative is employed to preserve the chemoselectivity of the S-alkylation process.

Conclusion

A relatively simple procedure for site-directed glycosylation of cysteine sulfhydryl group on peptides is described. A number of glycosylated peptide molecules were synthesized by em-

ploying carbohydrate electrophilic precursors and only activated MS as the base. The reaction occurs essentially through the S_N2 reaction mechanism, leading to S- β -linked glycopeptides. The maltose disaccharide moiety can be conjugated to a pre-formed Cys-containing peptide sequence with a yield comparable to the conjugation of a glucose unit. Although the glycosylation yields were moderate, it is worth noting that the peptide sequence does not require the incorporation of specific reactive nucleophilic substrates, nor the activation of the thiol group to achieve the chemo and stereoselective cysteine functionalization. Moreover, no manipulation of the sugar derivative is required, but commercially available (and relatively inexpensive) acetyl-protected carbohydrate halides are employed. This aspect has a relevant consequence, because it greatly facilitates the chemical procedure, making it accessible to chemists who are not specialized in carbohydrate synthesis.

Experimental Section

Materials and methods

Fmoc protected amino acids, Fmoc-Cys(Acm)-OH, Rink Amide MBHA resin, *N*-hydroxybenzotriazole (HOBT), and benzotriazol-1-yl-oxy-tris-pyrrolidino-phosphonium (PyBOP) were purchased from Calbiochem-Novabiochem (Laufelfingen, Switzerland); piperidine and diisopropylethylamine (DIPEA) were purchased from Fluka (Milwaukee, WI), and all remaining solvents were purchased from Aldrich (St Louis, MI) or Fluka (Milwaukee, WI) and were used without further purification, unless otherwise stated. Molecular sieves type 4 Å (beads, diameter 1.6 mm) were purchased from Aldrich and activated by heating at 280 °C for 4 h under vacuum. Acetobromo- α -D-glucose, acetobromo- α -D-galactose, 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-D-glucopyranosyl chloride (1-chloro-*N*-acetyl-

glucosamine triacetate), 2,3,6,2',3',4',6'-hepta-O-acetyl- α -D-maltosyl bromide (acetobromo-maltose) were purchased from Carbosynth (Berkshire, UK).

For all the RP-HPLC procedures the system solvent used was: H₂O 0.1% TFA (A) and CH₃CN 0.1% TFA (B), with a linear gradient starting from 5% to 70% B in 30 min or in 10 min and detection at 210 nm and 280 nm. Analytical RP-HPLC runs were carried out using a C18 column, 4.6×250 mm with a flow rate of 1.0 mL min⁻¹; preparative RP-HPLC was carried out using a C18 column, 22×250 mm with a flow rate of 20 mL min⁻¹.

For the LC-MS analysis, the system solvent used was: H₂O 0.05% TFA (A) and CH₃CN 0.05% TFA (B), with a linear gradient starting from 5% to 70% B in 10 min and detection at 210 nm and 280 nm. LC-ES-MS data were obtained using a C18 Kinetex column (5 μ m, 60×4.60 mm) with a flow rate of 1.0 mL min⁻¹ on a LCQ DECA XP MAX (Thermo/Finnigan) or using an Agilent 1290 Infinity LC System coupled to an Agilent 6230 TOF LC/MS System (Agilent Technologies, Cernusco Sul Naviglio, Italy). ESI-TOF-MS analyses were performed with an Agilent 1290 Infinity LC System coupled to an Agilent 6230 TOF LC/MS System (Agilent Technologies, Cernusco Sul Naviglio, Italy).

NMR spectra were acquired with a Bruker Avance III spectrometer operating at 11.7 T (corresponding to a proton Larmor frequency of 500 MHz), equipped with an inverse Z-gradient 5 mm BBFO probe. Temperature was set to 300.0 K, and controlled within \pm 0.1 K by means of the BTO2000 VTU system. Samples were dissolved in 600 μ L of [D₆]DMSO (99.9 atom %). The residual solvent resonance at 2.54 ppm was used as a secondary reference for chemical shift calibration. Resonance assignment was based on analysis of homonuclear 2D-TOCSY, 2D-NOESY, and 2D-COSY spectra. The 2D-COSY NMR spectra were acquired by means of the Bruker cosydfp pulse program (phase sensitive, double quantum filtered COSY). Presaturation of the broad water resonance was used if needed. Acquisition parameters included: 2 s relaxation delay, 32–48 scans, 16 dummy scans, 0.12 s acquisition time, 25 Hz bandwidth for the water suppression presaturation pulse (if required), 2048×256–400 data points, 17 ppm spectral width (in F2 and F1). Data were treated with squared cosine window functions (both along F2 and F1) prior to complex FT. 2D-TOCSY spectra were acquired with the Bruker mlevph pulse program (homonuclear Hartman–Hahn transfer by means of the MLEV17 sequence) in the phase-sensitive mode according to the States-TPPI Scheme. Typical acquisition parameters included: 2 s relaxation delay, 32–48 scans, 16 dummy scans, 0.12 s acquisition time, 25 Hz bandwidth for the water suppression presaturation pulse (if required), 2048×256–400 data points, 17 ppm spectral width (in F2 and F1), and 100 ms mixing time. Data were treated with squared cosine window functions (both along F2 and F1) prior to complex FT. 2D-NOESY spectra were acquired with the Bruker noesyph pulse sequence in the phase-sensitive mode by the States TPPI Scheme. Typical acquisition and processing parameters were as for 2D-TOCSY spectra, but with a mixing time of 350 ms and 48–64 scans. Spectra were processed by the Bruker Topspin 3.0 software package. Sequence specific assignment was carried out by the Computer Aided Resonance Assignment software package (CARA: R.L.J Keller "The Computer Aided Resonance Assignment", 2004 CANTINA Verlag, Goldau, Switzerland).

AcGlyCys(Aceto-Glucose)ValAlaNH₂ (1 a): To an 8 mL bottom-filtration reaction vessel, Rink Amide MBHA resin (70 mg, loading 0.7 mmol g⁻¹; scale 0.050 mmol) was transferred and swollen in DMF (4–5 mL) for 30 min. Each appropriate Fmoc-amino acid derivative (0.25 mmol) was activated in a separate vial by using PyBop (0.25 mmol)/HOBt (0.025 mmol)/DIPEA (0.5 mmol) activation proto-

col and DMF (3 mL) as solvent. The coupling step was monitored by ninhydrin test after 60 min of coupling cycle. Afterwards, the Fmoc-deprotection was performed by suspending the peptidyl-resin in a 4 mL solution of 20% piperidine in DMF (5 + 10 min). The N-terminal peptide was acetylated by treatment with a mixture (5 mL) of acetic anhydride (4.7%) and pyridine (4%) in DMF for 10 min. The peptide cleavage from the solid support and the simultaneous deprotection of all amino acid side chains were performed by suspending the peptidyl-resins in a 5 mL solution of TFA/H₂O/TIS (97:2:1) for 3 h at RT. The resin was filtered and washed with a small amount of cleavage cocktail and the filtrate was concentrated under a stream of nitrogen. The crude product was isolated by precipitation into cold diethyl ether and centrifuged to form a pellet (85% yield of crude peptide 1).

The obtained acetylated peptide (10 mg, 0.026 mmol) was dissolved in DMF (3 mL), under argon atmosphere, and placed in a round-bottom flask. Then, the alkylating reagent acetobromo- α -D-glucose (12.8 mg, 0.031 mmol) and, as final step, 4 Å molecular sieves (3–3.5 g), previously activated at 280 °C for 4 h under vacuum (10⁻⁴ mBar), were added. The obtained solution was stirred at RT for 3 h and the reaction was followed by analytical RP-HPLC. The mixture was centrifuged to eliminate the sieves and the supernatant was concentrated under vacuum. The obtained crude product was purified by RP-HPLC to give **1 a** in 24% yield (2.4 mg), as estimated after lyophilization, and analyzed by mass spectrometry. HPLC: *t*_R=20.95 min; ES-MS: *m/z* calculated 719.92 [M + H]⁺; found: 720.3.

Removal of the acetyl groups was achieved by dissolving the glycopeptide in the necessary volume of methanol, then sodium methoxide was added until a pH near 9 or 9.5. The solution was stirred for 1 h and the reaction was followed by mass spectrometry analysis. After acidification with Dowex resin, the glycopeptide was precipitated in cold diethyl ether and analyzed by mass spectrometry. ES-MS: *m/z* calculated 551.65 [M + H]⁺; found: 551.16

AcGlyValAlaCys(Aceto-Glucose)NH₂ (2 a): Peptide **2 a** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo- α -D-glucose was employed as alkylating agent (12.8 mg, 0.031 mmol). The final RP-HPLC purification provided compound **2 a** in 25% yield (2.5 mg). HPLC: *t*_R=21.09 min; ES-MS: *m/z* calculated: 719.92 [M + H]⁺; found: 720.5.

AcCys(Aceto-Glucose)GlyValAlaNH₂ (3 a): Peptide **3 a** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo- α -D-glucose was employed as alkylating agent (12.8 mg, 0.031 mmol). The final RP-HPLC purification provided compound **3 a** in 20% yield (2.0 mg). HPLC: *t*_R=20.84 min; ES-MS: *m/z* calculated: 719.92 [M + H]⁺; found: 720.4. ¹H NMR ([D₆]DMSO, 300 K): δ =8.21 (Gly H_N), 8.17 (Cys H_N), 7.99 (Ala H_N), 7.80 (Val H_N), 7.21/6.96 (C-term amide), 5.29 (Glc H3), 5.00 (Glc H1, ³J_{H1,H2}=9.1 Hz), 4.96 (Glc H4), 4.83 (Glc H2), 4.48 (Cys H α), 4.22 (Ala H α), 4.19 (Val H α), 4.18/4.07 (Glc H6), 4.02 (Glc H5), 3.82/3.76 (Gly H α), 3.06/2.75 (Cys H β), 2.05–1.98 (O-acetyl), 2.00 (Val H β), 1.89 (N-term acetyl), 1.25 (Ala H β), 0.88/0.84 (Val H γ 1/H γ 2).

AcGlyTrpCys(Aceto-Glucose)HisValAlaNH₂ (4 a): Peptide **4 a** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo- α -D-glucose was employed as alkylating agent (6.9 mg, 0.017 mmol). The final RP-HPLC purification provided compound **4 a** in 25% yield (2.5 mg). HPLC: *t*_R=24.13 min; ES-MS: *m/z* calculated: 1043.97 [M + H]⁺; found: 1043.7; ¹H NMR ([D₆]DMSO, 300 K): δ =10.85 (Trp H ϵ 1), 8.40 (Cys H_N), 8.31 (Ala H_N), 8.17 (His H ϵ 1), 8.08 (Gly H_N), 8.07 (Trp H_N), 8.01 (His H_N), 7.90 (Val H_N), 7.62 (Trp H ϵ 3), 7.35 (Trp H ζ 2), 7.20/6.98 (C-term amide), 7.17 (His H δ 2), 7.16 (Trp H δ 1), 7.09 (Trp H η 2), 7.00 (Trp H ζ 3), 5.31 (Glc H3), 5.03 (Glc H1, ³J_{H1,H2}=9.9 Hz), 4.95 (Glc H4),

4.84 (Glc H2), 4.63 (His H α), 4.58 (Trp H α), 4.48 (Cys H α), 4.25 (Ala H α), 4.19/4.07 (Glc H6), 4.13 (Val H α), 4.02 (Glc H5), 3.74/3.59 (Gly H α), 3.17/2.94 (Trp H β), 3.05/2.85 (Cys H β), 3.03/2.95 (His H β), 2.10–1.95 (O-acetyl), 2.06 (Val H β), 1.83 (N-term acetyl), 1.25 (Ala H β), 0.90/0.87 (Val H γ 1/H γ 2).

AcGlyTrpHisValAlaCys(Aceto-Glucose)NH₂ (5a): Peptide **5a** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo- α -D-glucose was employed as alkylating agent (6.9 mg, 0.017 mmol). The final RP-HPLC purification provided compound **5a** in 21% yield (2.1 mg). HPLC: t_R = 24.10 min; ES-MS: m/z calculated: 1043.97 [M + H]⁺; found: 1043.8; ¹H NMR ([D₆]DMSO, 300 K): δ = 10.84 (Trp H ϵ 1), 8.43 (His H ϵ), 8.35 (Ala H ϵ), 8.08 (Trp H ϵ), 8.07 (Gly H ϵ), 7.92 (Cys H ϵ), 7.77 (Val H ϵ), 7.59 (Trp H ϵ 3), 7.42/7.20 (C-term amide), 7.35 (Trp H ζ 2), 7.16 (Trp H δ 1), 7.10 (Trp H η 2), 7.00 (Trp H ζ 3), 5.29 (Glc H3), 4.98 (Glc H1, ³J_{H1,H2} = 9.9 Hz), 4.97 (Glc H4), 4.84 (Glc H2), 4.64 (His H α), 4.56 (Trp H α), 4.37 (Cys H α), 4.31 (Ala H α), 4.21 (Val H α), 4.19–4.06 (Glc H6), 4.01 (Glc H5), 3.75/3.62 (Gly H α), 3.14/2.96 (Trp H β), 3.07/2.99 (His H β), 3.05/2.86 (Cys H β), 2.05–1.95 (O-acetyl), 2.04 (Val H β), 1.86 (N-term acetyl), 1.27 (Ala H β), 0.90/0.87 (Val H γ 1/H γ 2).

AcCys(Aceto-Glucose)GlyTrpHisValAlaNH₂ (6a): Peptide **6a** was synthesized, characterized, and purified under the experimental conditions above described. The acetobromo- α -D-glucose was employed as alkylating agent (6.9 mg, 0.017 mmol). The final RP-HPLC purification provided compound **6a** in 23% yield (2.3 mg). HPLC: t_R = 24.04 min; ES-MS: m/z calculated: 1043.97 [M + H]⁺; found: 1043.5; ¹H NMR ([D₆]DMSO, 300 K): δ = 10.81 (Trp H ϵ 1), 8.45 (His H ϵ), 8.26 (Ala H ϵ), 8.17 (Cys H ϵ), 8.16 (Gly H ϵ), 8.13 (Trp H ϵ), 7.79 (Val H ϵ), 7.61 (Trp H ϵ 3), 7.35 (Trp H ζ 2), 7.25/7.01 (C-term amide), 7.15 (Trp H δ 1), 7.08 (Trp H η 2), 7.00 (Trp H ζ 3), 5.28 (Glc H3), 4.97 (Glc H1, ³J_{H1,H2} = 9.9 Hz), 4.96 (Glc H4), 4.83 (Glc H2), 4.62 (His H α), 4.56 (Trp H α), 4.46 (Cys H α), 4.24 (Ala H α), 4.16 (Val H α), 4.16/4.05 (Glc H6), 3.98 (Glc H5), 3.85/3.60 (Gly H α), 3.15/2.95 (Trp H β), 3.05/2.95 (His H β), 3.04/2.74 (Cys H β), 2.04 (Val H β), 2.02–1.94 (O-acetyl), 1.87 (N-term acetyl), 1.25 (Ala H β), 0.89/0.86 (Val H γ 1/H γ 2).

AcGlyTrpCys(Aceto-Galactose)HisValAlaNH₂ (4b): Peptide **4b** was synthesized, characterized, and purified under the experimental conditions above described. The acetobromo- α -D-galactose was employed as alkylating agent (6.9 mg, 0.017 mmol). The final RP-HPLC purification provided compound **4b** in 18% yield (1.8 mg). HPLC: t_R = 24.93 min; ES-MS: m/z calculated: 1043.97 [M + H]⁺; found: 1043.8.

AcGlyTrpHisValAlaCys(Aceto-Galactose)NH₂ (5b): Peptide **5b** was synthesized, characterized and purified under the experimental conditions above described. The acetobromo- α -D-galactose was employed as alkylating agent (6.9 mg, 0.017 mmol). The final RP-HPLC purification provided compound **5b** in 17% yield (1.7 mg). HPLC: t_R = 25.01 min; ES-MS: m/z calculated: 1043.97 [M + H]⁺; found: 1043.7.

AcCys(Aceto-Galactose)GlyTrpHisValAlaNH₂ (6b): Peptide **6b** was synthesized, characterized and purified under the experimental conditions described above. The acetobromo- α -D-galactose was employed as alkylating agent (6.9 mg, 0.017 mmol). The final RP-HPLC purification provided compound **6b** in 19% yield (1.9 mg). HPLC: t_R = 24.15 min; ES-MS: m/z calculated: 1043.97 [M + H]⁺; found: 1043.7.

AcGlyTrpCys(N-acetyl-glucosamine triacetate)HisValAlaNH₂ (4c): Peptide **4c** was synthesized, characterized and purified under the experimental conditions above described. The 1-chloro-N-acetyl-glucosamine triacetate was employed as alkylating agent (7.4 mg, 0.020 mmol). The final RP-HPLC purification provided compound

4c in 15% yield (1.5 mg). HPLC: t_R = 24.20 min; ES-MS: m/z calculated: 1042.53 [M + H]⁺; found: 1042.6.

Fmoc-Cys(Aceto-Maltose)NH₂ (7d): Rink Amide MBHA resin (0.7 mmol g⁻¹ substitution; 100 μ mol scale) was used as solid support, as it releases the amino acid amidated at C-terminus upon acid treatment. Fmoc-Cys(trt)-OH, activated in situ by PyBop/HOBt/DIPEA, was coupled for 60 min. The cleavage from the solid support and the simultaneous deprotection of the trityl group were performed by suspending the resin in TFA/H₂O/TIS (97:2:1) for 2 h. The Fmoc-Cys amidated at its C-terminus was isolated by precipitation into cold distilled water and centrifuged to form a pellet.

Acetylated amino acid (10 mg, 0.029 mmol) in DMF (5 mg mL⁻¹) was added, under argon atmosphere, to a round-bottom flask containing 4 Å molecular sieves, previously activated at 280 °C for 4 h under vacuum. After a few minutes, the acetobromo-maltose (24.6 mg, 0.035 mmol) was added. The mixture was stirred at RT for 3 h and followed by analytical RP-HPLC. The mixture was centrifuged and the precipitate was washed with DMF. The final RP-HPLC purification provided compound **7d** in 28% yield (2.8 mg). HPLC: t_R = 8.18 min; ES-MS: m/z calculated: 961.6 [M + H]⁺; found: 961.2; ¹H NMR ([D₆]DMSO, 300 K): δ = 7.93 (d), 7.80 (d), 7.76 (t), 7.66 (t), 7.5–7.4 (m, o), 7.4–7.3 (m, o), 7.21 (s), 5.34 (t, Glc H3), 5.31 (d, Mlt H1', ³J_{H1',H2'} = 3.6 Hz), 5.26 (t, Mlt H3'), 5.02 (t, Mlt H4'), 4.98 (d, Mlt H1, ³J_{H1,H2} = 10.1 Hz), 4.90 (dd, Mlt H2'), 4.79 (t, Mlt H2), 4.41 (m, o, Cys H α), 4.40–3.90 (Mlt H4, H5; H6A, H6B, H5', H6'A, H6'B), 3.04 (m, Cys H β 1), 2.76 (m, Cys H β 2), 2.08–1.93 (O-acetyl, signal manifold).

AcGlyTrpCys(Aceto-Maltose)HisValAlaNH₂ (4d): Peptide **4d** was synthesized, characterized and purified under the experimental conditions described above. The acetobromo-maltose was employed as alkylating agent (11.8 mg, 0.017 mmol). The final RP-HPLC purification provided compound **4d** in 23% yield (2.3 mg). HPLC: t_R = 27.02 min; ES-MS: m/z calculated: 1331.22 [M + H]⁺; found: 1331.4; ¹H NMR ([D₆]DMSO, 300 K): δ = 10.84 (Trp H ϵ 1), 8.96 (His H ϵ 1), 8.39 (Cys H ϵ), 8.12 (Ala H ϵ), 8.09 (Gly H ϵ), 8.07 (Trp H ϵ), 8.03 (His H ϵ), 7.92 (Val H ϵ), 7.62 (Trp H ϵ 3), 7.38 (His H δ 2), 7.35 (Trp H ζ 2), 7.28/7.00 (C-term amide), 7.18 (Trp H δ 1), 7.09 (Trp H η 2), 7.01 (Trp H ζ 3), 5.33 (Glc H3), 5.30 (Glc H1', ³J_{H1',H2'} = 3.8 Hz), 5.25 (Glc H3'), 5.01 (Glc H4'), 4.95 (Glc H1, ³J_{H1,H2} = 10.1 Hz), 4.89 (Glc H2'), 4.74 (Glc H2), 4.72 (His H α), 4.55 (Trp H α), 4.44 (Glc H6B), 4.43 (Cys H α), 4.24 (Ala H α), 4.19/4.05 (Glc H6'), 4.18 (Val H α), 4.16 (Glc H5,H6A), 4.01 (Glc H5'), 3.98 (Glc H4), 3.76/3.57 (Gly H α), 3.16/2.93 (Trp H β), 3.11/3.02 (His H β), 3.02/2.84 (Cys H β), 2.08–1.94 (O-acetyl), 2.04 (Val H β), 1.83 (N-term acetyl), 1.24 (Ala H β), 0.89–0.87 (Val H γ 1/H γ 2).

AcCys(Aceto-Maltose)GlyThrValAlaNH₂ (8d): Peptide **8d** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo-maltose was employed as alkylating agent (16.8 mg, 0.024 mmol). The final RP-HPLC purification provided compound **8d** in 20% yield (2.0 mg). HPLC: t_R = 7.47 min; ES-MS: m/z calculated: 1109.64 [M + H]⁺; found: 1109.20; ¹H NMR ([D₆]DMSO, 300 K): δ = 8.23 (t, Gly H ϵ), 8.19 (d, Cys H ϵ), 7.92 (d, Ala H ϵ), 7.78 (d, Thr H ϵ), 7.75 (d, Val H ϵ), 7.20/6.98 (C-term amide), 5.33 (t, Glc H3), 5.30 (d, Glc H1', ³J_{H1',H2'} = 3.8 Hz), 5.25 (t, Glc H3'), 5.02 (t, Glc H4'), 4.96 (d, Glc H1, ³J_{H1,H2} = 10.1 Hz), 4.90 (dd, Glc H2'), 4.74 (t, Glc H2), 4.45–4.35 (overlapping Glc H6B, Cys H α , Thr H α), 4.20–4.15 (overlapping Ala H α , Val H α , Glc H6'B, Glc H5, Glc H6A), 4.05–3.95 (overlapping Thr H β , Glc H5', Glc H6'A, Glc H4), 3.83 (m, Gly H α 1/ α 2), 3.02 (dd, Cys H β 2), 2.73 (dd, Cys H β 3), 2.10–1.94 (overlapping Val H β , O-acetyl), 1.89 (N-acetyl), 1.22 (d, Ala H β), 1.05 (d, Thr H γ 2), 0.89/0.87 (d, Val H γ 1/H γ 2).

AcCys(Aceto-Maltose)GlyArgValAlaNH₂ (9d): Peptide **9d** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo-maltose was employed as alkylating agent (15.4 mg, 0.022 mmol). The final RP-HPLC purification provided compound **9d** in 18% yield (1.8 mg). HPLC: $t_R = 8.00$ min; ES-MS: m/z calculated: 1164.70 $[M+H]^+$; found: 1164.25; ¹H NMR ([D₆]DMSO, 300 K): $\delta = 8.23$ (t, Gly H_N), 8.17 (d, Cys H_N), 7.96 (d, Arg H_N), 7.90 (d, Ala H_N), 7.85 (d, Val H_N), 7.42 (t, Arg H ϵ), 7.33/7.03 (C-term amide), 5.34 (t, Glc H3), 5.30 (d, Glc H1', ³J_{H1',H2} = 3.8 Hz), 5.25 (t, Glc H3'), 5.02 (t, Glc H4'), 4.95 (d, Glc H1, ³J_{H1,H2} = 10.1 Hz), 4.90 (dd, Glc H2'), 4.74 (t, Glc H2), 4.44–4.36 (overlapping Glc H6B, Cys H α Arg H α), 4.25–4.14 (overlapping Ala H α , Val H α , Glc H6'B, Glc H5, Glc H6A), 4.06–3.95 (Glc H5', Glc H6'A, Glc H4), 3.77 (m, Gly H α 1/ α 2), 3.12 (dt, Arg H δ), 3.02 (dd, Cys H β 2), 2.74 (dd, Cys H β 3), 2.10–1.94 (overlapping Val H β , O-acetyl), 1.89 (N-acetyl), 1.72/1.56 (m, Arg H β), 1.51 (m, Arg H γ), 1.24 (d, Ala H β), 0.89–0.87 (d, Val H γ 1/H γ 2).

AcCys(Aceto-Maltose)GlyGluValAlaNH₂ (10d): Peptide **10d** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo-maltose was employed as alkylating agent (16.2 mg, 0.023 mmol). The final RP-HPLC purification provided compound **10d** in 21% yield (2.1 mg). HPLC: $t_R = 7.98$ min; ES-MS: m/z calculated: 1137.63 $[M+H]^+$; found: 1137.19; ¹H NMR ([D₆]DMSO, 300 K): $\delta = 8.25$ (t, Gly H_N), 8.17 (d, Cys H_N), 7.92 (d, Glu H_N), 7.94 (d, Ala H_N), 7.80 (d, Val H_N), 7.27/6.97 (C-term amide), 5.32 (t, Glc H3), 5.29 (d, Glc H1', ³J_{H1',H2} = 3.8 Hz), 5.26 (t, Glc H3'), 5.06 (t, Glc H4'), 4.95 (d, Glc H1, ³J_{H1,H2} = 10.1 Hz), 4.88 (dd, Glc H2'), 4.74 (t, Glc H2), 4.45–4.30 (overlapping Glc H6B, Cys H α , Glu H α), 4.25–4.10 (overlapping Ala H α , Val H α , Glc H6'B, Glc H5, Glc H6A), 4.05–3.90 (Glc H5', Glc H6'A, Glc H4), 3.75 (m, Gly H α 1/ α 2), 3.02 (dd, Cys H β 2), 2.73 (dd, Cys H β 3), 2.24 (m, Glu H γ) 2.10–1.90 (overlapping Val H β , Glu H β 2, O-acetyl), 1.88 (N-acetyl), 1.77 (m, Glu H β 3), 1.23 (d, Ala H β), 0.88–0.86 (d, Val H γ 1/H γ 2).

Ac-Cys(Acm)AlaCys(Aceto-Glucose)SerProCys(Acm)NH₂ (11a): Peptide **11a** was synthesized, characterized, and purified under the experimental conditions described above. The acetobromo- α -D-glucose was employed as alkylating agent (6.4 mg, 0.016 mmol). The final RP-HPLC purification provided compound **8a** in 21% yield (2.1 mg). HPLC: $t_R = 6.43$ min; ES-MS: m/z calculated: 1097.07 $[M+H]^+$; found: 1096.4.

The removal of the acm groups was performed by dissolving the glycopeptide in the necessary volume of a mixture of TFA-anisole (99:1), then AgOTf (100 equiv) was added and the stirring was maintained for 2 h at 4 °C. The crude product was precipitated in cold diethyl ether and purified by RP-HPLC. ES-MS: m/z calculated: 953.05 $[M+H]^+$; found: 953.42547.

General procedure for MW-assisted S-glycosylation: Crude peptide **4** (10 mg) and Fmoc-Cys-CONH₂ (**7**) (10 mg) were dissolved in DMF (3 mL) under argon atmosphere and placed in 0.5–2 mL microwave vials. Acetobromo- α -D-glucose (Br-**a**) for peptide **4** and acetobromomaltose (Br-**d**), respectively, were then added by using a syringe. As a final step, 4 Å molecular sieves (3–3.5 g), previously activated at 280 °C for 4 h under vacuum (10⁻⁴ mBar), were placed into each microwave vial, which was quickly sealed. The obtained solutions were irradiated for 5 min at 40 °C in a microwave oven (Initiator, Biotage Sweden AB, Uppsala, Sweden) and stirred at 720 rpm. Direct temperature control was performed with infrared

sensors. The mixtures were centrifuged to eliminate the sieves and the supernatants were concentrated under vacuum. The final products were purified by RP-HPLC, analyzed by mass spectrometry and fully characterized by NMR spectroscopy.

Conflict of interest

The authors declare no conflict of interest.

Keywords: chemoselectivity · glycopeptides · glycosyl halide · peptide modifications · S-alkylation

- [1] M. R. Pratt, C. R. Bertozzi, *Chem. Soc. Rev.* **2005**, *34*, 58–68.
- [2] J. Stepper, S. Shastri, T. S. Loo, J. C. Preston, P. Novak, P. Man, C. H. Moore, V. Havlicek, M. L. Patchett, G. E. Norris, *FEBS Lett.* **2011**, *585*, 645–650.
- [3] H. Venugopal, P. J. B. Edwards, M. Schwalbe, J. K. Claridge, D. S. Libich, J. Stepper, T. Loo, M. L. Patchett, G. E. Norris, S. M. Pascal, *Biochemistry* **2011**, *50*, 2748–2755.
- [4] T. J. Oman, J. M. Boettcher, H. Wang, X. N. Okalibe, W. A. van der Donk, *Nat. Chem. Biol.* **2011**, *7*, 78–80.
- [5] H. Wang, W. A. van der Donk, *J. Am. Chem. Soc.* **2011**, *133*, 16394–16397.
- [6] H. Katayama, Y. Asahina, H. Hojo, *J. Pept. Sci.* **2011**, *17*, 818–821.
- [7] A. Novoa, S. Barluenga, C. Serba, N. Winssinger, *Chem. Commun.* **2013**, *49*, 7608–7610.
- [8] K. Pachamuthu, R. R. Schmidt, *Chem. Rev.* **2006**, *106*, 160–187.
- [9] D. Comegna, F. De Riccardis, *Org. Lett.* **2009**, *11*, 3898–3901.
- [10] D. Comegna, I. de Paola, M. Saviano, A. Del Gatto, L. Zaccaro, *Org. Lett.* **2015**, *17*, 640–643.
- [11] D. P. Galonić, W. A. Van Der Donk, D. Y. Gin, *Chem. Eur. J.* **2003**, *9*, 5997–6006.
- [12] D. P. Galonić, W. A. van der Donk, D. Y. Gin, *J. Am. Chem. Soc.* **2004**, *126*, 12712–12713.
- [13] X. Zhu, R. R. Schmidt, *Chem. Eur. J.* **2004**, *10*, 875–887.
- [14] M. I. Gutiérrez-Jiménez, C. Aydllo, C. D. Navo, A. Avenzoza, F. Corzana, G. Jiménez-Osés, M. M. Zurbano, J. H. Busto, J. M. Peregrina, *Org. Lett.* **2016**, *18*, 2796–2799.
- [15] V. Rojas-Ocáriz, I. Compañón, C. Aydllo, J. Castro-López, J. Jiménez-Barbero, R. Hurtado-Guerrero, A. Avenzoza, M.-M. Zurbano, J. M. Peregrina, J. H. Busto, F. Corzana, *J. Org. Chem.* **2016**, *81*, 5929–5941.
- [16] E. Calce, M. Leone, L. Monfregola, S. De Luca, *Org. Lett.* **2013**, *15*, 5354–5357.
- [17] E. Calce, M. Leone, L. Monfregola, S. De Luca, *Amino Acids* **2014**, *46*, 1899–1905.
- [18] E. Calce, M. Leone, F. A. Mercurio, L. Monfregola, S. De Luca, *Org. Lett.* **2015**, *17*, 5646–5649.
- [19] E. Calce, S. De Luca, *Amino Acids* **2016**, *48*, 2267–2271.
- [20] E. Calce, S. De Luca, *Chem. Eur. J.* **2017**, *23*, 224–233.
- [21] C. Aydllo, I. Compañón, A. Avenzoza, J. H. Busto, F. Corzana, J. M. Peregrina, M. M. Zurbano, *J. Am. Chem. Soc.* **2014**, *136*, 789–800.
- [22] J. Gervay-Hague, *Acc. Chem. Res.* **2016**, *49*, 35–47.
- [23] X. Zhu, K. Pachamuthu, R. R. Schmidt, *J. Org. Chem.* **2003**, *68*, 5641–5651.
- [24] V. J. Somovilla, N. Martínez-Sáez, A. Fernández-Tejada, B. G. de la Torre, D. Andreu, J. Jimenez-Barbero, J. L. Asensio, A. Avenzoza, J. H. Busto, F. Corzana, J. M. Peregrina, *Curr. Top. Med. Chem.* **2014**, *14*, 2712–2721.

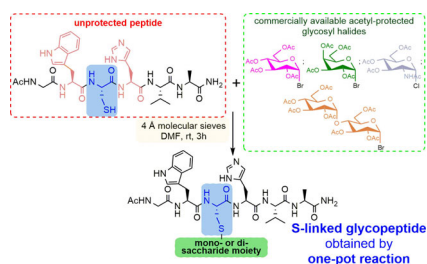
Manuscript received: January 18, 2018

Accepted manuscript online: February 19, 2018

Version of record online: ■ ■ ■ 0000

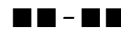
FULL PAPER

Peptide Glycosylation: A new synthetic approach to prepare S-glycopeptides is reported that uses a cysteine residue, inserted in a peptide sequence, as a sensitive substrate to bind carbohydrate derivatives through S-alkylation (see scheme). The method uses only activated molecular sieves as a base, and the appropriate glycosyl halide, both mono and di-saccharide, as the electrophilic reagent. The mild reaction conditions allow the selective functionalization of the cysteine thiol group.



Post-Synthetic Modifications

*E. Calce, G. Digilio, V. Menchise,
M. Saviano, S. De Luca**



Chemoselective Glycosylation of Peptides through S-Alkylation Reaction

