Probing Site-Selective Conjugation Chemistries for the Construction of Homogeneous Synthetic Glycodendriproteins

Isidro Cobo,[a, b] M. Isabel Matheu,[a] Sergio Castillón,[a] Benjamin G. Davis,*[b, c, d] and Omar Boutureira*[a, b]

Methods that site-selectively attach multivalent carbohydrate moieties to proteins can be used to generate homogeneous glycodendriproteins as synthetic functional mimics of glycoproteins. Here, we study aspects of the scope and limitations of some common bioconjugation techniques that can give access to well-defined glycodendriproteins. A diverse reactive platform was designed via use of thiol-Michael-type additions, thiol-ene reactions, and Cu(I)-mediated azide-alkyne cycloadditions from recombinant proteins containing the non-canonical amino acids dehydroalanine, homoallylglycine, homopropargylglycine, and azidohomoalanine.

Introduction

The use of synthetic glycosylated macromolecules or glycoconjugates, such as glycodendrimers and glyconanoparticles, can functionally mimic aspects of the sugar display of glycoproteins and glycolipids that decorate the outer surface of mammalian cells. For example, the blocking by decoys of carbohydrate-protein(lectin) interactions can represent an attractive anti-infective strategy for targeting pathogens.[1] Among glycoconjugates, synthetic glycoproteins[2] and particularly glycodendriproteins, resulting from the attachment of multivalent, antennary-like carbohydrate epitopes to a precise site of a protein scaffold (Figure 1a), have emerged as a class of mimics of naturally occurring N-linked glycoproteins with additional implications in vaccine design,[3] the development of bacterial/viral aggregation inhibitors,[4,5] ligands for the mannose-6-phosphate (M6PR)[7] and asialoglycoprotein (ASGPR)[8] receptors, and as glycomimetics of insulin,[9] human growth hormone, and the Fc region of human IgG.[10]

Whilst multivalent glycoprotein inhibitors of pathogen adhesion can be prepared by the non-selective attachment of dendrimeric glycans to proteins via standard amide, squaramide, or imine/amine formation,[11] recent advances in selective chemical protein modification[12] have allowed the precise attachment of multivalent carbohydrate moieties to predetermined sites of proteins using biologically-compatible reactions to generate single, well-defined glycodendriprotein glycoforms (Figure 1b,c). However, despite this progress, such methods remain scarce and have employed diverse linkages, including...
those derived enzymatically (amides),[9] those that are potentially cleavable/reversible (disulfides[4,5] and oximes[10]), and those that are stable under physiological conditions (e.g., 1,4-triazole linkages derived from Cu(I)-mediated azide-alkyne cycloadditions (CuAAC) with the alkyne tag homopropargylglycine (Hpg)[6] or N6-[(2-propynyloxy)carbonyl]-L-lysine (Lys(PA))[7]).

The aim of this proof-of-principle study is to comparatively evaluate the efficiency of some common, site-selective protein chemistries that could yield well-defined glycodendriproteins and so therefore may prove attractive in the design of putative synthetic protein therapeutics (synthetic biologics). By using comparable, representative tri-antennary, tri-galactosyl (β-D-Gal)3 carbohydrate dendron motifs, each equipped with corresponding reactive handles, we explored the generation of a series of recombinant glycodendriproteins using an approach of ‘tag-and-modify’ with several tag types.[13] These allowed the testing of thiol-Michael-type additions to dehydroalanine (Dha)-tagged proteins,[14] thiol-ene[15] radical additions at homoallylglycine (Hag) sites to generate thio-ether linkages,[27] and Cu(I)-mediated azide-alkyne cycloadditions to access 1,4-triazole linkages in two orientations, via azidohomoalanine (Aha) and homopropargylglycine (Hpg) tags (Figure 1d).[6,33]

Results and Discussion

Synthesis of glycodendron reagents

Model tri-β-D-galactosyl-containing (β-D-Gal)3 glycodendrons 1–3 derived from a 3,4,5-tris-(2-aminoethoxy)benzoic acid core and equipped with appropriate reactive handles (thiol, propargyl, and azide) were designed and synthesized as simple mimics of the asymmetric carbohydrate display observed in tri-antennary N-linked glycopolypeptides (Scheme 1).[16] These structures possess significant rigidity and useful distances between their β-D-Gal tip sugars, features that can prove favourable for multivalent ligand display.[17,18] Indeed, it has been noted that glycodendrimers can mimic the non-reducing termini as well as some secondary interactions using only imperfect structural analogues[19] of the branched carbohydrates found in glycoproteins, without the necessity of presenting the whole, synthetically challenging, natural complex oligosaccharide.[20] In addition, we selected thioglycosides as the glycan motifs in these glycodendrons as these typically confer greater stability under both basic and acidic aqueous conditions, as well as resistance to enzymatic hydrolysis.[21] Moreover, such thioglycoside mimetics, together with other chalcogen derivatives such as

Scheme 1. Synthesis of tri-antennary glycodendron reagents 1–3. Reagents and conditions: (a) dry K2CO3, 10 mol% TBAI, dry DMF, 80 °C, 6 d; (b) (i) dry CH2Cl2, 4 M HCl in dioxane, rt, 2 h; (ii) chloroacetyl chloride, NaHCO3, 2 : 1 Et2O/H2O, 0 °C to rt, 7 d; (c) 8, dry DMF, rt, 3 d; (d) (i) 9, dry DMF, rt, 24 h; (ii) 1 : 10 1 N NaOH (aq.)/EtOH, rt, 24 h; (f) (i) cystamine, HATU, DIPEA, dry DMF, 50 °C, 3 d; (ii) PBu3, H2O, rt, 2 h. (g) propargylamine hydrochloride, HATU, DIPEA, dry DMF, 45 °C, 27 h; (h) N-Boc-ethylenediamine, HATU, DIPEA, dry DMF, 45 °C, 27 h; (i) 1 : 2 Me2S/TFA, 0 °C, 3 h; (j) 0.4 M TFA in CH2Cl2, 10 mol% CuSO4, DMAP, MeOH, 0 °C to rt, 19 h. Boc = tert-butoxycarbonyl, TBAI = tetrabutylammonium iodide, DMF = N,N-dimethylformamide, HATU = 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate, DIPEA = N,N-diisopropylethylamine, TFA = trifluoroacetic acid, DMAP = 4-dimethylaminopyridine.
selenoglycosides, can maintain the intrinsic binding properties of the glycan towards the corresponding protein receptor (lectin).\(^{[26]}\)

We first prepared extended scaffold 4 as a key intermediate. After some preliminary attempts and reaction optimization (Supporting Information (SI), Schemes S1 and S2), 4 was obtained in 75% yield from 5 and 6 as previously described by Brouwer et al.\(^{[15]}\) Boc protection using 4 M HCl in dioxane and subsequent treatment of the resulting trihydrochloride salt with chloroacetyl chloride and NaHCO\(_3\) using a biphasic 2:1 Et\(_2\)O/H\(_2\)O solvent system afforded derivative 7 (97%). Boc removal using standard trifluoroacetic acid (TFA) and the use of chloroacetic anhydride as N-acetylating reagent led to lower overall yields (Supporting Information, Scheme S3). Next, the incorporation of the non-reducing β-D-galactose moiety (β-D-Gal) was first attempted using O-acetyl protected thioglycoside sodium salt 8\(^{[24]}\) in dry DMF. However, 10 was obtained in only 33% yield, and despite quantitative subsequent Zemplén deacetylation and methyl ester hydrolysis to 11 (99%), the reduced overall yield hampered the utilization of this route using protecting groups. Thus, an alternative protecting-group-free route was explored (Supporting Information, Table S1). Treatment of gallic acid core 7 with β-thiogalactoside sodium salt 9\(^{[25]}\) followed by methyl ester hydrolysis using aqueous NaOH in EtOH allowed the ready preparation of common glycodendron reagent precursor 11 in a superior 74% yield over two steps. Next, using this divergent intermediate, appropriate reactive handles (thiol, propargyl, and azide) were introduced through amide-coupling protocols (Scheme 1). A reactive thiol was incorporated by treating 11 and cystamine with HATU and DIPEA in dry DMF at 50 °C to afford a disulfide intermediate, which was subsequently reduced in situ to 1 (69%) with PBu\(_3\) in water at room temperature for 2 h. Similarly, an alkyne group was incorporated to obtain reagent 2 (75%) after stirring a mixture of 11, propargylamine hydrochloride, HATU, and DIPEA in dry DMF at 45 °C for 27 h. Finally, reactive azide was incorporated following a 3-step procedure. Similarly to propargyl 2, amide coupling with N-Boc-ethylenediamine afforded 12 in 67% yield. Subsequent Boc deprotection with 1:2 Me\(_2\)S/ TFA followed by diazo transfer to the resulting primary amine led to azide 3 in 95% yield over two steps.

Glycodendrprotein construction

The next step of the proposed strategy involved the use of these synthesized tri-antennary glycodendron reagents 1–3 to chemical modify a series of protein substrates bearing appropriate reactive tags; Dha 13\(^{[23–26]}\) (here chemically generated from Cys), Hag 14\(^{[23,27]}\), Aha 15\(^{[28]}\), 16\(^{[25,28]}\) and Hpg 17\(^{[29]}\) (the latter three were all generated through sense-codon reassignment of Met exploiting Met-auxotroph-mediated expression). In order to rapidly scope the generality of our tested methods to access well-defined glycodendrproteins, we used a multivariate selection of prototypical protein scaffolds featuring different residue sites/microenvironments, protein folds, as well as possessing different functional measures (catalytic activity or structural/self-assembling properties) of outcome. We first explored thiol-conjugate-addition chemistry, an approach absent from previous protocols for glycodendrprotein generation, using a single Dha protein mutant of the serine protease subtilisin from Bacillus lentus (SBL), a representative three-layer α/β-Rossman-fold protein with catalytic activity, quantitatively obtained from the corresponding Cys precursor using standard bisalkylation-elimination protocols.\(^{[20]}\) The identity, purity, and stability of the resulting glycodendrproteins was established by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) and SDS polyacrylamide gel electrophoresis (SDS-PAGE) (Supporting Information, Figures S1–S16). Thus, after generation from SBL-Cys156, incubation of SBL-Dha156 (13) with 1 in 50 mM sodium phosphate buffer (NaP) at pH 8.0 afforded pure, synthetic glycodendrprotein 18 in >95% conversion after 1.5 h at room temperature as determined by LC-ESI-MS (Scheme 2, left panel). Although CuAAC reactions are somewhat more established in the limited examples of glycodendrprotein generation, the use of this alternative method provides potential expansion of scope and also allows potential access to alternative reaction scope from the same glycodendron (here thiol 1) type.

Indeed, next, thiol-ene radical addition/ligation was explored as alternative/complementary thiol chemistry to this previous thiol-Michael addition. Here, we explored the use of homomultimer, virus-like bacteriophage particle Q\(_8\) 14, which self-assembles from 180 monomers, equipped with a Hag tag (Scheme 2, right panel)\(^{[23,27]}\). This icosahedral protein platform provides greatly differing dimensions (core diameter ~28 nm), and so potentially reactivities. It has also allowed the prior construction of multivalent systems with enhanced function (e.g., viral mimicry\(^{[46]}\)). Whilst direct comparison in this different protein scaffold would be inappropriate, use of the same thiol glycodendron 1 in thiol-ene reaction Q\(_8\)-Hag16 (14) was sluggish, and final glycodendrprotein nanoparticle 19 was obtained in only 25% conversion at pH 6.0 and 53% conversion at pH 4.0 (500 equiv. of glycodendron reagent 1, after 8 and 28 h at room temperature, respectively)\(^{[23,27]}\). Increased equivalents of 1 resulted in only similar conversion levels. Reactions carried out at pH 6.0 provided more homogeneous product (as judged, for example, by MS spectrum signal-to-noise (S/N)) than those at pH 4.0, albeit with lower conversions. Moreover, prolonged reaction times did not substantially improve the conversion. Next, to broaden the scope of reactions tested and protein scaffolds/sites/residue microenvironments, we explored well-established Cu(i)-mediated azide-alkyne cycloadditions (CuAAC) using a variety of protein scaffolds with azide tags; SsβG-Aha43 (15), SsβG-Hpg1-Hpg43/-Hpg43 (17 a/b) (as an example of a generic β-galactosidase, in an αβ-fold TIM barrel, that has been previously used to create synthetic glycoprotein probes for both in vitro and in vivo applications\(^{[13]}\)) and Np276-Aha61 (16) (in an all-β-helix, β-fold pentapeptide repeat protein scaffold from Nostoc punctiforme, fusion protein 275/276 also known as Npβ1) (Schemes 3 and 4). In this way, this allowed not only site-exploration within varied scaffolds with differing secondary structural features (β dominant vs. αβ mixed) but again again protein function variation (e.g., catalytic
Scheme 2. Synthesis of S-linked glycodendriproteins 18 and 19 from SBL-Dha156 (13) and Qβ-Hag16 (14), respectively. Note that whilst reactions were conducted under non-denatured conditions as shown masses, the intact protein masses shown for Qβ refer to monomer following denaturation to individual monomers for ESI-MS analysis.

Scheme 3. Synthesis of triazole-linked glycodendriproteins 20 and 21 from SsjiG-Aha43 (15) and Np276-Aha61 (16), respectively.
activity). Triazole-linked products 20, 21 were generated efficiently in conversions of >95% upon incubation with propargyl 2 glycodendron reagent in 50 mM NaP, at pH 8.2 for 1 h at room temperature as determined by LC-ESI-MS (Scheme 3). By contrast, under the same conditions, reaction of azide 3 with SspG-Hpg1-Hpg43 (17a) was more sluggish. Notably, when azide 3 was used to modify a mixture of SspG-Hpg1-Hpg43/Hpg43 (17a/b), crude product 22 was generated consistent with regioselective monomodification only at a more accessible position 1 (i.e., reaction only of SspG-Hpg1-Hpg43 and not of SspG-Hpg43). Such regioselectivity in the use of CuAAC on proteins is consistent with previous observations (Scheme 4).[30] This apparent dependency of reaction conversion upon protein site location (i.e., regioselectivity) mirrors previous observations[29,31,33] where correlation is observed with a combination of the intrinsic reactivity of the tag-reactant pair as well as the protein residue accessibility. The possible additional roles of residue microenvironment (e.g., charges, polar/hydrophobic interactions, etc.) may well also play a role but have typically proven less important in our hands. As a consequence, the presence at the same site of different reactive tags (Hpg43 vs. Aha43) or the same reactive tag at different site (Hpg1 vs. Hpg43) may react differently. This was previously rationalized according to a heuristic model (termed ‘reactive accessibility’, RA) to evaluate and predict site reactivity.[31] This model correlates the reactivity observed with predicted measures of protein residue accessibility[32] and is able to, in turn, guide the control of reaction conditions to achieve regioselective modification at a variety of protein sites.[31,33] The apparent contrast in the reactivity between Aha and Hpg at the same site 43 in protein SspG reinforces the importance of evaluating CuAAC reaction ‘orientations’ when deciding the choice of reactive handle location in two partners to be conjugated and the preferential use in our hands of Aha as a tag for less accessible sites and in proteins. Our observations (here and previously) consistently suggest lower protein reactivity in Hpg-tagged proteins compared to their Aha-tagged counterparts.[30,29,33]

**Conclusion**

In summary, a brief survey of the site-selective attachment of simple, multivalent (β-D-Gal)-dendrons to generate glycodendrproteins as N-linked glycoprotein mimics suggests that although well-defined, highly-valent structures can be generated through other methods,[6] those based on the use of Dha tags (for C-S-bond formation) and Aha tags (for triazole formation) may be the most applicable to the ready formation of well-defined constructs. A qualitative comparison of the use of glycodendrons to modify proteins both here and previously[30] suggests an apparent order of utility in these systems as follows, tags-via-linkage: Dha-via-triazole [Aha + R-C(=O)CH] > Hpg-via-triazole [Hpg + R-N3] > Hag-via-C-S [Hag + R-S]. With regard to application, it has been previously shown that the conjugation of glycodendron reagents to proteases, such as SBL used here, enables targeted protein degradation; this has been applied to, for example, bind and degrade bacterial adhesins.[31] As such, not only might glycodendrprotein glycoconjugates allow development of anti-infective therapeutics (by both direct blocking[30] and degradation[33]) but also other potentially broader clinical applications that may exploit the selective degradation of other sugar-binding proteins. Work to this goal is currently under investigation in our laboratories.

**Experimental Section**

**General remarks:** Proton (1H NMR) and carbon (13C NMR) nuclear magnetic resonance spectra were recorded on a Varian Mercury spectrometer (400 MHz for 1H) and (100.6 MHz for 13C) or a Bruker AVIV500 spectrometer (500 MHz for 1H) and (125.8 MHz for 13C). NMR spectra were assigned using COSY, DEPT 135, HSQC, HMBC, and NOESY and are subjective. All chemical shifts are quoted on the δ scale in ppm using the residual solvent as the internal standard (1H NMR: CDCl3 = 7.26, DMSO-d6 = 2.50, D2O = 4.79 and 13C NMR: CDCl3 = 77.0; DMSO-d6 = 39.3). Coupling constants (J) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet and app = appa-
rent. Melting points (mp) were recorded on a Leica Galen EA 1108 Analyser. Optical rotations were measured on a Perkin-Elmer 241 polarimeter with a path length of 1.0 dm and are reported with implied units of 10° deg cm² g⁻¹. Concentrations (c) are given in g/100 mL. Low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer using electrospray ionization (ESI) and high resolution mass spectra (HRMS) were recorded on a Bruker MicroTOF ES mass spectrometer. Nominal and exact m/z values are reported in Daltons (Da).

Chemical synthesis

Methyl 3,4,5-tris-(2-(2-butoxycarbonylamino)ethoxy)benzoate (4) [37,38]: A mixture of methyl 3,4,5-trihydroxybenzoate 5 (2 g, 10.9 mmol), 2-(2-butoxycarbonylamino)ethyl bromide 6 (9.7 g, 43.4 mmol), dry potassium carbonate (6.8 g, 48.9 mmol), and TBAI (10.9 mmol), 2-(NH₂)₄ (0.41 g, 1.1 mmol) in dry DMF (41 mL) was stirred at 80 °C for 6 days. The reaction mixture was cooled to room temperature, filtered through a short path of Celite®, washed with EtOAc, and then dried over anhydrous MgSO₄. The solvent was removed under reduced pressure to afford 4 as a white solid (4.9 g, 75 %), Rf (1:1 EtOAc/hexane) 0.33. 1H NMR (400 MHz, CDCl₃): δ 7.30 (s, 2H, Ar), 5.82 (bs, 1H, NH), 5.39 (bs, 2H, NH), 4.12 (m, 6H, OCH₂), 3.89 (s, 3H, OCH₃), 3.58 (m, 4H, CH₂NH), 3.42 (m, 2H, CH₂NH), 1.45 (s, 27H, CH₃), Boc. 13C NMR (100.6 MHz, CDCl₃): δ 166.0 (C=O, CO₂Me), 155.7, 155.6 (C=O, Boc), 152.0 (C=3, 140.8 (C=4), 125.5 (C-1), 108.1 (C-2,6), 79.5, 79.3 (C, Boc), 72.4, 68.4 (OCH₂), 52.1 (OCH₃), 40.4, 39.9 (CH₂NH), 28.2 (CH₃), Boc. HRMS (TOF ES +): m/z: [M + Na]+ Calcd for C₂₄H₂₄O₇Na⁺, 636.3103; found, 636.3104.

Methyl 3,4,5-tris(2-(2-chloroacetamido)ethoxy)benzoate (7): 4 (1.81 g, 2.946 mmol) was dissolved in dry CH₂Cl₂ (5.5 mL) and 4 M HCl in dioxane (18 mL) was added at room temperature under an atmosphere of argon. The reaction mixture was stirred at the same temperature for 2 h. After complete conversion, the solvent was evaporated and dried under high vacuum to afford methyl 3,4,5-tris(2-aminoethoxy)benzoate trihydrochloride as a white solid (1.23 g, 99 %). Used in the next step without further purification. A mixture of this intermediate and NaHCO₃ (1.48 g, 17.676 mmol) in 2:1 EtO/H₂O (8.8 mL) was cooled to 0 °C (ice/water) and chloroacetyl chloride (771 µL, 9.721 mmol) was slowly added over a period of 1 h. After complete addition, the mixture was allowed to warm to room temperature. After 7 days stirring at the same temperature, the reaction mixture was filtered. The precipitate was washed with water, 2 N aqueous HCl, water, and finally with Et₂O. The crude product was recrystallized from 1:1 ethanol/water and dried under reduced pressure to afford 7 as a white solid (1.5 g, 97 % over two steps). Rf (EtOAc): 0.23. mp (EtOH/water 1:1): 158–160 °C. 1H NMR (400 MHz, DMSO-d₆): δ 8.43 (bs, 2H, NH), 8.25 (bs, 1H, NH), 7.24 (s, 2H, Ar), 4.10–4.01 (m, 12H, CH₂Cl, OCH₃), 3.83 (s, 3H, OCH₃), 3.51–3.36 (m, 6H, CH₂NH), 13C NMR (100.6 MHz, DMSO-d₆): δ 166.4, 166.2 (C=O, CO₂CH₂), 165.7 (C=O, CO₂Me), 151.9 (C=3,5), 141.2 (C=4), 124.8 (C-1), 107.9 (C-2,6), 71.0, 67.2 (OCH₂), 52.3 (OCH₂), 42.6 (CH₂, OCH₂), 39.5, 38.8 (CH₃), FT-IR (KBr, vₙ, s), 3258, 2978, 2874, 1721, 1644.
3,4,5-Tris-(2-(2,3,4,6-tetra-O-acetyl-1-thio-D-galactopyranosyl)acetamido)ethoxy)benzoic acid (11): Sodium 1-thio-D-galactopyranosyl acetate (10 mg, 0.046 mmol) was added to a dispersion of 7 (7 mg, 0.014 mmol) in dry DMF (0.5 mL) at room temperature with an argon stream bubbling through the solution. The reaction mixture was stirred at the same temperature for 24 h. The solvent was then evaporated and the crude treated with 1 N aqueous NaOH (322 mL) in EtoAc (3.2 mL) at room temperature for 22 h. After neutralization with Dowex (H⁺ 50WX8-200), the ion exchanger was filtered off and washed with water. The crude was purified by gel permeation chromatography (Bio-gel P-2, H₂O) followed by lyophilization to afford 11 as a white solid (26 mg, 74% over two steps). Rₗ (7:3:1: IPiOH/H₂O/NH₄OH): 0.36. [α]₁₃⁰° +7.7 (c 0.02, H₂O).

1H NMR (500 MHz, CDOD): δ 7.11 (s, 2H, Ar), 4.38 (d, J = 9.7 Hz, 1H), 3.36 (d, J = 9.7 Hz, 2H), 4.21–4.15 (m, 18H, OCH), 3.90–3.83 (m, 3H, H-4'), 3.71–3.35 (m, 27H, H₂-3',5',6',αβ, CH₂NH, CH₂S). 13C NMR (125.8 MHz, CDOD): δ 172.5, 172.3 (C-3, C-5), 151.6 (C-3, C-5), 138.9 (C-4), 129.6 (C-1), 106.0 (C-2), 85.4, 85.1 (C-11), 79.0 (C-5'), 73.8 (C-3', 71) (CH₂O), 69.5 (C-2'), 68.7 (C-4'), 67.3 (OCH), 61.0 (CH), 60.4 (C-2'). FT-IR (KBr, νmax): 3355, 2870, 2831, 1619, 1527, 1253, 1176, 757, 573. HRMS (TOF ES) m/z: [M + Na]⁺ Calcd for C₉₁H₁₁₂N₂O₁₀S₉Na⁺, 1809.2631; found, 1809.2642.

3,4,5-Tris-(2-(1-thio-D-galactopyranosyl)acetamido)ethoxy)-N- (prop-2-ynyl)benzamide (2): A mixture of 11 (15 mg, 0.015 mmol), propargylamine hydrochloride (2 mg, 0.018 mmol), HATU (7.4 mg, 0.019 mmol), and DIPEA (9 µL, 0.052 mmol) in dry DMF (240 µL) was stirred at 45°C under an atmosphere of argon for 27 h. The solvent was evaporated and the crude was purified by gel permeation chromatography (Bio-gel P-2, H₂O) followed by lyophilization to afford 12 as a white solid (11.7 mg, 75%). Rₗ (7:3:1: IPiOH/H₂O/ NH₄OH): 0.23. mp 98–100°C. [α]₁₃⁰° +240 (c 0.03, H₂O).

1H NMR (500 MHz, CDOD): δ 7.15 (s, 2H, Ar), 4.45 (d, J = 9.5 Hz, 1H, H-1'), 4.40 (d, J = 9.5 Hz, 2H, H-1), 4.24 (m, 6H, OCH), 1.41–1.30 (m, 28H, 2H-3',5',6',αβ, CH₂NH, CH₂S), 13C NMR (125.8 MHz, CDOD): δ 173.7, 173.5 (C-3, C-5), 170.3 (C-1), 150.3 (C-3, C-5), 140.4 (C-4), 130.5 (C-1), 107.3 (C-2), 86.6, 86.3 (C-1'), 80.7 (C-3'), 75.0 (C-5'), 73.5 (C-3'), 71.5 (CH₂), 70.7, 69.9, 68.5 (C-2', CH₂O), 62.6 (C-2'), 60.4, 39.9 (CH₂NH), 34.3, 34.2 (CH₃), 30.5 (CH₂). FT-IR (KBr, νmax): 3583, 3406, 2995, 2122, 1646, 1636, 1558. Anal. Calcd for C₉₁H₁₁₂N₂O₁₀S₉Na⁺: 1807.2680 (100), 1068.28 (43), 1069.28 (30), 1069.27 (14), 1070.27 (6); found 1067.27 (100), 1068.28 (23), 1069.27 (8), 1070.28 (1).

3,4,5-Tris-(2-(1-thio-D-galactopyranosyl)acetamido)ethoxy)-N- (2-tert-butoxycarbonyl)amino)ethyl)benzamide (12): A mixture of 11 (15 mg, 0.015 mmol), N-Boc-ethylendiamine (3 µL, 0.018 mmol), HATU (7.4 mg, 0.019 mmol), and DIPEA (6 µL, 0.037 mmol) in dry DMF (417 µL) was stirred at 45°C under an atmosphere of argon for 27 h. The solvent was evaporated and the crude was purified by gel permeation chromatography (Bio-gel P-2, H₂O) followed by lyophilization to afford 12 as a white solid (11.5 mg, 67%). Rₗ (7:3:1: IPiOH/H₂O/NH₄OH): 0.36. mp 104–106°C. [α]₁₃⁰° +7.7 (c 0.02, H₂O).

1H NMR (500 MHz, CDOD): δ 7.13 (s, 2H, Ar), 4.36 (d, J = 9.5 Hz, 1H, H-1'), 4.42 (d, J = 9.5 Hz, 2H, H-1), 4.24–4.20 (m, 6H, OCH), 3.94–3.88 (m, 3H, H-4'), 3.73–3.40 (m, 31H, H₂-3',5',6',αβ, CH₂NH, CH₂S, CH₂CH₃), 13C NMR (125.8 MHz, CDOD): δ 173.7, 173.6 (C-3, C-5), 170.3 (C-1), 150.3 (C-3, C-5), 140.4 (C-4), 130.5 (C-1), 107.3 (C-2), 86.6, 86.3 (C-1'), 80.7 (C-3'), 75.0 (C-5'), 73.5 (C-3'), 71.5 (CH₂), 70.7, 69.9, 68.5 (C-2', CH₂O), 62.6 (C-2'), 60.4, 39.9 (CH₂NH), 34.3, 34.2 (CH₃), 30.5 (CH₂). FT-IR (KBr, νmax): 3583, 3406, 2995, 2122, 1646, 1636, 1558. Anal. Calcd for C₉₁H₁₁₂N₂O₁₀S₉Na⁺: 1807.2680 (100), 1068.28 (43), 1069.28 (30), 1069.27 (14), 1070.27 (6); found 1067.27 (100), 1068.28 (23), 1069.27 (8), 1070.28 (1).
temperature. After complete addition, the mixture was allowed to warm to room temperature. After 19 h stirring at the same temperature, the solvent was evaporated and the crude was purified by gel permeation chromatography (Bio-gel® P-2, H₂O) followed by lyophilization to afford 3 as a yellowish foam (3.5 mg, 95% over two steps). Rₓ: (7.3:1) 1PrOH/H₂O/0.1NH₃: 0.37. mp 106–109 °C. [α]s = 26.0 (c 0.02, H₂O). ⁱH NMR (500 MHz, CDCl₃): δ 7.15 (s, 2H, Ar), 4.45 (d, J=7.0 Hz, 1H, H-1), 4.41 (d, J=7.0 Hz, 2H, H-1), 4.23 (m, 6H, OCH₃), 4.08–3.87 (m, 3H, H-4'), 3.72–3.39 (m, 31H, H-2',3',4',5',6'ab, CH₂), 2.04 (s, 2H, Ar), 1.53 (s, 2H, Ar). FT-IR (KBr, ν): 3418, 3006, 2996, 2973, 2110, 1651, 1456, 1167. Anal. Calcd for C₈₆H₇₅N₂₀₂S₂: 435.49; Found, 435.51. The modified glycodendriprotein retained inherent peptide activity, as indicated by liberation of p-nitroaniline upon treatment with the chromogenic peptide suAAPPfN[30].

**Stability of Gal₃-G-triazole-435s[IG]** (20) in human plasma: A 10-μL aliquot of Gal₃-G-triazole-435s[IG] (20) (ca. 0.5 mg/mL) in 50 mM sodium phosphate buffer (pH 7.0) was transferred to a 0.5 mL Eppendorf tube. Reconstituted human plasma (0.5 mL; Sigma-Aldrich) was added at room temperature and the resulting mixture vortexed for 30 seconds. After 24 h of additional shaking at 37 °C, the reaction was analysed directly by LC-MS and starting protein 20 (calculated mass, 58276; observed mass, 58277) was detected unaltered.

Gal₃-G-triazole-61Np276 (21): Gal₃-G-propargyl 2 (1.86 mg, 1.78 μmol) was dissolved in 50 mM sodium phosphate buffer (pH 7.0) and the reaction mixture vortexed for 30 seconds. After 24 h of additional shaking at 37 °C, the reaction was analysed directly by LC-MS, and the product was characterized by LC–MS (calcd. 58276; found, 58285). Finally, the sample was flash frozen with liquid nitrogen and stored at –20 °C.

Gal₃-G-triazole-15s[IG] (22): Gal₃-G-N₃ (1.18 mg, 1.10 μmol) was dissolved in 50 mM sodium phosphate buffer (50 μL, pH 8.2). A freshly prepared solution of copper(l) bromide (99.999 %) in 50 mM sodium phosphate buffer (pH 8.2) and 5-column volume of buffer A (low imidazole concentration; 20 mM Tris HCl, 500 mM NaCl, 5 mM imidazole, pH 7.8) was added at room temperature and the resulting mixture vortexed for 30 seconds. After 24 h of additional shaking at 37 °C, the reaction was analysed directly by LC-MS, and the product was characterized by LC–MS (calcd. 21796; found, 21794). Finally, the sample was flash frozen with liquid nitrogen and stored at –20 °C.

Gal₃-G-triazole-435s[IG] (20) was synthesized by coupling of Gal₃-G-propargyl 2 (1.95 mg, 1.87 μmol) was dissolved in 50 mM sodium phosphate buffer (50 μL, pH 8.2). A freshly prepared solution of copper(l) bromide (99.999 %) in acetonitrile solution of tris-triazolyl amine ligand tris[(1-ethylacrylate-1H,1,2,3-triazol-4-y)] methylamine (63 μL, 127 mg/mL) was added to the above solution and mixed thoroughly. Ss[IG-Aha43 (15) (100 μL, 0.5 mg/mL) was added to the mixture and the reaction was agitated on a rotator at room temperature for 1 h. 0.2 mL of High Affinity Ni-Charged resin was then added to the mixture and the reaction was agitated on a rotator at 4 °C for 1 h. The sample was then placed in a syringe and eluted with 5-column volume of buffer B (high imidazole concentration; 20 mM Tris HCl, 500 mM NaCl, 15 mM imidazole, pH 7.8). The eluted fraction (0.25 mg/mL) was premixed with a Vivaspin™ membrane concentrator (10 kDa molecular weight cut off) and washed with 50 mM sodium phosphate buffer (3 × 200 μL, pH 7.0). The solution was concentrated to 100 μL and the product was characterized by LC–MS (calcd. 58276; found, 58285). Finally, the sample was flash frozen with liquid nitrogen and stored at –20 °C.
sample was flash frozen with liquid nitrogen and stored at −20 °C. Note: The modified glycoproteins retained galactosidase activity as evidenced by X-Gal stain. In addition, regioselective monomodification only at position 1 was observed for dialkynic protein SβG-Hpg1-Hpg43 (17a) together with residual, post-translationally modified SβG-Hpg3 (17b) resulting from N-terminal Hpg excision in 15a (calcd. 57214; found, 57218). This result is in agreement with results previously reported. For a detailed discussion on reactive accessibility of sites 1 and 43 on SβG-Hpg1-Hpg43 (17a), see work by van Kasteren et al. and corresponding Supporting Information.

Acknowledgements

We thank the Spanish Government – MCIU, the national agency of investigation (AEI/10.13039/501100011033), and the European Regional Development Fund – ERDF (projects; CTQ2017-90088-R and PID2020-120584RB-I00 to O.B., CTQ2017-90088-R to M.I.M and S.C.) for financial support. The Rosalind Franklin Institute is supported by the EPSRC. O.B. was a Ramón y Cajal Fellow (RYC-2015-17705). O.B. thanks Paula and Júlia Boutureira Regla for inspiration.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: bioconjugations · bioorthogonal reactions · dendrimers · glycoproteins · protein modifications

(0.20 M in DMSO, Bachem) was added to the protein sample, which turned yellow immediately upon addition of the peptide substrate. A solution colour change from colourless to yellow was considered a positive result. The yellow solution indicates liberation of -nitroaniline (pNA), confirming peptidase activity.

X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (40 mg) dissolved in DMSO (1 mL) was added to PBS buffer (40 mL) containing K$_2$FeCN$_6$ (5 mM) and K$_3$FeCN$_6$ (5 mM). β-Galactosidase activity was measured qualitatively. X-Gal solution (100 μL) was added to protein sample (50 μL) and the mixture was incubated at 37 °C for ca. 1 h. A solution colour change from colourless/yellowish to blue was considered a positive result.