Coiled coil type neoglycoproteins presenting three lactose residues

Sweeney, Sinclair; Bullen, Gemma; Gillis, Richard; Adams, Gary ; Rowe, Arthur; Harding, Stephen ; Tucker, James; Peacock, Anna; Murphy, Paul

DOI:
10.1016/j.tetlet.2016.02.005

Document Version
Peer reviewed version

Citation for published version (Harvard):
Sweeney, S, Bullen, G, Gillis, R, Adams, G, Rowe, A, Harding, S, Tucker, J, Peacock, A & Murphy, P 2016, 'Coiled coil type neoglycoproteins presenting three lactose residues', *Tetrahedron Letters*, vol. 57, no. 13, pp. 1414–1417. https://doi.org/10.1016/j.tetlet.2016.02.005

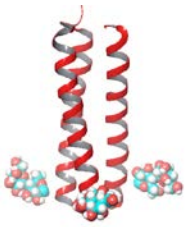
Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.

Coiled coil type neoglycoproteins presenting three lactose residues

Leave this area blank for abstract info.

Sinclair M. Sweeney,[a] Gemma A. Bullen,[b] Richard B. Gillis,[c,d] Gary G. Adams,[c,d] Arthur J. Rowe,[d] Stephen E. Harding,[d] James H. R. Tucker,[b] Anna F. A. Peacock,[b,] and Paul V. Murphy[*a]*



Tetrahedron Letters
journal homepage: www.elsevier.com

Coiled coil type neoglycoproteins presenting three lactose residues

Sinclair M. Sweeney,[a] Gemma A. Bullen,[b] Richard B. Gillis,[c,d] Gary G. Adams,[c,d] Arthur J. Rowe,[c,d] Stephen E. Harding,[d] James H. R. Tucker,[b] Anna F. A. Peacock,*[,b] and Paul V. Murphy*[,a]

[a] *School of Chemistry, National University of Ireland, Galway, Galway, Ireland*
[b] *School of Chemistry, University of Birmingham, Edgbaston, B15 2TT, United Kingdom*
[c] *Faculty of Medicine and Health Science, University of Nottingham, Queens Medical Centre, Nottingham, NG7 2RD, United Kingdom*
[d] *National Centre for Macromolecular Hydrodynamics, University of Nottingham, Nottingham, LE12 5RD, United Kingdom*
Corresponding author email: A.F.A.Peacock@bham.ac.uk; paul.v.murphy@nuigalway.ie

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
Scaffold;
Glycosylated;
Coiled Coil;
Glycocluster.

ABSTRACT

Scaffold design, synthesis and application is relevant for biomedical research. For example, multivalent interactions, such as those between cell surface glycoproteins and lectins can influence the potency and duration of signalling. The spacing between carbohydrates on their native protein scaffold could be important. Herein, the coiled coil design principle is used to generate synthetic coiled coil type glycoproteins, where three lactose residues are grafted to the coil via N-linkages to asparagine. Molecular modelling indicates that the distance between the galactose anomeric carbon atoms on the neoglycoproteins is ~30 Å. The inclusion of lactose was accommodated in both the final heptad towards the N-terminus, or more centrally in the penultimate heptad. In either case, neither the helicity nor the assembly to the trimeric form was unduly altered by the presence of the disaccharide.

Bioactive molecules can be considered to be comprised of core scaffolds displaying pharmacophoric groups to enable them to interact with their target receptor. Glycoproteins, for example, can contain multiple carbohydrate groups on a protein surface which are involved in multivalent interactions with lectins that can alter a variety of cellular and physiological events.[1] Multivalent carbohydrates can attain high affinity for lectins through the glycoside cluster effect[2] or they can be involved in promoting the cross-linking of glycoproteins which can lead to amplification as well as the protraction of signalling events.[3] The design and synthesis of multivalent carbohydrates is relevant in lectin ligand development and a number of factors are considered important. These include the valency (bi, tri, tetra etc.) and also the nature of scaffold[4] used to display the carbohydrates. Other groups can be appended to the scaffold to improve the affinity of the molecule, or multiple carbohydrates can be strategically located on scaffolds to influence geometric properties.[5]

As part of an interest in glycoclusters, which contain at least two carbohydrate residues, we have been involved in the synthesis of glycoclusters that have geometric constraints and have investigated a number of scaffolds in this respect. These include glycophanes, terephthalamides[6] and tetraphenylethylene,[7] and has led to the identification of inhibitors of biomedically important lectins such as galectins[8] and the macrophage galactose-C-type lectin.[6] Herein, we have taken advantage of coiled coil design principles to generate self-assembled coiled coil type glycoproteins that display three lactose residues and are thus trivalent. Left handed α-helical trimeric coiled coils, on which our designs are based, have previously been designed with a metal binding centre to self-assemble in the presence of a trivalent lanthanide for MRI applications.[9] However, self-assembling trimeric coiled coils which do not need a metal ion for coordination, are also known and are attractive as scaffolds as they mirror native protein architectures.[10] Recent reports on multiply *O*-glycosylated dimeric coiled coils,[11,12,13] including some that interact with asialoglycoprotein receptors, has prompted us to report our synthesis of *N*-lactosylated α-helical trimeric coiled coils. The designed α-helical coiled coil scaffold presents the lactose residues[14] on the outermost surface, either in the final heptad towards the N-terminus, where there is greater flexibility, or more centrally in the penultimate heptad. We have recently shown that coiled coil functionality can be highly dependent on which heptad is modified.[15]

Firstly, a monomeric peptide was designed using the heptad repeat approach (a seven amino acid repeating sequence with positions identified by letters *a*-*g*).[16] The sequence Ac-$G(I_aA_bA_cI_dE_eQ_fK_g)_4G$-$NH_2$ was used as the peptide template. Isoleucine (Ile, I) residues were positioned in *a* and *d* sites creating the hydrophobic core whilst at the same time encouraging the formation of a trimeric coiled coil.[10] Alanine (Ala, A) residues in the *b* and *c* locations encourage helical formation.[17] Glutamate (Glu, E) and lysine (Lys, K), in *e* and *g* positions respectively, form favourable interhelical electrostatic interactions between their side chains.[9] Glutamine (Gln, Q), a solubilising neutral residue, is generally introduced at position *f*. The *f* site is solvent exposed and was identified as the ideal location to incorporate a glycosylated amino acid. Thus lactosylated asparagine was incorporated at this location.

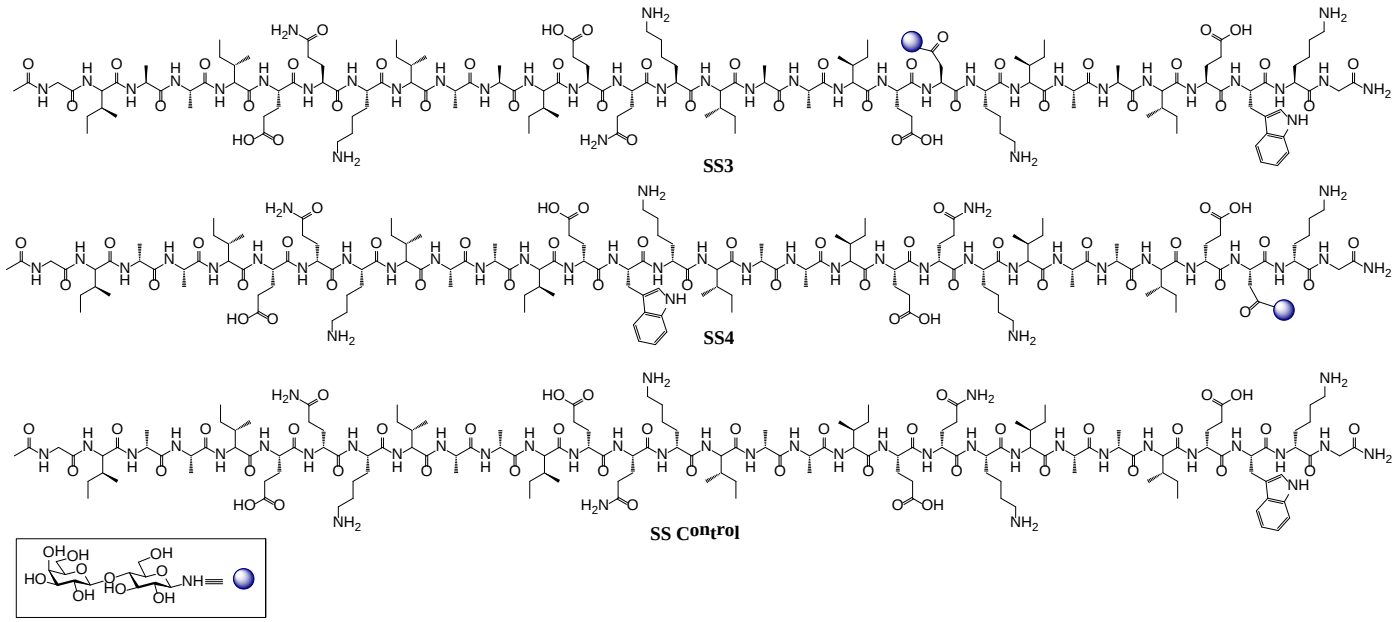
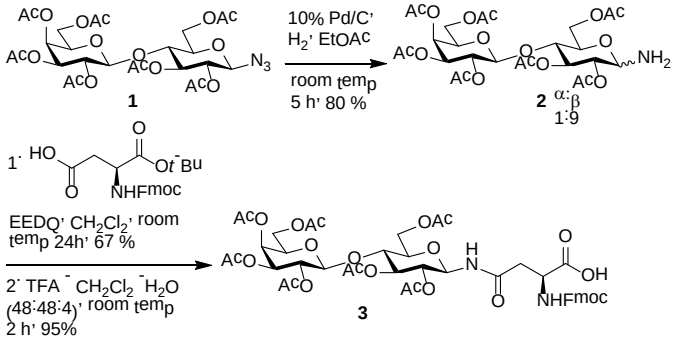


Figure 1. Structure of glycosylated peptides, **SS3** and **SS4** and non-glycosylated peptide, **SS Control**.

In addition, one of the glutamine residues at position *f* was replaced with a tryptophan residue (Trp, W). The presence of which enables the concentration of the peptide to readily be determined by UV spectroscopy as Trp is a natural chromophore absorbing light at 280 nm (ε_{280} = 5690 M^{-1} cm^{-1}).



Scheme 1. Synthesis of **3** from azide **1**

The lactosylated asparagine **3**, needed for solid phase peptide synthesis (SPPS), was prepared using a modification of the procedure from Liskamp[18] (Scheme 1). The known beta azide **1** was reduced to amine **2** using palladium on carbon in ethyl acetate in the presence of hydrogen. This reduction gave a 9:1 mixture of anomers (β:α) that could not be separated by chromatography. The use of the Adams catalyst in THF for the hydrogenation reaction gave the same product mixture, but in lower yield. Coupling of the mixture of amines, with appropriately protected aspartic acid (Scheme 1) in the presence of 2-ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline (EEDQ) as the coupling agent afforded two anomeric amides that could be separated by chromatography. This gave the desired β-lactosamide, which was treated with TFA to cleave the *t*-Bu ester and afforded **3**. This preparation was carried out on an approximately 7 g scale.

Next the synthesis of the 30 amino acid containing glycopeptide was performed by automated microwave assisted SPPS using Fmoc protected amino acids. Rink amide MBHA resin was used as the solid phase for the SPPS, performed on a 0.25 mM scale. Automated synthesis was carried out up to the point of introduction of the lactose containing residue. The introduction of the modified lactosylated amino acid **3** into the growing peptide was performed 'offline' in order to be able to carry out the coupling reaction of the lactose derivative over a longer reaction time and to monitor the incorporation of **3** into the peptide using the Kaiser test protocol.[19] Acetylation of the uncoupled free amines after the addition of the modified amino acid was carried out so as to render any free amines unreactive. After incorporation of the lactosylated amino acid, the remaining peptide synthesis was accomplished by automated SPPS.

The first glycosylated peptide (**SS3**, Ac-G IAAIEQK IAAIEQK IAAIE**X**K IAAIEWK G-NH_2, where X = lactosylated asparagine) was synthesised with the tryptophan positioned in the fourth and final heptad and with the lactose positioned in the third heptad (Figure 1). The second glycosylated peptide (**SS4**, Ac-G IAAIEQK IAAIEWK IAAIEQK IAAIE**X**K G-NH_2, X = lactosylated asparagine) had tryptophan positioned in the second heptad whilst the lactose was located in the fourth. The non-glycosylated peptide (**SS Control**, Ac-G IAAIEQK IAAIEQK IAAIEQK IAAIEWK G-NH_2) had tryptophan located in the fourth heptad, analogous to **SS3**, as shown in Figure 1. It was found that coupling of the lactosylated amino acid in **SS3** was slow and the yields were lower, based on HPLC analysis, when compared to amino acids that did not contain the lactose residue. Peptide aggregation and/or the bulky nature of the lactose moiety may have impaired the efficiency of the coupling. Therefore, **SS4** was synthesised, in which the lactosylated asparagine was introduced earlier on in the peptide synthesis (peptides are synthesised from the N-terminus to the C-terminus), as it was the first *f* site and the third amino acid residue, to be introduced. This led to enhanced coupling efficiency and an improved yield with respect to **SS3**.

After cleavage of the glycosylated peptide from the solid support, both glycosylated peptides were subjected to reaction with NaOMe in methanol (pH 10), under stirring for 1 h. These conditions facilitated efficient removal of the acetate protecting groups from the lactose moieties. Amberlite H^+ resin was used to acidify the resulting mixture to pH 7 during work up. The glycopeptides were purified via preparative C18 RP-HPLC and subsequently analysed by ESI-MS (Figure 2).

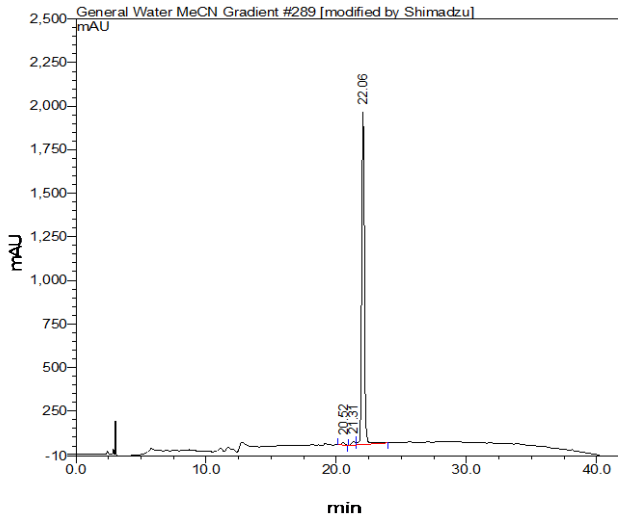


Figure 2. Analytical C18 HPLC trace of **SS3** after RP-HPLC. Electrospray ionisation MS showed a peak for the peptide at $[M+H]^+$ 3556.72. Similar HPLC and MS analysis was carried out for **SS4** and is included in the ESI.

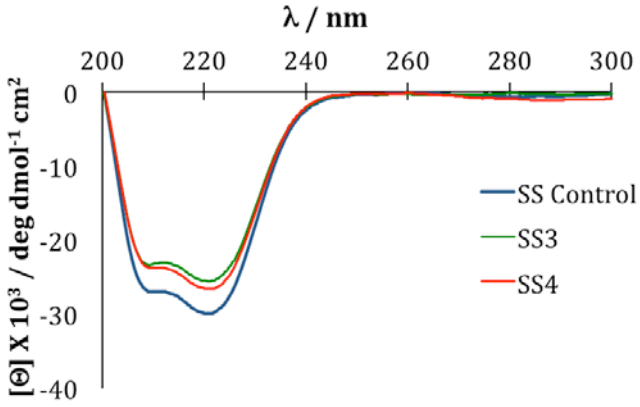


Figure 3. Circular dichroism spectroscopy of 30 μM solutions of **SS3**, **SS4** and **SS Control** monomer, in 5 mM HEPES buffer pH 7.

Both the glycosylated peptides and the control peptide were analysed using circular dichroism (CD) spectroscopy and the data obtained are presented in Figure 3. The CD spectra showed characteristic minima at 222 and 208 nm, consistent with formation of the α-helical coiled coil. The % folding of these coiled coils was determined from the intensity at 222 nm and was consistent with **SS Control** being 87% folded. Both glycosylated peptides are slightly less folded, but still consistent with a well folded coiled coil (**SS4** 78%, and **SS3** 75%). The self-assembling characteristics of the glycoclusters and the non-glycosylated peptide were studied using analytical ultracentrifugation (AUC). Sedimentation equilibrium curves were analysed using a modified form of the INVEQ algorithm,[20] adapted to estimate monomer-dimer-trimer stoichiometry. These data showed the presence of a small amount of monomer for both the **SS Control** and **SS4** (circa 5 %). The glycosylated peptide **SS3** with the sugar moiety positioned in the third heptad showed exclusive trimer formation, with no monomer or dimer species detected. The data obtained from the analytical ultracentrifugation shows that the coiled coils can tolerate the disaccharide moieties without disrupting the self-assembly into trimers,

Molecular modelling was used to estimate possible distances between the β-galactose anomeric carbon atoms in both **SS3** and **SS4** (Figure 4). The distance between these carbon atoms was ~30 Å in both structures, consistent with the different structures involving a ~10 Å linear translation (one heptad) along the coiled coil scaffold. These inter-galactose distances are greater than those found in other glycoclusters prepared in our group to date and greater than those found in tri-antennary glycans (~20 Å). In **SS3**, as expected there is a larger portion of peptide flanking either side of the lactose residues, while in **SS4** this is reduced, which may in turn influence lectin binding to the disaccharide residue.

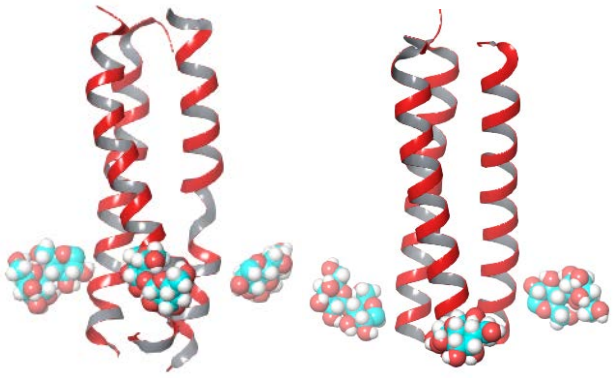


Figure 4. Graphical representations of **SS3** (left) and **SS4** (right) are shown, where the lactose residues project towards the exterior of the coiled coil. The peptide component is depicted in ribbons with the conjugated lactose residues shown as space filling models The distance between the β-galactose anomeric carbon atoms in both **SS3** and **SS4** is ~30 Å. The models of **SS3** and **SS4** were built using Maestro and structures were energy minimised in Macromodel using OPLSAA force field (www.schrodinger.com). Angle constraints (C1-C1-C1) were applied between the galactose anomeric carbons of 60º during energy minimisation.

In summary, this research has provided glycoclusters based on protein scaffolds for biological evaluation, the results of which will be presented in due course. These agents will be compared with glycoclusters based on other synthetic scaffolds. The N-glycoslated lactose residues have been found to be well tolerated in *de novo* peptide design and not to prevent helix formation or alter trimerisation of the coils. This will inform future design[21] of neoglycoproteins with a view to developing applications.

Acknowledgments

The material described herein was funded by the Irish Research Council (scholarship to SS), COST CM1102 (short term scientific mission for SS to Birmingham) and the University of Birmingham and EPSRC (studentships for GAB).

Supplementary data

HPLC, MS and AUC analysis

References and notes

[1] André, S.; Kaltner, H.; Manning, J. C.; Murphy, P. V.; Gabius, H.-J. *Molecules* **2015**, *20*, 1788-1823.
[2] Lee, Y. C. *J. Biochem.* **1997**, *121*, 818-825.
[3] Garner, O. B.; Baum, L. G. *Biochem. Soc. Trans*. **2008**, *36*, 1472-1477.
[4] Murphy, P. V. *Eur. J. Org. Chem.* **2007**, 4177-4187.

[5] (a) Murphy, P. V.; André S.; H.-J. Gabius, *Molecules* **2013**, *18*, 4026-4053. (b) Wittmann, V.; Pieters, R. J. *Chem. Soc. Rev.* **2013**, *42*, 4492-4503.
[6] (a) Leyden, R.; Velasco-Torrijos, T. André, S; Gouin, S.; Gabius H.-J.; Murphy, P. V. *J. Org. Chem.* **2009**, *74*, 9010-9026. (b) André, S.; Velasco-Torrijos, T.; Leyden, R.; Gouin, S.; Tosin, M.; Murphy, P. V.; Gabius, H. J. *Org. Biomol. Chem.* **2009**, *7*, 4715-4725.
[7] André, S.; O'Sullivan, S.; Koller, C.; Murphy, P. V.; Gabius, H. *J. Org. Biomol. Chem.* **2015**, *13*, 4190–4203.
[8] André, S., Jarikote, D. V.; Yan, D.; Vincenz, L.; Wang, G. N.; Kaltner, H.; Murphy P. V.; Gabius, H. J. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 313-318.
[9] Berwick, M. R.; Lewis, D. J. Pikramenou, Z.; Jones, A. W.; Cooper, H. J.; Wilkie, J.; Britton, M. M.; Peacock, A. F. A. *J. Am. Chem. Soc.* **2014**, *136*, 1166-1169.
[10] Fletcher, J. M.; Boyle, A. L.; Bruning, M.; Bartlett, G. J.; Vincent, T. L.; Zaccai, N. R.; Armstrong, C. T.; Bromley, E. H. C.; Booth, P. J.; Brady, R. L.; Thomson, A. R.; Woolfson, D. N., *ACS Synth. Biol.* **2012**, *1*, 240-250.
[11]. Mehta, S.; Meldal, M.; Ferro, V.;.Duus, J. Ø.; Bock, K. *J. Chem. Soc., Perkin Trans. 1*, **1997**, 1365-74.
[12] Falenski, J. A.; Gerling, U. I. M.; Koksch, B. *Bioorg Med. Chem.* **2010**, *18*, 3703-3706.
[13] Zacco, E.; Hütter, J.; Heier, J. L.; Mortier, J. ; Seeberger, P. H.; Lepenies, B.; Koksch, B. *ACS Chem. Biol.* **2015**, *10*, 2065-2072.
[14] For work on other types of non-natural amino acids in peptide design see (a) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proc. Nat. Acad. Sci., USA*, **2008**, *105*, 16566-71 (b) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angew. Chem. Int. Ed.* **2009**, 48, 7371-4.
[15] Berwick, M. R.; Slope, L. N.; Smith, C. F.; King, S. M.; Newton, S. L.; Gillis, R. B.; Adams, G. G.; Rowe, A. J.; Harding, S. E.; Britton, M. M.; Peacock, A. F. A. *Chem. Sci.* **2015** DOI: 10.1039/c5sc04101e.
[16] Crick, F. H. C. *Acta Cryst.* **1953**, *6*, 689-697.
[17] Pace, C. N.; Scholtz, J.M.; *Biophys. J.* **1998**, *75*, 422-427.
[18] Van Ameijde, J.; Albada, H. B.; Liskamp, R. M. J. *J. Chem. Soc. Perkin Trans. 1* **2002**, 1042-1049.
[19] Coin, I.; Beyermann, M.; Bienert, M., Nat. Protocols 2007, 2, 3247-3256.
[20] Rowe, A. J. *Methods* **2011**, *54*, 157-166.
[21] For a recent review on xenopeptide design see Slope, L. N.; Peacock, A. F. A.; *Chem. Asian J.* **2015**, DOI: 10.1002/asia.201501173.