Tetrahedron Letters 57 (2016) 1414–1417

Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

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 [d], Stephen E. Harding [d], James H. R. Tucker [b], Anna F. A. Peacock [b,*], 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

ARTICLE INFO

Article history:
Received 11 January 2016
Revised 28 January 2016
Accepted 1 February 2016
Available online 6 February 2016

Keywords:
Scaffold
Glycosylated
Coiled coil
Glycocluster

ABSTRACT

Scaffold design, synthesis and application are 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.

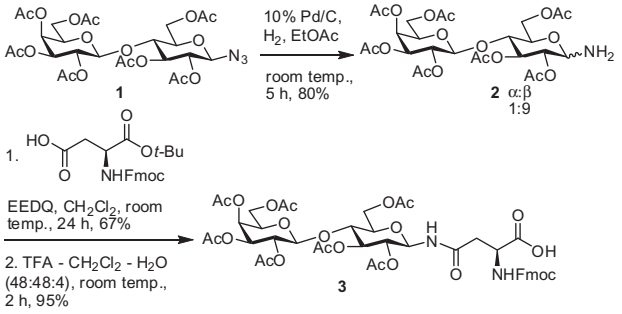
Bioactive molecules can be considered to be comprised of core scaffolds displaying pharmacophoric groups to enable them to interact with their target receptor. Glycoproteins, for example, can contain multiple carbohydrate groups on a protein surface which are involved in multivalent interactions with 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 are 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 Letters on preparation of *O*-glycosylated dimeric coiled coils,[11–13] including some that interact with asialoglycoprotein receptors, have 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 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($I_aA_bA_cI_dE_eQ_fK_g$)$_4$G–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

* Corresponding authors.
E-mail addresses: A.F.A.Peacock@bham.ac.uk (A.F.A. Peacock), paul.v.murphy@nuigalway.ie (P.V. Murphy).

http://dx.doi.org/10.1016/j.tetlet.2016.02.005
0040-4039/



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

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.

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}).

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 β-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 (Fig. 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 C-terminus to the N-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

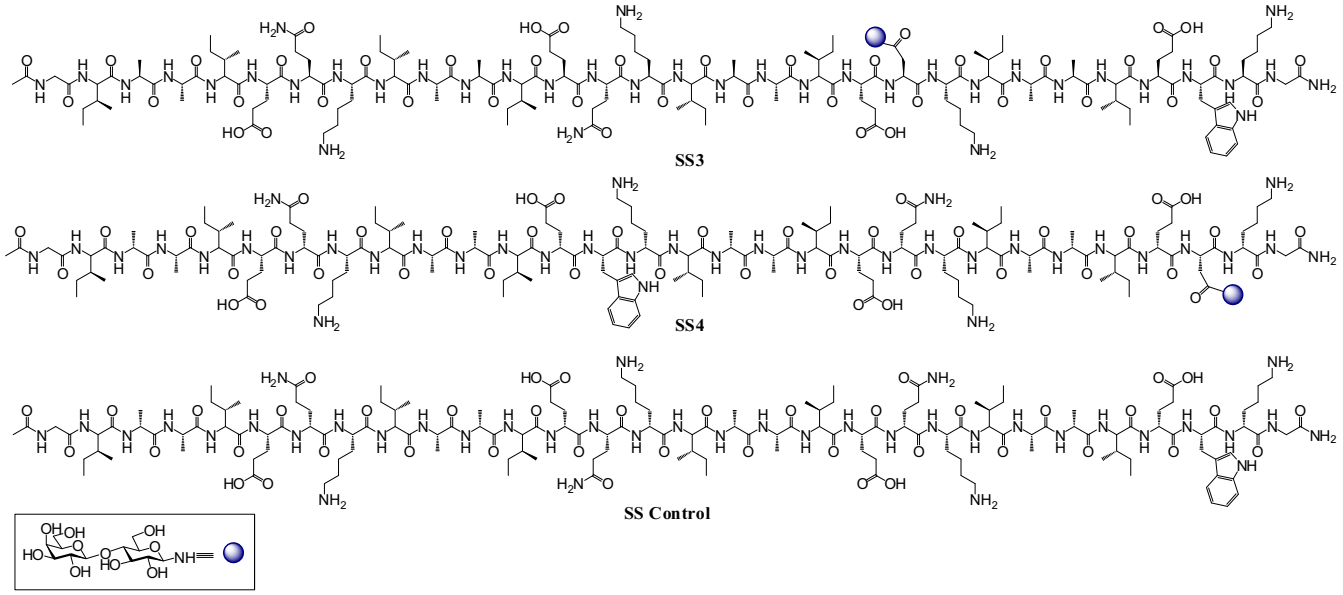


Figure 1. Structure of glycosylated peptides, **SS3** and **SS4** and non-glycosylated peptide, **SS Control**.