O-Glycosylation of glycine-serine linkers in recombinant Fc-fusion proteins
Attachment of glycosaminoglycans and other intermediates with phosphorylation at the xylose sugar subunit

Chris Spahr*, Stone D-H Shi and Hsieng S Lu
Biologics Optimization; Therapeutic Discovery; Amgen Inc.; Thousand Oaks, CA USA

Keywords: glycosaminoglycans, xylose, Fc-fusion protein, linkers, serine, phosphorylation, sulfation

Abbreviations: Xyl, xylose; Gal, galactose; GlcA, glucuronic acid; HexNAc, N-acetylhexosamine; Fuc, fucose; Neu5NAc, N-acetyl neuraminic acid or sialic acid; GAG, glycosaminoglycan; G0F, Fuc1Man3GlcNAc2Gal1; G2F, Fuc1Man3GlcNAc2Gal2; CID, collision-induced dissociation; PNGase F, Peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidase; HS, heparan sulfate; CS, chondroitin sulfate; CHO, Chinese hamster ovary; DTT, dithiothreitol; PTM, post-translational modification; ScFvs, single chain variable fragments; LCAT, lecithin-cholesterol acyltransferase; UPLC, ultra-pressure liquid chromatography; IAM, iodoacetamide; PBS, phosphate buffer saline

A xylose-based glycosaminoglycan (GAG) core was recently identified at a Ser residue in the linker sequence of a recombinant Fc fusion protein. The linker sequence, G-S-G-G-G-G, and an upstream acidic residue were serving as a substrate for o-xylosyltransferase, resulting in a major glycan composed of Xyl-Gal-Gal-GlcA and other minor intermediates. In this paper, a portion of an unrelated protein was fused to the C-terminus of an IgG Fc domain using the common (G4S)n linker repeat. This linker resulted in a heterogeneous population of xylose-based glycans all containing at least a core Xyl. Commonly observed glycan structures include GAG-related di-, tri-, tetra-, and penta-saccharides (e.g., Xyl-Gal, Xyl-Gal-Gal, Xyl-Gal-Gal-GlcA, and Xyl-Gal-Gal-GlcA-HexNAc), as well as Xyl-Gal-Neu5Ac. Following alkaline phosphatase or sialidase treatment combined with CID fragmentation, low-level glycans with a mass addition of 79.9 Da were confirmed to be a result of phosphorylated xylose. A minute quantity of phosphorylated GAG pentasaccharides may also be sulfated (also 79.9 Da), possibly at the HexNAc moeity due to non-reactivity to alkaline phosphatase. The xylose moiety may be randomly incorporated in one of the three G-S-G sequence motifs; and the linker peptide shows evidence for multiple additions of xylose at very low levels.

Introduction

To extend their serum half-life and in vivo efficacy, many protein or polypeptide chains can be linked to an Fc moiety (usually IgG1 Fc) via protein engineering to generate protein/peptide Fc chimeric molecules. Several of these engineered molecules, such as TNFR2-Fc (etanercept), CTLA4-Fc (abatacept), VEGF Trap-Eye (aflibercept), and the thrombopoietin mimetic peptibody (romiplostim), have been demonstrated to be clinically efficacious in various therapeutic indications. Similarly, to extend the clinical use of antibodies, a broad variety of antibody derivatives can also be engineered to improve biological and pharmacological characteristics, or to exhibit dual or multiple target specificities. Examples include single-chain variable fragments (scFvs), bispecific scFvs, bi- and tri-specific antibodies, double-variable domain antibodies, bispecific Zybodies (peptide linked to antibodies) and bispecific trimerbodies. Protein-Fc fusion and antibody derivatives usually require linkers of various lengths to connect different protein domains. The most frequently used linkers are peptides containing repeating units of Gly-Gly-Gly-Gly-Ser, (G4S)n. These linkers are considered to be universal due to their flexibility and resistance to proteases.

In addition to critical functional and pharmacological tests, therapeutic development of protein-Fc fusion and antibody derivatives typically involves extensive structural characterization of candidate molecules. Linker sequences may need to be characterized, inter alia, to determine the presence or absence of post-translational modifications (PTMs). Many of the marketed Fc fusion products have short linkers for which no PTMs have been reported. However, recent papers have reported that PTMs can indeed occur at the Ser residue in linkers containing
the repeating linker sequence with (G4S)₄.¹⁵⁻¹⁷ As reported by Wen et al., xylose attached at the GSG motif in the linker regions is the predominant modification in bispecific engineered antibodies produced by Chinese hamster ovary (CHO) cells, and the level of xylosylation is clone-dependent.¹⁵ At the peptide level, very low levels of GAGs and intermediates, e.g., bi-, tri-, tetra-, and penta-saccharides together with a pentasaccharide sulfated at glucuronic acid (GlcA) can also be identified. Spencer et al. reported O-glycosylation at two different Gly-Ser linker sequences [either (G4S), or (G4S),] in an engineered protein containing linkers that connect two fibronectin Tn3 domains and a non-Tn3 domain.¹⁶ The level of glycosylation is significantly higher in engineered protein produced in HEK293 cells compared with that produced in CHO cells. The predominant glycans observed only in the first linker include the non-sialylated and sialylated glycosaminoglycan (GAG) tetrasaccharide core (Xyl-Gal-Gal-GlcA) as well as a pentasaccharide, both of which are also sulfated.

We recently reported that xylose glycans were attached to a Ser residue when a G-S-G-G-G-G (GSG4) peptide linker was used in a human lecithin-cholesterol acyltransferase-Fc (LCAT-Fc) fusion construct expressed in CHO cells.¹⁷ Up to 50% of the molecule was detected to contain a GAG tetrasaccharide core due to the presence of the LCAT C-terminal acidic glutamic acid penultimate to the linker. Other GAG-related sugars and smaller polysaccharide units such as Xyl, Xyl-Gal, and Xyl-Gal-Gal were also detected as minor species at the peptide level. We also observed that the phosphorylated GAG core is present as a minor species, whereby the phosphate moiety is directly attached to the xylose moiety and can be easily removed by alkaline phosphatase treatment. Discrepancies between our data and those in the Spencer and Wen reports led us to perform additional experiments described here to confirm the GAG and GAG-related species, and to verify that small portions of these species containing an additional 79.9 Da mass are phosphorylated at the xylose sugar unit. Potential sulfation was only found in a phosphorylated or non-phosphorylated pentasaccharide at a level even lower than phosphorylation.

**Results**

**Intact mass analysis of the deglycosylated Fc-fusion protein**

In the design of the Fc-protein fusion, the (G4S)₄ linker (peptide sequence of GGGS(G4S)GGGS(G4S)GGGS) was initially used to provide more separation between Fc and the protein. Following DTT reduction and LC-MS, the Fc-protein fusion construct expressed in CHO cells yielded unexpected heterogeneity in the Fc-protein fusion chain beyond the expected Fc glycans (G0F, G1F, G2F, and sialylated G2F forms) as shown in Figure 1 (top panel). In addition to the assigned glycans, a mass increase of 132 Da was detected in all respective glycoforms. Upon PNGase F deglycosylation, much of the Fc glycan heterogeneity was eliminated as confirmed by LC-MS of the DTT-reduced sample (Fig. 1, bottom panel). The molecular mass with the addition of 132 Da, accounts for -15–20% of the remaining detectable signal at the protein chain level. Other higher masses with increases of 294.1 Da, 585.5 Da and 712 Da that cannot be identified in the non-deglycosylated sample also became visible at lower level (Fig. 1, bottom panel). Distribution of these forms relative to the unglycosylated form is shown in panel A of Table 1. From our previous work with xylose glycans in LCAT-Fc¹⁷ and in other reports,¹⁵,¹⁶ these observed forms with unique mass addition are clearly related to Ser xylosylation, which can be further characterized at the peptide level as described below.

**Heterogeneity of the linker tryptic-derived peptide**

The (G4S)₄ linker sequence shown in bold resides within the tryptic-derived peptide SLSLSPGGG SGGGGGGG SGGGGGSGAR, bridges the C-terminus of Fc (SLSLSPG), and includes the first two residues (AR) of the fused protein sequence. Three Ser residues within the linker are underlined as GSG motifs and suggest possible sites for xylose placement by a xylosyltransferase. High resolution LC-MS/MS of tryptic digests revealed the unmodified (G4S)₄ linker peptide elutes at 20.6 min (data not shown). Extracting the doubly charged peptide masses of the peptide with either a +132 Da or + 585 Da mass increase (m/z = 1141.01 and 1367.58) revealed that modified/glycosylated forms all eluted within 1 min of the unmodified peptide. To ensure that no glycosylated forms of the linker peptide are missed, the full MS data from T19.5–22.5 min shown in Figure 2 was averaged and then deconvoluted using the Xtract component of the Xcalibur software. The resulting deconvoluted masses were displayed as monoisotopic (M+H)⁺ with m/z (2130–3000) in the top panel and m/z (2800–3900) in the bottom panel with an expanded ion intensity scale. Focusing on the m/z range of 2800–3900 allows observation of many of the lower abundance glycosylated species that would be overlooked in the larger (2130–3000) mass range. In addition to the expected +132 Da and +585 Da mass increases, many other ions that were not observed at the protein chain level were identified at the peptide level.

The unmodified tryptic peptide containing the (G4S)₄ linker has a monoisotopic [M+H]⁺ mass of 2147.97 Da. Addition of +132 Da (2280.01 Da) corresponds to addition of a xylose (Xyl), while + 585 Da (2733.16 Da) corresponds to addition of Xyl-Gal-Neu5Ac. The 2412.05 Da mass, about 264 Da higher than the unmodified peptide, is derived from double addition of xylose, likely at two Ser residues within three of the GSG motifs. The actual sites of Ser xylosylation were not determined because xylose may or may not be evenly incorporated, although localization was demonstrated by Wen using ETD. Other glycosylated species of the tryptic linker peptide include masses 2442.07 Da (Xyl-Gal), 2604.12 Da (Xyl-Gal-Gal), 2684.08 Da (Xyl-[PO₃]-Gal-Gal), 2780.15 Da (Xyl-Gal-Gal-GlcA), 2813.13 (Xyl-[PO₃]-Gal-Neu5Ac), 2860.11 Da (Xyl-[PO₃]-Gal-Gal-GlcA), 3063.19 Da (Xyl-[PO₃]-Gal-Glca-GlcNAc), and 3160.19 Da (Xyl-Gal-GlcA-GlcNAc-GlcA). Based on the calculated glycan mass, the following masses shown in Figure 2 may contain two distinct glycans in the linker: 2865.20 Da (Xyl and Xyl-Gal-Neu5Ac), 2945.17 Da (Xyl and Xyl-[PO₃]-Gal-Neu5Ac), 2992.16 Da (Xyl and Xyl-[PO₃]-Gal-GlcA), 3027.26 Da (Xyl-Gal and Xyl-Gal-Neu5Ac), 3318.34 Da (Xyl-Gal-Neu5Ac and Xyl-Gal-Neu5Ac), 3365.34 Da (Xyl-Gal-GlcA and
Xyl-Gal-Neu5Ac), and 3444.31 Da (Xyl-[PO$_3$]-Gal-Gal-GlcA and Xyl-Gal-Neu5Ac). Molecular ions at 3143.16 Da, 3195.22 Da, and 3275.19 Da may also contain sulfate or phosphorylated glycans that will be described in detail later in the results section. Table 1 (panel B) summarizes all of the observed xylose glycosylated species in the Fc-(G4S)$_4$-fusion protein sample. There are other low-level xylosylated peptide species present at up to ~5 kDa, likely combinations of the various observed glycans species distributed across all three GSG motif Ser residues (data not shown). Each of these glycosylated species is related to the unmodified tryptic linker peptide and the detailed structural assignment of many are described in the following paragraphs.

Analysis of tryptic linker peptides after sialidase or alkaline phosphatase digestion

To confirm the existence of terminal sialylation in the various glycan species described above, desalted tryptic digest was analyzed by LC-MS/MS both before and after sialidase treatment. Figure 3 compares full MS spectra averaged from 19.5–22.5 min that were then deconvoluted and displayed as monoisotopic (M+H)$^+$ from (2400–3200). In the Figure 3 (top panel), sialylated glycopeptide masses of 2733.16 Da (Xyl-Gal-Neu5Ac), 2813.13 Da (Xyl-[PO$_3$]-Gal-Neu5Ac), 2865.20 Da (Xyl and Xyl-Gal-Neu5Ac), and 2945.17 Da (Xyl and Xyl-[PO$_3$]-Gal-Neu5Ac) are present in blue, while other non-sialylated glycan masses are shown in red. Following sialidase treatment (Fig. 3, bottom panel), these masses are eliminated and result in an increase in intensity for 2442.06 Da (Xyl-Gal) and the appearance of new peptide ions at 2522.03 Da (Xyl-[PO$_3$]-Gal), 2574.10 Da (Xyl and Xyl-Gal), and 2654.07 Da (Xyl and Xyl-[PO$_3$]-Gal), suggesting that the peptide species represented by 2733.16 Da 2813.13 Da, and 2945.17 Da all originally contained terminal sialic acid. Other non-sialylated peptide ions shown in both ion panels of Figure 3 retain their respective ion intensity and are apparently not affected by sialidase treatment.

Assignment of glycan structure can be obtained by LC-MS/MS analysis of selected glycopeptide ions in which mono or oligosaccharide subunits are preferentially fragmented under CID conditions. Figure 4 (top panel) displays the fragment ion spectra of the unmodified tryptic linker peptide at m/z = 1074.99 Da (doubly charged ion). Fragmentation of the peptide backbone can be readily obtained for assignment of the peptide sequence. The middle and bottom panels of Figure 4 show MS/MS spectra
Table 1. Relative distribution of xylosylated glycans from an Fc-fusion protein after removal of Fc N-linked glycans using PNGase F (panel A) and from the linker tryptic peptide (panel B).

| Table 1A | Mass (Da) | Glycoforms on the Reduced Protein Chain | Relative Distribution (%) |
|----------|-----------|----------------------------------------|---------------------------|
| 37877.2  | Deglycosylated* | ~72                                    |
| 38009.2  | Deglycosylated + 132.1 Da | ~18                                   |
| 38171.2  | Deglycosylated + 294.1 Da | ~4                                    |
| 38464.1  | Deglycosylated + 585.5 Da | ~4                                    |
| 38590.6  | Deglycosylated + 712 Da | ~2                                    |

| Table 1B | Mass (Da) | Glycan Structure(s) on the Linker Peptide** | Relative Distribution (%) |
|----------|-----------|-----------------------------------------------|---------------------------|
| 2147.97  | none      |                                               | 36.5                      |
| 2280.01  | Xyl       |                                               | 19.9                      |
| 2412.05  | Xyl and Xyl |                                               | 1                         |
| 2442.07  | Xyl-Gal   |                                               | 2.5                       |
| 2604.12  | Xyl-Gal-Gal | < 1                                      |
| 2684.08  | Xyl-[PO3]-Gal-Gal | 1.2                              |
| 2733.16  | Xyl-Gal-Neu5Ac | 19.3                      |
| 2780.15  | Xyl-Gal-GlcA | 2.5                                    |
| 2813.13  | Xyl-[PO3]-Gal-Neu5Ac | 5.9                              |
| 2860.11  | Xyl-[PO3]-Gal-GlcA | 3.3                               |
| 2865.20  | Xyl and Xyl-[PO3]-Gal-Neu5Ac | 1.7                              |
| 2945.17  | Xyl and Xyl-[PO3]-Gal-Neu5Ac | < 1                              |
| 2992.16  | Xyl and Xyl-[PO3]-Gal-GlcA | < 1                              |
| 3027.26  | Xyl-Gal and Xyl-Neu5Ac | < 1                              |
| 3063.19  | Xyl-[PO3]-Gal-GlcA-GlcNAc | 1.5                              |
| 3143.16  | Xyl-[PO3]-Gal-GlcA-GlcNAc-(SO3) | < 1                             |
| 3160.19  | Xyl-Gal-GlcA-GlcNAc-GlcA | < 1                              |
| 3195.22  | Xyl and Xyl-Gal-GlcA-GlcNAc-(SO3) | < 1                             |
| 3275.19  | Xyl and Xyl-[PO3]-Gal-GlcA-GlcNAc-(SO3) | < 1                            |
| 3318.34  | Xyl-Gal-Neu5Ac and Xyl-Gal-Neu5Ac | < 1                            |
| 3365.34  | Xyl-Gal-Neu5Ac and Xyl-Gal-GlcA | < 1                              |
| 3444.31  | Xyl-Gal-Neu5Ac and Xyl-[PO3]-Gal-GlcA | < 1                             |

*Followed by treatment with PNGase F; **Attached to Ser(s) in the linker peptide.

for doubly charged glycopeptide ions of m/z 1222.04 Da and 1367.58 Da, respectively. Following CID activation, the preservation of the m/z 1074.8 fragment ion as the base peak was observed for both spectra, while the major fragmentation products are sugar units released from dissociation of the 294 Da and 585 Da glycan structures. The mass increases of 294 Da and 585 Da observed for the PNGase-digested and DTT-reduced sample (Fig. 1 bottom panel) correlate with the masses 2442.07 Da and 2733.16 Da shown in Figure 2 and result from attachment of Xyl-Gal and Xyl-Gal-Neu5Ac, respectively, to the linker peptide.

Many GAGS and intermediate species were observed to contain a 79.9 Da increase in glycan mass (Fig. 2); phosphate and sulfate, which share a similar nominal mass, are the best chemical “fits” for the moieties that are attached to these glycans. To confirm that phosphate, and not sulfate, is being directly attached to the glycan, desalted tryptic digest was analyzed by LC-MS/MS with and without alkaline phosphatase treatment. Figure 5 compares deconvoluted full MS data from retention time of 19.5–22.5 min of an LC-MS/MS tryptic peptide map. The top panel displays only deconvoluted glycopeptide masses from (2900–3450) with no alkaline phosphatase treatment, while the bottom panel displays those after alkaline phosphatase treatment. Glycopeptide masses believed to be phosphorylated are present at 2684.08 Da (Xyl-[PO3]-Gal-Gal), 2813.13 Da (Xyl-[PO3]-Gal-Neu5Ac), 2860.11 Da (Xyl-[PO3]-Gal-GlcA), 2945.17 Da (Xyl and Xyl-[PO3]-Gal-Neu5Ac), 2992.16 Da (Xyl and Xyl-[PO3]-Gal-GlcA), 3063.19 Da (Xyl-[PO3]-Gal-GlcA-GlcNAc), 3239.22 Da (Xyl-[PO3]-Gal-GlcA-GlcNAc-GlcA), and 3444.31 Da (Xyl-Gal-Neu5Ac and Xyl-[PO3]-Gal-GlcA). Additional phosphorylated or sulfated masses of 3143.16 Da, 3195.22 Da, and 3275.19 Da will be discussed in detail below.

Following alkaline phosphatase treatment, the masses mentioned above are eliminated; however, the intensity of the non-phosphorylated versions of the same glycopeptides have now increased or have newly appeared: 2604.12 Da (Xyl-Gal-Gal), 2733.16 Da (Xyl-Gal-Neu5Ac), 2780.16 Da (Xyl-Gal-GlcA), 2865.21 Da (Xyl and Xyl-Gal-Neu5Ac), 2912.20 Da (Xyl and Xyl-Gal-GlcA), 2945.17 Da (Xyl-Gal-Gal-GlcA), 3027.26 Da (Xyl-Gal and Xyl-Neu5Ac), 3063.19 Da (Xyl-[PO3]-Gal-GlcA-GlcNAc), 3143.16 Da (Xyl-[PO3]-Gal-GlcA-GlcNAc-GlcA), 3195.22 Da (Xyl and Xyl-Gal-GlcA-GlcNAc-GlcA), and 3275.19 Da (Xyl-Gal-Neu5Ac and Xyl-Gal-GlcA). The data indicate that these glycans are sensitive to alkaline phosphatase treatment and that the extra 79.9 Da of mass is due to phosphorylation. Alkaline phosphatase reactivity can also be used to differentiate that glycans have been phosphorylated vs. sulfated. In our data, most glycoforms with mass additions of 79.9 Da can be removed by alkaline phosphatase treatment and were thus phosphorylated. However, there are exceptions, such as the molecular ions containing at least GAG pentasaccharides (e.g., Xyl-Gal-GlcA-GlcNAc-GlcA) at 3143.16 Da, 3195.22 Da, and 3275.19 Da (Fig. 5, top panel). The molecular ions corresponding to 3143.16 Da and 3275.19 Da both have a mass increase of 159.8 Da in addition to the estimated GAG pentasaccharide mass, suggesting either two phosphate moieties, two sulfates, or both a phosphate and a sulfate are attached. Following phosphatase digestion, these two ions disappear with the concomitant generation of peptides at masses 3063.19 Da and 3195.24 Da (Fig. 5, bottom panel). Both
masses still exhibit an increase of 79.9 Da, suggesting the presence of sulfate. We postulate that this sulfate moiety is attached at the terminal sugar unit, HexNAc, in the GAG pentasaccharide (Xyl-Gal-Gla-GlcA-HexNAc), a common sulfate incorporation site for heparan sulfate (HS) and chondroitin sulfate (CS).

Confirmation of xylose phosphorylation

CID fragmentation of the doubly charged glycopeptide ions of m/z 1391.08, 1431.06, and 1532.60, corresponding to masses 2780.15 Da, 2860.11 Da, and 3063.19 Da were obtained during LC-MS/MS tryptic mapping. The 2780.15 Da mass was confirmed to be the tetrasaccharide GAG core, Xyl-Gal-Gal-GlcA, based on the CID fragmentation pattern (Fig. 6, top panel). The 2860.11 Da mass contains an identical GAG core structure with the addition of 79.9 Da associated with xylose (middle panel). The 3063.19 Da mass is a pentasaccharide containing a GAG core with terminal extension of N-acetyl hexosamine (presumably N-acetyl glucosamine, GlcNAc) and an additional 79.9 Da mass that is also associated with xylose (bottom panel). The alkaline phosphatase digestion experiments described in the previous section confirm that the addition of 79.9 Da is due to incorporation of phosphate and the MS/MS data prove that the phosphate moiety is attached at the xylose moiety. Phosphorylated xylose is relatively stable; upon CID activation the phosphate moiety remains associated with xylose as evidenced by the prominent doubly charged 1180.8 Da fragment ion shown in the middle and bottom panels of Figure 6. Phosphorylated glycans generate a loss of the Xyl-(PO₃) as a single entity following the loss of the other glycans. In contrast, sulfation is well known to be extremely labile to electrospray ionization and would result in a complete mass loss of sulfate independent of the xylose under CID conditions.¹⁸

Evidence for trace sulfation on certain pentasaccharides

The 3063.19 Da masses present in both panels of Figure 7 (and also Fig. 5) were fragmented both before and after alkaline phosphatase treatment. Prior to treatment, the structure is a + 915 Da pentasaccharide composed of Xyl-(PO₃)₂-Gal-GlcA-HexNAc. Under CID conditions (Fig. 5, top panel), it yields a fragment ion spectra showing the sequential loss of all glycans including the final loss of Xyl-(PO₃) as a single unit. Following alkaline phosphatase treatment, a similar 3063.19 Da mass is
generated (Xyl-Gal-Gal-GlcA-HexNAc + SO₃), the product of alkaline phosphatase treatment of a phosphorylated and sulfated 3143.16 Da pentasaccharide structure Xyl-(PO₃)-Gal-Gal-GlcA-HexNAc + SO₃. However, under CID conditions (Fig. 7, bottom panel), this 3063.19 Da mass repeatedly yields non-productive spectra with the labile loss of sulfate. Trace levels of this loss and the precursor are shown as fragment ions of \( m/z \) 1492.85 and 1074.61. We suspect the sulfate is attached to the HexNAc, as it was not observed in any of the shorter glycan forms; however, this has not been experimentally determined.

**Discussion**

The GAG tetrasaccharide, also known as the GAG-protein linkage region, has been identified as a core structure required for initiating chain polymerization in the biosynthesis of heparan sulfate (HS), as well as chondroitin sulfate (CS).

Many reports have indicated that modifications of the GAG tetrasaccharide core do exist and novel modified structures have been identified. For example, sulfates have been found to be attached at two of the galactose sugar units via different linkages. Sulfated galactose was only found in CS, but not in HS, and may play a regulatory role in CS biosynthesis. Phosphate has also been identified to be attached at xylose by a specific xylose kinase, FAB20B. In a recent report, Nadanaka et al. hypothesized that the phosphorylation of xylose by FAM20B facilitates the formation of the linkage region and thereby leads to enhanced GAG biosynthesis. Our observation of xylose phosphorylation at the GAG core and intermediates suggests that recombinant CHO expression system may contain residual xylose kinase activity. Wakabayashi et al. reported the attachment of sulfate at the terminal GlcA of GAG core in human urine α-thrombomodulin (αTM) that contains a typical sequence motif for syntheses of proteoglycans. However, in addition to sulfated galactose as described above, sulfation typically occurs at N-acetyl glucosamine-glucuronic acid (GlcNAc-GlcA) disaccharide repeating units extended off of the GAG core. Therefore, sulfated GlcA found in urinary αTM may be unique, as soluble human molecules recombinantly produced in CHO cells only generate immature, non-sulfated
GAG tetrasaccharide and other shorter intermediates based on analysis of the released glycans from recombinant αTM. This early report on the immature GAGs found in recombinant αTM is consistent with that for recombinant LCAT-Fc and other fusion molecules containing linker sequence with GSG sequence motif. Our current observation further concludes that low level xylose phosphorylation, but not sulfation, in the GAG core is a prominent modification event and that sulfation may also exist at extremely low levels on the GAG pentasaccharide presumably at the terminal HexNAc moiety.

The diverse functions ascribed to HS, CS, and characteristic polysaccharide chains are expressed in a tissue- and cell-type-specific manner, suggesting that the sulfation and polymerization of the sulfate polysaccharide chains are strictly regulated. In addition, the fact that GAG polymerization and modification is very rapid and involves over 10 concerted enzymatic reactions implies the requirement of highly organized enzymatic machinery during HS or CS biosynthesis. In contrast, despite the fact that GSG motifs in the linker region provide sites for Ser xylosylation and GAG oligosaccharide incorporation, in recombinant CHO cell expression this PTM event is probably not controlled due to the lack of required enzymatic machinery. As a result, this biosynthesis process immately stops with accumulation of partial GAG-intermediate products from xylose to a hexasaccharide, followed by low level phosphorylation at xylose and minute sulfation at the terminal HexNAc.

Materials and Methods

Protein expression and purification
Engineered fusion proteins were constructed using the peptide linker (G4S) to connect a portion of a protein to Fc (IgG1) with deletion of the hinge sequence. Engineered fusion proteins were then produced using a stable CHO cell expression system in 1–3 L shaker flasks or 5-L wave bag bioreactor. The Fc-fusion proteins were first purified by protein A affinity chromatography, followed by ion-exchange chromatography using an SP column.
Denaturation, reduction, deglycosylation, and LC-MS of proteins

Prior to LC-MS analysis, the Fc-fusion proteins were denatured and reduced at 0.5 μg/μL in 8M guanidine hydrochloride (GuHCl)/50 mM Tris, pH 8.3/20 mM DTT for 1 h at 37 °C, followed by acidification. For PNGase F digestion to remove N-linked glycans, about 20 μg of protein was diluted with PBS and 0.6 μL of PNGase F (New England Biolab), incubated overnight at 37 °C, then dried, denatured, reduced, and acidified as described above. About 5μg of protein was analyzed using an Agilent Technologies 1100 capillary HPLC connected to an Agilent Technologies 6224 ESI-TOF mass spectrometer. The HPLC used a 1.0 mm X 50 mm Agilent Zorbax 300SB-C8 column with a flow rate of 50 μL/min and a column temperature of 75 °C. Buffer A consisted of 0.1%TFA/ H₂O, while Buffer B was 0.1%TFA/ H₂O/ 90% n-propanol. The gradient consisted of initial conditions at 2% B for 5 min, up to 45% B over 20 min, up to 95% B over 3 min, isocratic at 95% B for 4 min, and back down to 2% B over 1 min. The MS method scanned m/z (750–6000) at rate of 1 spectra/second. Other instrumental parameters include: VCap = 3200V, fragmentor = 225V, skimmer = 60V, and OCT 1 RF Vpp = 800V. Using the Agilent MassHunter software, appropriate LC-MS spectra were combined and deconvoluted from [35,000–45,000] using a mass step of 1.0 Da and a S/N threshold of 30.0.

Proteolytic digestion and LC-MS/MS

About 50 μg of the various Fc-fusion protein samples were dried down, resuspended in 25 μL of 150 mM Tris, pH 7.5/8M urea/40 mM hydroxylamine/10 mM DTT, and then incubated for 1 h at 37 °C. The samples were alkylated with 20 mM iodoacetamide (IAM) for 30 min at room temperature in the dark. Finally, the samples were diluted to 100 μL with water and 2 μg of trypsin (1:25) and digested overnight at 37 °C.

The digests were acidified, followed by injection onto a Waters NanoAcquity UPLC system. Samples were first loaded onto a 180 μm X 20 mm Symmetry C18 trapping column at 15 μL/min, followed by peptide separation also at 15 μL/min on an Agilent Zorbax 0.5mm X 250 mm 300SB-C18 column. Buffer

![Figure 5. Deconvoluted LC-MS/MS of the linker tryptic peptide [210–238] and associated glycosylated species without and with alkaline phosphatase treatment (top and bottom spectra, respectively). Glycans with potential sulfation are designated with +S. Note that many masses shift, indicating the loss of phosphate from the xylose group upon alkaline phosphatase treatment. Certain pentasaccharide structures contain sulfation that is retained.](image)
A was 0.1% formic acid/ water, while buffer B was 0.1% formic acid/ 99.9% acetonitrile. The gradient consisted of initial conditions at 1% B, followed by an increase to 45% B over 85 min, to 97% B over 1 min, isocratic at 97% B for 6 min, to 1% B over 3 min, and then isocratic at 1% B for 20 min. The UPLC column effluent was sprayed into a Thermo Fisher Scientific Orbitrap Velos Pro mass spectrometer using the standard HESI II ionization source. The mass spectrometer method consisted of a full MS scan of \( m/z \) (300–2000) at 30K resolution in the Orbitrap, followed by MS/MS (CID activation) of the top 10 most abundant precursor ions acquired in the ion trap. The following instrument parameters were used for the analysis: source voltage = 3.5 kV; capillary temperature = 275 °C; S-lens RF level = 50%; activation time = 10 msec.; normalized collision energy = 35; isolation width = 2.0 Da; and threshold = 1.0E4. The Xtract component of the Thermo Xcalibur 2.1 software was used for deconvolution of high-resolution MS data. Averaged data from (300–2000) were deconvoluted using Xtract default settings: a S/N threshold of 1.2 and a resolution of 100,000 at \( m/z \) 400. It was observed that deconvolution of the high resolution MS data generated identical spectra regardless of processing using a 30K or 100K setting of Xtract. Deconvoluted peptide masses (glycosylated and non-glycosylated) were displayed as monoisotopic (M+H)+. The various glycosylated species were confirmed by the stepwise loss of glycan subunits and the presence of the unglycosylated precursor ion as the most intense fragment following CID activation.

**Sialidase and alkaline phosphatase treatment**

To confirm some of the different glycan components, additional tryptic digestion was performed, and about 300 μg of digest was desalted over a Waters Oasis HLB 1cc (30mg) cartridge. To confirm the phosphate attachment to xylose, 30 μg of tryptic digest was combined with 16 μL of 75 mM Tris, pH 8.0, and 4 μL of an alkaline phosphatase solution from bovine intestinal mucosa (P6774, Sigma-Aldrich) and incubated for 1 h.

---

**Figure 6.** MS spectra obtained from CID activation of different forms of the linker tryptic peptide [210–238]. Top panel: Peptide + 632 Da: Xyl-Gal-Gal-GlcA. Middle panel: Peptide + 712 Da: Xyl-[PO₃]-Gal-Gal-GlcA. Bottom panel: Peptide + 915 Da: Xyl-[PO₃]-Gal-Gal-GlcA-HexNAc. Direct attachment of phosphate to xylose is shown by the loss of Xyl-[PO₃] as a single entity (indicated by the intense 1180.8 fragment ion). As shown in Figure 5, following alkaline phosphatase treatment, a similar mass of the peptide + 915 Da can appear; however, the MS/MS properties are different, indicating sulfation rather than phosphorylation.
at 37 °C. To confirm the presence of sialic acid on the termini of glycans, 30 μg of tryptic digest was combined with 50 mM sodium phosphate, pH 6.0 and 2 μL of sialidase (QA Bio) and incubated for 1 h at 37 °C. A control digest that did not include any sialidase or alkaline phosphatase treatment was also included. Approximately 5 μg of each digest was analyzed by LC-MS/MS using the conditions described above under “Proteolytic Digestion and LC-MS/MS.”

Disclosure of Potential Conflicts of Interest

All the authors are or were Amgen’s employees or stockholders.

Acknowledgments

The authors would like to thank Dr. Michael Hall and John Hui for critical review of this manuscript.

References

1. Chamow SM, Ryll T, Lowman HB, Farson D, eds. Therapeutic Fc-fusion proteins. Wiley-Blackwell; 2014. 400 p.
2. Peppel K, Crawford D, Beutler B. A tumor necrosis factor (TNF) receptor-IgG heavy chain chimeric protein as a bivalent antagonist of TNF activity. J Exp Med 1991; 174:1483-9; PMID:1660525; http://dx.doi.org/10.1084/jem.174.6.1483
3. Moreland L, Bate G, Kirkpatrick P. Abatacept. Nat Rev Drug Discov 2006; 5:185-6; PMID:16557658; http://dx.doi.org/10.1038/nrd1989
4. Semeraro F, Morescalchi F, Duse S, Parmeggiani F, Gambicorti E, Costagliola C. Affibebrept in wet AMD: specific role and optimal use. Drug Des Devel Ther 2013; 7:711-22; PMID:23999075; http://dx.doi.org/10.2147/DDDTP.S40215
5. Shimamoto G, Gegg C, Boone T, Queva C. Peptibodies: A flexible alternative format to antibodies. MAb 2012; 4:586-91; PMID:22820181; http://dx.doi.org/10.4161/mabs.21024
6. Huston JS, Levinson D, Mudgett-Hunter M, Tai MS, Novotny J, Margolies MN, Ridge RJ, Brucoleri RE, Haber E, Crea R, et al. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. Proc Natl Acad Sci U S A 1988; 85:5879-83; PMID:3045807; http://dx.doi.org/10.1073/pnas.85.16.5879
7. Baeuerle PA, Reinhardt C. Bispecific T-cell engaging antibodies for cancer therapy. Cancer Res 2009; 69:4941-4; PMID:19509221; http://dx.doi.org/10.1158/0008-5472.CAN-09-0547
8. Chames P, Baty D. Bispecific antibodies for cancer therapy: the light at the end of the tunnel? MAbs 2009; 1:59-77; PMID:20073127; http://dx.doi.org/10.4161/mabs.1.1.10015
9. Kontermann R. Dual targeting strategies with bispecific antibodies. MAb 2012; 4:181-72; PMID:22453100; http://dx.doi.org/10.4161/mabs.4.2.19000

Figure 7. Mass spectra obtained from CID activation of linker peptide + 915 Da (m/z 1532.60). Top panel: No alkaline phosphatase treatment. Phosphorylated xylose is stable and lost as a single entity as indicated by the intense 1180.8 fragment ion. Bottom panel: Alkaline phosphatase treatment. Phosphate has been removed, however sulfate from the 3143.16 Da mass in Figure 4 remains. The resulting ion is labile and yields non-productive spectra.

www.landesbioscience.com mAbs 913
10. Castoldi R, Jucknischke U, Pradel LP, Arnold E, Klein C, Scheiblich S, Niederfellner G, Sustmann C. Molecular characterization of novel trispecific ErbB-Cmet-IGFR1 antibodies and their antigen-binding properties. Protein Eng Des Sel 2012; 25:551-9; PMID:22996099; http://dx.doi.org/10.1093/protein/gst048

11. Jakob CG, Edalji R, Judge RA, DiGiammarino E, Li Y, Gu J, Ghaury T. Structure reveals function of the dual variable domain immunoglobulin (DvD-IgM) molecule. Mabs 2013; 5:358-63; PMID:23549062; http://dx.doi.org/10.4161/mabs.23977

12. Blanco-Toribio A, Sainz-Pastor N, Álvarez-Cienfuegos A, Merino N, Cuesta AM, Sánchez-Martin D, Bonet J, Santos-Valle P, Sanz L, Oliva B, et al. Generation and characterization of monospecific and bispecific hexavalent trimers. Mabs 2013; 5:70-9; PMID:23221741; http://dx.doi.org/10.4161/mabs.22698

13. Kanakaraj P, Puffer BA, Yao X-T, Cenkkanala S, Boyd E, Shah RR, Wang G, Patel D, Krishnamurthy B, Kathamana S, et al. Simultaneous targeting of TNF and Ang2 with a novel bispecific antibody enhances efficacy in an in vivo model of arthritis. Mabs 2012; 4:600-13; PMID:22864384; http://dx.doi.org/10.4161/mabs.21227

14. Michaelson JS, Demaree SJ, Miller B, Amatucci A, Cienfuegos A, Merino N, Cuesta AM, Sánchez-Martin D, Bonet J, Santos-Valle P, Sanz L, Oliva B, et al. Anti-tumor activity of stability-engineered IgG-Snyder WB, Wu X, Huang F, Phan S, Gao S, Doern A, PMID:20061822; http://dx.doi.org/10.1002/pro.2373

15. Wen D, Foley SF, Huckerby TN, Nieduszynski IA. Discovery and investigation of O-xylosylation in a "part-time" proteoglycan. J Biol Chem 1998; 273:6072-35; PMID:956114

16. Koike T, Izumikawa T, Yamamoto T, Izumikawa K, Sugihara K, Asano M, Kitagawa H. EXTL2, a member of the EXT family of tumor suppressors, controls glycosaminoglycan biosynthesis in a syloose kinase-dependent manner. J Biol Chem 2013; 288:9321-35; PMID:23935820; http://dx.doi.org/10.1074/jbc.M112.416909

17. Nakanaoka P, Puffer BA, Yao X-T, Kankanala S, Boyd E, Shah RR, Wang G, Patel D, Krishnamurthy B, Kathamana S, et al. Simultaneous targeting of TNF and Ang2 with a novel bispecific antibody enhances efficacy in an in vivo model of arthritis. Mabs 2012; 4:600-13; PMID:22864384; http://dx.doi.org/10.4161/mabs.21227

18. Men W, Foley SF, Huckerby TN, Nieduszynski IA. Discovery and investigation of O-xylosylation in engineered proteins containing a (GGGGS)n linker. J Biol Chem 2000; 275:6027-35; PMID:10026155; http://dx.doi.org/10.1074/jbc.275.25.6027

19. Men D, Foley SF, Huckerby TN, Nieduszynski IA. Discovery and investigation of O-xylosylation in engineered proteins containing a (GGGGS)n linker. J Biol Chem 2000; 275:6027-35; PMID:10026155; http://dx.doi.org/10.1074/jbc.275.25.6027

20. Wilson IB. The never-ending story of peptide O-sulfation. Cell Mol Life Sci 2004; 61:794-809; PMID:15959094; http://dx.doi.org/10.1007/s00018-003-2378-2

21. Sugahara K, Yamashina I, De Waard P, Van Halbeek H, Vliegenthart JFG. Structural studies on sulfated glycopeptides from the carbohydrate-protein linkage region of chondroitin 4-sulfate proteoglycans of swann rat chondrosarcoma. Demonstration of the structure Gal(4-O-sulfate)beta 1-3Gal beta 1-4XYL beta 1-O-Ser. J Biol Chem 1988; 263:10168-74; PMID:3134349

22. Sugahara K, Usy H, Harada T, de Waard P, Vliegenthart JFG. Structural studies on sulfated oligosaccharides derived from the carbohydrate-protein linkage region of chondroitin 6-sulfate proteoglycans of shark cartilage. I. Six compounds containing 0 or 1 sulfate and/or phosphate residues. J Biol Chem 1992; 267:6072-35; PMID:1556114

23. Lauder RM, Huckerby TN, Nieduszynski IA. Increased incidence of unsulfated and 4-sulfated residues in the chondroitin sulfate linkage region of recombinant rat chondrosarcoma. demonstration of the linkage tetrasaccharide GlcAbeta1-3Galbeta1-4Xyl on recombinant soluble human a-thrombomodulin. An oligosaccharide structure on a "part-time" proteoglycan. J Biol Chem 1998; 273:33728-34; PMID:9837960; http://dx.doi.org/10.1074/jbc.273.50.33728

24. De Beer T, Inui A, Tsuda H, Sugahara K, Vliegenthart JFG. Prolidipaserity in sulfation profile of oligosaccharide aldolts isolated from the protein-linkage region and the repeating disacidate region of chondroitin 4-sulfate of bovine nasal septal cartilage. Eur J Biochem 1996; 240:789-97; PMID:8856085; http://dx.doi.org/10.1111/j.1432-1033.1996.0789h.x

25. Tone Y, Pedersen LG, Yamamoto T, Isumikawa T, Kitagawa H, Nishihara J, Tamura J, Negishi M, Sugahara K. 2-o-phosphorylation of xylose and 6-o-sulfation of galactose in the protein linkage region of glycosaminoglycans influence the glucuronotransferase-I activity involved in the linkage region synthesis. J Biol Chem 2008; 283:16001-7; PMID:18409750; http://dx.doi.org/10.1074/jbc.M709556200

26. Koike T, Isumikawa T, Tamura J, Kitagawa H. FAM42B is a kinase that phosphorylates xylose in the glycosaminoglycan-protein linkage region. Biochem J 2009; 421:157-62; PMID:19473117; http://dx.doi.org/10.1042/BJ20090474