GENOVIS PRESENTS

Novel Enzymes for O-glycan Analysis

AUTHORS

Malin Mejàre*, Maria Nordgren, Fredrik Leo, Rolf Lood, Stephan Björk, Jonathan Sjögren, Helén Nyhlén and Fredrik Olsson

Genovis AB, Lund, Sweden

INTRODUCTION

The degree and composition of glycosylation of proteins are critical for a wide range of biological processes. Alterations of the glycan structures may impact the function and effect of the glycoprotein and thus, close monitoring of the glycan profile is required during development and manufacturing of biopharmaceuticals. The analyses of O-glycans have suffered from a lack of specific enzymes and there is a great need for novel tools.

We here describe new enzymes that are useful for analysis of O-glycosylated proteins. These include an O-glycan-specific endoprotease (O-protease), an O-glycosidase and sialidases that were all discovered in bacteria and cloned for expression in *E. coli.* We have developed LC-MS workflows and analytical strategies for site-specific O-glycan characterization using these new enzymatic tools, and we here demonstrate these by analyzing etanercept, an Fc-fusion protein with a highly O-glycosylated hinge region.

SUMMARY

- Aim: Characterize new enzymes with applications in glycoprotein analysis
- The O-protease has no activity on non- or only N-glycosylated proteins
- The O-protease displays site-specific digestion N-terminally of O-glycosylated serine and threonine residues
- The Sialidase mix efficiently removes sialic acids at α2-3, α2-6 and α2-8 bonds
- The O-glycosidase efficiently removes O-glycans on both native and denatured proteins
- New tools and workflows for improved O-glycoprotein analysis are demonstrated

RESULTS

Unique O-glycan Enzymes

A unique **O-protease** was discovered in *Akkermansia muciniphila*, a commensal bacterium in the human microbiota. The bacteria colonize distal ileum to rectum and degrade and metabolize highly O-glycosylated mucin. The O-protease was recombinantly expressed in *E. coli* and purified to homogeneity. The activity and specificity of the protease was evaluated by digestion of several protein substrates at native conditions. The O-protease only digests proteins carrying O-glycans. No activity was observed on non-glycosylated proteins or proteins carrying only N-glycan structures (*Fig. 1*). The O-protease displayed activity on sialylated proteins, but the activity was markedly increased on O-glycan proteins pre-treated with sialidases (*A. muciniphila*).

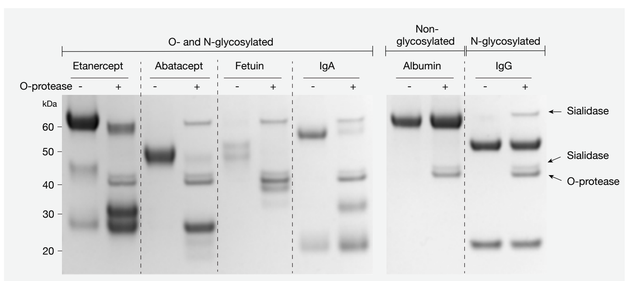


Figure 1. Specificity for O-glycosylated proteins. The O-protease digests only O-glycosylated proteins. No activity was observed on a non-glycosylated protein or on a protein carrying only N-glycans. The proteins were pre-treated with sialidases and incubated o/n with the O-protease.

To further verify the specificity of the O-protease, a protein with a single core 1 O-glycan, erythropoietin (EPO), was digested under native conditions. A specific digestion site N-terminally of the O-glycosylated serine[126] was defined (*Fig. 2*).

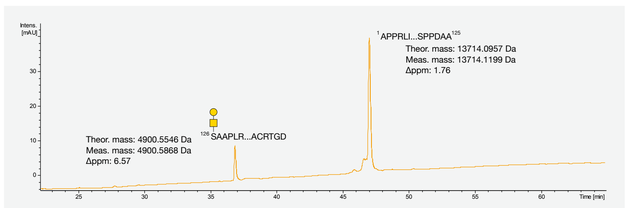


Figure 2. Specific digestion N-terminally of the O-glycosylation site. The reduced fragments were separated on a reversed phase C4 column followed by ESI-QTOF MS detection. The EPO protein carrying one core 1 O-glycan was hydrolyzed at a single specific site N-terminally of the O-glycosylated serine.

The **sialidases** from *Akkermansia muciniphila,* with high activity on α2-3, α2-6 and α2-8 bonds, were expressed in *E. coli*. Using a mix of sialidases (Sialidase 1 and 2), the activity on the different linkages of the sialic acid bonds was evaluated on synthetic substrates (*Fig. 3a*). The activity of the sialidases was further characterized on native glycoproteins by determining the relative amount of sialic acids released (*Fig. 3b*) and by a shift in molecular weight (*Fig. 3c*). The sialidases are active on both N- and O-linked glycans.

A new endo-**O-glycosidase** from *Streptococcus oralis*, with activity on native O-glycoproteins, was expressed in *E. coli*. The enzyme is active on asialylated core 1 and core 3 O-linked disaccharides (*Fig. 4a*). The combined activities of the enzymes, i.e. removal of sialic acids using the Sialidase mix (*A. muciniphila*) and subsequent hydrolysis of O-glycans using the O-glycosidase (*S. oralis*), was compared to that of a sialidase from *C. perfringens* and an O-glycosidase from *E. faecalis* (*Fig. 4b*). The *S. oralis* O-glycosidase removed the O-glycans more efficiently than the *E. faecalis* O-glycosidase on the highly glycosylated TNFR.

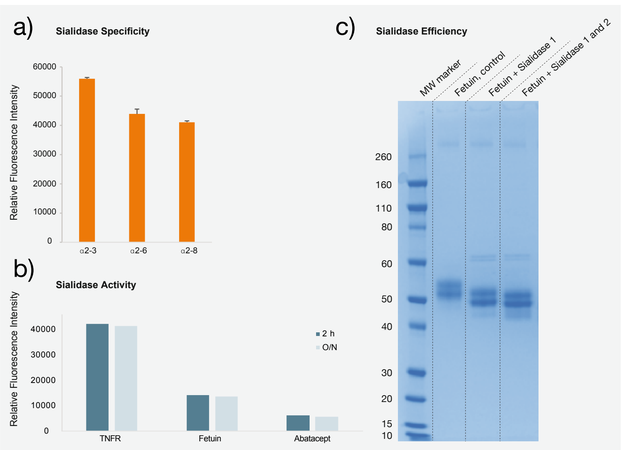


Figure 3. Specificity, activity and efficiency of the Sialidase mix. **a)** Activity of the Sialidase mix on sialic acid linkages of the substrates 3'-sialyllactose (α2-3 bonds), 6'-sialyllactose (α2-6 bonds), and colominic acid (α2-8 bonds). **b)** Activity on the glycosylated proteins TNFR, fetuin and abatacept at native conditions. Released free sialic acids in a) and b) were determined using the NANA kit (Abcam). **c)** Fetuin was treated with Sialidase 1 or Sialidase 1 and 2 (Sialidase mix) at native conditions. The asialylation was more efficient using the Sialidase mix.

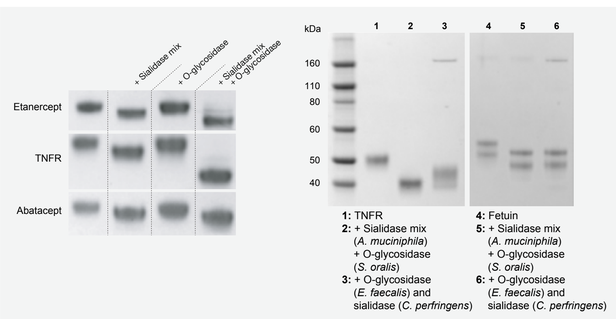


Figure 4a) Efficient removal of O-glycans. To investigate the endo-O-glycosidase (*S. oralis*) activity on native glycoproteins, the enzyme was incubated with etanercept, TNFR and abatacept for 1 h at 37°C with or without the Sialidase mix. Asialylation of the O-glycoproteins was necessary for hydrolysis of the O-glycans by the O-glycosidase.

Figure 4b) Improved performance for deglycosylation of O-glycans. Comparison of the combined enzymatic activities of *S. oralis* O-glycosidase and *A. muciniphila* Sialidases to O-glycosidases and sialidases of other bacterial origins. All incubations were performed for 4 h at 37 °C. The reactions were performed at native (Lane 2 and 5) and at denatured (Lane 3 and 6) conditions.

Applications of the O-protease in LC-MS

The unique specificity for O-glycosylated residues on glycoproteins opens for a variety of applications using this new O-protease in the glycoproteomic field. A schematic illustration of the enzymatic activity and a workflow example/suggestion including the Sialidase mix and the O-glycosidase is outlined in *Fig. 5*.

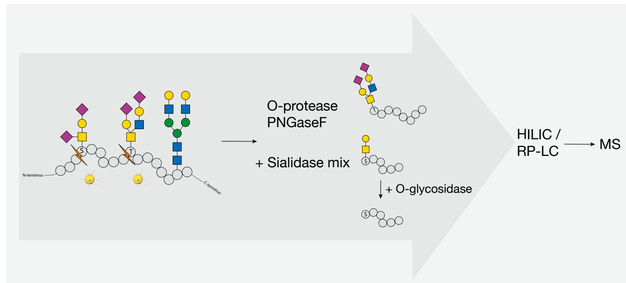
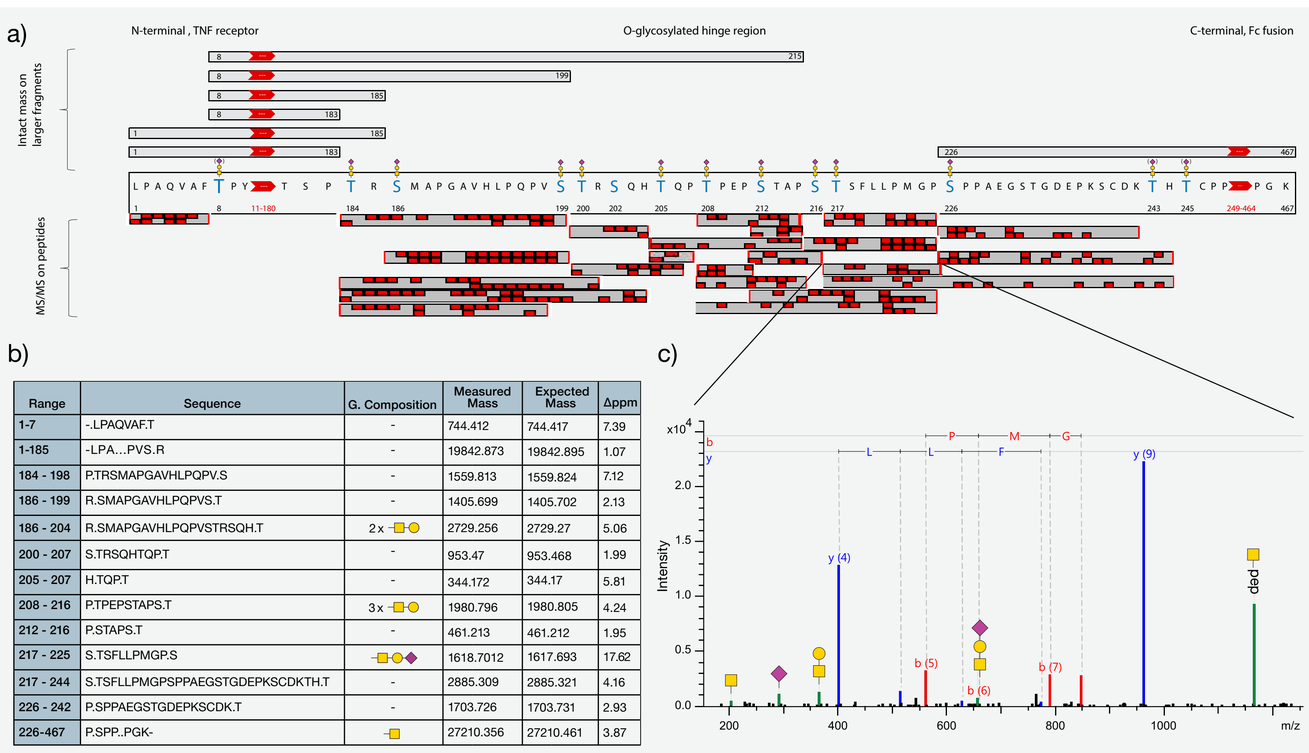


Figure 5. Workflow from intact glycoprotein to LC-MS and LC-MS/MS. In a native one pot o/n digestion reaction, the glycoprotein is hydrolyzed N-terminally to O-glycosylated sites using the O-protease. Addition of Sialidase mix is optional but improves digestion. In a second step, the remaining GalNAc/Gal can be removed from the asialylated glycopeptides using the O-glycosidase. After reduction, the samples are analyzed on RP or HILIC-LC ESI-QTOF MS and MS/MS. PNGaseF is added in the o/n digestion step to reduce heterogeneity.

The O-protease was used to map glycosylation sites of etanercept, an Fc-fusion protein with a highly core 1 O-glycosylated hinge region. The workflow illustrated in *Fig. 5* generated peptides with intact O-glycans, glycopeptides without sialic acids and peptides lacking sugar residues. The intact mass fragments and MS/MS peptides completes the amino acid sequence of etanercept and defines the O-glycosylated serine and threonine residues. A summary of the data is presented in *Fig. 6*. The RP-LC MS method was performed as in *Fig. 5*, HILIC separation was performed on a Waters Acquity BEH Amide column after buffer exchange on Graphite carbon.



Range	Sequence	G. Composition	Measured Mass	Expected Mass	Δppm
1-7	-.LPAQVAF.T	-	744.412	744.417	7.39
1-185	-LPA…PVS.R	-	19842.873	19842.895	1.07
184 - 198	P.TRSMAPGAVHLPQPV.S	-	1559.813	1559.824	7.12
186 - 199	R.SMAPGAVHLPQPVS.T	-	1405.699	1405.702	2.13
186 - 204	R.SMAPGAVHLPQPVSTRSQH.T	2 x	2729.256	2729.27	5.06
200 - 207	S.TRSQHTQP.T	-	953.47	953.468	1.99
205 - 207	H.TQP.T	-	344.172	344.17	5.81
208 - 216	P.TPEPSTAPS.T	3 x	1980.796	1980.805	4.24
212 - 216	P.STAPS.T	-	461.213	461.212	1.95
217 - 225	S.TSFLLPMGP.S		1618.7012	1617.693	17.62
217 - 244	S.TSFLLPMGPSPPAEGSTGDEPKSCDKTH.T	-	2885.309	2885.321	4.16
226 - 242	P.SPPAEGSTGDEPKSCDK.T	-	1703.726	1703.731	2.93
226-467	P.SPP..PGK-		27210.356	27210.461	3.87

Figure 6. O-protease digestion of etanercept. a) The O-glycosylated sites of etanercept are primarily located in the hinge region, therefore both larger C- and N-terminal fragments as well as smaller hinge region fragments were generated using the O-protease. The heterogeneity in O-glycosylation led to overlapping peptides. **b)** Selected mass data of peptides covering the entire amino acid sequence. **c)** Typical MS/MS spectrum with the GalNAc still attached to the peptide (in ESI-Q-TOF-MS, the glycan is often lost from the peptide before fragmentation, and found as oxonium ions).

Contact email: malin.mejare@genovis.com