GENOVIS PRESENTS

Site-specific O-glycan Analysis using a Novel O-glycan Protease and LC-MS/MS

AUTHORS | Maria Nordgren, Fredrik Leo, Rolf Lood, Stephan Björk and Fredrik Olsson | Genovis AB, Lund, Sweden

INTRODUCTION

Glycoproteins are found in almost all living organisms and 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 of the glycoprotein and close monitoring of the glycan profile is therefore required during development and manufacturing of biopharmaceuticals. The analyses of O-glycans have suffered from lack of specific enzymes and there has been a great need for novel tools.

An O-protease (an O-glycan-specific endoprotease) discovered in *Akkermansia muciniphila* and cloned in *E. coli*, is described and characterized in this work. The enzyme binds specifically to O-glycans on native glycoproteins and digests the peptide backbone N-terminally of serine and threonine residues. To demonstrate the workflow of the O-protease, we characterized the O-glycan sites of etanercept, an Fc-fusion protein with an O-glycosylated hinge region, using LC-MS/MS. Taken together, this work presents analytical workflows for site-specific O-glycan characterization of proteins using a new O-glycan-specific protease.

SUMMARY

- Aim: Characterize a novel O-protease with applications in glycoprotein analysis
- O-glycan-specific endoprotease - no activity on non- or N-glycosylated proteins
- Site-specific digestion N-terminally of O-glycosylated serine and threonine residues
- Increased performance on asialylated glycoproteins
- Recognizes core 1 and core 2 glycan structures and requires galactose residues for activity
- A unique new tool improving the workflow for O-glycoprotein analysis

RESULTS

Unique O-glycan-specific Protease

A unique endo 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 displayed activity on an asialylated protein, but the activity was markedly increased on O-glycan proteins pretreated with sialidases (*A. muciniphila*). It was found that the protease only digests proteins decorated with O-glycans. No activity was observed on non-glycosylated proteins nor on proteins with only N-glycan structures (*Fig. 1*).

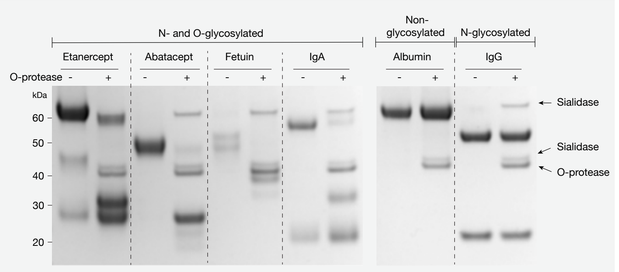

Figure 1. Specificity for O-glycosylated proteins. The O-protease digests only O-glycosylated proteins. No activity was observed on a non-glycosylated protein, or on a protein with only N-glycans. The proteins were pre-treated with sialidases and incubated o/n with the O-protease.

To address the dependency of the composition of O-glycans for O-protease activity, modifications of the glycosylation of a highly core 1-glycosylated protein were performed using different exoglycosidases before addition of the O-protease. Removal of the galactose residue resulted in a loss of O-protease activity (*Fig. 2*).

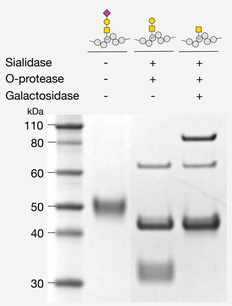

Figure 2. The O-protease requires galactose for activity. The impact of different sugar residues in the core 1 O-glycan was investigated by pre-treatment of TNFαR with exoglycosidases prior to digestion with the O-protease. The activity was lost upon removal of the galactose.

To further verify the specificity of the O-protease, a protein with a single core 1-glycan, erythropoietin (EPO), was digested at native conditions. A specific digestion site N-terminally of the O-glycosylated Serine[126] was defined (*Fig. 3*).

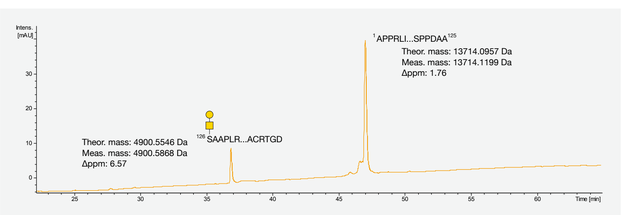

Figure 3. Specific digestion N-terminally of the O-glycosylation site. The reduced fragments were separated on a reversed phase C4 column followed by ESI-QTOF MS detection. The EPO protein carrying one core 1 O-glycan was hydrolyzed at a single specific site N-terminally of the O-glycosylated serine.

Human chorionic gonadotropin beta chain (hCGβ) is a glycoprotein with four O- and two N-glycans. At Serine[121] there is a constant core 2[1]. Digestion of this protein was done in native conditions with and without sialidase and the RP-LC-MS result confirms digestion also at core 2 (*Fig. 4*).

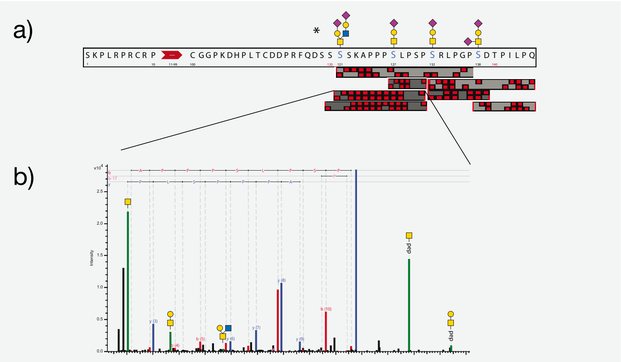

Figure 4. Digestion of human chorionic gonadotropin beta chain (hCGβ). a) The generated glycopeptides from hCGβ confirms glycosylation at S121, S127, S132, S138 and also indicate glycosylation at S120 and T140. RP separation was performed on an Advance Bio-Peptide map column from Agilent and samples were desalted in-line prior to ESI Q-TOF on a Bruker Impact II MS. *Glycan structures according to Valmu et al. (2006). **b)** MS/MS spectra of fragment S121-P131 confirms digestion at the core 2 structure by the presence of GalNAc/GlcNAc/Gal as an oxonium ion.

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 enzyme activity and a workflow example/suggestion is outlined in *Fig. 5*.

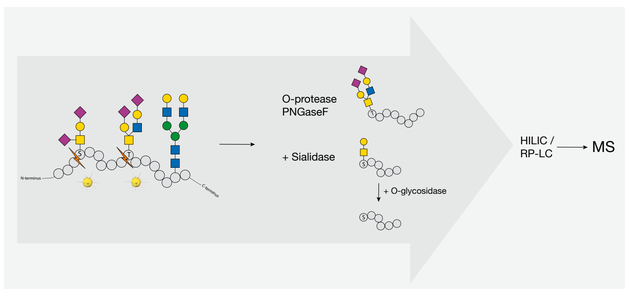

Figure 5. Workflow from intact glycoprotein to LC-MS and LC-MS/MS. In a native one pot o/n digestion reaction, the glycoprotein is hydrolyzed N-terminally to O-glycosylated sites using the O-protease. Addition of sialidase is optional but improves digestion. In a second step, the remaining GalNAc/Gal can be removed from the asialylated glycopeptides using an 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).

REFERENCES

1) Valmu L, Alfthan H, Hotakainen K, Birken S, Stenman UH. Site-specific glycan analysis of human chorionic gonadotropin beta-subunit from malignancies and pregnancy by liquid chromatography–electrospray mass spectrometry. Glycobiology. 2006;16:1207–18

Contact email: maria.nordgren@genovis.com