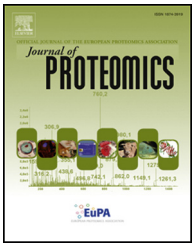
JOURNAL OF PROTEOMICS XX (2013) XXX–XXX

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot



Technical note

N-glycan occupancy of *Arabidopsis* N-glycoproteins ☆

Wei Song[a,b], *Remco A. Mentink*[c], *Maurice G.L. Henquet*[b], *Jan H.G. Cordewener*[b], *Aalt D.J. van Dijk*[b], *Dirk Bosch*[b], *Antoine H.P. America*[b], *Alexander R. van der Krol*[a,*]

[a]*Laboratory of Plant Physiology, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands*
[b]*Plant Research International, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands*
[c]*Hubrecht Intstitue, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands*

ARTICLE INFO

Article history:
Received 15 February 2013
Accepted 27 July 2013

Keywords:
N-glycosylation
N-glycoproteomics
Arabidopsis
LC–MS
MS/MS
Data independent acquisition (MS^E)

ABSTRACT

Most secreted proteins in eukaryotes are modified on the amino acid consensus sequence NxS/T by an N-glycan through the process of N-glycosylation. The N-glycans on glycoproteins are processed in the endoplasmic reticulum (ER) to different mannose-type N-glycans or, when the protein passes through the Golgi apparatus, to different complex glycan forms. Here we describe the capturing of N-glycopeptides from a trypsin digest of total protein extracts of *Arabidopsis* plants and release of these captured peptides following Peptide N-glycosidase (PNGase) treatment for analysis of N-glycan site-occupancy. The mixture of peptides released as a consequence of the PNGase treatment was analyzed by two dimensional nano-LC–MS. As the PNGase treatment of glycopeptides results in the deamidation of the asparagine (N) in the NxS/T site of the released peptide, this asparagine (N) to aspartic acid (D) conversion is used as a glycosylation 'signature'. The efficiency of PNGase F and PNGase A in peptide release is discussed. The identification of proteins with a single glycopeptide was limited by the used search algorithm but could be improved using a reference database including deamidated peptide sequences. Additional stringency settings were used for filtering results to minimize false discovery. This resulted in identification of 330 glycopeptides on 173 glycoproteins from *Arabidopsis*, of which 28 putative glycoproteins, that were previously not annotated as secreted protein in The *Arabidopsis* Information Resource database (TAIR). Furthermore, the identified glycosylation site occupancy helped to determine the correct topology for membrane proteins. A quantitative comparison of peptide signal was made between wild type and *complex-glycan-less* (*cgl*) *mutant Arabidopsis* from three replicate leaf samples using a label-free MS peak comparison. As an example, the identified membrane protein SKU5 (AT4G12420) showed differential glycopeptide intensity ratios between WT and *cgl* indicating heterogeneous glycan modification on single protein.

Abbreviations: ER, Endoplasmatic Reticulum; *cgl*, complex glycan less *Arabidopsis* mutant; OST, oligosaccharyltransferase; PNGase, peptide-N-glycosidase; PAGE, polyacrylamide gel electrophoresis; LC–MS/MS, liquid chromatography tandem mass spectrometry; UPLC, Ultra performance liquid chromatography; ACN, acetonitrile; DDA, data dependent acquisition; DIA (MS^E), data independent acquisition; Q-TOF, quadrupole time of flight; GO, Gene Ontology; TAIR, The *Arabidopsis* Information Resource database.

☆ This article is part of a Special Issue entitled: Translational Plant Proteomics.

* *Corresponding author*. Tel.: +31 317 482448; fax: +31 317 418094.
E-mail address: sander.vanderkrol@wur.nl (A.R. van der Krol).

1874-3919/$ – see front matter
http://dx.doi.org/10.1016/j.jprot.2013.07.032

Biological significance

Proteins that enter the secretory pathway are mostly modified by N-glycans. The function of N-glycosylation has been well studied in mammals. However, in plants the function of N-glycosylation is still unclear, because glycosylation mutants in plants often do not have a clear phenotype. Here we analyzed which proteins are modified by N-glycans in plants by developing a glycopeptide enrichment method for plant proteins. Subsequently, label free comparative proteomics was employed using protein fractions from wild type and from a mutant which is blocked in modification of the N-glycan into complex glycans. The results provide new information on N-glycosylation sites on numerous secreted proteins. Results allow for specific mapping of multiple glycosylation site occupancy on proteins, which provides information on which glycosylation sites are protected or non-used from downstream processing and thus presumably are buried into the protein structure. Glycoproteomics can therefore contribute to protein structure analysis. Indeed, mapping the glycosylation sites on membrane proteins gives information on the topology of protein folds over the membrane. We thus were able to correct the topology prediction of three membrane proteins. Besides, these studies also identified limitations in the software that is used to identify single modified peptide per protein.

This article is part of a Special Issue entitled: Translational Plant Proteomics.

1. Introduction

Glycosylation is one of the most common and essential post translational modifications of proteins in eukaryotic cells [1]. In eukaryotes proteins may be modified by N-glycans, O-glycans or a glycosyl phosphatidyl inositol (GPI) lipid anchor [2]. Here we are mainly concerned with N-glycosylation which takes place on proteins secreted into the ER. In eukaryotes N-glycans are involved in protein folding, protein folding quality control, polar protein localization and stability, ligand binding, endocytosis, immune recognition, inflammation and pathogenicity, cell signaling and cell motility [3–6]. The importance of N-glycan modifications in mammalian cells is exemplified by the strong congenital disorders resulting from mutations in N-glycan processing [5,7]. In contrast, in plants the structure of mannose type glycans in the ER seems to be non-essential as for example in the *alg3-1* mutant, where no clear phenotype can be observed [8]. Also the complex glycan less (*cgl*) mutant of *Arabidopsis*, which completely lacks complex N-glycans has no obvious growth phenotype and only shows some altered salt stress tolerance [9]. The lack of a clear phenotype from N-glycosylation mutants may be the reason that this process is less well studied in plants. However, techniques that can determine NxS/T N-glycan site occupancy of proteins and follow these under different (a)biotic stress conditions or through development may eventually allow for wider elucidation of their biological function in plants. Moreover, because N-glycoproteins are part of the secretome, they make an interesting sub-class of the proteome for characterization and targeted functionality studies in developmental and (a)biotic stress response studies. In order to study the role of N-glycosylation and N-glycan occupancy on plant glycoproteins we optimized the previously described procedure for isolation of glycoproteins from yeast or mammalian cells and we applied it to capture and analyze glycopeptides from plants [10].

1.1. *Complex N-glycan modifications in plants hinder glycoproteomics*

The core N-glycosylation pathway is evolutionary strongly conserved in all eukaryotes. Upon import into the ER, secreted proteins are glycosylated by the OST complex on the N-glycosylation consensus sequence site NxS/T, where x can be any amino acid except proline. N-glycosylation starts with transfer of a glycosylated high mannose N-glycan [11] to the asparagine side chain. In the ER the mannose glycans are slowly trimmed as part of a protein folding quality control mechanism [12]. This processing in the ER results in different high mannose type N-glycans for ER resident proteins, while glycoproteins that are transported to and through the Golgi apparatus are further modified by Golgi resident glycosyltransferases and glycosidases into complex type N-glycans [11]. In general, the distinction between mannose and complex type N-glycans on a glycoprotein may therefore give information on its sub-cellular location history. Plant specific complex N-glycans contain α-1,3 fucose attached to the glycan core and β-1,2 xylose attached to the central mannose residue [13]. The plant specific α-1,3 fucose core modification prevents cleavage by PNGase F, an enzyme broadly used for deglycosylating glycopeptides. Procedures developed for yeast or mammalian glycoproteomics involving the usage of PNGase F are therefore severely limited in plant glycoproteomics. However, for full glycoproteomics in *Arabidopsis* we can make a use of the glycosylation mutant *cgl* [14,15]. This mutant lacks a functional N-acetylglucosaminyltransferase 1 (GnT1) gene, which is essential for the formation of complex glycans. Consequently, the glycoproteins in *cgl* are all of the mannose type, which can be cleaved by PNGase F. Under normal growth conditions this *cgl* mutation has no visible effect on the growth phenotype of the *Arabidopsis* plants.

Please cite this article as: Song W, et al, N-glycan occupancy of *Arabidopsis* N-glycoproteins, J Prot (2013), http://dx.doi.org/10.1016/j.jprot.2013.07.032