

frontiers RESEARCH TOPICS

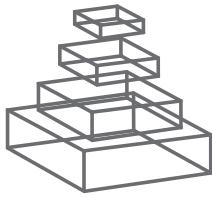
PLANT GLYCOBIOLOGY – A SWEET
WORLD OF LECTINS, GLYCOPROTEINS,
GLYCOLIPIDS AND GLYCANS

Topic Editors

Nausicaä Lannoo, Els J. M. Van Damme,
Cécile Albenne and Elisabeth Jamet



frontiers in
PLANT SCIENCE



FRONTIERS COPYRIGHT STATEMENT

© Copyright 2007-2015
Frontiers Media SA.
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

Cover image provided by Iblb sarl,
Lausanne CH

ISSN 1664-8714

ISBN 978-2-88919-421-6

DOI 10.3389/978-2-88919-421-6

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

PLANT GLYCOBIOLOGY – A SWEET WORLD OF LECTINS, GLYCOPROTEINS, GLYCOLIPIDS AND GLYCANS

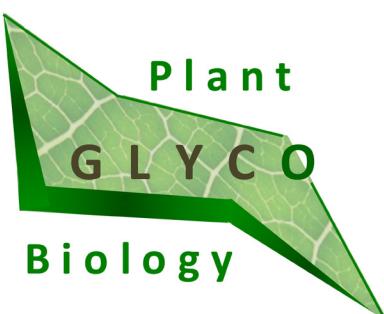
Topic Editors:

Nausicaä Lannoo, Ghent University, Belgium

Els J. M. Van Damme, Ghent University, Belgium

Cécile Albenne, Université Paul Sabatier-Toulouse 3/CNRS, Castanet-Tolosan, France

Elisabeth Jamet, Université Paul Sabatier-Toulouse 3/CNRS, Castanet-Tolosan, France



The figure was made by Nausicaä Lannoo (copyright).

Plants synthesize a wide variety of unique glycan structures which play essential roles during the life cycle of the plant. Being omnipresent throughout the plant kingdom, ranging from simple green algae to modern flowering plants, glycans contribute to many diverse processes. Glycans can function as structural components in the plant cell wall, assist in the folding of nascent proteins, act as signaling molecules in plant defense responses or (ER) stress pathways, or serve within the energy metabolism of a plant. In most cases, glycans are attached to other macromolecules to form so-called glycoconjugates (e.g. glycoproteins, proteoglycans and glycolipids),

but they can also be present as free entities residing in the plant cell. Next to the broad, complex set of glycans, plants also evolved an elaborate collection of lectins or proteins with a lectin-like domain, which can recognize and bind to endogenous (plants-own) or exogenous (foreign) glycans. Though still poorly understood in plants, the dynamic interactions between lectins and carbohydrate structures are suggested to be involved in gene transcription, protein folding, protein transport, cell adhesion, signaling as well as defense responses. As such, a complex and largely undetermined glycan-interactome is established inside plant cells, between cells and their surrounding matrix, inside the extracellular matrix, and even between organisms. Studying the biological roles of plant glycans will enable to better understand plant development and physiology in order to fully exploit plants for food, feed and production of pharmaceutical proteins.

In this Research Topic, we want to provide a platform for articles describing the latest research, perspectives and methodologies related to the fascinating world of plant glycobiology, with a focus on following subjects:

1. Identification and characterization of plant glycans, their biosynthetic and degradation enzymes
2. Characterization of plant lectins and glycoproteins
3. Plant glycans in the plant's energy metabolism
4. Role of plant glycans in plant defense signaling
5. Use of plant lectins in pest control
6. Plant lectins as new tools in human medicine
7. Glyco-engineering in plants

Table of Contents

- 06 Plant Glycobiology – A Diverse World of Lectins, Glycoproteins, Glycolipids and Glycans**
Nausicaä Lannoo, Els J. M. Van Damme, Cécile Albenne and Elisabeth Jamet
- 08 Cell Wall O-Glycoproteins and N-Glycoproteins: Aspect of Biosynthesis and Function**
Eric Nguema-Ona, Maïté Vicré-Gibouin, Maxime Gotté, Barbara Plancot, Patrice Lerouge, Muriel Bardor and Azeddine Driouch
- 20 Structural Features of Free N-Glycans Occurring in Plants and Functional Features of De-N-Glycosylation Enzymes, ENGase and PNGase: The Presence of Unusual Plant Complex Type N-Glycans**
Megumi Maeda and Yoshinobu Kimura
- 29 Biological Significance of Complex N-Glycans in Plants and their Impact on Plant Physiology**
Richard Strasser
- 35 Sugars take a Central Position in Plant Growth, Development and Stress Responses. A Focus on Apical Dominance**
Wim Van den Ende
- 38 Arabinogalactan Proteins: Focus on Carbohydrate Active Enzymes**
Eva Knoch, Adiphil Dilokpimol and Naomi Geshi
- 47 An Update on Post-Translational Modifications of Hydroxyproline-Rich Glycoproteins: Towards a Model Highlighting their Contribution to Plant Cell Wall Architecture**
May Hijazi, Silvia M. Velasquez, Elisabeth Jamet, Jose M. Estevez and Cécile Albenne
- 57 Loss of *Arabidopsis GAUT12/IRX8* Causes Anther Indehiscence and Leads to Reduced G Lignin Associated with Altered Matrix Polysaccharide Deposition**
Zhangying Hao, Utku Avci, Li Tan, Xiang Zhu, John Glushka, Sivakumar Pattathil, Stefan Eberhard, Tipton Scholes, Grace E. Rothstein, Wolfgang Lukowitz, Ron Orlando, Michael G. Hahn and Debra Mohnen
- 78 Lectin Domains at the Frontiers of Plant Defense**
Nausicaä Lannoo and Els J. M. Van Damme
- 94 Recombinant Production of Plant Lectins in Microbial Systems for Biomedical Application – the Frutalin Case Study**
Carla Oliveira, José A. Teixeira and Lucília Domingues

100 *Protein N-Glycosylation in Eukaryotic Microalgae and its Impact on the Production of Nuclear Expressed Biopharmaceuticals*

Elodie Mathieu-Rivet, Marie-Christine Kiefer-Meyer, Gaetan Vanier, Clément Ovide, Carole Burel, Patrice Lerouge and Muriel Bardor

113 *Glyco-Engineering for Biopharmaceutical Production in Moss Bioreactors*

Eva L. Decker, Juliana Parsons and Ralf Reski

120 *Trafficking of Endoplasmic Reticulum-Retained Recombinant Proteins is Unpredictable in *Arabidopsis Thaliana**

Thomas De Meyer and Ann Depicker

127 *Plant Glyco-Biotechnology on the Way to Synthetic Biology*

Andreas Loos and Herta Steinkellner



Plant Glycobiology—a diverse world of lectins, glycoproteins, glycolipids and glycans

Nausicaä Lannoo¹, Els J. M. Van Damme^{1*}, Cécile Albenne² and Elisabeth Jamet²

¹ Department of Molecular Biotechnology, Ghent University, Ghent, Belgium

² Laboratoire de Recherche en Sciences Végétales, UMR 5546, Université Paul Sabatier-Toulouse 3/CNRS, Castanet-Tolosan, France

*Correspondence: elsjm.vandamme@ugent.be

Edited and Reviewed by:

Andreas P. M. Weber, Heinrich-Heine-Universität, Germany

Keywords: arabinogalactan proteins, cell wall, hydroxyproline-rich glycoproteins, glycans and glycoconjugates, glycoengineering, lectins, protein-carbohydrate interactions, sugar signaling

Glycosylation is essential for the growth, development or survival of every organism (Varki and Lowe, 2009). Defects in glycan signaling often lead to abnormal development and severe diseases. Glycosylation is ubiquitous and the tremendous structural complexity of glycans makes it quite impossible to predict the biological importance of individual structures. Nowadays, glycans are no longer regarded solely as an energy reservoir, but are associated with storage and transfer of biological information as part of a highly complicated multidimensional coding system (Rüdiger and Gabius, 2009; Gabius et al., 2011; Solís et al., 2014). Plants synthesize a wide variety of unique glycan structures and glycan-binding proteins which play pivotal roles during their life cycle. The increasing number of excellent publications, both in primary and applied plant glycobiology research, demonstrates the great promise and importance of this area for current and future plant science. With 13 original contributions, this Research Topic is a nice compilation of Mini Reviews and Reviews, an Original research paper, and an Opinion Article, highlighting important aspects of plant glycobiology.

In plant glycobiology, *N*-glycans constitute core structures which are grafted on polypeptide backbones. Complex *N*-glycans are ubiquitously present in plants (Wilson et al., 2001), yet their biological function is virtually unknown. Nguema-Ona et al. (2014) provide an overview of the biosynthesis of *N*-glycans. Maeda and Kimura nicely review the group of free *N*-glycans that are released from misfolded proteins or originate from fully processed and secreted proteins by the action of the *N*-glycan releasing enzymes ENGase and PNGase. They discuss the impact of these plant complex *N*-glycans in terms of plant development and fruit ripening (Maeda and Kimura, 2014). The paper from Strasser continues this discussion and focuses on recent developments with respect to *N*-glycan signaling in transgenic *A. thaliana* and rice plants with disabled *N*-glycan processing, which ultimately could lead to the development of some new glyco-engineering tools (Strasser, 2014). Next to *N*-glycans, photosynthesis-derived small sugars such as sucrose, fructose, glucose, trehalose, and derived oligosaccharides, which are generally accepted to be involved in plant energy metabolism and plant growth, have very recently been suggested to act as signal molecules in important plant developmental programs (Ruan, 2014; Smeekens and

Hellmann, 2014). In his Opinion Article, Van den Ende (2014) focuses on this intimate communication between plant hormones and small sugars, better-known as the sugar sensing mechanism, and the putative role of small sugars in apical dominance.

Plant cell walls are formed of complex interlaced networks of polysaccharides (cellulose, hemicellulose and pectins) and hydroxyproline-rich O-glycoproteins (HRGPs) which are considered as structural proteins (Carpita and Gibeaut, 1993). However, the way these macromolecules are arranged in supramolecular scaffolds is still poorly understood. Knoch et al. (2014) focus on the recent discoveries of carbohydrate active enzymes (CAZy) (Lombard et al., 2014) that are involved in the synthesis as well as in the degradation of arabinogalactan proteins (AGPs), i.e., a highly diverse class of cell surface HRGPs found in most plant species. They discuss the role of these enzymes in plant development. Nguema-Ona et al. (2014) and Hijazi et al. (2014) broaden this discussion and present an overview of the enzymes not only involved in the synthesis of AGPs, but also of extensins, another type of HRGPs, and discuss the importance of both AGPs and extensins for proper cell wall development and morphology as well as their role in biotic stress responses. Hijazi et al. (2014) propose a new model to explain how all types of HRGPs could contribute to a continuous glyco-network with their respective partners including polysaccharides to form a complex architecture in plant cell walls. In the case of secondary cell walls, lignin, and different types of hemicelluloses are found. Hao et al. (2014) present an Original Research paper in which they identified a galacturonosyltransferase (GAUT12) from *A. thaliana* as a new glycosyltransferase possibly contributing to the synthesis of a polysaccharidic structure including pectins allowing the deposition of xylan and lignin.

Plant cell walls not only have a structural function, but also play a critical role in the perception of invading pathogens and the activation of specific plant defense responses, as discussed by Lannoo and Van Damme (2014). This review elaborates how plants can recognize plant pathogens or predators upon perception of characteristic epitopes or damage-associated patterns, using protein-protein interactions as well as protein-glycan interactions mediated by lectins. In addition, they highlight that protein-glycan interactions mediated by different types of

nucleocytoplasmic lectins are part of signaling pathways implicated in plant defense responses. Plant lectins not only attracted a lot of attention due to their phytoprotective properties, they are also of interest for medical applications and use in biomedical diagnosis. They can be purified from natural resources, but with the increasing demand for biopharmaceuticals, different expression platforms are being exploited for their recombinant production. Oliveira et al. (2014) describe how they can produce recombinant frutalin, a lectin from *Artocarpus incisa* (breadfruit) which possesses immuno-modulatory, anti-tumor, and tumor biomarker properties, in distinct microbial systems. Since the presence and quality of glycosylation plays a crucial role for the pharmacological properties of the therapeutic protein, also plants have received growing attention for molecular farming. In this Research Topic, several papers review the humanization of the plant glycosylation pathway allowing the production of human proteins with optimized glycosylation profiles in eukaryotic microalgae (Mathieu-Rivet et al., 2014), lower plants (mosses) (Decker et al., 2014) and in higher plants (De Meyer and Depicker, 2014; Loos and Steinkellner, 2014).

The major aim of this Research Topic was to provide the reader an overview of the latest progress in plant glycobiology research. All contributions demonstrate recent and exciting breakthroughs and present the intrinsic capacity of this particular scientific research area for further improvement of plant biotechnology. We hope that this e-book can provide useful information to readers and stimulate future research in the dynamic plant glycobiology community.

ACKNOWLEDGMENTS

We would like to thank all the authors for their excellent contributions to this e-book as well as all the reviewers for their timely efforts in judging the manuscripts and for giving their valuable constructive comments. This work was supported by the Fund for Scientific Research—Flanders and Ghent University (N. Lannoo and EJM Van Damme) and by Université Paul Sabatier-Toulouse, CNRS and TULIP Laboratoire d'Excellence (ANR-10-LABX-41; ANR-11-IDEX-0002-02) (C. Albenne and E. Jamet).

REFERENCES

- Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 1, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x
- Decker, E. L., Parsons, J., and Reski, R. (2014). Glyco-engineering for pharmaceutical production in moss bioreactors. *Front. Plant Sci.* 5:346. doi: 10.3389/fpls.2014.00346
- De Meyer, T., and Depicker, A. (2014). Trafficking of endoplasmic reticulum-retained recombinant proteins is unpredictable in *Arabidopsis thaliana*. *Front. Plant Sci.* 5:473. doi: 10.3389/fpls.2014.00473
- Gabius, H. J., André, S., Jiménez-Barbero, J., Romero, A., and Solís, D. (2011). From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.* 36, 298–313. doi: 10.1016/j.tibs.2011.01.005
- Hao, Z., Avci, U., Tan, L., Zhu, X., Glushka, J., Pattathil, S., et al. (2014). Loss of *Arabidopsis* GAUT12/IRX8 causes anther indehiscence and leads to reduced G lignin associated with altered matrix polysaccharide deposition. *Front. Plant Sci.* 5:357. doi: 10.3389/fpls.2014.00357
- Hijazi, M., Velasquez, S. M., Jamet, E., Estevez, J. M., and Albenne, C. (2014). An update on post-translational modifications of hydroxyproline-rich glycoproteins: toward a model highlighting their contribution to plant cell wall architecture. *Front. Plant Sci.* 5:395. doi: 10.3389/fpls.2014.00395
- Knoch, E., Dilokpimol, A., and Geshi, N. (2014). Arabinogalactan proteins: focus on carbohydrate active enzymes. *Front. Plant Sci.* 5:198. doi: 10.3389/fpls.2014.00198
- Lannoo, N., and Van Damme, E. J. M. (2014). Lectin domains at the frontiers of plant defense. *Front. Plant Sci.* 5:397. doi: 10.3389/fpls.2014.00397
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Loos, A., and Steinkellner, H. (2014). Plant glyco-biotechnology on the way to synthetic biology. *Front. Plant Sci.* 5:523 doi: 10.3389/fpls.2014.00523
- Maeda, M., and Kimura, Y. (2014). Structural features of free N-glycans occurring in plants and functional features of de-N-glycosylation enzymes, ENGase, and PNGase: the presence of unusual plant complex type N-glycans. *Front. Plant Sci.* 5:429. doi: 10.3389/fpls.2014.00429
- Mathieu-Rivet, E., Kiefer-Meyer, M.-C., Vanier, G., Ovide, C., Burel, C., Lerouge, P., et al. (2014). Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Front. Plant Sci.* 5:359. doi: 10.3389/fpls.2014.00359
- Nguema-Ona, E., Vicré-Gibouin, M., Gotté, M., Plancot, B., Lerouge, P., Bardor, M., et al. (2014). Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. *Front. Plant Sci.* 5:499. doi: 10.3389/fpls.2014.00499
- Oliveira, C., Teixeira, J. A., and Domingues, L. (2014). Recombinant production of plant lectins in microbial systems for biomedical application – the frutalin case study. *Front. Plant Sci.* 5:390. doi: 10.3389/fpls.2014.00390
- Ruan, Y.-L. (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu. Rev. Plant Biol.* 65, 26.1–26.35. doi: 10.1146/annurev-aplant-050213-040251
- Rüdiger, H., and Gabius, H.-J. (2009). “The biochemical basis and coding capacity of the sugar code,” in *The Sugar Code. Fundamentals of Glycosciences*. Gabius, ed. Hans-Joachim Gabius (Weinheim, Germany: Wiley-VCH), 3–13.
- Smeekens, J., and Hellmann, H. A. (2014). Sugar sensing and signaling in plants. *Front. Plant Sci.* 5:113. doi: 10.3389/fpls.2014.00113
- Solís, D., Bovin, N. V., Davis, A. P., Jiménez-Barbero, J., Romero, A., Roy, R., et al. (2014). A guide into glycosciences: how chemistry, biochemistry and biology cooperate to crack the sugar code. *Biochim. Biophys. Acta.* doi: 10.1016/j.bbagen.2014.03.016. [Epub ahead of print].
- Strasser, R. (2014). Biological significance of complex N-glycans in plants and their impact on plant physiology. *Front. Plant Sci.* 5:363. doi: 10.3389/fpls.2014.00363
- Van den Ende, W. (2014). Sugars take a central position in plant growth, development and stress responses. A focus on apical dominance. *Front. Plant Sci.* 5:313. doi: 10.3389/fpls.2014.00313
- Varki, A., and Lowe, J. B. (2009). “Biological roles of glycans,” in *Essentials of Glycobiology*, 2nd Edn., Cold Spring Harbor, eds A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, M. E. Etzler (New York, NY: Cold Spring Harbor Laboratory Press), Chapter 6.
- Wilson, I. B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core α1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 October 2014; accepted: 16 October 2014; published online: 04 November 2014.

*Citation: Lannoo N, Van Damme EJM, Albenne C and Jamet E (2014) Plant Glycobiology—a diverse world of lectins, glycoproteins, glycolipids and glycans. *Front. Plant Sci.* 5:604. doi: 10.3389/fpls.2014.00604*

This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Lannoo, Van Damme, Albenne and Jamet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function

Eric Nguema-Ona^{1*}, Maïté Vicré-Gibouin¹, Maxime Gotté¹, Barbara Plancot¹, Patrice Lerouge¹, Muriel Bardor^{1,2†} and Azeddine Driouich^{1,3†}

¹ Laboratoire de Glycobiologie et Matrice Extracellulaire Végétale, UPRES EA 4358, Institut de Recherche et d'Innovation Biomédicale, Grand Réseau de Recherche-Végétal, Agronomie, Sol, Innovation, UFR des Sciences et Techniques, Normandie Université – Université de Rouen, Mont-Saint-Aignan, France

² Institut Universitaire de France, Paris, France

³ Plate-Forme de Recherche en Imagerie Cellulaire de Haute-Normandie, Institut de Recherche et d'Innovation Biomédicale, Faculté des Sciences et Techniques, Normandie Université, Mont-Saint-Aignan, France

Edited by:

Elisabeth Jamet, Laboratoire de Recherche en Sciences Végétales, France

Reviewed by:

Ian S. Wallace, University of Nevada, Reno, USA

Richard Strasser, University of Natural Resources and Life Sciences, Austria

***Correspondence:**

Eric Nguema-Ona, Laboratoire de Glycobiologie et Matrice Extracellulaire Végétale, UPRES EA 4358, Institut de Recherche et d'Innovation Biomédicale, Grand Réseau de Recherche-Végétal, Agronomie, Sol, Innovation, UFR des Sciences et Techniques, Normandie Université – Université de Rouen, Bâtiment Gadeau de Kerville, 1 rue Tesnière, 76821 Mont-Saint-Aignan, France
e-mail: eric.nguema-on@univ-rouen.fr

[†]Muriel Bardor and Azeddine Driouich have contributed equally to this work.

Cell wall O-glycoproteins and N-glycoproteins are two types of glycomolecules whose glycans are structurally complex. They are both assembled and modified within the endomembrane system, i.e., the endoplasmic reticulum (ER) and the Golgi apparatus, before their transport to their final locations within or outside the cell. In contrast to extensins (EXTs), the O-glycan chains of arabinogalactan proteins (AGPs) are highly heterogeneous consisting mostly of (i) a short oligo-arabinoside chain of three to four residues, and (ii) a larger β -1,3-linked galactan backbone with β -1,6-linked side chains containing galactose, arabinose and, often, fucose, rhamnose, or glucuronic acid. The fine structure of arabinogalactan chains varies between, and within plant species, and is important for the functional activities of the glycoproteins. With regards to N-glycans, ER-synthesizing events are highly conserved in all eukaryotes studied so far since they are essential for efficient protein folding. In contrast, evolutionary adaptation of N-glycan processing in the Golgi apparatus has given rise to a variety of organism-specific complex structures. Therefore, plant complex-type N-glycans contain specific glyco-epitopes such as core β ,2-xylose, core α 1,3-fucose residues, and Lewis^a substitutions on the terminal position of the antenna. Like O-glycans, N-glycans of proteins are essential for their stability and function. Mutants affected in the glycan metabolic pathways have provided valuable information on the role of N/O-glycoproteins in the control of growth, morphogenesis and adaptation to biotic and abiotic stresses. With regards to O-glycoproteins, only EXTs and AGPs are considered herein. The biosynthesis of these glycoproteins and functional aspects are presented and discussed in this review.

Keywords: arabinogalactan protein, cell wall, endoplasmic reticulum, extensin, glycan, glycosyltransferase, Golgi apparatus, plants

INTRODUCTION

Plants synthesize glycoconjugates that are structurally diverse and complex reflecting the diversity of plant physiological functions. The glycomolecules are usually assembled and modified within the plant endomembrane system, including the endoplasmic reticulum (ER), the Golgi apparatus and secretory vesicles responsible for their transport to different cell compartments/organelles including the cell wall. Their synthesis involves a number of steps, beginning with the formation of activated nucleotide sugars such as NDP-sugars or NMP-sugars (Bar-Peled and O'Neill, 2011). After their synthesis in the cytosol, the nucleotide sugars are then actively transported into the ER and Golgi stacks where they serve as donor substrates during glycan synthesis. Glycosyltransferases (GTs) transfer specific sugars from activated nucleotide sugars to a specific glycan acceptor leading to the extension of the glycomolecule involved. This occurs through a stepwise and sequential process which involves a number of different GTs of the secretory system. It is worth noting that 1.8% of *Arabidopsis thaliana*'s genome currently encode GT

genes representing more than 462 GTs in total (Ulvskov et al., 2013).

Among the different plant organelles, the plant cell wall is a polysaccharide-rich extracellular compartment (Albersheim et al., 2011). In addition to polysaccharides, the plant cell wall also contains a significant percentage (~10–15%) of N- and O-glycosylated proteins that are relatively less studied with regards to their biosynthesis and function. Both the N- and the O-glycosylation of proteins has a significant impact on both their structural properties and biological activities (Varki, 1993). Glycosylation and glycan processing are major *post-translational modifications* (PTMs) that cell wall proteins undergo inside the cell, and are considered important for their proper function. Indeed, in general, glycans are involved in the control of protein folding, cellular targeting and mobility, as well as signaling for regulation of plant growth, defense and different interactions with the surrounding environment (Varki and Lowe, 2009; Larkin and Imperiali, 2011; Cannesan et al., 2012; Nguema-Ona et al., 2013; Chen et al., 2014).

The *N*-/*O*- glycosylation of cell wall proteins is critical for plant development and responses to stress. Understanding and controlling *O*- and *N*-glycosylation of secreted proteins is also important in plant biotechnological applications.

***N*-GLYCOSYLATED PROTEINS: SYNTHESIS AND FUNCTION**

The *N*-glycosylation of proteins starts in the ER. ER-synthesizing events for *N*-glycans are highly conserved in all eukaryotes studied so far since they are instrumental for efficient protein folding (Aebi, 2013). The *N*-glycosylation pathway starts by the transfer *en bloc* of a lipid linked preassembled precursor (Glc₃Man₉GlcNAc₂) by the oligosaccharyltransferase (OST) onto the *N*-glycosylation sites (Asn-X-Ser/Thr and/or Asn-X-Cys) of the nascent proteins (Burda and Aebi, 1999; Gil et al., 2009; Zielinska et al., 2010; Matsui et al., 2011). The α -glucosidases I and II then remove two

glucose residues from the *N*-glycan resulting in the presence of only one terminal glucose on the glycoprotein. This allows its entry into the ER control quality cycle (Aebi, 2013). Once the glycoprotein is correctly folded, the last glucose residue is removed by the α -glucosidase II prior to its transport into the Golgi apparatus where further modifications occur including removal of mannose residues and sequential addition of specific sugars through the action of GTs resulting in the formation of complex-type *N*-glycans. In plants, many genes encoding for Golgi GTs have already been identified (Table 1). These include, for example, *N*-acetylglucosaminyltransferase I (GnT I; Bakker et al., 1999; Strasser et al., 1999a; Wenderoth and von Schaewen, 2000), *N*-acetylglucosaminyltransferase II (GnT II; Strasser et al., 1999b), core α -1,3-fucosyltransferase (α 1,3-FuT; Leiter et al., 1999; Wilson et al., 2001a), β -1,2-xylosyltransferase (β 1,2-XylT; Strasser

Table 1 | Known enzymes involved in plant *N*- glycans and *O*- cell wall glycan biosynthesis.

AGP glycan biosynthetic enzymes	CAZy family	Protein name	Origin	Reference
Hydroxyproline	GT31	AtGALT2	<i>Arabidopsis thaliana</i>	Basu et al. (2013)
<i>O</i> -galactosyltransferase				
β -1,3-galactosyltransferase	GT31	At1g77810	<i>Arabidopsis thaliana</i>	Qu et al. (2008)
β -1,6-galactosyltransferase	GT31	AtGalT31A	<i>Arabidopsis thaliana</i>	Geshi et al. (2013)
–	GT29	AtGalT29A	<i>Arabidopsis thaliana</i>	Dilokpimol et al. (2014)
Arabinofuranosyltransferase	GT77	RAY1	<i>Arabidopsis thaliana</i>	Gille et al. (2013)
β -glucuronosyltransferase	GT14	AtGlcAT14A	<i>Arabidopsis thaliana</i>	Knoch et al. (2013)
α -1,2-fucosyltransferase	GT37	AtFUT4	<i>Arabidopsis thaliana</i>	Wu et al. (2010)
–	GT37	AtFUT6	<i>Arabidopsis thaliana</i>	Wu et al. (2010)
Extensin glycan biosynthetic enzymes	CAZy family	Protein name	Origin	Reference
Serine <i>O</i> -galactosyltransferase	unknown	SGT1	<i>Chlamydomonas reinhardtii</i> ; <i>Arabidopsis thaliana</i>	Saito et al. (2014)
Arabinosyltransferase	GT77	RRA3	<i>Arabidopsis thaliana</i>	Velasquez et al. (2011)
–	GT77	XEG113	<i>Arabidopsis thaliana</i>	Gille et al. (2009)
<i>N</i>-glycan biosynthetic enzymes	CAZy family	Protein name	Origin	Reference
Oligosaccharyltransferase		OST	<i>Arabidopsis thaliana</i> ; <i>Oryza sativa</i>	Farid et al. (2013), Qin et al. (2013)
α -glucosidase I		GCS I	<i>Arabidopsis thaliana</i>	Boisson et al. (2001)
α -glucosidase II		GCS II	<i>Solanum tuberosum</i>	Taylor et al. (2000)
α -mannosidase I		MNS 1-3	<i>Arabidopsis thaliana</i>	Liebminger et al. (2009)
<i>N</i> -acetylglucosaminyltransferase I	GT13	GnT I	<i>Arabidopsis thaliana</i> ; <i>Nicotiana tabacum</i> ; <i>Solanum tuberosum</i>	Bakker et al. (1999), Strasser et al. (1999a), Wenderoth and von Schaewen (2000)
α -mannosidase II		GM II	<i>Arabidopsis thaliana</i>	Strasser et al. (2006)
<i>N</i> -acetylglucosaminyltransferase II	GT16	GnT II	<i>Arabidopsis thaliana</i>	Strasser et al. (1999b)
α -1,3 fucosyltransferase	GT10	α -1,3-FuT	<i>Vigna radiata</i> ; <i>Arabidopsis thaliana</i> ; <i>Medicago sativa</i>	Leiter et al. (1999), Wilson et al. (2001a), Sourrouille et al. (2008)
β -1,2-xylosyltransferase	GT61	β -1,2-XylT	<i>Arabidopsis thaliana</i>	Strasser et al. (2000), Pagny et al. (2003), Bencur et al. (2005)
β -1,3-galactosyltransferase	GT31	β -1,3-GalT	<i>Arabidopsis thaliana</i>	Strasser et al. (2007)

et al., 2000; Pagny et al., 2003), Lewis-type α -1,4-fucosyltransferase (α 1,4-FuT Bakker et al., 2001; Wilson et al., 2001b; Léonard et al., 2002), and β -1,3-galactosyltransferase (β 1,3-GalT; Strasser et al., 2007). In contrast to the ER steps, evolutionary adaptation of *N*-glycan processing in the Golgi apparatus has given rise to a

variety of organism-specific complex structures (Varki, 2011). Therefore, more complex plant *N*-glycans consist of specific glyco-epitopes such as core β -1, 2-xylose, core α -1,3-fucose residues, and Lewis^a substitutions on the terminal position of the antenna (Figure 1A; Lerouge et al., 1998; Bardor et al., 2003, 2011; Strasser

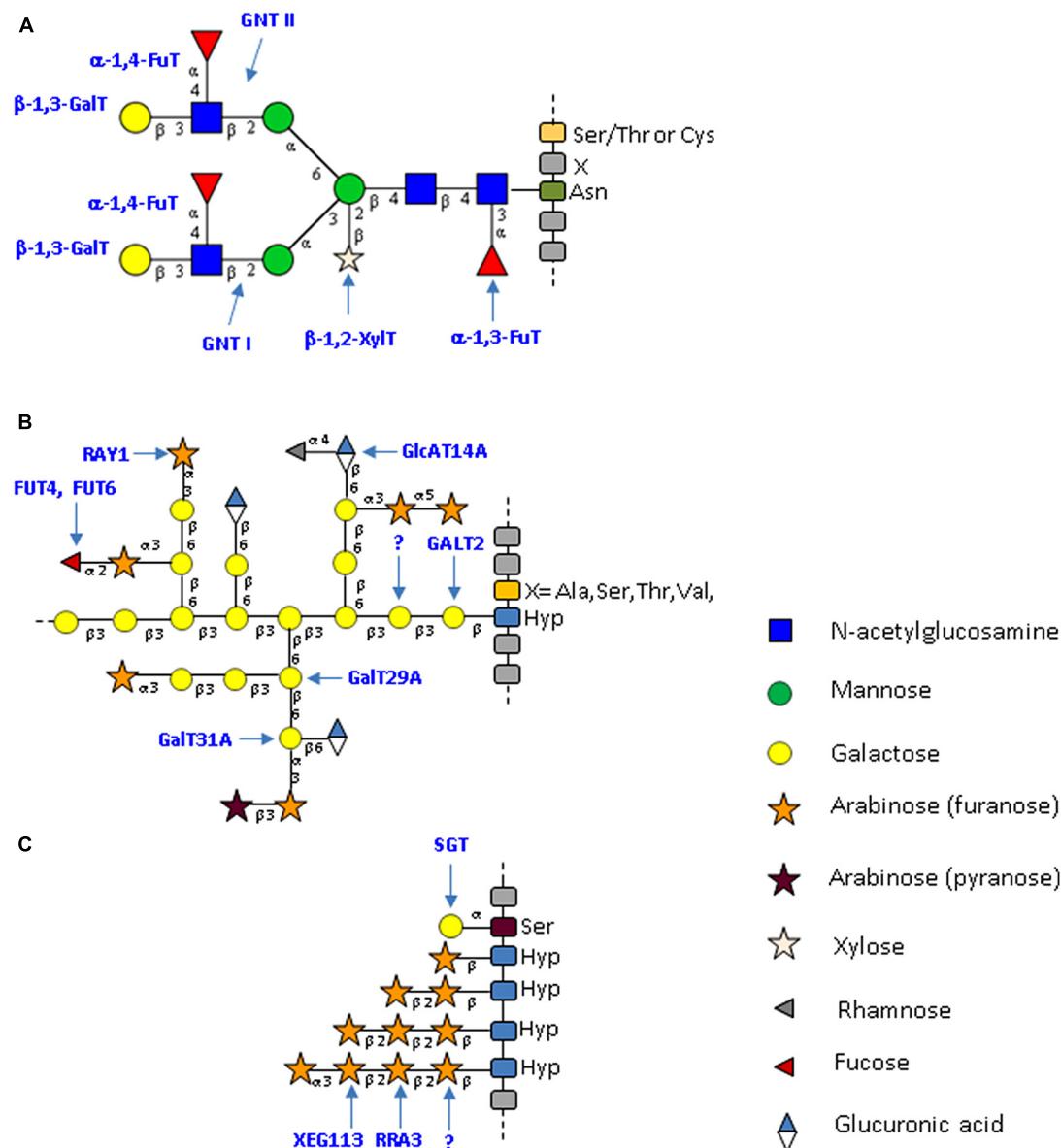


FIGURE 1 | Typical structure of plant *N*- and *O*-glycans from cell wall proteins. (A) Specific complex-type *N*-glycans attached to plant glycoproteins. This *N*-glycan results from the action of a plant-specific repertoire of glycosyltransferases that lead to the formation of a glycan bearing plant-specific glyco-epitopes such as a core β -1,2-xylose; a core α -1,3-fucose and a Lewis^a antennae (Lerouge et al., 1998; Wilson et al., 2001b; Bardor et al., 2003, 2011). The *N*-glycan structures presented here are drawn according to the symbolic nomenclature adopted by the Consortium for Functional Glycomics (Varki et al., 2009). (B) Schematic representation of *O*-glycans (type II arabinogalactan) attached to AGCs. These glycans predominantly consist of arabinose and galactose. Minor

sugars, such as glucuronic acid, fucose or rhamnose, are also present. The *O*-glycans are attached to non-contiguous Hyp residues. The model presented is modified from Tan et al. (2010) and Tryfona et al. (2012). (C) Schematic representation of *O*-glycans attached to plant EXT. These glycans consist of short chains of arabinose and one galactose residues. The *O*-glycans are attached to contiguous Hyp residues. The model presented is modified from Saito et al. (2014). Yellow circle: galactose; green circle: mannose; blue square: N-acetylglucosamine; star: white xylose and red triangle: fucose; gray triangle: rhamnose; orange star: arabinose (furanose); purple star: arabinose (pyranose); blue/white diamond: glucuronic acid.

et al., 2004). When abnormally processed, N-glycosylated proteins cause major developmental disorders and are usually associated to diseases in mammals (Ioffe and Stanley, 1994; Metzler et al., 1994; Lowe and Marth, 2003; Hennet, 2012). In plants, abnormal N-glycosylated proteins rarely present developmental disorders under normal growth conditions (von Schaewen et al., 1993; Strasser et al., 2004). However, the cellulose-deficient *Arabidopsis* mutant *rsw3* which is defective in the catalytic subunit of the α -glucosidase II presents radially swollen roots and a deficiency in cellulose content (Burn et al., 2002). Moreover, under stress conditions (e.g., salt), modified phenotypes such as abnormal plant growth (Strasser et al., 2007) or altered root growth in the *Arabidopsis* *cgl* mutants (Kang et al., 2008; von Schaewen et al., 2008), have been observed. Indeed, in these studies, reduced root growth and abnormal root morphology were observed for *Arabidopsis* plants cultivated on media containing high NaCl concentration. In contrast to *Arabidopsis*, a severe phenotype with arrested seedling development and premature death before reaching the reproductive stage has been reported recently for rice *gntI* mutant (Fanata et al., 2013). Such plants also present defects in cell wall composition, especially reduced cell wall thickness, and decreased in cellulose content as well as reduced sensitivity to cytokinin. Plant complex-type N-glycans are ascribed to many biological functions in relation with plant development that have been recently reviewed by Strasser (2014; this issue). These include effects on plant innate immunity, tolerance to abiotic stress and root development. Therefore these functional aspects will not be further described in this review.

0-GLYCOSYLATED CELL WALL PROTEINS, ARABINOGALACTAN PROTEINS, AND EXTENSINS

Plant O-glycosylated cell wall proteins belong to the superfamily of hydroxyproline-rich glycoproteins (HRGPs). This superfamily of plant cell wall proteins which account for nearly 10% of the dry weight of the wall, is characterized by a high proline (Pro) content. Furthermore many of these Pro residues become hydroxylated (hydroxyproline, Hyp) during synthesis and consequently become glycosylated in various ways. Pro residues are distributed at different sites within the sequence and these patterns have suggested different classifications of HRGP members into different groups. EXTs and arabinogalactan proteins (AGPs) are two O-glycosylated HRGP subfamilies which have gained much attention (Kieliszewski and Shpak, 2001; Showalter, 2001; Schultz et al., 2002; Showalter et al., 2010; Kieliszewski et al., 2011; Lamport et al., 2011; Nguema-Ona et al., 2012, 2013; Tan et al., 2012; Velasquez et al., 2012). The nature of sugars being incorporated and the level of glycosylation vary between these two families, but also within the members of these subfamilies. For example, Kieliszewski et al. (2011) have shown that occurrence of contigs of 3–5 Hyp, preceded by a serine residue (Ser-Hyp₄) led to the synthesis of a short arabinoside of 3–5 residues. Serine residue in the Ser-Hyp₄ contig is often O-glycosylated with a single galactose (Velasquez et al., 2012; Saito et al., 2014). This action is performed by serine-O-galactosyltransferases (Ser-O-Gal-T), specific to plants (Saito et al., 2014). Non-contiguous Hyp residues rather lead to the synthesis of a large arabino-galactosylated glyco-epitope on the

protein (Kieliszewski and Lamport, 1994; Shpak et al., 1999; Kieliszewski et al., 2011).

Arabinogalactan proteins and EXTs have been studied for decades, and shown to fulfill many functions related to development, and responses to biotic and abiotic stresses in plants (Hall and Cannon, 2002; Motose et al., 2004; Lee et al., 2005; Nguema-Ona et al., 2007, 2013; Seifert and Roberts, 2007; Cannon et al., 2008; Ellis et al., 2010; Lamport et al., 2011; Velasquez et al., 2011; Cannesan et al., 2012; Moore et al., 2014a,b). These studies have emphasized the importance of their O-glycan structures. Indeed, AGPs and EXTs are decorated with complex to simple carbohydrate-chains (Figures 1B,C) that are required for functionality of these glycomolecules. Until recently, the enzymes, as well as the molecular mechanisms controlling the synthesis of HRGP O-glycans, were poorly understood. A recent effort in the identification of the genes involved in the biosynthesis of HRGP O-glycans has considerably improved our understanding of the molecular events controlling the addition of sugars on these Hyp-rich proteins. The aim of this section is to bring together recent advances in the biosynthesis of HRGP O-glycans, with a focus on AGPs and EXTs. Structural and biological functions are also discussed.

AGPs AND EXTs: THE SYNTHESIS OF O-GLYCANS PRO HYDROXYLATION OF HRGPs

Pro hydroxylation of plant cell wall HRGPs occurs predominantly on Hyp that are formed in the secretory pathway through the action of proline hydroxylases (P4Hs). In *Arabidopsis*, 13 P4Hs have been identified (Hieta and Myllyharju, 2002; Vlad et al., 2007; Velasquez et al., 2011). P4Hs are membrane-anchored enzymes (Yuasa et al., 2005). It is likely that Pro hydroxylation begins in the ER and continues in the Golgi apparatus. Detailed investigations of substrate affinity of two *Arabidopsis* P4-Hs, AtP4H1 and AtP4H2, showed that both AtP4H1 and AtP4H2 hydroxylate AGP-like and EXT-like synthetic peptides (Hieta and Myllyharju, 2002; Tiainen et al., 2005). However, the substrate specificity of the enzymes towards the two classes of synthetic peptides differed. Additional data showed that AtP4H2 poorly hydroxylated animal collagen, but did not hydroxylate animal hypoxia-inducible transcription factor (HIF); while AtP4H1 hydroxylates both animal Hyp-containing proteins collagen and HIF (Hieta and Myllyharju, 2002). Similarly, Velasquez et al. (2011) showed that some root hair-specific P4Hs are able to hydroxylate EXTs, and displayed almost no activity toward AGP-like peptides. Root hair morphology of *Arabidopsis* *p4h* mutants was dramatically altered. Complementation of these mutants with wild type genes restored the phenotype. In addition to being substrate-specific, Velasquez et al. (2011) also showed that some P4Hs were also cell type-specific: P4H2 and P4H5 being confined to trichoblast cells, while P4H13 being present in both trichoblast and atrichoblast cells. Recently, it has also been shown that different tomato P4Hs played a role in plant growth, and exhibited substrate- and tissue-specific activities (Fragkostefanakis et al., 2014). The authors have shown that silencing individual P4Hs result on an increased expansion of root and leaf cells in tomato. This increase correlated with a reduction in the amount of AGPs and possibly EXTs. Plants are therefore likely to regulate the secretion of various classes of

HRGPs at different stages of development and/or responses to stress to perform specific functions in a given cell type or organ. After their synthesis and secretion, HRGPs may be modified and/or re-arranged in the cell wall, but this aspect has received little attention so far, particularly in the case of AGPs.

GLYCOSYLTRANSFERASES INVOLVED IN AGP O-GLYCAN BIOSYNTHESIS

Liang et al. (2010) has suggested that ~15 GTs are involved in AGP glycan biosynthesis. Initiation of the biosynthesis requires the action of specific AGP O-Hyp Gal-T, able to initiate the galactosylation of hydroxylated residues on AGP backbone. Recently, an *Arabidopsis* Gal-T (*AtGalT2*), belonging to CAZy GT family 31 and containing a pfam 01762 domain encoding a Gal-T catalytic domain, able to add one galactosyl residue to Hyp residues of synthetic AGP-like peptides, has been identified (Basu et al., 2013). The authors showed that *AtGalT2* was able to add one galactose residue to synthetic AGP-like peptide, and not to synthetic EXT-like peptides. *AtGalT2* was also harboring a GAL-LECTIN binding domain pfam 00337. This domain, previously identified as a *N*-acetylgalactosaminyl GT (CAZy GT family 27), was involved in catalyzing the first steps of the glycosylation of mammalian mucins (Hassan et al., 2000; Wandall et al., 2007). *AtGAIT2* was found to be located in the ER and in the Golgi apparatus, a pattern similar to the one displayed by P4Hs (Yuasa et al., 2005; Velasquez et al., 2011). It is possible that these two enzymes co-operate in plants to hydroxylate Pro residues and add the first galactosyl residue of the newly synthesized β -1,3- galactan chain. In addition to *AtGalT2*, Qu et al. (2008), using a combination of bioinformatic approaches, identified several additional Gal-Ts belonging to the GT family 31, and showed their putative involvement in the elongation of the β -1,3- galactan backbone of AGPs. For instance, the protein encoded by the gene *At1g77810* was demonstrated to exhibit a specific β -1,3-Gal-T activity. An additional Gal-T activity that adds the second galactose to the Gal-Hyp nascent chains has also been partially characterized (Liang et al., 2010).

In addition to β -1,3-Gal-T, AGP glycan synthesis also requires the action of different other β -1,6-Gal-T, α -1,3- and α -1,5-arabinosyltransferase (Ara-T), β -glucuronosyltransferase (GlcA-T), and α -1,2-fucosyltransferase (FuT; Wu et al., 2010). Recently, two *Arabidopsis* Gal-Ts showing a β -1,6-Gal-T activity have been identified: *AtGalT31A*, a β -1,6-Gal-T which is classified into the CAZy GT family 31, is required for the addition of Gal residues to existing β -1,6- galactan chains (Geshi et al., 2013) while *AtGalT29A* (CAZy GT family 29) is required for the addition of galactose residues to β -1,3- and β -1,6- galactan chains (Dilokpimol et al., 2014). Both *AtGalT31A* and *AtGalT29A* are type II transmembrane proteins located in the Golgi apparatus. Traces of ER-localization previously observed with P4Hs and *AtGalT2* were not found, suggesting that addition of β -1,6-galactose residues to the side chains of AGPs occurs later during the transit of nascent HRGPs into Golgi stacks. Using subcellular co-localization approaches, FRET acceptor photo-bleaching techniques as well as immuno-precipitation techniques, the authors showed that (i) *AtGalT31A* and *AtGalT29A* were organized into heterodimer complexes, and (ii) this heterodimer had an enhanced enzymatic activity than the homodimer *AtGalT31A/AtGalT31A*,

or *AtGalT29A/AtGalT29A*. AGP arabinogalactan chains are also modified with glucuronic acid (GlcA) residues. Knoch et al. (2013) have identified an *Arabidopsis* transferase belonging to the CAZy GT family 14, named *AtGlcAT14A*, exhibiting an AGP-specific GlcA-T activity, able to transfer GlcA residues both onto β -1,3- and β -1,6-galactan chains (see also Zhou et al., 2009; Ye et al., 2011). Interestingly, *AtGlcAT14A* was localized to the Golgi apparatus. *AtGlcAT14A* is co-expressed with *AtGalT31A* and co-localize in the Golgi apparatus. However, the FRET photo-bleaching acceptor technique showed that both enzymes did not physically interact. These findings suggest that all the enzymes involved in AGP glycan synthesis, although probably co-regulated, are not necessarily part of a unique multi-protein complex. *Arabidopsis* AGP glycans were also shown to contain fucose residues (Tryfona et al., 2012). Two *Arabidopsis* FuT *AtFUT4* and *AtFUT6*, belonging to the CAZy GT family 37, were shown to specifically add fucose residues to tobacco arabinogalactosylated AGP glycan chains (Wu et al., 2010). Interestingly, de-arabinosylation of tobacco AGP glycans (using arabinofuranosidase) prevented the addition of fucose residue to the glycan, suggesting that arabinosylation was required for further addition of fucose by *AtFUT4* and *AtFUT6*, supporting the arguments for sequential synthesis of AGP glycans along the Golgi cisternae. Biochemical data showed that both FuTs fucosylate AGP glycan in a different manner, most likely on different arabinose residues (Wu et al., 2010). Finally, arabinose, along with galactose, is the more abundant sugar found in AGP glycans. Recently, Gille et al. (2013) have identified an AGP-altered mutant of *Arabidopsis* named *reduced arabinose yariv1* (*ray1-1*). Monosaccharide composition of a root AGP fraction precipitated with β -glucosyl Yariv, showed a significant decrease in arabinose content in the *ray1-1* mutant, as compared to the wild type. In addition, the *ray1-1* mutant showed a reduction in the length of its primary roots. *RAY1-1* gene was found to encode for a CAZy GT family 77 Ara-T, localized in the Golgi apparatus (Gille et al., 2013). It is however, unknown if *RAY1* is able to add arabinosyl residues to short oligo-arabinosides also found on AGPs.

EXT O-GLYCAN BIOSYNTHESIS

Extensin O-glycans consists of short arabinoside chains with single galactose residues, linked respectively to Hyp residues, and serine residues of the Ser-Hyp₄ motifs. In contrast to the length and the molecular weight of arabinogalactan chains found in AGPs; EXT arabinoside chains are limited to 4–5 arabinosyl residues, predominantly β -1,2-linked. The number of enzymes required for their biosynthesis is also reduced to Ara-T initiating and elongating the arabinoside chains, and to the Ser-O-Gal-T, adding the single galactose residue to serine (Velasquez et al., 2012; Saito et al., 2014). Ser-O-Gal-T are type I transmembrane proteins, located in the ER and possibly in the *cis*-Golgi cisternae (Saito et al., 2014), and prior hydroxylation of Pro residues is required for galactosylation of serine residues on EXTs. It is unknown if this initial galactosylation is required for further EXT arabinosylation.

While the enzyme adding the first galactose residue to AGPs is now identified, the enzyme transferring the first arabinosyl-residue to O-Hyp EXT (and maybe on Ser-Hyp₃ domains of certain AGPs; Qi et al., 1991), remains unidentified. However, Ara-Ts adding the second, the third, and then the fourth arabinose

residue to O-Hyp EXTs have been identified. Indeed, *Arabidopsis* RRA1-3, XEG113, and ExAD were shown (or proposed for ExAD; Velasquez et al., 2012) to transfer respectively, the second, the third and the fourth arabinose residue β -1,2-linked to EXT (Egelund et al., 2007; Gille et al., 2009; Velasquez et al., 2011, 2012). XEG113 belongs to the CAZy GT family 77 and *xeg113* mutants exhibited abnormally elongated hypocotyls under stress conditions. XEG113 was found to be associated with Golgi membranes (Gille et al., 2009). RRA3 is also a type II transmembrane protein, member of the CAZy GT family 77, localized in the Golgi apparatus, and shown to reduce root hair growth. *Rra3* *Arabidopsis* mutants exhibited impaired root hairs. Using a base-mediated hydrolysis of the peptide backbone, followed by mass spectrometry analyses, the authors elegantly showed that XEG113 was responsible for the addition of the second arabinose residue to an elongating β -1,2- arabinan chain, while RRA3 was responsible for the addition of the third residue. RRA3, XEG113, P4H2, and P4H5 were co-expressed, also indicating that EXT glycan synthesis must be tightly regulated within the endomembrane system. Moreover, different cell wall related proteins including *AtRSH1* (a classical EXT HRGP), *AtLRX1* (a hybrid EXT HRGP), *AtPRP1* (a proline-rich protein), and several peroxidase genes were also co-expressed with Ara-T and P4Hs, in *Arabidopsis* root hairs (Velasquez et al., 2011).

Together, these studies suggest that AGP and EXT glycan synthesis is initiated in the ER and continues in the Golgi apparatus, similarly to the *N*-glycosylation pathway. P4Hs and *AtGalT2* may co-operate during hydroxylation and galactosylation of AGP in the ER and in the Golgi apparatus. Elongation and ramification of the AGP glycans would probably take place in different Golgi subcompartments before their export to the cell surface. But specific compartmentalization of the enzymes involved in HRGPs synthesis within specific Golgi cisternae is not yet established and requires further investigations. Such an arrangement has already been described for enzymes involved in the *N*-glycosylation of secreted proteins (Saint-Jore-Dupas et al., 2006) and for the synthesis of the hemicellulosic polysaccharide xyloglucan (Chevalier et al., 2010; Driouich et al., 2012).

AGPs AND EXTs: ROLE IN MORPHOLOGY AND DEVELOPMENT

Cell wall components are organized into networks of polysaccharides and glycoproteins which, apart from operating individually, are strongly interconnected (Carpita and Gibeaut, 1993; Burton et al., 2010; Albersheim et al., 2011). Indeed, it is widely acknowledged that cellulose microfibrils and hemicellulose constitute a primary network of polysaccharides, embedded into a second network made of pectic polysaccharides. Less often referred as such, O-Hyp cell wall proteins constitute the third network of the wall component. Bridges between these three networks do also exist and structural alteration occurring on a single cell wall component often affects overall cell wall architecture and integrity. Thus, structurally altered AGPs/EXTs weaken cell wall architecture (both covalently and non-covalently), and affect biological processes controlled by the cell wall compartment.

Indeed, most of the *Arabidopsis* mutants defective in one or more enzymes described above presented various developmental

and morphological alterations. van Hengel and Roberts (2002) showed that the lack of fucose residue in the *Arabidopsis mrl1* mutant caused their roots to be shortened. This growth defect was due to structural modification of root AGPs (van Hengel and Roberts, 2002), as well as a result of altered rhamnogalacturonan-II synthesis, since the disorder was partially rescued by exogenous application of boric acid (O'Neill et al., 2001). Liang et al. (2013) showed that a deficiency in the genes *AtFUT4* and *AtFUT6* caused a reduction of root growth under saline stress conditions (see also Tryfona et al., 2014). The lack of fucose residue was proposed to affect intramolecular interactions between AGPs and other wall components. Similarly reduced galactosylation of AGPs in *reb1-1* mutant of *Arabidopsis* caused strong swelling of trichoblast cells as well as reduced root growth (Andème-Onzighi et al., 2002; Nguema-Ona et al., 2006). The *Arabidopsis* mutant *atglcat14a*, deficient in an AGP-specific GlcA GT, showed an abnormal increase in root and hypocotyl length, when compared to the wild type (Knoch et al., 2013). Biochemical analysis in this mutant showed an alteration in the AGP composition and associated glycosidic linkages as compared to the wild type, suggesting that biochemical phenotype indirectly impacts cell elongation via an overall change in cell wall architecture and integrity. The mutation in *AtGalT31A* caused the arrest of the embryo development at the globular stage, while complementation of the mutant with *AtGALT31A* restored the wild type phenotype, thus linking the requirement of correctly glycosylated AGPs with the progression of embryogenesis beyond the globular stage (Geshi et al., 2013). However, a study of *atgalT2* deficient *Arabidopsis* mutants showed that allelic mutant lines contained less Gal-T activity when compared to the wild type without displaying any significant alteration of the phenotype. Basu et al. (2013) suggested that other Hyp-O Gal-Ts may compensate for the loss of *AtGalT2*, and that examination of these mutants under non-physiological conditions, or the production of multigene mutants within this gene family may reveal novel phenotypes. Recently, an unusual AGP (named APAP1) was found to be covalently linked to pectin rhamnogalacturonan-I and to arabinoxylans (Tan et al., 2013). Absence of APAP1 in the corresponding mutant led to an increased extractability of pectins and xylans, thus suggesting the alteration of its overall wall architecture. *Apap1* mutants exhibited a significant increase in the height inflorescence stem, although the overall morphology was comparable to that of the wild type.

Arabidopsis EXT-deficient or EXT-altered mutants also presented various developmental and morphological alterations, due to an alteration in their overall wall architecture. Velasquez et al. (2011) showed that disrupting the Pro hydroxylation and/or improper O-glycosylation impacted EXT ability to form covalent intra and inter-molecular network in the wall. Indeed, secondary helix conformation found in EXTs, required for normal catalysis of the di-isodityrosine bondages by wall peroxidases (Held et al., 2004), was altered in P4H-deficient and Ara-T-defective *Arabidopsis* mutants. The authors concluded that the absence, or the alteration of their Hyp-O-arabinosides, destabilized the EXT helical secondary structure, altering their ability to interact in the wall with other cell wall components, thus altering their structural function *in muro*. Interestingly, unlike in plants, the hydroxylated Pro residues of animal proteins are not glycosylated. Pro

hydroxylation itself is sufficient for the conformational stability of animal Hyp-rich proteins such as collagen. This PTM is generally sufficient for proper functioning of such proteins. Indeed, Hyp stabilizes their triple helical structure at body temperature (Kivirikko and Pihlajaniemi, 1998; Myllyharju, 2003), and to date, no animal Hyp-containing proteins have been found to be glycosylated. *O*-glycosylation of Hyp is a rather plant-specific PTM, required for the proper functioning of plant Hyp-containing proteins including HRGPs. While the enzymes hydroxylating Pro residues on AGPs and EXTs are similar to mammalian systems, the ones that initiate and elongate the glycan chains of these HRGPs are unique to plants.

Furthermore, the study of the *Arabidopsis rsh* mutant, deficient in an EXT (RSH/EXT3) has also shown the importance of EXTs for normal plant cell wall architecture and function in development. RSH/EXT3 is an *Arabidopsis* EXT which was shown to play a key role during cytokinesis, by controlling cell plate formation (Hall and Cannon, 2002; Cannon et al., 2008). RSH/EXT3 positively charged was proposed to interact with negatively charged pectins to create a template for newly synthesized cell walls.

AGPs AND EXTs: ROLE IN BIOTIC STRESS

In addition to their role in morphology and growth, AGPs and EXTs were shown to play key roles in plant responses to biotic stress. Esquerre-Tugayé (1979) and Esquerre-Tugayé et al. (1979) have shown that plants respond to fungal infection by an increased secretion of HRGPs. Both AGPs (reviewed in Nguema-Ona et al., 2013) and EXTs were later on shown to play various roles in this response to pathogens. More specifically root apices and exudates were found to be enriched in AGPs (**Figure 2**), their chemical composition being different depending on both root tissues and plant species (Dolan et al., 1995; Durand et al., 2009; Cannesan et al., 2012). AGPs have long been suspected to be involved in root-microorganisms interaction including symbiotic associations (Scheres et al., 1990; Balestrini et al., 1996; Berry et al., 2002). For instance, alteration of AGP synthesis or secretion was shown to inhibit *Rhizobium* sp. YAS34 attachment to the root surface of *Arabidopsis thaliana* (Vicré et al., 2005; **Figure 2**). Xie et al. (2012) further demonstrated that AGPs from pea root exudates promote polar orientation and adhesion of *Rhizobium leguminosarum*. However, AGP functioning in root defense remained speculative until recently, as demonstrated by the study of Cannesan et al. (2012) on pea roots. The authors have shown that AGPs isolated from root cap (RC) and border cells are strong attractants of zoospores of the pathogenic oomycete *Aphanomyces euteiches* *in vitro* (Cannesan et al., 2012). The AGPs also inhibited *in vitro* cyst germination and the subsequent mycelium growth and propagation. These findings highlight the important contribution of AGPs in *Aphanomyces euteiches* root infection and show for the first time that AGPs are involved in controlling root-pathogenic oomycete interaction (see also Nguema-Ona et al., 2013).

Extensins have also been shown to play a significant role in plant defense and protection against bioaggressors. Immunolocalization studies using the mAbs JIM 20 and JIM 11 revealed the abundant presence of EXT epitopes in cell walls of the resistant wax gourd cultivar to *Fusarium oxysporum* as compared to susceptible cultivar (Xie et al., 2011). In addition, elicitation with fusaric acid

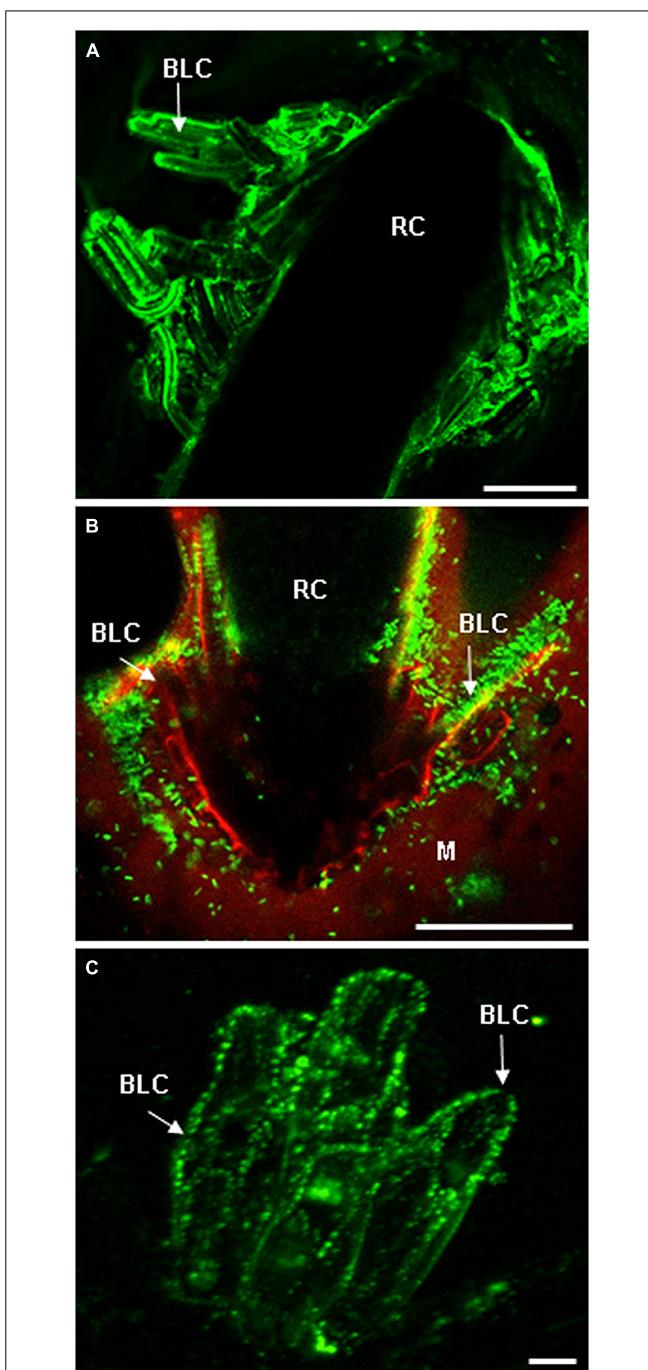


FIGURE 2 | Root cap (RC) and border cells are both enriched in AGP and EXT epitopes. **(A)** Immunostaining of AGP epitopes at the surface of RC and border-like cells of *Brassica napus* with the mAb JIM8 (from Cannesan et al., 2012 with permission). Root border-like cells are produced and released from the RC. **(B)** Micrographs showing the association between root border-like cells from *Arabidopsis thaliana* and *Rhizobium* sp. YAS34-GFP. The GFP-expressing bacteria appear green at the root surface (from Vicré et al., 2005 with permission). This association is AGP-dependant as demonstrated in Vicré et al. (2005). **(C)** Fluorescent micrographs of root border-like cells from flax (*Linum usitatissimum*) immunostained with the monoclonal antibody LM1 specific for EXT epitopes (from Plancot et al., 2013 with permission). Bars = 20 μ m **(A)**, 50 μ m **(B)**, and 8 μ m **(C)**. BLCs, border-like cells; M, mucilage; RC, root cap.

or infection with *F. oxysporum* caused important decrease of the immunofluorescence in both resistant and susceptible cultivars. Also, elicitation of grapevine callus cultures resulted in both the insolubilization of a specific 89.9 kD EXT and the induction of the catalytic activity of an EXT peroxidase (Jackson et al., 2001). Furthermore, EXTs have been shown to accumulate in response to the pathogenic oomycete *Sclerospora graminicola* in resistant pearl millet cultivar (Deepak et al., 2007). The high content of EXTs was tightly correlated with an increase in the levels of isodityrosine and H₂O₂ suggesting cell wall strengthening in the resistant cultivar presumably to limit *Sclerospora graminicola* penetration and tissue infection.

More recently, the implication of EXTs as part of the innate immune response of root border-like cells (BLCs) of *Arabidopsis thaliana* and *Linum usitatissimum* has been investigated by Plancot et al. (2013). Root border cells from plants such as pea, soybean, or cotton are highly specialized in root protection and production of various anti-microbial compounds (Hawes et al., 2000, 2003). Although such a function still needs to be clearly established for BLCs, a class of border cells that is relatively less studied. Recent work suggests a role for these BLCs in root defense (Driouich et al., 2013). Plancot et al. (2013) have also demonstrated that, in response to elicitors (e.g., flagellin 22), a significant increase in the production of H₂O₂ was detected in root BLCs together with a strong activation of genes involved in EXT biosynthesis and cross-linking. This is consistent with the finding of Velasquez et al. (2011) which showed that EXTs biosynthesis genes were co-expressed with peroxidase genes. Interestingly, treatment with elicitors also caused modifications in the distribution of EXT epitopes within cell walls of root BLCs (**Figure 2**). The effect of elicitation on the pattern of labeling with the mAb LM1 was shown to depend on both the nature of elicitors and plant species. Elicitation with flagellin 22 almost abolished immunostaining of LM1-recognized epitopes reflecting reorganization of the EXT network within the cell wall due to extensive cross-linking. Such an oxidative cross-linking of EXTs may result in a reinforced glyco-network that enhances physical properties of the cell wall in both *Arabidopsis thaliana* and *L. usitatissimum* (Plancot et al., 2013). This reinforcement of the cell wall would in turn limit/prevent penetration and progression of pathogens within root tissues.

Together these findings strongly suggest that AGPs and EXTs are key components of root protection, and more specifically of root border cells. However, further investigations where root border cells are directly challenged with specific pathogens are needed to provide a biological context for these observations. So far, the immune response in roots remains poorly understood and appears to be highly complex and cell-type specific (Millet et al., 2010; Cannesan et al., 2012; Balmer and Mauch-Mani, 2013). To our knowledge, the only study that clearly demonstrated the relationship between the production of EXT and plant resistance to pathogens was performed in leaf tissues (Wei and Shirsat, 2006). In this study, over-expression of the *EXT1* gene in leaves of *Arabidopsis thaliana* clearly limits the spreading of the pathogenic bacteria *Pseudomonas syringae* DC3000 within the tissues. Subsequently, the infection symptoms are significantly reduced. It is clear that the implication of EXT and AGP populations in root protection is far from being fully understood and more studies are

needed to elucidate the role of individual HRGPs/or their glycans in resistance to biotic stress.

CONCLUSION AND OUTLOOK

Like N-glycoproteins, cell wall O-glycoproteins, AGPs and EXTs, are synthesized, assembled and modified within the secretory system. Their glycans, although structurally different and diverse, play a major role in their stability, activity and function. Both types of glycoproteins were shown to be involved in the control of many biological activities and physiological processes in various plant species. However, the specific role of each glycan type and the associated oligosaccharides in biological processes is not known. One of the important challenges for the future is to elucidate the contribution of each of these glycans (and associated sugars) in regulating cell growth, development and adaptation of plants to environmental stresses, either biotic or abiotic. Even more challenging is the search for potential relationships between a given glycan/oligosaccharide structure and a given function in a given tissue. For instance, how specific O-glycan structures regulate morphology, growth or biotic interactions of certain root cell types with microbes is a major issue that deserves further attention.

Recently, a number of the carbohydrate active enzymes involved in N- and O-glycan metabolism have been identified and have advanced our understanding of the biosynthetic machineries of these glycoproteins. How these enzymes are spatially organized and assembled within different compartments of the endomembrane system (i.e., specifically within Golgi subcompartments and Golgi-derived secretory vesicles) and how these are regulated during development is not fully understood and remains an exciting research opportunity for the future.

ACKNOWLEDGMENTS

Part of the work described in this review was supported by the University of Rouen, and the GRR-Végétal-Agronomie-Sol-Innovation of Haute Normandie, Le Fonds Européen de Développement Regional (FEDER), and the ANR. We are grateful to our lab members for their active implication in various “Plant Glycobiology” projects and to Dr. C. Santaella (CNRS Cadarache-Université de Marseille) for her stimulating discussions during the course of our root-microbe interaction studies. The authors are thankful to Dr. J. P. Moore (University of Stellenbosch-Republic of South Africa) for careful reading of the manuscript.

REFERENCES

- Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta* 1833, 2430–2437. doi: 10.1016/j.bbamcr.2013.04.001
- Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., and Staehelin, A. (2011). *Plant Cell Walls: from Chemistry to Biology*. New York: Garland Science, 365–407.
- Andème-Onzighi, C., Sivaguru, M., Judy-March, J., Baskin, T. I., and Driouich, A. (2002). The *reb1-1* mutation of *Arabidopsis* alters the morphology of trichoblasts, the expression of arabinogalactan-proteins and the organization of cortical microtubules. *Planta* 215, 949–958. doi: 10.1007/s00425-002-0836-z
- Bakker, H., Lommen, A., Jordi, W., Stiekema, W., and Bosch, D. (1999). An *Arabidopsis thaliana* cDNA complements the N-acetylglucosaminyltransferase I deficiency of CHO Lec1 cells. *Biochem. Biophys. Res. Commun.* 261, 829–832. doi: 10.1006/bbrc.1999.1117
- Bakker, H., Schijlen, E., de Vries, T., Schiphorst, W. E., Jordi, W., Lommen, A., et al. (2001). Plant members of the α-1-3/4-fucosyltransferase gene family encode an α-1-4-fucosyltransferase, potentially involved in Lewis(a) biosynthesis, and two

- core α -1,3-fucosyltransferases. *FEBS Lett.* 507, 307–312. doi: 10.1016/S0014-5793(01)02999-4
- Balestrini, R., Hahn, M. G., Faccio, A., Mendgen, K., and Bonfante, P. (1996). Differential localization of carbohydrate epitopes in plant cell walls in the presence and absence of arbuscular mycorrhizal fungi. *Plant Physiol.* 111, 203–213.
- Balmer, D., and Mauch-Mani, B. (2013). More beneath the surface? Root versus shoot antifungal plant defenses. *Front. Plant Sci.* 4:256. doi: 10.3389/fpls.2013.00256
- Bardor, M., Crémata, J., and Lerouge, P. (2011). Glycan engineering in transgenic plants. *Annu. Plant Rev.* 41, 409–424. doi: 10.1002/9781444391015.ch17
- Bardor, M., Faveeuw, C., Fitchette, A., Gilbert, D., Galas, L., Trottein, F., et al. (2003). Immunoreactivity in mammals of two typical plant glycoepitopes, core α -(1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. doi: 10.1093/glycob/cwg024
- Bar-Peled, M., and O'Neill, A. (2011). Plant nucleotide sugar formation, interconversion and salvage by sugar recycling. *Annu. Rev. Plant Biol.* 62, 127–155. doi: 10.1146/annurev-arplant-042110-103918
- Basu, D., Liang, Y., Liu, X., Himmeldirk, K., Faik, A., Kieliszewski, M. J., et al. (2013). Functional identification of a hydroxyproline-O-galactosyltransferase specific for arabinogalactan protein biosynthesis in *Arabidopsis*. *J. Biol. Chem.* 288, 10132–10143. doi: 10.1074/jbc.M112.432609
- Bencúr, P., Steinkellner, H., Svoboda, B., Mucha, J., Strasser, R., Kolarich, D., et al. (2005). *Arabidopsis thaliana* β -1,2-xylosyltransferase: an unusual glycosyltransferase with the potential to act at multiple stages of the plant N-glycosylation pathway. *Biochem. J.* 388, 515–525. doi: 10.1042/BJ20042091
- Berry, A. M., Rasmussen, U., Bateman, K., Huss-Danell, K., Lindwall, S., and Bergman, B. (2002). Arabinogalactan proteins are expressed at the symbiotic interface in root nodules of *Alnus* spp. *New Phytol.* 155, 469–479. doi: 10.1046/j.1469-8137.2002.00466.x
- Boisson, M., Gomord, V., Audran, C., Berger, N., Dubreucq, B., Granier, F., et al. (2001). *Arabidopsis* glucosidase I mutants reveal a critical role of N-glycan trimming in seed development. *EMBO J.* 20, 1010–1019. doi: 10.1093/emboj/20.5.1010
- Burda, P., and Aebl, M. (1999). The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta* 1426, 239–257. doi: 10.1016/S0304-4165(98)00127-5
- Burn, J. E., Hurley, U. A., Birch, R. J., Arioli, T., Cork, A., and Williamson, R. E. (2002). The cellulose-deficient *Arabidopsis* mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. *Plant J.* 32, 949–960. doi: 10.1046/j.1365-313X.2002.01483.x
- Burton, R. A., Gidley, M. J., and Fincher, G. B. (2010). Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat. Chem. Biol.* 6, 724–732. doi: 10.1038/nchembio.439
- Canesan, M. A., Durand, C., Burel, C., Gangneux, C., Lerouge, P., Ishii, T., et al. (2012). Effect of arabinogalactan proteins from the root caps of pea and *Brassica napus* on *Aphanomyces euteiches* zoospore chemotaxis and germination. *Plant Physiol.* 159, 1658–1670. doi: 10.1104/pp.112.198507
- Cannon, M. C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B. L., et al. (2008). Self-assembly of the plant cell wall requires an extensin scaffold. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2226–2231. doi: 10.1073/pnas.0711980105
- Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x
- Chen, X. L., Shi, T., Yang, J., Shi, W., Gao, X., Chen, D., et al. (2014). N-glycosylation of effector proteins by an α -1,3-mannosyltransferase is required for the rice blast fungus to evade host innate immunity. *Plant Cell* 26, 1360–1376. doi: 10.1105/tpc.114.123588
- Chevalier, L., Bernard, S., Ramdani, Y., Lamour, R., Bardor, M., Lerouge, P., et al. (2010). Subcompartment localization of the side chain xyloglucan-synthesizing enzymes within Golgi stacks of tobacco suspension-cultured cells. *Plant J.* 64, 977–989. doi: 10.1111/j.1365-313X.2010.04388.x
- Deepak, S., Shailasree, S., Kini, R. K., Hause, B., Shetty, S. H., and Mithöfer, A. (2007). Role of hydroxyproline-rich glycoproteins in resistance of pearl millet against downy mildew pathogen *Sclerospora graminicola*. *Planta* 226, 323–333. doi: 10.1007/s00425-007-0484-4
- Dilokpimol, A., Poulsen, C. P., Vereb, G., Kaneko, S., Schulz, A., and Geshi, N. (2014). Galactosyltransferases from *Arabidopsis thaliana* in the biosynthesis of type II arabinogalactan: molecular interaction enhances enzyme activity. *BMC Plant Biol.* 14:90. doi: 10.1186/1471-2229-14-90
- Dolan, L., Linstead, P., and Roberts, K. (1995). An AGP epitope distinguishes a central metaxylem initial from other vascular initials in the *Arabidopsis* root. *Protoplasma* 189, 149–155. doi: 10.1007/BF01280168
- Driouich, A., Follet-Gueye, M.-L., Bernard, S., Kousar, S., Chevalier, L., Vicré-Gibouin, M., et al. (2012). Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Front. Plant Sci.* 3:79. doi: 10.3389/fpls.2012.00079
- Driouich, A., Follet-Gueye, M.-L., Vicré-Gibouin, M., and Hawes, M. (2013). Root border cells and secretions as critical elements in plant host defense. *Curr. Opin. Plant Biol.* 16, 489–495. doi: 10.1016/j.pbi.2013.06.010
- Durand, C., Vicré-Gibouin, M., Follet-Gueye, M.-L., Duponchel, L., Moreau, M., Lerouge, P., et al. (2009). The organization pattern of root border-like cells of *Arabidopsis* is dependent on cell wall homogalacturonan. *Plant Physiol.* 150, 1411–1421. doi: 10.1104/pp.109.136382
- Egelund, J., Obel, N., Ulvskov, P., Geshi, N., Pauly, M., Bacic, A., et al. (2007). Molecular characterization of two *Arabidopsis thaliana* glycosyltransferase mutants, rra1 and rra2, which have a reduced residual arabinose content in a polymer tightly associated with the cellulosic wall residue. *Plant Mol. Biol.* 64, 439–451. doi: 10.1007/s11103-007-9162-y
- Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan protein: key regulators at the cell surface. *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000
- Esquerre-Tugayé, M. T. (1979). Cell surfaces in plant-microorganism interactions: I. A structural investigation of cell wall hydroxyproline-rich glycoproteins which accumulate in fungus-infected plants. *Plant Physiol.* 64, 314–319. doi: 10.1104/pp.64.2.314
- Esquerre-Tugayé, M. T., Lafitte, C., Mazau, D., Toppan, A., and Touzé, A. (1979). Cell surfaces in plant-microorganism interactions: II. Evidence for the accumulation of hydroxyproline-rich glycoproteins in the cell wall of diseased plants as a defense mechanism. *Plant Physiol.* 64, 320–326. doi: 10.1104/pp.64.2.320
- Fanata, W. I., Lee, K. H., Son, B. H., Yoo, J. Y., Harmoko, R., Ko, K. S., et al. (2013). N-glycan maturation is crucial for cytokinin-mediated development and cellulose synthesis in *Oryza sativa*. *Plant J.* 73, 966–979. doi: 10.1111/tpj.12087
- Farid, A., Malinovsky, F. G., Veit, C., Schoberer, J., Zipfel, C., and Strasser, R. (2013). Specialized roles of the conserved subunit OST3/6 of the oligosaccharyltransferase complex in innate immunity and tolerance to abiotic stresses. *Plant Physiol.* 162, 24–38. doi: 10.1104/pp.113.215509
- Fragkostefanakis, S., Sedeek, K. E. M., Raad, M., Zaki, M. S., and Kalaitzis, P. (2014). Virus induced gene silencing of three putative prolyl 4-hydroxylases enhances plant growth in tomato (*Solanum lycopersicum*). *Plant Mol. Biol.* 85, 459–471. doi: 10.1007/s11103-014-0197-6
- Geshi, N., Johansen, J. N., Dilokpimol, A., Rolland, A., Belcram, K., Verger, S., et al. (2013). A galactosyltransferase acting on arabinogalactan protein glycans is essential for embryo development in *Arabidopsis*. *Plant J.* 76, 128–137. doi: 10.1111/tpj.12281
- Gil, G.-C., Velander, W. H., and Van Cott, K. E. (2009). N-glycosylation microheterogeneity and site occupancy of an Asn-X-Cys sequon in plasma-derived and recombinant Protein C. *Proteomics* 9, 2555–2567. doi: 10.1002/pmic.200800775
- Gille, S., Hänsel, U., Ziemann, M., and Pauly, M. (2009). Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14699–14704. doi: 10.1073/pnas.0905434106
- Gille, S., Sharma, V., Baidoo, E. E. K., Keasling, J. D., Scheller, H. V., and Pauly, M. (2013). Arabinosylation of a Yariv-precipitable cell wall polymer impacts plant growth as exemplified by the *Arabidopsis* glycosyltransferase mutant ray1. *Mol. Plant* 6, 1369–1372. doi: 10.1093/mp/sst029
- Hall, Q., and Cannon, M. C. (2002). The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in *Arabidopsis*. *Plant Cell* 14, 1161–1172. doi: 10.1105/tpc.010477
- Hassan, H., Reis, C. A., Bennett, E. P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M. A., et al. (2000). The Lectin domain of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalacto-saminyltransferase-T4 directs its glycopeptide specificities. *J. Biol. Chem.* 275, 38197–38205. doi: 10.1074/jbc.M005783200
- Hawes, M. C., Bengough, G., Cassab, G., and Ponce, G. (2003). Root caps and rhizosphere. *J. Plant Growth Regul.* 21, 352–367. doi: 10.1007/s00344-002-0035-y

- Hawes, M. C., Gunawardena, U., Miyasaka, S., and Zhao, X. (2000). The role of root border cells in plant defense. *Trends Plant Sci.* 5, 128–133. doi: 10.1016/S1360-1385(00)01556-9
- Held, M. A., Tan, L., Kamyab, A., Hare, M., Shpak, E., and Kieliszewski, M. J. (2004). Di-isodityrosine is the intermolecular cross-link of isodityrosine-rich extensin analogs cross-linked in vitro. *J. Biol. Chem.* 279, 55474–55482. doi: 10.1074/jbc.M408396200
- Hennet, T. (2012). Diseases of glycosylation beyond classical congenital disorders of glycosylation. *Biochim. Biophys. Acta* 1820, 1306–1317. doi: 10.1016/j.bbagen.2012.02.001
- Hieta, R., and Myllyharju, J. (2002). Cloning and characterization of a low molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*. *J. Biol. Chem.* 277, 23965–23971. doi: 10.1074/jbc.M201865200
- Ioffe, E., and Stanley, P. (1994). Mice lacking N-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proc. Natl. Acad. Sci. U.S.A.* 91, 728–732. doi: 10.1073/pnas.91.2.728
- Jackson, P. A. P., Galinha, C. I. R., Pereira, C. S., Fortunato, A., Soares, N. C., Amâncio, S. B. Q., et al. (2001). Rapid deposition of extension during the elicitation of grapevine callus cultures is specifically catalyzed by a 40-kilodalton peroxidase. *Plant Physiol.* 127, 1065–1076. doi: 10.1104/pp.010192
- Kang, J., Frank, J., Kang, C., Kajiura, H., Vikram, M., Ueda, A., et al. (2008). Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5933–5938. doi: 10.1073/pnas.0800237105
- Kieliszewski, M. J., Lampert, D. T. A. (1994). Extensins: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J.* 5, 157–172. doi: 10.1046/j.1365-313X.1994.05020157.x
- Kieliszewski, M. J., Lampert, D. T. A., and Cannon, M. C. (2011). Hydroxyproline-rich glycoproteins: form and function. *Annu. Plant Rev.* 41, 321–342.
- Kieliszewski, M. J., and Shpak, E. (2001). Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins. *Cell. Mol. Life Sci.* 58, 1386–1398. doi: 10.1007/PL00000783
- Kivirikko, K. I., and Pihlajaniemi, T. (1998). Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. *Adv. Enzymol. Relat. Areas Mol. Biol.* 72, 325–398.
- Knoch, E., Dilokpimol, A., Tryfona, T., Poulsen, C. P., Xiong, G., Harholt, J., et al. (2013). A β-glucuronosyltransferase from *Arabidopsis thaliana* involved in biosynthesis of type II arabinogalactan has a role in cell elongation during seedling growth. *Plant J.* 76, 1016–1029. doi: 10.1111/tpj.12353
- Lampert, D. T. A., Kieliszewski, M. J., Chen, Y., and Cannon, M. C. (2011). Role of the extensin superfamily in primary cell wall architecture. *Plant Physiol.* 156, 11–19. doi: 10.1104/pp.110.169011
- Larkin, A., and Imperiale, B. (2011). The expanding horizons of asparagine-linked glycosylation. *Biochemistry* 50, 4411–4426. doi: 10.1021/bi200346n
- Lee, K. J. D., Sakata, Y., Mau, S.-L., Pettolino, F., Bacic, A., Quatrano, R. S., et al. (2005). Arabinogalactan proteins are required for apical cell extension in the moss *Physcomitrella patens*. *Plant Cell* 17, 3051–3065. doi: 10.1105/tpc.105.034413
- Leiter, H., Mucha, J., Staudacher, E., Grimm, R., Glössl, J., and Altmann, F. (1999). Purification, cDNA cloning, and expression of GDP-L-Fuc:Asn-linked GlcNAc α-1,3-fucosyltransferase from mung beans. *J. Biol. Chem.* 274, 21830–21839. doi: 10.1074/jbc.274.31.21830
- Léonard, R., Costa, G., Darrambide, E., Lhernould, S., Fleurat-Lessard, P., Carlué, M., et al. (2002). The presence of Lewis a epitopes in *Arabidopsis thaliana* glycoconjugates depends on an active α-4-fucosyltransferase gene. *Glycobiology* 12, 299–306. doi: 10.1093/glycob/12.5.299
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.-C., Gomord, V., and Faye, L. (1998). N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* 38, 31–48. doi: 10.1023/A:1006012005654
- Liang, L., Basu, D., Pattathil, S., Xu, W.-L., Venetos, A., Martin, S. L., et al. (2013). Biochemical and physiological characterization of fut4 and fut6 mutants defective in arabinogalactan-protein fucosylation in *Arabidopsis*. *J. Exp. Bot.* 64, 5537–5551. doi: 10.1093/jxb/ert321
- Liang, L., Faik, A., Kieliszewski, M. J., Tan, L., Xu, W. L., and Showalter, A. M. (2010). Identification and characterization of in vitro galactosyltransferase activities involved in the arabinogalactan protein glycosylation in tobacco and *Arabidopsis*. *Plant Physiol.* 154, 632–642. doi: 10.1104/pp.110.160051
- Liebminger, E., Hüttner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., et al. (2009). Class I α-mannosidases are required for N-glycan pr 398 ocessing and root development in *Arabidopsis thaliana*. *Plant Cell* 21, 3850–3867. doi: 10.1105/tpc.109.072363
- Lowe, J. B., and Marth, J. D. (2003). A genetic approach to mammalian glycan function. *Annu. Rev. Biochem.* 72, 643–691. doi: 10.1146/annurev.biochem.72.121801.161809
- Matsui, T., Takita, E., Sato, T., Kinjo, S., Aizawa, M., Sugiura, Y., et al. (2011). N-glycosylation at noncanonical Asn-X-Cys sequences in plant cells. *Glycobiology* 21, 994–999. doi: 10.1093/glycob/cwq198
- Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., and Marth, J. D. (1994). Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *EMBO J.* 13, 2056–2065.
- Millet, Y. A., Danna, C. H., Clay, N. K., Songnuan, W., Simon, M. D., Werck-Reichhart, D., et al. (2010). Innate immune responses activated in *Arabidopsis* roots by microbe-associated molecular patterns. *Plant Cell* 22, 973–990. doi: 10.1105/tpc.109.069658
- Moore, J. P., Fangel, J. U., Willats, W. G. T., and Vivier, M. A. (2014a). Pectic β(1,4)-galactan, extensin and arabinogalactan–protein epitopes differentiate ripening stages in wine and table grape cell walls. *Ann. Bot.* doi: 10.1093/aob/mcu053 [Epub ahead of print].
- Moore, J. P., Nguema-Ona, E., Fagerström, A., Fangel, J. U., Willats, W. G. T., Hugo, A., et al. (2014b). Profiling the main cell wall polysaccharides of grapevine leaves using high throughput techniques. *Carbohydr. Polym.* 99, 190–198. doi: 10.1016/j.carbpol.2013.08.013
- Motose, H., Sugiyama, M., and Fukuda, H. (2004). A proteoglycan mediates inductive interaction during plant vascular development. *Nature* 429, 873–878. doi: 10.1038/nature02613
- Myllyharju, J. (2003). Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol.* 22, 15–24. doi: 10.1016/S0945-053X(03)00006-4
- Nguema-Ona, E., Andème-Onzighi, C., Aboughe-Angone, S., Bardor, M., Ishii, T., Lerouge, P., et al. (2006). The reb1-1 mutation of *Arabidopsis*. Effect on the structure and localization of galactose-containing cell wall polysaccharides. *Plant Physiol.* 140, 1406–1417. doi: 10.1104/pp.105.074997
- Nguema-Ona, E., Bannigan, A., Chevalier, L., Baskin, T. I., and Driouch, A. (2007). Disruption of arabinogalactan proteins disorganizes cortical microtubules in the root of *Arabidopsis thaliana*. *Plant J.* 52, 240–251. doi: 10.1111/j.1365-313X.2007.03224.x
- Nguema-Ona, E., Coimbra, S., Vicré-Gibouin, M., Mollet, J.-C., and Driouch, A. (2012). Arabinogalactan-proteins in root and pollen tube cells: distribution and functional properties. *Ann. Bot.* 110, 383–404. doi: 10.1093/aob/mcs143
- Nguema-Ona, E., Vicré-Gibouin, M., Cannesan, M.-A., and Driouch, A. (2013). Arabinogalactan-proteins in root microbe interactions. *Trends Plant Sci.* 18, 440–449. doi: 10.1016/j.tplants.2013.03.006
- O'Neill, M. A., Eberhard, S., Albersheim, P., and Darvill, A. G. (2001). Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. *Science* 294, 846–849. doi: 10.1126/science.1062319
- Pagny, S., Bouissonne, F., Sarkar, M., Follet-Gueye, M.-L., Driouch, A., Schachter, H., et al. (2003). Structural requirements for *Arabidopsis* β-1,2-xylosyltransferase activity and targeting to the Golgi. *Plant J.* 33, 189–203. doi: 10.1046/j.0960-7412.2002.01604.x
- Plancton, B., Santaella, C., Jaber, R., Kiefer-Meyer, M.-C., Follet-Gueye, M.-L., Leprinse, J., et al. (2013). Deciphering the responses of root border-like cells of *Arabidopsis thaliana* and *Linum usitatissimum* to pathogen derived elicitors. *Plant Physiol.* 163, 1583–1597. doi: 10.1104/pp.113.222356
- Qin, C., Li, Y., Gan, J., Wang, W., Zhang, H., Liu, Y., et al. (2013). OsDGL1, a homolog of an oligosaccharyltransferase complex subunit, is involved in N-glycosylation and root development in rice. *Plant Cell Physiol.* 54, 129–137. doi: 10.1093/pcp/pcs159
- Qi, W., Fong, C., and Lampert, D. T. A. (1991). Gum arabic glycoprotein is a twisted hairy rope: new model based on O-galactosylhydroxyproline as a polysaccharide attachment site. *Plant Physiol.* 96, 848–855. doi: 10.1104/pp.96.3.848
- Qu, Y., Egelund, J., Gilson, P. R., Houghton, F., Gleeson, P. A., Schultz, C. J., et al. (2008). Identification of a novel group of putative *Arabidopsis thaliana*

- β -(1,3)-galactosyltransferases. *Plant Mol. Biol.* 68, 43–59. doi: 10.1007/s11103-008-9351-3
- Saint-Jore-Dupas, C., Nebenführ, A., Boulaflous, A., Follet-Gueye, M.-L., Plisson, C., Hawes, C., et al. (2006). Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. *Plant Cell* 18, 3182–3200. doi: 10.1105/tpc.105.036400
- Saito, F., Suyama, A., Oka, T., Yoko-o, T., Matsuoka, K., Jigami, Y., et al. (2014). Identification of novel peptidyl serine α -galactosyltransferase/lectosyltransferase gene family in plants. *J. Biol. Chem.* 289, 20405–20420. doi: 10.1074/jbc.M114.553933
- Scheres, B., van Engelen, F., van der Knaap, E., van de Wiel, C., van Kammen, A., and Bisseling, T. (1990). Sequential induction of nodulin gene expression in the developing pea nodule. *Plant Cell* 2, 687–700. doi: 10.1105/tpc.2.8.687
- Schultz, C. J., Rumsewicz, M. P., Johnson, K. L., Jones, B. J., Gaspar, Y. M., and Bacic, A. (2002). Using genomic resources to guide research directions: the arabinogalactan protein gene family as a test case. *Plant Physiol.* 129, 1448–1463. doi: 10.1104/pp.003459
- Seifert, G., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161. doi: 10.1146/annurev.arplant.58.032806.103801
- Showalter, A. M. (2001). Arabinogalactan proteins: structure, expression and function. *Cell. Mol. Life Sci.* 58, 1399–1417. doi: 10.1007/PL00000784
- Showalter, A. M., Keppler, B., Lichtenberg, J., Gu, D., and Welch, L. R. (2010). A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol.* 153, 485–513. doi: 10.1104/pp.110.156554
- Shipak, E., Leykam, J. F., and Kieliszewski, M. J. (1999). Synthetic genes for glycoprotein design and the elucidation of hydroxyproline-O-glycosylation codes. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14736–14741. doi: 10.1073/pnas.96.26.14736
- Sourrouille, C., Marquet-Blouin, E., D'Aoust, M. A., Kiefer-Meyer, M.-C., Seveno, M., Pagny-Salehabadi, S., et al. (2008). Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnol. J.* 6, 702–721. doi: 10.1111/j.1467-7652.2008.00353.x
- Strasser, R. (2014). Biological significance of complex N-glycans in plants and their impact on plant physiology. *Front. Plant Sci.* 5:363. doi: 10.3389/fpls.2014.00363
- Strasser, R., Altmann, F., Mach, L., Glössl, J., and Steinkellner, H. (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β -1,2-linked xylose and core α -1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4
- Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glossl, J., et al. (2007). A unique 1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* 19, 2278–2292. doi: 10.1105/tpc.107.052985
- Strasser, R., Mucha, J., Mach, L., Altmann, F., Wilson, I., Glössl, J., et al. (2000). Molecular cloning and functional expression of β -1,2-xylosyltransferase cDNA from *Arabidopsis thaliana*. *FEBS Lett.* 472, 105–108. doi: 10.1016/S0014-5793(00)01443-5
- Strasser, R., Mucha, J., Schwiha, H., Altmann, F., Glössl, J., and Steinkellner, H. (1999a). Molecular cloning and characterization of cDNA coding for β -1,2N-acetylglucosaminyltransferase I (GlcNAc-TI) from *Nicotiana tabacum*. *Glycobiology* 9, 779–785. doi: 10.1093/glycob/9.8.779
- Strasser, R., Steinkellner, H., Borén, M., Altmann, F., Mach, L., Glössl, J., et al. (1999b). Molecular cloning of cDNA encoding N-acetylglucosaminyltransferase II from *Arabidopsis thaliana*. *Glycoconj. J.* 16, 787–791. doi: 10.1023/A:1007127815012
- Strasser, R., Schoberer, J., Jin, C., Glossl, J., Mach, L., and Steinkellner, H. (2006). Molecular cloning and characterization of *Arabidopsis thaliana* Golgi α -mannosidase II, a key enzyme in the formation of complex N-glycans in plants. *Plant J.* 45, 789–803. doi: 10.1111/j.1365-313X.2005.02648.x
- Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., et al. (2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25, 270–287. doi: 10.1105/tpc.112.107334
- Tan, L., Showalter, A. M., Egelund, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci.* 3:140. doi: 10.3389/fpls.2012.00140
- Tan, L., Varnai, P., Lamport, D. T. A., Yuan C., Xu J., Qui F., et al. (2010). Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated side chains. *J. Biol. Chem.* 285, 24575–24583. doi: 10.1074/jbc.M109.100149
- Taylor, M., Ross, H., Mcrae, D., Stewart, D., Roberts, I., Duncan, G., et al. (2000). A potato α -glucosidase gene encodes a glycoprotein processing α -glucosidase II-like activity. Demonstration of enzyme activity and effects of down-regulation in transgenic plants. *Plant J.* 24, 305–316. doi: 10.1046/j.jds.2000.00873.x
- Tiainen, P., Myllyharju, J., and Koivunen, P. (2005). Characterization of a second *Arabidopsis thaliana* prolyl-4-hydroxylase with distinct substrate specificity. *J. Biol. Chem.* 280, 1142–1148. doi: 10.1074/jbc.M411109200
- Tryfona, T., Liang, H.-C., Kotake, T., Tsumuraya, Y., Stephens, E., and Dupree, P. (2012). Structural characterization of *Arabidopsis* leaf arabinogalactan polysaccharides. *Plant Physiol.* 160, 653–666. doi: 10.1104/pp.112.202309
- Tryfona, T., Theys, T. E., Wagner, T., Stott, K., Keegstra, K., and Dupree, P. (2014). Characterization of FUT4 and FUT6 α -(1-2)-Fucosyltransferases reveals that absence of root arabinogalactan fucosylation increases *Arabidopsis* root growth salt sensitivity. *PLoS ONE* 9:e93291. doi: 10.1371/journal.pone.0093291
- Ulvskov, P., Paiva, D. S., Domozych, D., and Harholt, J. (2013). Classification, naming and evolutionary history of glycosyltransferases from sequenced green and red algal genomes. *PLoS ONE* 8:e76511. doi: 10.1371/journal.pone.0076511
- van Hengel, A. J., and Roberts, K. (2002). Fucosylated arabinogalactan-proteins are required for full root cell elongation in *Arabidopsis*. *Plant J.* 32, 105–113. doi: 10.1046/j.jds.2002.01406.x
- Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3, 97–130. doi: 10.1093/glycob/3.2.97
- Varki, A. (2011). Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells. *Cold Spring Harb. Perspect. Biol.* 3:a005462. doi: 10.1101/csdperspect.a005462
- Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Martth, J. D., et al. (2009). Symbol nomenclature for glycan representation. *Proteomics* 9, 5398–5399. doi: 10.1002/pmic.200900708
- Varki, A., and Lowe, J. B. (2009). “Biological roles of glycans,” in *Essentials of Glycobiology*, 2nd Edn, Chap. 6, eds A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, et al. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Velasquez, M., Salter, J. S., Dorosz, J. G., Petersen, B. L., and Estevez, J. M. (2012). Recent advances on the posttranslational modifications of EXTs and their roles in plant cell walls. *Front. Plant Sci.* 3:93. doi: 10.3389/fpls.2012.00093
- Velasquez, S. M., Ricardi, M. M., Dorosz, J. G., Fernandez, P. V., Nadra, A. D., Pol-Fachin, L., et al. (2011). O-glycosylated cell wall proteins are essential in root hair growth. *Science* 332, 1401–1403. doi: 10.1126/science.1206657
- Vicré, M., Santaella, C., Blanchet, S., Gateau, A., and Driouich, A. (2005). Root border-like cells of *Arabidopsis*: microscopical characterization and role in the interaction with rhizobacteria. *Plant Physiol.* 138, 998–1008. doi: 10.1104/pp.104.051813
- Vlad, F., Spano, T., Vlad, D., Boudaher, F., Ouelhadj, A., and Kalaitzis, P. (2007). *Arabidopsis* prolyl-4 hydroxylases are differentially expressed in response to hypoxia, anoxia and mechanical wounding. *Physiol. Plant.* 130, 471–483. doi: 10.1111/j.1399-3054.2007.00915.x
- von Schaewen, A., Frank, J., and Koiba, H. (2008). Role of complex N-glycans in plant stress tolerance. *Plant Signal. Behav.* 3, 871–873. doi: 10.4161/psb.10.6227
- von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M. (1993). Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol.* 102, 1109–1118. doi: 10.1104/pp.102.4.1109
- Wandall, H. H., Irazoqui, E., Tarp, M. A., Bennett, E. P., Mandel, U., Takeuchi, H., et al. (2007). The lectin domains of polypeptide GalNAc-transferases exhibit carbohydrate binding specificity for GalNAc. Lectin bind to GalNAc-glycopeptide substrates is required for high density GalNAc O-glycosylation. *Glycobiology* 17, 374–387. doi: 10.1093/glycob/cwl082
- Wei, G., and Shirsat, A. H. (2006). Extensin over-expression in *Arabidopsis* limits pathogen invasiveness. *Mol. Plant Pathol.* 7, 579–592. doi: 10.1111/j.1364-3703.2006.00363.x
- Wenderoth, I., and von Schaewen, A. (2000). Isolation and characterization of plant N-acetyl glucosaminyltransferase I (GntI) cDNA sequences. Functional analyses in the *Arabidopsis* cgl mutant and in antisense plants. *Plant Physiol.* 123, 1097–1108. doi: 10.1104/pp.123.3.1097

- Wilson, I. B., Rendić, D., Freilinger, A., Dumić, J., Altmann, F., Mucha, J., et al. (2001a). Cloning and expression of cDNAs encoding α -1,3-fucosyltransferase homologues from *Arabidopsis thaliana*. *Biochim. Biophys. Acta* 1527, 88–96. doi: 10.1016/S0304-4165(01)00151-9
- Wilson, I. B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C., Kamerling, J., et al. (2001b). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core α -1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Wu, Y., Williams, M., Bernard, S., Driouich, A., Showalter, A. M., and Faik, A. (2010). Functional identification of two non-redundant *Arabidopsis* α -(1,2)-fucosyltransferases specific to arabinogalactan proteins. *J. Biol. Chem.* 285, 13638–13645. doi: 10.1074/jbc.M110.102715
- Xie, D., Ma, L., Šamaj, J., and Xu, C. (2011). Immunohistochemical analysis of cell wall hydroxyproline-rich glycoproteins in the roots of resistant and susceptible wax gourd cultivars in response to *Fusarium oxysporum* f. sp. *Benincasae* infection and fusaric acid treatment. *Plant Cell Rep.* 30, 1555–1569. doi: 10.1007/s00299-011-1069-z
- Xie, F., Williams, A., Edwards, A., and Downie, J. A. (2012). A plant arabinogalactan-like glycoprotein promotes a novel type of polar surface attachment by *Rhizobium leguminosarum*. *Mol. Plant Microbe Interact.* 25, 250–258. doi: 10.1094/MPMI-08-11-0211
- Ye, C. Y., Li, T., Tuskan, G. A., Tschaplinski, T. J., and Yang, X. H. (2011). Comparative analysis of GT14/GT14-like gene family in *Arabidopsis*, *Oryza*, *Populus*, *Sorghum* and *Vitis*. *Plant Sci.* 181, 688–695. doi: 10.1016/j.plantsci.2011.01.021
- Yuasa, K., Toyooka, K., Fukuda, H., and Matsuoka, K. (2005). Membrane-anchored prolyl hydroxylase with an export signal from the endoplasmic reticulum. *Plant J.* 41, 81–94. doi: 10.1111/j.1365-313X.2004.02279.x
- Zhou, Y. H., Li, S. B., Qian, Q., Zeng, D., Zhang, M., Guo, L., et al. (2009). BC10, a DUF266-containing and Golgi-located type II membrane protein, is required for cell wall biosynthesis in rice (*Oriza sativa* L.). *Plant J.* 57, 446–462. doi: 10.1111/j.1365-313X.2008.03703.x
- Zielinska, D. F., Gnad, F., Wiśniewski, J. R., and Mann, M. (2010). Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141, 897–907. doi: 10.1016/j.cell.2010.04.012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 July 2014; accepted: 08 September 2014; published online: 02 October 2014.

*Citation: Nguema-Ona E, Vicré-Gibouin M, Gotté M, Plancot B, Lerouge P, Bardor M and Driouich A (2014) Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. *Front. Plant Sci.* 5:499. doi: 10.3389/fpls.2014.00499*

*This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.*

Copyright © 2014 Nguema-Ona, Vicré-Gibouin, Gotté, Plancot, Lerouge, Bardor and Driouich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Structural features of free *N*-glycans occurring in plants and functional features of de-*N*-glycosylation enzymes, ENGase, and PNGase: the presence of unusual plant complex type *N*-glycans

Megumi Maeda and Yoshinobu Kimura*

Functional Glycobiology, Department of Biofunctional Chemistry, Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Kazuhiro Fujiyama, Osaka University, Japan

Renaud Léonard, University Lille 1, France

***Correspondence:**

Yoshinobu Kimura, Functional Glycobiology, Department of Biofunctional Chemistry, Graduate School of Environmental and Life Science, Okayama University, 1-1-1, Tsushima-Naka, Kita-ku, Okayama, Japan
e-mail: yosh8mar@cc.okayama-u.ac.jp

Free *N*-glycans (FNGs) are present at micromolar concentrations in plant cells during their differentiation, growth, and maturation stages. It has been postulated that these FNGs are signaling molecules involved in plant development or fruit ripening. However, the hypothetical biochemical and molecular function of FNGs has not been yet established. The structure of FNGs found ubiquitously in plant tissues such as hypocotyls, leaves, roots, developing seeds, or fruits can be classified into two types: high-mannose type and plant complex type; the former, in most cases, has only one GlcNAc residue at the reducing end (GN1 type), while the latter has the chitobiosyl unit at the reducing end (GN2 type). These findings suggest that endo- β -*N*-acetylglucosaminidase (ENGase) must be involved in the production of GN1 type FNGs, whereas only peptide:*N*-glycanase (PNGase) is involved in the production of GN2 type FNGs. It has been hypothesized that cytosolic PNGase (cPNGase) and ENGase in animal cells are involved in the production of high-mannose type FNGs in order to release *N*-glycans from the misfolded glycoproteins in the protein quality control systems. In the case of plants, it is well known that another type of PNGase, the acidic PNGase (aPNGase) is involved in the production of plant complex type FNGs in an acidic organelle, suggesting the de-*N*-glycosylation mechanism in plants is different from that in animal cells. To better understand the role of these FNGs in plants, the genes encoding these *N*-glycan releasing enzymes (ENGase and PNGase) were first identified, and then structure of FNGs in ENGase knocked-out plants were analyzed. These transgenic plants provide new insight into the plant-specific de-*N*-glycosylation mechanism and putative physiological functions of FNGs. In this review, we focus on the structural features of plant FNGs, as well as functional features of cPNGase/ENGase and plant specific PNGase, and putative functions of FNGs are also discussed.

Keywords: free *N*-glycans, PNGase, ENGase, knockout plant, glycochaperone

STRUCTURAL FEATURES OF *N*-GLYCANS LINKED TO THE PLANT *N*-GLYCOPROTEINS

It is well known that almost all the secreted-type proteins produced in eukaryotic cells are *N*-glycosylated, suggesting that the *N*-glycosylation of proteins and the subsequent modification of the glycan moiety are ubiquitous and pivotal biological process in eukaryotes, especially the multicellular organisms (Lerouge et al., 1998; Kobata, 2007). In the protein quality-control mechanism working in the endoplasmic reticulum (ER), it is believed that these *N*-glycans are linked to the nascent glycoproteins in the ER and play a critical role as ligands for calnexin or calreticulin, molecular chaperons that assist the protein-folding system in the ER (Suzuki et al., 2007). In plant cells, for example, deletion of ER α -glucosidase I, an enzyme that hydrolyzes α -1,3-glucosidic linkages in the nascent and premature *N*-glycans and controls the interaction of the unfolded nascent glycoproteins and the molecular chaperons, results in death (Boisson-

et al., 2001; Gillmor et al., 2002). *N*-Glycans that are biosynthesized in eukaryotic cells (regardless of animal, plant, insect, fungi, and yeast) share a common trimannosyl core structure [$\text{Man}\alpha 1\text{-}6(\text{Man}\alpha 1\text{-}3)\text{Man}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4\text{GlcNAc}$], but the final structure of these *N*-glycans varies by species. In the case of animals, plants, and insects, the *N*-glycan structures are classified into three subgroups: high-mannose type, hybrid type, and complex type depending on the location of additional sugar residues transferred to the trimannosyl core. The high-mannose type *N*-glycans are common in all eukaryotic cells, but the structural features of complex-type *N*-glycans vary widely depending on the species. The structural features of plant *N*-glycans are as follows: (1) the occurrences of β 1,2-xylosyl (Xyl) residue linked to β 1,4-mannosyl (Man) residue and α 1,3-fucosyl (Fuc) residue linked to the reducing end *N*-acetylglucosaminyl (GlcNAc) residue, (2) the occurrence of Lewis a (Le^a) epitope [$\text{Gal}\beta 1\text{-}3(\text{Fuc}\alpha 1\text{-}4)\text{GlcNAc}$] at the non-reducing end, and (3) the lack of *N*-acetylneuraminic

acid residue and β 1,4-linkage galactosyl (Gal) residue. The combination of α 1,3-Fuc and β 1,2-Xyl linked to the trimannosyl core is an outstanding characteristic of the plant complex type N-glycans. The most abundant complex type N-glycan linked to storage glycoproteins in seeds or vacuole-accumulated glycoproteins is $\text{Man}\alpha 1\text{-}6(\text{Man}\alpha 1\text{-}3)(\text{Xyl}\beta 1\text{-}2)\text{Man}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4(\text{Fuc}\alpha 1\text{-}3)\text{GlcNAc}$, which is sometimes described as a truncated-type, pauci-mannose-type, or vacuolar-type structure. Furthermore, the secreted plant glycoproteins sometimes carry large complex type N-glycans with Le^a epitope. The plant N-glycans harboring the Le^a epitope, which are sometimes referred to as secreted type structures, have been found in many foodstuffs and pollen allergens (Fichette-Lainé et al., 1997; Alisi et al., 2001; Wilson et al., 2001; Kimura et al., 2005; Maeda et al., 2005), although their physiological function in plants remains to be elucidated. The conversion of N-glycan structures from high-mannose type to complex type seems to play an important role in adaptation to environmental changes. It has been reported that the maturation of N-glycans to the complex type N-glycans in the Golgi apparatus is a prerequisite for the sufficient cell-wall formation under salt stress (Kang et al., 2008). Typical structures of N-glycans linked to plant glycoproteins are shown in Figure 1, and the possible processing pathway of plant N-glycans is outlined in Figure 2.

STRUCTURAL FEATURES OF FREE N-GLYCANS IN PLANTS

As the name implies, N-glycans are linked to a specific asparagine (Asn) residue in eukaryotic proteins, but free forms of these glycans, namely free N-glycans (FNGs), have been found in seedlings, stems, developing fruits, and culture cells (Priem et al., 1990b, 1993; Faugeron et al., 1997a,b; Kimura et al., 1997; Kimura and

Matsuo, 2000; Kimura and Kitahara, 2000). Since similar FNGs have been found in animal cells, the occurrence of these FNGs is not just limited to plants. FNGs found in both plants and animals are classified into two types, GN1 type and GN2 type, based on the reducing terminal structure; GN1 type FNGs with one GlcNAc residue and GN2 type FNGs with $\text{GlcNAc}\beta 1\text{-}4\text{GlcNAc}$ (*N*-acetylchitobiosyl unit). The structures of typical FNGs found in plants are summarized in Figure 3.

It is generally believed that FNGs are produced from the misfolded glycoproteins by cytosolic peptide:N-glycanase (cPNGase), which hydrolyzes the amide linkage in the glycosylated Asn residue, prior to degradation of the misfolded proteins by proteasomes in the cytosol. The disposal of misfolded/miss-associated proteins is known as ER-associated degradation (ERAD), and the resulting FNGs produced by cPNGase always belong to high-mannose type structure with the *N*-acetylchitobiosyl unit at their reducing ends (GN2 type FNGs). As for plant cPNGase, it was confirmed that the product of putative *Arabidopsis* PNGase gene when transformed in budding yeast facilitates the glycoprotein ERAD activity, suggesting that the gene product probably possesses PNGase activity and is likely involved in deglycosylation of misfolded glycoproteins (Diepold et al., 2007; Masahara-Negishi et al., 2012). In the cytosolic metabolism of FNGs released from misfolded proteins by cPNGase, GN2 type FNGs are rapidly modified by cytosolic endo- β -N-acetylglucosaminidase (cENGase), which hydrolyzes the glycosidic linkage in the *N*-acetylchitobiosyl unit to form the GN1 type FNGs (Suzuki and Funakoshi, 2006). Since almost all high-mannose type FNGs belong to the GN1 type, it is clear that ENGase is involved in the production of the high-mannose type FNGs found in plants, but there is no experimental evidence suggesting that such high-mannose type

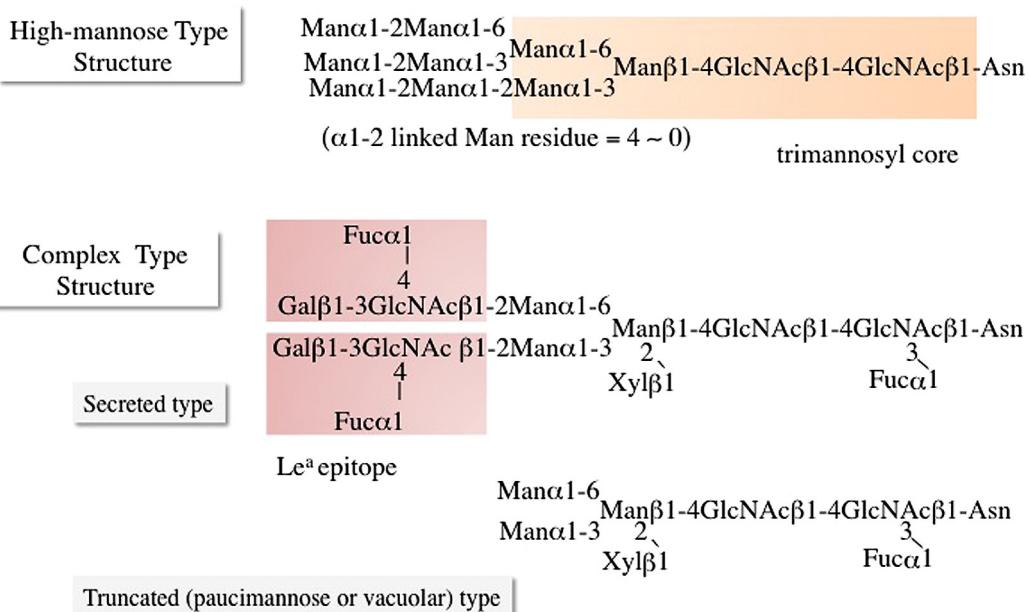


FIGURE 1 | Typical structures of plant N-glycans: high-mannose type structure and two kinds of complex type N-glycans.

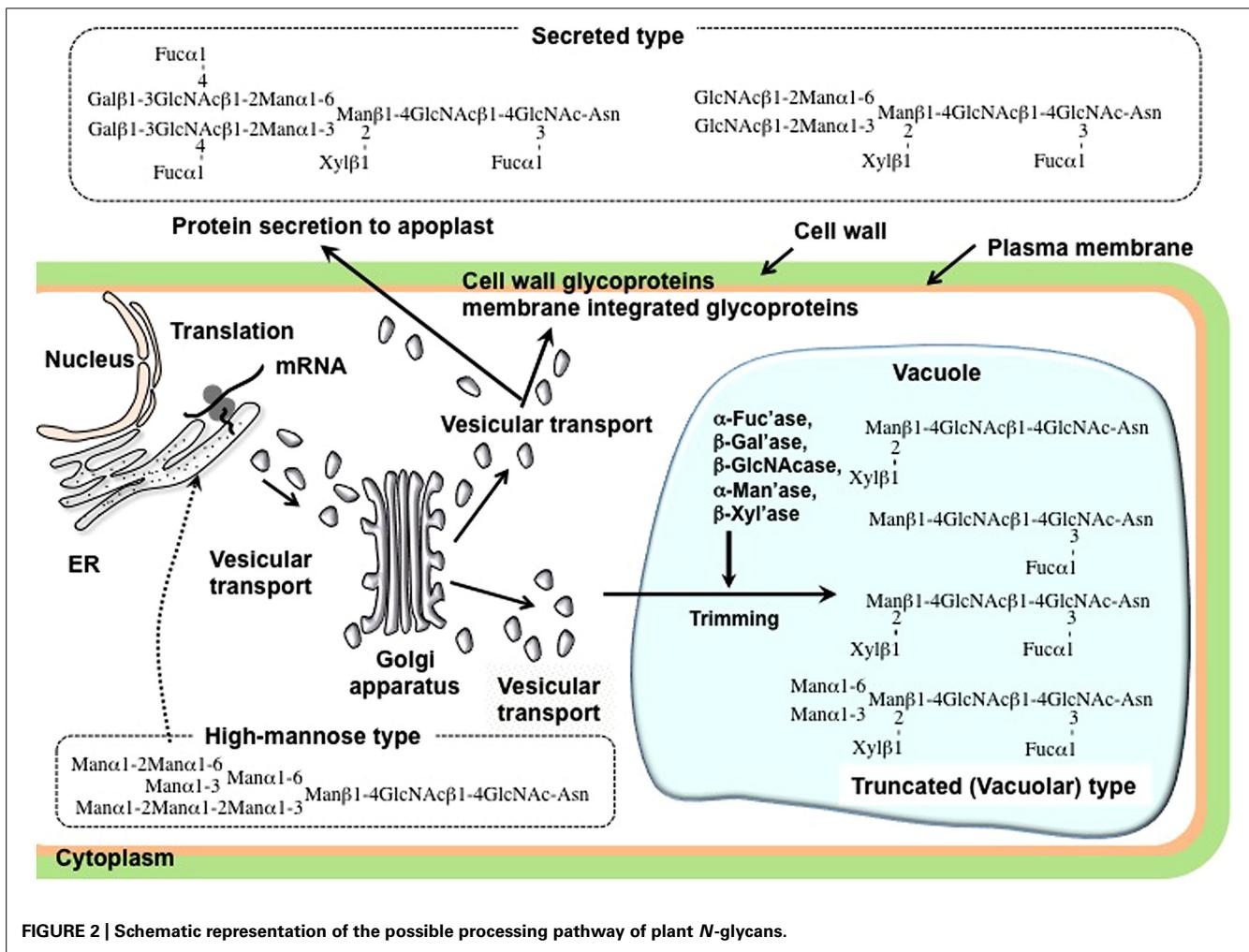


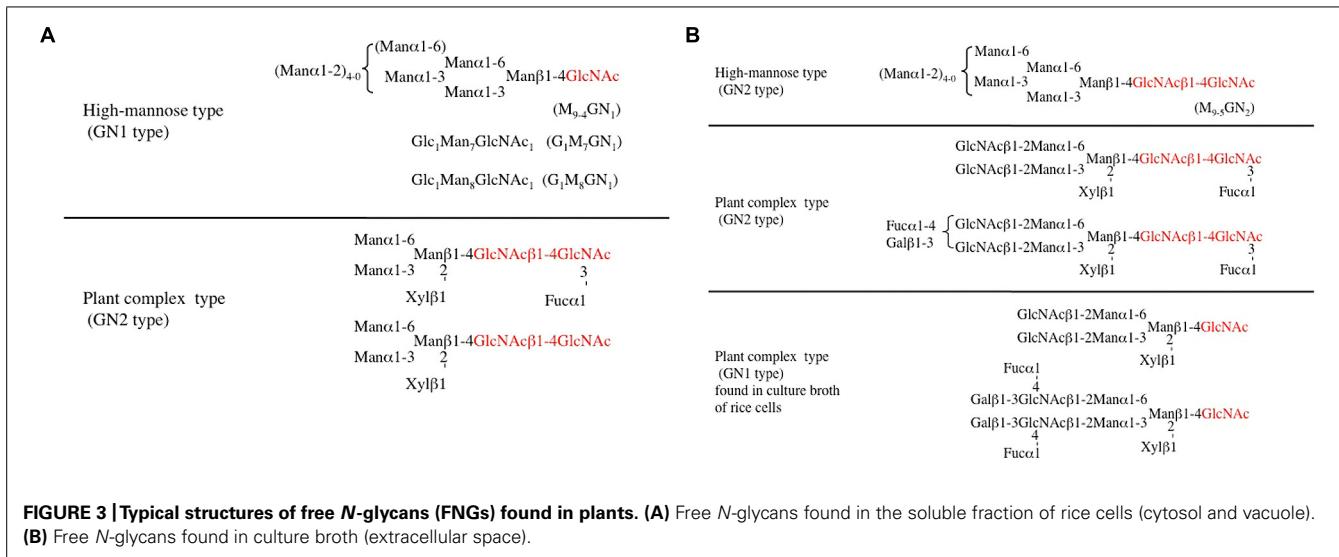
FIGURE 2 | Schematic representation of the possible processing pathway of plant N-glycans.

GN1-FNGs are exclusively generated from GN2 FNGs (cPNGase products) and not from the misfolded glycoproteins by the action of ENGase.

On the other hand, almost all plant complex type and truncated type FNGs belong to GN2 type, suggesting that only PNGase, but not ENGase, must be involved in the production. As we discuss later, the plant ENGase shows strong activity against high-mannose type N-glycans having Man α 1-2Man α 1-3Man β 1-unit, but no activity against plant complex and truncated structures having GlcNAc β 1-4(Fuc α 1-3)GlcNAc-unit at their reducing ends. It is, therefore, reasonable to conclude that almost all plant complex type FNGs have the GN2 type structure. Based on the biosynthesis or processing mechanism of N-glycans linked to plant glycoproteins, as shown in Figure 2, it can be considered that these complex type FNGs must be released from matured and secreted glycoproteins fully processed in the Golgi apparatus, but not from the misfolded glycoproteins that are fated to be degraded in the cytosol. Plant specific aPNGase involved in the production of plant complex type FNGs is genetically different from cPNGase localized in the cytosol and has the optimum activity in the acidic environment. The occurrence of two subtypes of GN2-FNGs, with the truncated structure and large

size structure as shown in Figure 3, suggests that there may be two kinds of aPNGase in plants; one works in the vacuole and the other in the extracellular space such as cell wall or the apoplast.

As described above, the structural features of FNGs found in plants indicate their origins (misfolded proteins in ER or fully processed and secreted proteins) and the enzyme, PNGase or ENGase, that is involved in the final stage of the FNGs-production. For example, structural analysis of FNGs found in cultured rice cells and culture broth provided valuable information on the deglycosylation mechanism in plant cells (Maeda et al., 2010). In the soluble fraction (a mixture of cytosolic and vacuolar fractions) of cultured cells, only high-mannose type GN1-FNGs and truncated complex type GN2-FNGs were found, suggesting that the former FNGs were produced by ENGase and the latter by aPNGase. It is believed that the plant complex type N-glycans linked to vacuolar-targeting glycoproteins are modified to the truncated type structure in the vacuole (Vitale and Chrispeels, 1984; Lerouge et al., 1998), and it is reasonable to assume that truncated complex type FNGs must be produced from deteriorated glycoproteins (or glycopeptides generated by endopeptidases) by vacuolar PNGase. On the other hand, high-mannose type



GN2-FNGs ($M_{9-5}GN_2$) and two longer complex type FNGs having *N*-acetylglucosamine residue(s) or Le^a unit(s) at the non-reducing end (GN1 and GN2 species) have been identified from the culture broth (Figure 3B; Maeda et al., 2010). The occurrence of GN2-FNGs (both high-mannose type and longer complex type structures) in the culture broth suggests that the secreted-type aPNGase in either the cell wall or apoplastic area plays a critical role in the release of *N*-glycans in the extracellular space. In the structural analysis of rice FNGs, some plant complex type FNGs with unique structural feature were found in the culture broth from rice cell (Maeda et al., 2010). As shown in Figure 3B, the unusual FNGs carry the complex type structure at their non-reducing end and only one GlcNAc residue at their reducing end, suggesting that these FNGs were fully processed in the Golgi apparatus and the ENGase were involved in the process. As described previously, ENGases that have been purified and characterized to date cannot hydrolyze the glycosidic linkage of the *N*-acetylchitobiosyl unit of plant complex type *N*-glycans, and no other ENGase that is active against plant complex type *N*-glycans has been found. In summary, the occurrence of the plant complex type GN1-FNGs cannot be explained by the established scheme of deglycosylation and processing mechanism of *N*-glycoproteins. On the other hand, in the case of animal cells, the complex type GN1-FNGs were recently found in the stomach cancer cultured cells (Ishizuka et al., 2008), suggesting that the occurrence of complex type GN1-FNGs is not limited to plants. However, the animal complex type *N*-glycans usually carry α 1-6 fucosyl (but not α 1-3 fucosyl) residue linked to the innermost GlcNAc residue, and this type of animal complex type glycans can be a modest substrate for some types of ENGase. At present, therefore, the possibility of involvement of ENGase in the production of complex type GN1-FNGs in animal cells cannot be excluded.

A putative processing scheme for the formation of the unusual FNGs, the plant complex type GN1-FNGs, has been proposed on the basis of their chemical structure (Figure 4), the substrate specificity of plant ENGase, and the *N*-glycan processing mechanism

in the ER and Golgi apparatus (Maeda et al., 2010). Prior to discussion of the production mechanism and physiological functions of plant complex type FNGs, we would like to describe some properties of plant ENGase and PNGase, including localization, substrate specificity, and gene structure, in the following section.

FUNCTIONAL CHARACTERIZATION OF CYTOSOLIC PNGase, ACIDIC PNGase, AND ENGase IN PLANTS

CYTOSOLIC PNGase

Cytosolic PNGases in budding yeast (Suzuki et al., 2000), nematodes (Kato et al., 2002), and mammals (Suzuki et al., 2002) are well characterized and their gene structures identified. It is widely accepted that these cPNGase are localized in the cytosol and involved in the release of high-mannose type *N*-glycans from misfolded glycoproteins found in the ER. These cPNGase have the optimum activity in the neutral pH region and are active against denatured *N*-glycoproteins and *N*-glycopeptides bearing both high-mannose and animal complex type structures. At the same time, the glycosylated Asn residues are converted to Asp residues by the action of PNGase. It is noteworthy, however, that animal complex type *N*-glycoproteins cannot be substrates *in vivo* (or cPNGase cannot encounter these glycoproteins in cells), because cPNGase resides in the cytosol and *N*-glycoproteins carrying complex type *N*-glycans are secreted to the extracellular space through the ER and Golgi apparatus. In the case of de-*N*-glycosylation of misfolded glycoprotein in animal cytosol, the misfolded glycoproteins that retrotranslocated to the cytosol from the ER are first ubiquitinated by lectin-ubiquitin ligase Fbx2 complex, and then deglycosylated by cPNGase associated with 26S proteasome through HR23B (Rad23p in budding yeast; Suzuki et al., 2001, 2007). The resulting deglycosylated misfolded proteins are then degraded in the cytosolic proteasome complex. In contrast, the role of plant cPNGase involved in deglycosylation mechanism functioning in plant cells remains to be fully elucidated. The *Arabidopsis* cPNGase gene was identified (Diepold et al., 2007), and the cPNGase activity was recently

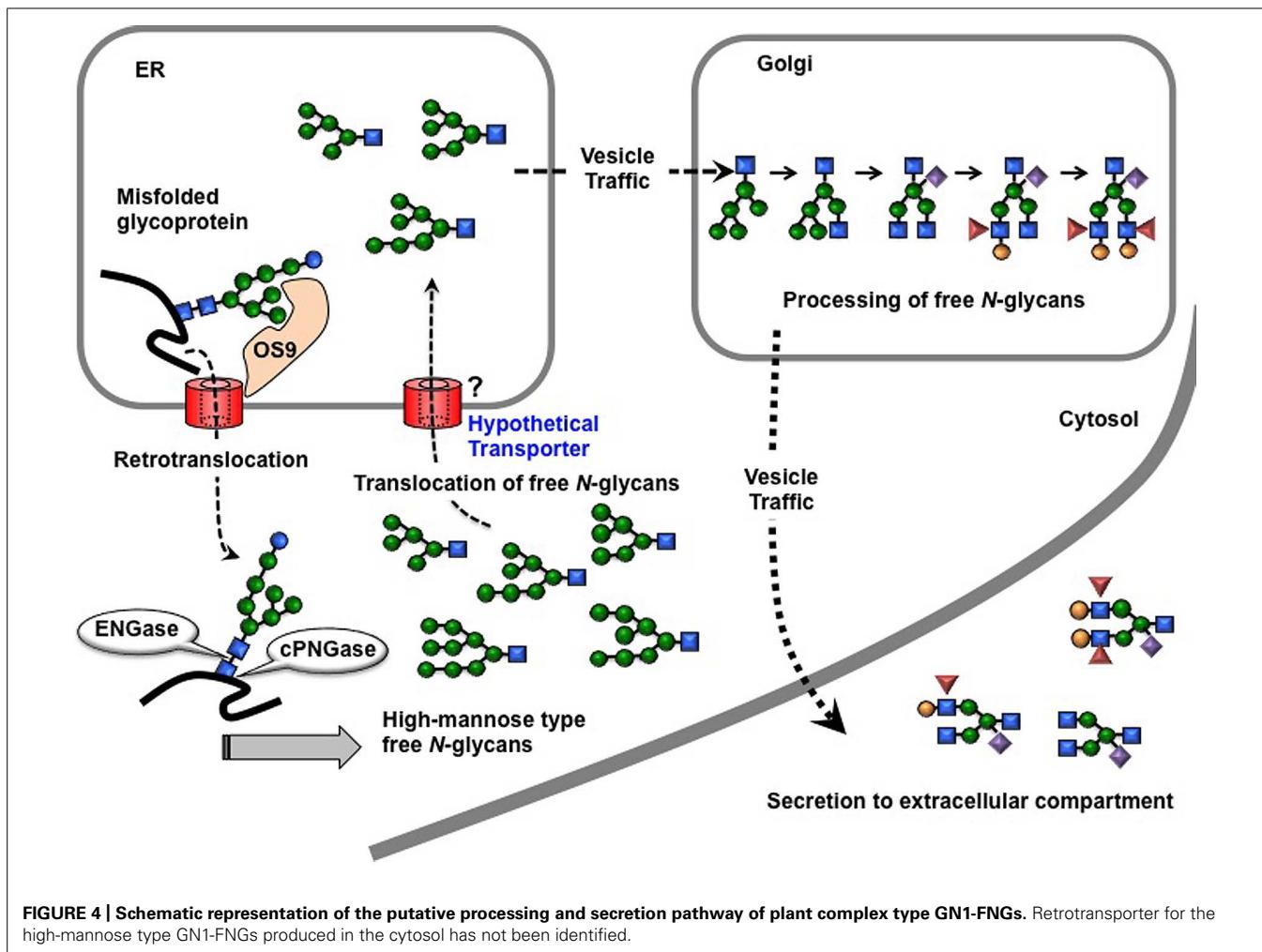


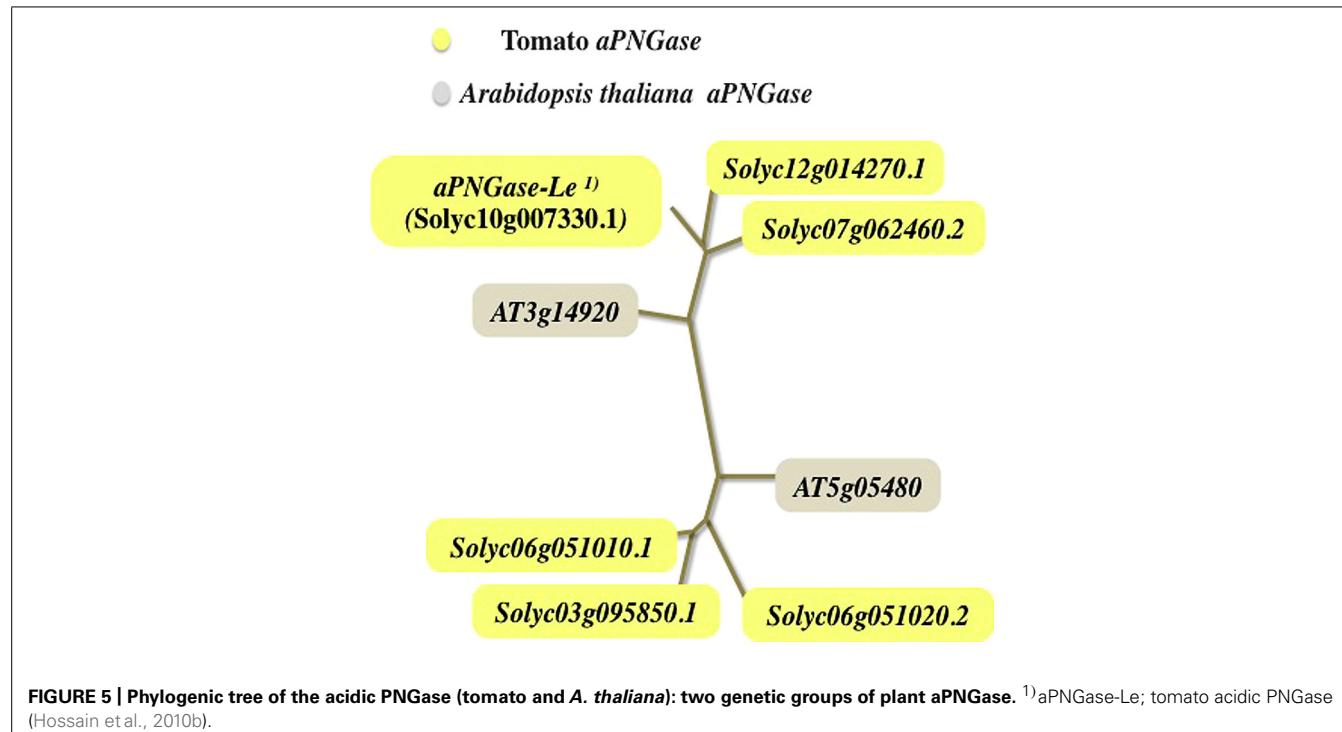
FIGURE 4 | Schematic representation of the putative processing and secretion pathway of plant complex type GN1-FNGs. Retrotransporter for the high-mannose type GN1-FNGs produced in the cytosol has not been identified.

confirmed by the transformation of the putative plant cPNGase gene into a mutant yeast strain lacking its own PNGase activity (Masahara-Negishi et al., 2012). In the transformed mutant yeast, the degradation of ERAD substrates (or misfolded glycoproteins) was facilitated in the N-glycan-dependent manner, indicating that the product of the putative cPNGase gene was involved in the de-N-glycosylation of misfolded glycoproteins. However, cPNGase activity has not been directly detected *in vitro* using misfolded glycoproteins or glycopeptides bearing high-mannose type N-glycans as substrates. Given that the cPNGases purified from animals or yeast can easily release the N-glycans from denatured N-glycoproteins or N-glycopeptides *in vitro*, the plant cPNGase may require some other specific factors to display the full amidase activity.

PLANT SPECIFIC ACID PNGase

It is well known that another species of PNGase occurs in plants, and this kind of PNGase (aPNGase) was first purified and characterized from almond seeds (Takahashi, 1977; Takahashi and Nishibe, 1978). The plant aPNGase has the optimum activity in the acidic pH region and is active against glycopeptides carrying high-mannose type or plant complex type N-glycans

(Takahashi, 1977; Takahashi and Nishibe, 1978; Lhernould et al., 1995; Altmann et al., 1998; Kimura and Ohno, 1998; Chang et al., 2000). However, both native N-glycoproteins and denatured N-glycoproteins are not suitable substrates for the aPNGase, suggesting that the plant aPNGase may play a role in the release of N-glycans from glycopeptides generated by the proteolysis of denatured glycoproteins in acidic environment such as vacuole or extracellular space. In 2010, a gene encoding tomato aPNGase was identified and the gene product when expressed in yeast showed the typical aPNGase activity (Hossain et al., 2010b). Since the orthologs of the tomato aPNGase gene were found among plants and fungi but not in animals, aPNGase must have certain specific physiological function in the development or proliferation of plants and fungi. To elucidate the plant physiological role of aPNGase, construction of transgenic plants in which the aPNGase gene is suppressed or overexpressed is necessary as the first step. Furthermore, the localization of aPNGase has not been clarified to date. Since the plant complex type FNGs ($\text{GlcNAc}_2\text{Man}_3\text{Xyl}_1\text{Fuc}_1\text{GlcNAc}_2$ as an example) and the truncated type FNGs ($\text{Man}_3\text{Xyl}_1\text{Fuc}_1\text{GlcNAc}_2$ as an example) have been found in several plant materials, it can be assumed that there may be at least two types of



aPNGase; one working in the vacuole to release the truncated type N-glycans and the other in extracellular space to release the longer complex type glycans. Indeed, at least two genetic groups of aPNGase have been found in the tomato genome as shown in **Figure 5**. If two types of aPNGase reside in different space (vacuole and extracellular space), it is important to reveal the molecular mechanism that directs the two types of PNGase to the adequate organelle or extracellular space in plants.

PLANT ENGase

Endo- β -N-acetylglucosaminidase (ENGase) is another glycoenzyme involved in the degradation pathway of misfolded glycoproteins, and hydrolyzes the β 1-4 GlcNAc linkage in high-mannose type N-glycans linked to glycoproteins (glycopeptides) or GN2-FNGs. Bacterial or fungal ENGases have been used as tools for structural analysis of N-glycans over a period of years. Especially, a fungus ENGase (Endo-M, ENGase from *Mucor hiemalis*) that has high transglycosylation activity has been applied for remodeling of biological active glycoproteins or glycopeptides (Yamamoto, 2006, 2012). On the other hand, relatively recently the animal and plant ENGases were purified and genetically characterized, although it has been reported that ENGase occurs in both animals and plants. The animal ENGase was first purified and characterized from hen oviduct (Kato et al., 1997), and then the gene structures of nematode and human ENGase were identified (Kato et al., 2002; Suzuki et al., 2002). The animal ENGases as well as the fungal ENGases (Endo-M) belong to glycoside hydrolase (GH) family 85 and these eukaryotic ENGases are genetically different from actinomycetal or bacterial ENGase such as Endo-H or Endo-F (GH Family 18).

As for plant ENGase, the functional characterization and the gene identification of the plant ENGase were reported, although the ENGase activities had been detected from several plant materials. The plant ENGase was first purified and characterized from cultured rice cells and the genes of rice, tomato, and *Arabidopsis* ENGases were successively identified (Kimura, 2007; Nakamura et al., 2009; Kimura et al., 2011). Plant ENGases have molecular weight of about 70 kDa and the optimum activity at near neutral pH (\sim pH 6.5). The gene structures show that plant enzymes belong to GH family 85, suggesting that they have the transglycosylation activity like Endo-M. Plant ENGases are highly active against high-mannose type N-glycans bearing the following trimannosyl unit, Man α 1-2Man α 1-3Man1 β -, such as Man₉-₆GlcNAc₂, but the activity dramatically decreases against the core structure of the high-mannose type structure (Man α 1-6(Man α 1-3)Man α 1-6(Man α 1-3)Man β 1-4GlcNAc β 1-4GlcNAc) lacking in α 1-2 Man residue. The plant complex type N-glycans bearing β 1-2Xyl and/or α 1-3Fuc residues cannot be a substrate for plant ENGase, suggesting that plant ENGase is involved in the degradation mechanism of misfolded glycoproteins carrying exclusively high-mannose type N-glycans in the cytosol. Subcellular localization analysis using ultracentrifugation and immunofluorescence techniques have revealed that plant ENGase resides primarily in the cytosol (Kimura, 2007; Kimura et al., 2002). However, when the subcellular localization was analyzed by ultracentrifugation technique (Kimura et al., 2002), the ENGase activity was slightly detected in the ER fraction, suggesting that small amount of the ENGase may reside in the ER or ENGase in the cytosol and may interact weakly with the cytosolic side of the ER membrane. As described earlier, almost all high-mannose type FNGs in plants belong to GN1 type, indicating that the reaction of plant ENGase against

high-mannose type GN2-FNGs (cPNGase products) or misfolded glycoproteins carrying high-mannose type *N*-glycans must be highly efficient.

BIOLOGICAL FUNCTIONS OF FREE *N*-GLYCANS IN PLANTS AS SIGNALING MOLECULES

In the 1990s, auxin-like activity of plant FNGs to stimulate elongation of stems in seedlings or maturation of tomato fruit was proposed (Priem et al., 1990a; Priem and Gross, 1992; Yunovitz and Gross, 1994). As part of the study to elucidate the physiological function of FNGs involved in fruit maturation process, Nakamura et al. analyzed changes in the amount of FNGs, ENGase activities, and ENGase gene expression during tomato fruit maturation (Nakamura et al., 2008, 2009). They found that the amount of high-mannose type GN1-FNGs increased significantly with fruit maturation (mature green, breaker, pink, and mature red), but the enzyme activity (as total activity) and gene expression were nearly unchanged. These results suggested that tomato ENGase is constantly expressed during tomato fruit maturation step to produce high-mannose type GN1-FNGs. However, the amount of substrates (GN2-FNGs or misfolded glycoproteins) may increase or α -mannosidase activity (Hossain et al., 2010a; Meli et al., 2010) responsible for degradation of the GN1-FNGs may decrease. To examine the putative auxin-like activity to stimulate plant development, the transgenic *A. thaliana* plants, in which two ENGase genes were knocked out, were constructed (Kimura et al., 2011; Fischl et al., 2011). The high-mannose type GN1-FNGs found predominately in the wild plants were completely converted to GN2-FNGs in the double knockout plants and the ENGase activity was completely lost in the mutant plants, clearly indicating that plant cPNGase, as well as the animal cPNGase, is involved in the production of FNGs prior to the action of ENGase. Two single knockout (*At3g11040* or *At5g05460*) plants produced the high-mannose type GN1-FNGs, and the ENGase activities found in two single knockout plants and wild plants were comparable to each other, but the structural features of FNGs found in these single knockout mutants were slightly different between each other, suggesting that these two ENGases (*At3g11040* or *At5g05460*) may reside in the cytosol of different tissues (roots, stems, seeds, or leaves). No apparent morphological changes were observed among the wild-type plants, two single knockout lines, and double knockout plants under the normal cultivation conditions (Fischl et al., 2011; Kimura et al., 2011). The construction of double knockout plants, in which both ENGase and cPNGase genes are completely suppressed, is the prerequisite to elucidate physiological function(s) of FNGs involved in plant differentiation, growth, and fruit maturation.

PUTATIVE MECHANISM RESPONSIBLE FOR GENERATION OF THE PLANT COMPLEX TYPE GN1-FNGS

As described previously, the plant complex type FNGs bearing one GlcNAc residue (GN1-FNGs) at their reducing ends were first found as secreted FNGs in the rice-cell culture broth. Since these FNGs were not the truncated type and some of them carried the Le^a epitope, these FNGs must have been fully processed in the Golgi apparatus and secreted to extracellular space (but not

sorted to the vacuole) through the trans Golgi network. If special ENGase that are active against plant complex type *N*-glycans would occur in the plant extracellular space, the occurrence of such kind of FNGs (plant complex type GN1-FNGs) could be reasonably explained. However, such special ENGase activity has not been found in plants to date. Considering the structural features of the unique GN1-FNGs, subcellular localization of ENGase and PNGase, and *N*-glycan processing pathway, a putative mechanism of the unique GN1-FNGs has been proposed (Figure 4; Maeda et al., 2010). In this scheme, it has been assumed that (1) the high-mannose type GN1-FNGs may be first formed by a combination of cPNGase and ENGase in the cytosol, (2) some part of the resulting GN1-FNGs may be transported from the cytosol to the ER through a putative translocator (similar to TAP, one of ABC transporters involved in the cellular immune system) and transported to the Golgi apparatus, and then (3) the high-mannose type *N*-glycans may be processed to the longer plant complex type structures and secreted to the extracellular space together with fully folded and processed proteins.

It is noteworthy to consider another interesting function of *N*-glycan involved in protein folding or reconstruction of oligomeric proteins. In a previous study (Kimura et al., 1999), it was reported that high-mannose type *N*-glycans (Man₉GlcNAc₂ or Glc1Man₉GlcNAc₂) linked to the jack bean α -mannosidase stimulated or induced the reconstruction from denatured and dissociated structure to functional oligomeric structure to recover the enzyme activity. Furthermore, Kimura et al. (1998) and Jitsuhara et al. (2002) reported that glycosylated Asn (Asn-glycan) functioned as a kind of molecular chaperon to reconstruct the functional conformation of some enzymes from denatured state, suggesting that at least the high-mannose type FNGs can play a critical role in the protein-refolding. Since it is well known that sugar molecules or polyethylene glycol stabilize protein structure and prevent protein aggregations, it is not a total surprise that the FNGs produced from glycoproteins show similar function. On the other hand, in the context of generation and degradation mechanisms of FNGs accepted to date, it is difficult to assume that unfolded or prefolded proteins in the ER encounter the high-mannose type FNGs that are formed in the cytosol. However, if there is a possibility that the high-mannose type FNGs produced in the cytosol can be transported to the ER, as postulated above, the unfolded or premature proteins, both glycoproteins and non-glycoproteins, in the ER may be able to receive the sweet benefit from the free sugar chains holding the latent chaperon-like function. To prove the hypothetical chaperon-like function of FNGs, it must be essential to confirm the occurrence of FNGs in the ER and identify the putative transporter responsible for the translocation of FNGs. Furthermore, on the platform of the biological activity of FNGs involved in the protein folding, it will be possible to develop a new glyco-technology in which FNGs are used to promote refolding of recombinant proteins *in vitro*.

REFERENCES

- Alisi, C., Afferni, C., Iacovacci, P., Barletta, B., Tinghino, R., Butteroni, C., et al. (2001). Rapid isolation, characterization, and glycan analysis of Cup a 1, the major allergen of Arizona cypress (*Cupressus arizonica*) pollen. *Allergy* 56, 978–984. doi: 10.1034/j.1365-9995.2001.103125.x

- Altmann, F., Paschinger, K., Dalik, T., and Vorauer, K. (1998). Characterisation of peptide-N4-(*N*-acetyl- β -glucosaminyl)asparagine amidase A and its N-glycans. *Eur. J. Biochem.* 252, 118–123. doi: 10.1046/j.1432-1327.1998.25201118.x
- Boisson, M., Gomord, V., Audran, C., Berger, N., Dubreucq, D., Granier, F., et al. (2001). *Arabidopsis glucosidase I* mutants reveal a critical role of N-glycan trimming in seed development. *EMBO J.* 20, 1010–1019. doi: 10.1093/emboj/20.5.1010
- Chang, T., Kuo, M. C., Khoo, K. H., Inoue, S., and Inoue, Y. (2000). Developmentally regulated expression of a peptide:N-glycanase during germination of rice seeds (*Oryza sativa*) and its purification and characterization. *J. Biol. Chem.* 275, 129–134. doi: 10.1074/jbc.275.1.129
- Diepold, A., Li, G., Lennarz, W. J., Nürnberg, T., and Brunner, F. (2007). The *Arabidopsis AtPNG1* gene encodes a peptide:N-glycanase. *Plant J.* 52, 94–104. doi: 10.1111/j.1365-313X.2007.03215.x
- Faugeron, C., Lhernould, S., Lemoine, J., Costa, G., and Morvan, H. (1997a). Identification of unconjugated N-glycans in strawberry plants. *Plant Physiol. Biochem.* 35, 891–895.
- Faugeron, C., Lhernould, S., Maes, E., Lerouge, P., Strecker, G., and Morvan, H. (1997b). Tomato plant leaves also contain unconjugated N-glycans. *Plant Physiol. Biochem.* 35, 73–79.
- Fichette-Lainé, A. C., Gomord, V., Cabanes, M., Michalski, J. C., Macary, M. S., Foucher, B., et al. (1997). N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J.* 12, 1411–1417. doi: 10.1046/j.1365-313X.1997.12061411.x
- Fischl, R. M., Stadlmann, J., Grass, J., Altman, F., and Léonard, R. (2011). The two endo- β -N-acetylglucosaminidase genes from *Arabidopsis thaliana* encode cytoplasmic enzymes controlling free N-glycan levels. *Plant Mol. Biol.* 77, 275–284. doi: 10.1007/s11103-011-9808-7
- Gillmor, S. C., Poindexter, P., Lorieau, J., Palcic, M. M., and Somerville, C. (2002). α -Glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. *J. Cell. Biol.* 156, 1003–1013. doi: 10.1083/jcb.200111093
- Hossain, M. A., Nakano, R., Nakamura, K., Hossain, M. T., and Kimura, Y. (2010a). Molecular characterization of plant acidic α -mannosidase, a member of glycosyl-hydrolase family 38, involved in the turnover of N-glycans during tomato fruit ripening. *J. Biochem.* 148, 603–616. doi: 10.1093/jb/mvq094
- Hossain, M. A., Nakano, R., Nakamura, K., and Kimura, Y. (2010b). Molecular identification and characterization of an acidic peptide:N-glycanase (PNGase) from tomato (*Lycopersicum esculentum*) fruits. *J. Biochem.* 147, 157–165. doi: 10.1093/jb/mvq157
- Ishizuka, A., Hashimoto, Y., Naka, R., Kinoshita, M., Kakehi, K., Seino, J., et al. (2008). Accumulation of free complex-type N-glycans in MKN7 and MKN45 stomach cancer cells. *Biochem. J.* 413, 227–237. doi: 10.1042/BJ20071562
- Jitsuhara, Y., Toyoda, T., Itai, T., and Yamaguchi, H. (2002). Chaperone-like functions of high-mannose type and complex-type N-glycans and their molecular basis. *J. Biochem.* 132, 803–811. doi: 10.1093/oxfordjournals.jbchem.a003290
- Kang, J. S., Frank, J., Kang, C. H., Kajiru, H., Vikram, M., Ueda, A., et al. (2008). Salt tolerance of *Arabidopsis thaliana* requires mutation of N-glycosylated proteins in the Golgi apparatus. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5933–5938. doi: 10.1073/pnas.0800237105
- Kato, T., Fujita, K., Takeuchi, M., Kobayashi, K., Natsuka, S., Ikura, K., et al. (2002). Identification of an endo- β -N-acetylglucosaminidase gene in *Caenorhabditis elegans* and its expression in *Escherichia coli*. *Glycobiology* 12, 581–587. doi: 10.1093/glycob/cwf073
- Kato, T., Hatanaka, K., Mega, T., and Hase, S. (1997). Purification and characterization of endo- β -N-acetylglucosaminidase from hen oviduct. *J. Biochem.* 122, 1167–1173. doi: 10.1093/oxfordjournals.jbchem.a021877
- Kimura, N., Uchida, M., Nishimura, S., and Yamaguchi, H. (1998). Promotion of polypeptide folding by interactions with Asn-Glycans. *J. Biochem.* 124, 857–862. doi: 10.1093/oxfordjournals.jbchem.a022190
- Kimura, Y. (2007). “Structural and functional features of plant glycoprotein glycans,” in *Comprehensive Glycoscience—From Chemistry to Systems Biology*, Vol. 3, ed. H. Kamerling (Oxford: Elsevier), 61–76.
- Kimura, Y., Hess, D., and Strum, A. (1999). The N-glycans of jack bean α -mannosidase. Structure, topology and function. *Eur. J. Biochem.* 264, 168–175. doi: 10.1046/j.1432-1327.1999.00598.x
- Kimura, Y., Kamamoto, M., Maeda, M., Okano, M., Yokoyama, M., and Kino, K. (2005). Occurrence of Lewis a epitope in N-glycans of a glycoallergen, Jun a 1, from mountain cedar (*Juniperus ashei*) pollen. *Biosci. Biotechnol. Biochem.* 69, 137–144. doi: 10.1271/bbb.69.137
- Kimura, Y., and Kitahara, E. (2000). Structural analysis of free N-glycans occurring in soybean seedlings. *Biosci. Biotechnol. Biochem.* 64, 1847–1855. doi: 10.1271/bbb.64.1847
- Kimura, Y., and Matsuo, S. (2000). Free N-glycans already occur at an early stage of seed development. *J. Biochem.* 127, 1013–1019. doi: 10.1093/oxfordjournals.jbchem.a022692
- Kimura, Y., Matsuo, S., Tsurusaki, S., Kimura, M., Nishimura-Hara, I., and Nishimura, M. (2002). Subcellular localization of endo- β -N-acetylglucosaminidase and high-mannose type free N-glycans in plant cell. *Biochim. Biophys. Acta* 1570, 38–46. doi: 10.1016/S0304-4165(02)00149-6
- Kimura, Y., and Ohno, A. (1998). A new peptide-N⁴-(acetyl- β -glucosaminyl)asparagine amidase from soybean (*Glycine max*) seeds: purification and substrate specificity. *Biosci. Biotech. Biochem.* 62, 412–418. doi: 10.1271/bbb.62.412
- Kimura, Y., Takagi, S., and Shiraishi, T. (1997). Occurrence of free N-glycans in pea (*Pisum sativum* L.) seedlings. *Biosci. Biotechnol. Biochem.* 61, 924–926. doi: 10.1271/bbb.61.924
- Kimura, Y., Takeoka, Y., Inoue, M., Maeda, M., and Fujiyama, K. (2011). Double-knockout of putative endo- β -N-acetylglucosaminidase (ENGase) genes in *Arabidopsis thaliana*: loss of ENGase activity induced accumulation of high-mannose type free N-glycans bearing N,N'-acetylchitobiosyl unit. *Biosci. Biotech. Biochem.* 75, 1019–1021. doi: 10.1271/bbb.110148
- Kobata, A. (2007). “Glycoprotein glycan structures,” in *Comprehensive Glycoscience—From Chemistry to Systems Biology*, ed. H. Kamerling (Oxford: Elsevier), 1, 39–72.
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A. C., Gomord, V., and Faye, L. (1998). N-Glycoprotein biosynthesis in plants: recent developments and future trends. *Plant. Mol. Biol.* 38, 31–48. doi: 10.1023/A:1006012005654
- Lhernould, S., Karamanos, Y., Lerouge, P., and Morvan, H. (1995). Characterization of the peptide-N⁴-(N-acetylglucosaminyl)asparagine amidase (PNGase Se) from *Silene alba* cells. *Glycoconjugate J.* 12, 94–98. doi: 10.1007/BF00731874
- Maeda, M., Kamamoto, M., Yamamoto, S., Kimura, M., Okano, M., and Kimura, Y. (2005). Glycoform analysis of Japanese cedar pollen allergen, Cry j1. *Biosci. Biotechnol. Biochem.* 69, 1700–1705. doi: 10.1271/bbb.69.1700
- Maeda, M., Kimura, M., and Kimura, Y. (2010). Intracellular and extracellular free N-glycans produced by plant cells: occurrence of unusual plant complex-type free N-glycans in extracellular spaces. *J. Biochem.* 148, 681–692. doi: 10.1093/jb/mvq102
- Masahara-Negishi, Y., Hosomi, A., Della Mea, M., Serafini-Fracassini, D., and Suzuki, T. (2012). A plant peptide: N-glycanase orthologue facilitates glycoprotein ER-associated degradation in yeast. *Biochim. Biophys. Acta* 1820, 1457–1462. doi: 10.1016/j.bbagen.2012.05.009
- Meli, V. S., Ghosh, S., Prabha, T. N., Chakraborty, N., Chakraborty, S., and Datta, A. (2010). Enhancement of fruit shelf life by suppressing N-glycan processing enzymes. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2413–2418. doi: 10.1073/pnas.0909329107
- Nakamura, K., Inoue, M., Nakano, R., Hosoi, K., and Kimura, Y. (2009). Molecular cloning and gene expression analysis of tomato endo- β -N-acetylglucosaminidase, an endoglycosidase involved in the production of high-mannose type free N-glycans during tomato fruit ripening. *Biosci. Biotechnol. Biochem.* 73, 461–464. doi: 10.1271/bbb.80889
- Nakamura, K., Inoue, M., Yoshiie, T., Hosoi, K., and Kimura, Y. (2008). Changes in structural features of free N-glycan and endoglycosidase activity during tomato fruit ripening. *Biosci. Biotechnol. Biochem.* 72, 2936–2945. doi: 10.1271/bbb.80414
- Priem, B., Gitti, R., Bush, C. A., and Gross, K. C. (1993). Structure of ten free N-glycans in ripening tomato fruit. Arabinose is a constituent of a plant N-glycan. *Plant Physiol.* 102, 445–458. doi: 10.1104/pp.102.2.445
- Priem, B., and Gross, K. C. (1992). Mannosyl- and xylosyl-containing glycans promote tomato (*Lycopersicon esculentum* Mill.) fruit ripening. *Plant Physiol.* 98, 399–401. doi: 10.1104/pp.98.1.399
- Priem, B., Morvan, H., Hafez, A. M. A., and Morvan, C. (1990a). Influence d'un xylmannoside d'origine végétale sur l'elongation de l'hypocotyle de Lin. *C. R. Acad. Sci. Paris.* 311, 411–416.
- Priem, B., Solokwan, J., Wierszeski, J. M., Strecker, G., Nazih, H., and Morvan, H. (1990b). Isolation and characterization of free glycans of the oligomannoside

- type from the extracellular medium of a plant cell suspension. *Glycoconjugate J.* 7, 121–132. doi: 10.1007/BF01050375
- Suzuki, T., and Funakoshi, Y. (2006). Free N-linked oligosaccharide chains: formation and degradation. *Glycoconjugate J.* 23, 291–302. doi: 10.1007/s10719-006-6975-x
- Suzuki, T., Park, H., Hollingsworth, N. M., Sternglanz, R., and Lennarz, W. J. (2000). *PNG1*, a yeast gene encoding a highly conserved peptide:N-glycanase. *J. Cell. Biol.* 29, 1039–1052. doi: 10.1083/jcb.149.5.1039
- Suzuki, T., Park, H., Kwofie, M. A., and Lennarz, W. J. (2001). Rad23 provides a link between the Png1 deglycosylating enzyme and the 26 S proteasome in yeast. *J. Biol. Chem.* 276, 21601–21607. doi: 10.1074/jbc.M100826200
- Suzuki, T., Tanabe, K., and Funakoshi, Y. (2007). “Folding and quality control of glycoproteins,” in *Comprehensive Glycoscience—From Chemistry to Systems Biology*, Vol. 3, ed. H. Kamerling (Oxford: Elsevier), 129–149.
- Suzuki, T., Yano, K., Sugimoto, S., Kitajima, K., Lennarz, W. J., Inoue, S., et al. (2002). Endo- β -N-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9691–9696. doi: 10.1073/pnas.152333599
- Takahashi, N. (1977). Demonstration of a new amidase acting on glycopeptides. *Biochem. Biophys. Res. Commun.* 76, 1194–1201. doi: 10.1016/0006-291X(77)90982-2
- Takahashi, N., and Nishibe, H. (1978). Some characteristics of a new glyCOPEPTIDASE acting on aspartylglycosylamine linkages. *J. Biochem.* 84, 1467–1473.
- Vitale, A., and Chrispeels, M. J. (1984). Transient N-acetylglucosamine in the biosynthesis of phytohemagglutinin: attachment in the golgi apparatus and removal in protein bodies. *J. Cell Biol.* 99, 133–140. doi: 10.1083/jcb.99.1.133
- Wilson, I. B. H., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J. M., Kamerling, J. P., et al. (2001). Analysis of Asn linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core α 1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Yamamoto, K. (2006). “Endoglycosidases that relate to N-glycans,” in *Endoglycosidase*, eds M. Endo, S. Hase, K. Yamamoto, and K. Takagaki (Tokyo: Kodansha-Springer), 55–74.
- Yamamoto, K. (2012). Biological analysis of the microbial metabolism of heterooligosaccharides in application to glycotechnology. *Biosci. Biotechnol. Biochem.* 76, 1815–1827. doi: 10.1271/bbb.120401
- Yunovitz, H., and Gross, K. C. (1994). Delay of tomato fruit ripening by an oligosaccharide N-glycan. Interactions with IAA, galactose and lectins. *Physiol. Plantarum* 90, 152–156. doi: 10.1111/j.1399-3054.1994.tb02205.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 June 2014; accepted: 12 August 2014; published online: 04 September 2014.

Citation: Maeda M and Kimura Y (2014) Structural features of free N-glycans occurring in plants and functional features of de-N-glycosylation enzymes, ENGase, and PNGase: the presence of unusual plant complex type N-glycans. *Front. Plant Sci.* 5:429. doi: 10.3389/fpls.2014.00429

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Maeda and Kimura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Biological significance of complex *N*-glycans in plants and their impact on plant physiology

Richard Strasser*

Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Muriel Bardor, Université de Rouen, France

Kazuhito Fujiyama, Osaka University, Japan

***Correspondence:**

Richard Strasser, Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria

e-mail: richard.strasser@boku.ac.at

Asparagine (*N*)-linked protein glycosylation is a ubiquitous co- and post-translational modification which can alter the biological function of proteins and consequently affects the development, growth, and physiology of organisms. Despite an increasing knowledge of *N*-glycan biosynthesis and processing, we still understand very little about the biological function of individual *N*-glycan structures in plants. In particular, the *N*-glycan-processing steps mediated by Golgi-resident enzymes create a structurally diverse set of protein-linked carbohydrate structures. Some of these complex *N*-glycan modifications like the presence of β 1,2-xylose, core α 1,3-fucose or the Lewis a-epitope are characteristic for plants and are evolutionary highly conserved. In mammals, complex *N*-glycans are involved in different cellular processes including molecular recognition and signaling events. In contrast, the complex *N*-glycan function is still largely unknown in plants. Here, in this short review, I focus on important recent developments and discuss their implications for future research in plant glycobiology and plant biotechnology.

Keywords: endoplasmic reticulum, Golgi apparatus, protein glycosylation, *N*-glycosylation, glycoprotein, *N*-acetylglucosaminyltransferase

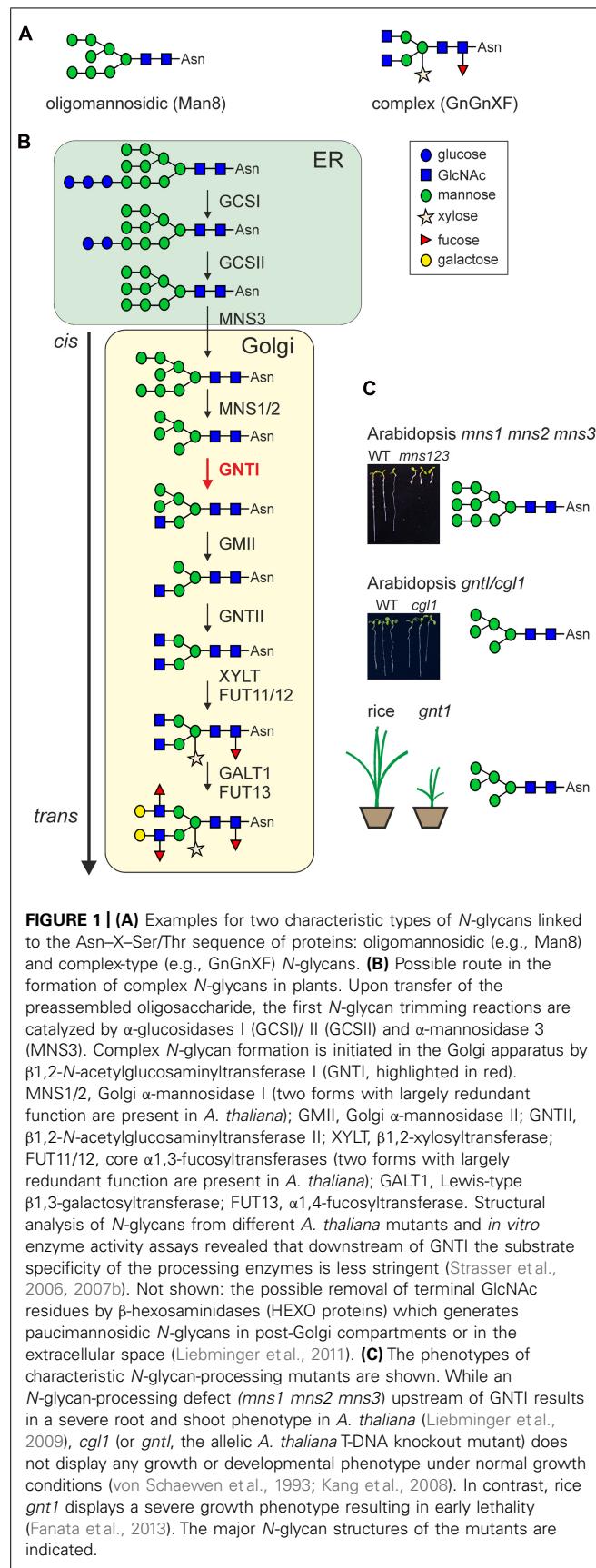
INTRODUCTION

N-Glycosylation is a major co- and post-translational modification of proteins in all eukaryotes. It has been estimated that the majority of all secretory proteins are *N*-glycosylated (Apweiler et al., 1999). *N*-Glycosylation is initiated in the ER by transfer of a preassembled oligosaccharide (Glc₃Man₉GlcNAc₂) precursor to asparagine residues within the sequence motif Asn-X-Ser/Thr (X represents any amino acid except proline) on nascent polypeptide chains. In addition, *N*-glycosylation at the unusual Asn-X-Cys site has also been described for some proteins (Matsui et al., 2011; Zielinska et al., 2012). Upon transfer of the oligosaccharide, the *N*-glycan is rapidly processed by highly specific α -glucosidases and α -mannosidases that remove terminal glucose and mannose residues, respectively. Incompletely trimmed *N*-glycans (Glc_{0–3}Man_{5–9}GlcNAc₂) that contain different amounts of mannose residues (also called oligomannosidic *N*-glycans) are mainly found on ER-resident proteins (Figure 1A). The mannose trimming reactions are carried out by α -mannosidases (MNS1–MNS3) that act in the ER and Golgi (Liebminger et al., 2009). The Man₅GlcNAc₂ oligosaccharide, which is the final product of these early *N*-glycan-processing steps is used by GNTI as a acceptor substrate for the transfer of a single *N*-acetylglucosamine (GlcNAc) residue to the exposed α 1,3-mannose of the *N*-glycan (Strasser et al., 1999). This enzymatic reaction is absolutely required for all further *N*-glycan modifications and results in the formation of complex *N*-glycans in the Golgi apparatus. In particular, GNTI generates the GlcNAc₁Man₅GlcNAc₂ *N*-glycan that is further

processed by Golgi- α -mannosidase II (GMII), GNTII, XYLT, and FUT11/12 (Figure 1B). All these enzymes are absolutely dependent on GNTI activity and reside in the *cis*/medial-Golgi apparatus of plants where they might form a multi-protein complex that could play a role for the organization of the glycosylation enzymes within the Golgi and subsequently also for the controlled processing of *N*-glycans (Schöberer and Strasser, 2011; Schöberer et al., 2013). GNTI is evolutionary highly conserved and present in land plants including mosses as well as in some microalgae (Strasser et al., 1999; Koprivova et al., 2003; Baïet et al., 2011). Due to its central function in initiation of complex *N*-glycan formation, GNTI controls the final *N*-glycosylation pattern on individual glycoproteins which can influence their biological function. XYLT and FUT11/12 attach β 1,2-xylose and core α 1,3-fucose residues, respectively, to different acceptor substrates and create common complex plant *N*-glycans like GlcNAc₂XylFucMan₃GlcNAc₂ (GnGnXF, Figure 1A). Such complex *N*-glycans are not present in mammals and thus can elicit an unwanted anti-carbohydrate immune response when for example present on plant-produced recombinant glycoproteins (Bardor et al., 2003; Jin et al., 2008).

The final *N*-glycan modification steps take place in the *trans*-Golgi and are carried out by the Lewis-type β 1,3-galactosyltransferase (GALT1) and the α 1,4-fucosyltransferase (FUT13) which generate the Lewis a-trisaccharide [Fuc α 1-4(Gal β 1-3)GlcNAc-R] on complex *N*-glycans (Strasser et al., 2007b). The Lewis a-type structures seem ubiquitous in the plant kingdom (Fitchette et al., 1999; Wilson et al., 2001), but they are only present on a small number of still widely unknown glycoproteins and the biological function of these large complex *N*-glycans remains to be shown.

Abbreviations: EMS, ethyl methanesulfonate; ERAD, ER-associated degradation; GNTI, β 1,2-*N*-acetylglucosaminyltransferase I; FUT11/12, core α 1,3-fucosyltransferase; XYLT, β 1,2-xylosyltransferase



Truncated N-glycans are generated by removal of terminal GlcNAc residues in post-Golgi compartments. These paucimannosidic N-glycans have been found on vacuolar and extracellular glycoproteins (Strasser et al., 2007a; Liebminger et al., 2011).

THE FUNCTION OF OLIGOMANNOSIDIC N-GLYCANs

Early N-glycan-processing reactions mediated by α -glucosidase I and II are essential for *Arabidopsis* and presumably also for other plant species (Taylor et al., 2000; Boisson et al., 2001; Gillmor et al., 2002; Soussilane et al., 2009; Farid et al., 2011; Wang et al., 2014). The generated oligomannosidic N-glycans are implicated in folding of nascent polypeptides and play an important role during ER-quality control processes and ERAD of misfolded or incompletely assembled glycoproteins (Aebi, 2013). The overall principles of these processes are conserved in eukaryotes. Recent findings suggest that monoglycosylated N-glycans in the ER are important for association with the lectins calreticulin or calnexin also in plants. For example, the pattern recognition receptor EFR involved in innate immunity and a misfolded variant of the brassinosteroid receptor BRI1 displayed a selective interaction with the plant-specific calreticulin 3 (Jin et al., 2009; Li et al., 2009). Additional data suggest that N-glycans present on these heavily glycosylated leucine-rich repeat receptor kinases are subjected to re-glucosylation by the folding sensor UDP-glucose:glycoprotein glucosyltransferase (UGGT) and glucosidase-mediated de-glucosylation followed by a release from the calreticulin/calnexin quality control cycle (reviewed in Liu and Li, 2014; Tintor and Sajio, 2014). Moreover, specific mannose residues present on terminally misfolded glycoproteins play also a crucial role for the selective disposal via ERAD (Hong et al., 2009, 2012; Hüttner et al., 2012, 2014) and a complete block of mannose removal in the *Arabidopsis mns1 mns2 mns3* triple mutant causes also a severe root growth phenotype (Figure 1C; Liebminger et al., 2009).

THE FUNCTION OF COMPLEX N-GLYCANS

In all higher eukaryotes, GNTI is the central enzyme that initiates complex N-glycan formation on secreted and membrane-bound proteins that are trafficking through the Golgi to their final destination. Early studies in mice revealed that GNTI is essential for the development of embryos (Ioffe and Stanley, 1994; Metzler et al., 1994), but cultured mammalian cells can survive in the absence of complex N-glycans (Stanley et al., 1975). More recent genetic approaches revealed that the structurally diverse complex N-glycans on mammalian proteins participate in many different biological processes and distinct alterations are often associated with diseases (Lowe and Marth, 2003; Dennis et al., 2009). *Drosophila melanogaster* deficient in GNTI activity are viable, but display distinct phenotypes like abnormal brain development and a reduced life span (Sarkar et al., 2006). *Caenorhabditis elegans* GNTI-null mutants develop normally but are more susceptible to bacterial pathogens (Schachter, 2010). Together these findings highlight the importance of complex N-glycan modifications in various organisms.

In spite of the fact that complex N-glycans are ubiquitously present in plants (Wilson et al., 2001), their biological function is

virtually unknown. The first mutant lacking complex N-glycans was isolated more than two decades ago by EMS mutagenesis of *Arabidopsis* and subsequent screening for lines that lack β 1,2-linked xylose and core α 1,3-fucose residues (von Schaewen et al., 1993). The isolated *complex glycan 1* (*cgl1*) mutants displayed a defect in the formation of complex N-glycans due to a point mutation in the gene coding for GNTI (Strasser et al., 2005). Consequently, in *cgl1* all endogenous glycoproteins carry exclusively oligomannosidic N-glycans with Man₅GlcNAc₂ as predominant oligosaccharide. Remarkably, the *Arabidopsis* *cgl1* plants are viable, fertile and do not display any obvious phenotype under different growth conditions including heat (30°C) and cold (8°C) stress or increased light conditions (von Schaewen et al., 1993; Figure 1C). Related studies identified various other *Arabidopsis* mutants with distinct defects in N-glycan-processing steps downstream of GNTI. In line with data for *cgl1*, no clear growth or developmental phenotypes were observed for *Arabidopsis* mutants that produce hybrid structures (Strasser et al., 2006) or complex N-glycans devoid of β 1,2-xylose and core α 1,3-fucose residues (Strasser et al., 2004). In agreement with these findings, neither the complete elimination nor the overexpression of the Lewis a-type structures on complex N-glycans caused a substantial change in *Arabidopsis* growth or development when grown under long day conditions (16 h-light/8 h-dark) at 22°C (Strasser et al., 2007b). Up to now, the only evidence for a biological function of complex N-glycans in *Arabidopsis* was found when *cgl1* and other mutants were subjected to osmotic and salt stress (Kang et al., 2008). Reduced root growth on media containing high NaCl concentrations indicated that complex N-glycans are implicated in tolerance to salt stress. However, a deeper understanding of complex N-glycan function in *Arabidopsis* and studies that associate distinct complex N-glycan structures on individual glycoproteins with the enhanced salt sensitivity are completely missing.

Based on the aforementioned studies, it has been suggested that N-glycan processing in the Golgi is dispensable for the normal development of plants and plays only a role under certain stress conditions. A recent study by Fanata et al. (2013) challenges our current view and provides strong evidence that complex N-glycans play indeed an essential role in some plant species. A homozygous *Oryza sativa* line (*gnt1*) with a T-DNA insertion in the single rice GNTI gene was identified that completely abolished GNTI mRNA expression. As a consequence of missing GNTI transcripts and in accordance with the central function of GNTI in the formation of complex N-glycans, the rice *gnt1* mutant displayed only oligomannosidic N-glycans with approximately 75% Man₅GlcNAc₂ structures. Almost the same amounts of Man₅GlcNAc₂ N-glycans were found for *Arabidopsis* *cgl1* (von Schaewen et al., 1993; Strasser et al., 2005). However, in marked contrast to *Arabidopsis* *cgl1*, a severe phenotype with arrested seedling development and lethality before reaching the reproductive stage was reported for rice *gnt1* (Fanata et al., 2013; Figure 1C). In addition, rice *gnt1* plants displayed defects in cell wall composition and cytokinin insensitivity. Although the final confirmation that the observed severe phenotypes are indeed linked to defects in *gnt1* is missing as the cytokinin defect caused problems with complementation of the *gnt1* plants, all other data

are convincing and indicate that complex N-glycans are essential in some plant species. How can we explain this discrepancy between *Arabidopsis* and rice? Based on data from total N-glycan analysis and annotation of the rice genome, it is quite likely that the N-glycan-processing steps in the Golgi are very similar between the two species (Figure 1B). However, there might be subtle differences in cell-/tissue-specific expression of certain N-glycan-processing enzymes that might have been missed by total N-glycan analysis from whole plant organs. Interestingly, the rice genome contains more than one glycosyltransferase with homology to *Arabidopsis* GALT1 (Strasser et al., 2007b) and it seems that the formation of Lewis a-type structures occurs more frequently in rice than in *Arabidopsis* (Léonard et al., 2004; Strasser et al., 2004, 2007b). The rice GALT1 homologs belong to Carbohydrate-Active enzyme glycosyltransferase-family 31, which contains a large number of enzymes with quite diverse functions (Strasser et al., 2007b; Basu et al., 2013). These GALT1 candidates have not been characterized and in the absence of data from plants devoid of Lewis a-type structures, their contribution to the development of rice remains an open question. Moreover, N-glycosylation defects are generally pleiotropic and affect numerous secretory as well as membrane-anchored proteins. Consequently, the observed phenotype in rice *gnt1* could arise from several different glycoproteins that are dysfunctional in the absence of Golgi-mediated N-glycan processing. As rice *gnt1* displays reduced cellulose contents, glycoproteins involved in cellulose biosynthesis could be affected (Fanata et al., 2013). While impaired N-glycosylation or N-glycan processing has also been linked to changes in cellulose contents in *Arabidopsis* (Burn et al., 2002; Gillmor et al., 2002; Zhang et al., 2009) *gnt1/cgl1* does not contain significantly altered cellulose contents compared to wild-type *Arabidopsis* (Kang et al., 2008). Recently, it was also shown that the heavily glycosylated endoglucanase KORRIGAN1, whose enzymatic activity is important for efficient cellulose formation, does not need complex N-glycans for its function (Liebminger et al., 2013).

Based on the detected cytokinin insensitivity it was speculated that members of the cytokinin-receptor family are N-glycosylated and their function might be impaired in the rice *gnt1* line (Fanata et al., 2013). These histidine sensor kinases contain an extracellular domain of approximately 280 amino acids with putative N-glycosylation sites (Caesar et al., 2011; Steklov et al., 2013). The degree of N-glycosylation and the N-glycan structures of cytokinin receptors are not very well known, but for *Arabidopsis* AHK3 N-glycosylation could be shown by transient expression in tobacco (Caesar et al., 2011). Notably, in *Arabidopsis* as well as in maize these receptors were primarily found in the ER implying that cytokinin binding takes place in this compartment (Caesar et al., 2011; Lomin et al., 2011; Wulfetange et al., 2011). If so, then Golgi-processed complex N-glycans are very likely not present on cytokinin receptors and consequently these receptors are not directly affected in GNTI-deficient rice.

IMPLICATIONS FOR PLANT GLYCobiOLOGY

To understand the mechanisms underlying the observed defects in rice *gnt1* and compare them with data from other plants species a number of key experimental approaches have to be

explored: (i) It is very important to isolate other rice N-glycan-processing mutants to pin down the complex N-glycan structure or individual sugar residue that is crucial for the growth and development of rice. (ii) There is an urgent need for high-throughput glycoproteome approaches that enable the isolation of a large number of glycoproteins and mapping of the corresponding N-glycan structures from different plant species. Advances in this field will be crucial for structure–function analysis and identification of target glycoproteins. Plant N-glycoproteome studies have been reported recently (Zhang et al., 2011; Zielinska et al., 2012; Song et al., 2013), but compared to other posttranslational modifications these approaches are still too limited (Albenne et al., 2013). (iii) Information on regulation of glycosylation enzymes as well as information on cell-type or organ-specific occurrence of certain glycan structures is almost completely missing. Tools that have been used for the cell-type-specific analysis of protein expression (Petricka et al., 2012) should also be applied to unravel the N-glycoproteome in different plant species. (iv) Up to now, null mutants devoid of N-glycan processing have been characterized from *Arabidopsis* and rice, but information on the significance of N-glycosylation and complex N-glycan formation in other vascular plants is missing. Together, the highlighted experimental approaches will enable us to decode the biological function of the so far largely unknown complex N-glycan modifications like the attachment of β 1,2-xylose, core α 1,3-fucose, and the formation of the Lewis a-type structures.

IMPLICATIONS FOR PLANT BIOTECHNOLOGY

Plants are emerging hosts for the manufacturing of valuable recombinant proteins. Recently, the first plant-produced recombinant biopharmaceutical, a recombinant human glucocerebrosidase, has been approved for enzyme replacement therapy in humans and is commercially available in the United States (Grabowski et al., 2014). Many biopharmaceutical proteins like human immunoglobulins or hormones are glycosylated and the composition of the glycans very often affect protein–protein interactions leading to altered efficacies of the recombinant drugs or unwanted side-effects like fast clearance from the blood or increased immunogenicity. Consequently, for the pharmaceutical industry as well as for structure–function studies, there is a growing demand to modify and control protein glycosylation of expression hosts. The ultimate aim of these approaches is the production of recombinant glycoproteins with defined and homogenous glycan structures (Rich and Withers, 2009; Dalziel et al., 2014). Developments during the last 10 years have shown that plants are amenable to glyco-engineering and capable of producing valuable recombinant glycoproteins with defined human-like structures (Castilho and Steinellner, 2012; Nagels et al., 2012; Bosch et al., 2013). The absence of any growth phenotype in *Arabidopsis cgl1* laid the foundation for N-glycan engineering of other species like *Nicotiana benthamiana* and *Lemna minor* as well as of rice suspension cells (Cox et al., 2006; Strasser et al., 2008; Shin et al., 2011). In these studies, gene silencing of XYLT and FUT11/12 was used to eliminate the non-human and potentially immunogenic β 1,2-xylose and core α 1,3-fucose residues from complex N-glycans of recombinant proteins. Overall, these glyco-engineering efforts

were quite successful, but the plants still produced low amounts of complex N-glycans like GnGnXF. A detailed characterization of null mutants for XYLT and FUT11/12 will reveal whether these and other plant species tolerate the absence of β 1,2-xylose and core α 1,3-fucose residues on endogenous glycoproteins during their whole life cycle. In addition, further studies are necessary to investigate in detail the consequences on growth, development, reproduction and stress response of stable engineered plants that carry human-type complex N-glycan modifications. So far, these knock-in approaches were limited to a small number of plant species and mainly to stable expression of single mammalian glycosyltransferases (Bakker et al., 2001; Rouwendal et al., 2007; Castilho et al., 2008; Sourrouille et al., 2008; Frey et al., 2009; Nagels et al., 2011). In contrast, most of the more advanced glyco-engineering approaches that require the concerted action of several mammalian enzymes were done by simultaneous transient expression of whole glycosylation pathways (Castilho et al., 2010, 2012, 2013). The stable expression of the proteins and enzymes involved in multi-step N-glycan processing like the formation of highly sialylated complex N-glycans without any negative effects on plant growth and development remains to be shown.

In the light of the recent findings from rice, glyco-engineering in some plant species might require new strategies and implementation of more elaborate tools to overcome adverse phenotypes linked with extensive N-glycan remodeling. In terms of plant glycobiology, the new findings from rice open the door for an exciting new era.

ACKNOWLEDGMENTS

The author apologizes to colleagues whose work is not cited due to space limitations. This work was supported by a grant from Austrian Science Fund (FWF): P23906-B20.

REFERENCES

- Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta* 1833, 2430–2437. doi: 10.1016/j.bbamer.2013.04.001
- Albenne, C., Canut, H., and Jamet, E. (2013). Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*. *Front. Plant Sci.* 4:111. doi: 10.3389/fpls.2013.00111
- Apweiler, R., Hermjakob, H., and Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8. doi: 10.1016/S0304-4165(99)00165-8
- Baiet, B., Burel, C., Saint-Jean, B., Louvet, R., Menu-Bouaouiche, L., Kiefer-Meyer, M. C., et al. (2011). N-glycans of *Phaeodactylum tricornutum* diatom and functional characterization of its N-acetylglucosaminyltransferase I enzyme. *J. Biol. Chem.* 286, 6152–6164. doi: 10.1074/jbc.M110.175711
- Bakker, H., Bardor, M., Molthoff, J., Gomord, V., Elbers, I., Stevens, L., et al. (2001). Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2899–2904. doi: 10.1073/pnas.031419998
- Bardor, M., Faveeuw, C., Fitchette, A., Gilbert, D., Galas, L., Trottein, F., et al. (2003). Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. doi: 10.1093/glycob/cwg024
- Basu, D., Liang, Y., Liu, X., Himmeldirk, K., Faik, A., Kieliszewski, M., et al. (2013). Functional identification of a hydroxyproline-o-galactosyltransferase specific for arabinogalactan protein biosynthesis in *Arabidopsis*. *J. Biol. Chem.* 288, 10132–10143. doi: 10.1074/jbc.M112.432609
- Boisson, M., Gomord, V., Audran, C., Berger, N., Dubreucq, B., Granier, F., et al. (2001). *Arabidopsis* glucosidase I mutants reveal a critical role of N-glycan trimming in seed development. *EMBO J.* 20, 1010–1019. doi: 10.1093/emboj/20.5.1010

- Bosch, D., Castilho, A., Loos, A., Schots, A., and Steinkellner, H. (2013). N-glycosylation of plant-produced recombinant proteins. *Curr. Pharm. Des.* 19, 5503–5512. doi: 10.2174/1381612811319310006
- Burn, J., Hurley, U., Birch, R., Arioli, T., Cork, A., and Williamson, R. (2002). The cellulose-deficient *Arabidopsis* mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. *Plant J.* 32, 949–960. doi: 10.1046/j.1365-313X.2002.01483.x
- Caesar, K., Thamm, A. M., Witthöft, J., Elgass, K., Huppenberger, P., Grefen, C., et al. (2011). Evidence for the localization of the *Arabidopsis* cytokinin receptors AHK3 and AHK4 in the endoplasmic reticulum. *J. Exp. Bot.* 62, 5571–5580. doi: 10.1093/jxb/err238
- Castilho, A., Neumann, L., Daskalova, S., Mason, H. S., Steinkellner, H., Altmann, F., et al. (2012). Engineering of sialylated mucin-type O-glycosylation in plants. *J. Biol. Chem.* 287, 36518–36526. doi: 10.1074/jbc.M112.402685
- Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., et al. (2013). Generation of biologically active multi-sialylated recombinant human EPO-Fc in plants. *PLoS ONE* 8:e54836. doi: 10.1371/journal.pone.0054836
- Castilho, A., Pabst, M., Leonard, R., Veit, C., Altmann, F., Mach, L., et al. (2008). Construction of a functional CMP-sialic acid biosynthesis pathway in *Arabidopsis*. *Plant Physiol.* 147, 331–339. doi: 10.1104/pp.108.117572
- Castilho, A., and Steinkellner, H. (2012). Glyco-engineering in plants to produce human-like N-glycan structures. *Biotechnol. J.* 7, 1088–1098. doi: 10.1002/biot.201200032
- Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gattinger, P., et al. (2010). In planta protein sialylation through overexpression of the respective mammalian pathway. *J. Biol. Chem.* 285, 15923–15930. doi: 10.1074/jbc.M109.088401
- Cox, K., Sterling, J., Regan, J., Gasdaska, J., Frantz, K., Peele, C., et al. (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* 24, 1591–1597. doi: 10.1038/nbt1260
- Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N., and Dwek, R. A. (2014). Emerging principles for the therapeutic exploitation of glycosylation. *Science* 343, 1235681. doi: 10.1126/science.1235681
- Dennis, J. W., Nabi, I. R., and Demetriou, M. (2009). Metabolism, cell surface organization, and disease. *Cell* 139, 1229–1241. doi: 10.1016/j.cell.2009.12.008
- Fanata, W. I., Lee, K. H., Son, B. H., Yoo, J. Y., Harmoko, R., Ko, K. S., et al. (2013). N-glycan maturation is crucial for cytokinin-mediated development and cellulose synthesis in *Oryza sativa*. *Plant J.* 73, 966–979. doi: 10.1111/tpj.12087
- Farid, A., Pabst, M., Schoberer, J., Altmann, F., Glössl, J., and Strasser, R. (2011). *Arabidopsis thaliana* alpha1,2-glucosyltransferase (ALG10) is required for efficient N-glycosylation and leaf growth. *Plant J.* 68, 314–325. doi: 10.1111/j.1365-313X.2011.04688.x
- Fitchette, A., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., et al. (1999). Biosynthesis and immunolocalization of Lewis a-containing N-glycans in the plant cell. *Plant Physiol.* 121, 333–344. doi: 10.1104/pp.121.2.333
- Frey, A. D., Karg, S. R., and Kallio, P. T. (2009). Expression of rat beta(1,4)-N-acetylglucosaminyltransferase III in *Nicotiana tabacum* remodels the plant-specific N-glycosylation. *Plant Biotechnol. J.* 7, 33–48. doi: 10.1111/j.1467-7652.2008.00370.x
- Gillmor, C., Poindexter, P., Lorieau, J., Palcic, M., and Somerville, C. (2002). Alpha-glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. *J. Cell Biol.* 156, 1003–1013. doi: 10.1083/jcb.200111093
- Grabowski, G. A., Golembio, M., and Shahtiel, Y. (2014). TAGlipidase alpha: an enzyme replacement therapy using plant cell expression technology. *Mol. Genet. Metab.* 112, 1–8. doi: 10.1016/j.ymgme.2014.02.011
- Hong, Z., Jin, H., Fitchette, A., Xia, Y., Monk, A., Faye, L., et al. (2009). Mutations of an alpha1,6 mannosyltransferase inhibit endoplasmic reticulum-associated degradation of defective brassinosteroid receptors in *Arabidopsis*. *Plant Cell* 21, 3792–3802. doi: 10.1105/tpc.109.070284
- Hong, Z., Kajiura, H., Su, W., Jin, H., Kimura, A., Fujiyama, K., et al. (2012). Evolutionarily conserved glycan signal to degrade aberrant brassinosteroid receptors in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11437–11442. doi: 10.1073/pnas.1119173109
- Hüttner, S., Veit, C., Schoberer, J., Grass, J., and Strasser, R. (2012). Unraveling the function of *Arabidopsis thaliana* OS9 in the endoplasmic reticulum-associated degradation of glycoproteins. *Plant Mol. Biol.* 79, 21–33. doi: 10.1007/s11103-012-9891-4
- Hüttner, S., Veit, C., Vavra, U., Schoberer, J., Liebminger, E., Maresch, D., et al. (2014). *Arabidopsis* class I mannosidases MNS4 and MNS5 are involved in ER-associated degradation of misfolded glycoproteins. *Plant Cell* 26, 1712–1728. doi: 10.1105/tpc.114.123216
- Ioffe, E., and Stanley, P. (1994). Mice lacking N-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proc. Natl. Acad. Sci. U.S.A.* 91, 728–732. doi: 10.1073/pnas.91.2.728
- Jin, C., Altmann, F., Strasser, R., Mach, L., Schähs, M., Kunert, R., et al. (2008). A plant-derived human monoclonal antibody induces an anti-carbohydrate immune response in rabbits. *Glycobiology* 18, 235–241. doi: 10.1093/glycob/cwm137
- Jin, H., Hong, Z., Su, W., and Li, J. (2009). A plant-specific calreticulin is a key retention factor for a defective brassinosteroid receptor in the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13612–13617. doi: 10.1073/pnas.0906144106
- Kang, J., Frank, J., Kang, C., Kajiura, H., Vikram, M., Ueda, A., et al. (2008). Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5933–5938. doi: 10.1073/pnas.0800237105
- Koprivova, A., Altmann, F., Gorr, G., Kopriva, S., Reski, R., and Decker, E. (2003). N-Glycosylation in the moss *Physcomitrella patens* is organized similarly to higher plants. *Plant Biol.* 5, 582–591. doi: 10.1055/s-2003-44721
- Léonard, R., Kolarich, D., Paschinger, K., Altmann, F., and Wilson, I. (2004). A genetic and structural analysis of the N-glycosylation capabilities of rice and other monocotyledons. *Plant Mol. Biol.* 55, 631–644. doi: 10.1007/s11103-004-1558-3
- Li, J., Zhao-Hui, C., Batoux, M., Nekrasov, V., Roux, M., Chinchilla, D., et al. (2009). Specific ER quality control components required for biogenesis of the plant innate immune receptor EFR. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15973–15978. doi: 10.1073/pnas.0905532106
- Liebminger, E., Grass, J., Altmann, F., Mach, L., and Strasser, R. (2013). Characterizing the link between glycosylation state and enzymatic activity of the endo-β1,4-glucanase KORRIGAN1 from *Arabidopsis thaliana*. *J. Biol. Chem.* 288, 22270–22280. doi: 10.1074/jbc.M113.475558
- Liebminger, E., Hüttner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., et al. (2009). Class I alpha-mannosidases are required for N-glycan processing and root development in *Arabidopsis thaliana*. *Plant Cell* 21, 3850–3867. doi: 10.1105/tpc.109.072363
- Liebminger, E., Veit, C., Pabst, M., Batoux, M., Zipfel, C., Altmann, F., et al. (2011). N-Acetylhexosaminidases HEXO1 and HEXO3 are responsible for the formation of paucimannosidic N-glycans in *Arabidopsis thaliana*. *J. Biol. Chem.* 286, 10793–10802. doi: 10.1074/jbc.M110.178020
- Liu, Y., and Li, J. (2014). Endoplasmic reticulum-mediated protein quality control in *Arabidopsis*. *Front. Plant Sci.* 5:162. doi: 10.3389/fpls.2014.00162
- Lomin, S. N., Yonekura-Sakakibara, K., Romanov, G. A., and Sakakibara, H. (2011). Ligand-binding properties and subcellular localization of maize cytokinin receptors. *J. Exp. Bot.* 62, 5149–5159. doi: 10.1093/jxb/err220
- Lowe, J. B., and Marth, J. D. (2003). A genetic approach to mammalian glycan function. *Annu. Rev. Biochem.* 72, 643–691. doi: 10.1146/annurev.biochem.72.121801.161809
- Matsui, T., Takita, E., Sato, T., Kinjo, S., Aizawa, M., Sugiura, Y., et al. (2011). N-glycosylation at noncanonical Asn-X-Cys sequences in plant cells. *Glycobiology* 21, 994–999. doi: 10.1093/glycob/cwq198
- Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., and Marth, J. D. (1994). Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *EMBO J.* 13, 2056–2065.
- Nagels, B., Van Damme, E. J., Pabst, M., Callewaert, N., and Weterings, K. (2011). Production of complex multiantennary N-glycans in *Nicotiana benthamiana* plants. *Plant Physiol.* 155, 1103–1112. doi: 10.1104/pp.110.168773
- Nagels, B., Weterings, K., Callewaert, N., and Van Damme, E. J. M. (2012). Production of plant made pharmaceuticals: from plant host to functional protein. *Crit. Rev. Plant Sci.* 31, 148–180. doi: 10.1080/07352689.2011.616075
- Petricka, J. J., Schauer, M. A., Megraw, M., Breakfield, N. W., Thompson, J. W., Georgiev, S., et al. (2012). The protein expression landscape of the *Arabidopsis* root. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6811–6818. doi: 10.1073/pnas.1202546109

- Rich, J. R., and Withers, S. G. (2009). Emerging methods for the production of homogeneous human glycoproteins. *Nat. Chem. Biol.* 5, 206–215. doi: 10.1038/nchembio.148
- Rouwendal, G. J., Wührer, M., Florack, D. E., Koeleman, C. A., Deelder, A. M., Bakker, H., et al. (2007). Efficient introduction of a bisecting GlcNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase III. *Glycobiology* 17, 334–344. doi: 10.1093/glycob/cwl078
- Sarkar, M., Leventis, P. A., Silvescu, C. I., Reinhold, V. N., Schachter, H., and Boulian, G. L. (2006). Null mutations in *Drosophila* N-acetylglucosaminyltransferase I produce defects in locomotion and a reduced life span. *J. Biol. Chem.* 281, 12776–12785. doi: 10.1074/jbc.M512769200
- Schachter, H. (2010). Mgat1-dependent N-glycans are essential for the normal development of both vertebrate and invertebrate metazoans. *Semin. Cell Dev. Biol.* 21, 609–615. doi: 10.1016/j.semcd.2010.02.010
- Schoberer, J., Liebminger, E., Botchway, S. W., Strasser, R., and Hawes, C. (2013). Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in plants. *Plant Physiol.* 161, 1737–1754. doi: 10.1104/pp.112.210757
- Schoberer, J., and Strasser, R. (2011). Sub-compartmental organization of Golgi-resident N-glycan processing enzymes in plants. *Mol. Plant* 4, 220–228. doi: 10.1093/mp/ssq082
- Shin, Y. J., Chong, Y. J., Yang, M. S., and Kwon, T. H. (2011). Production of recombinant human granulocyte macrophage-colony stimulating factor in rice cell suspension culture with a human-like N-glycan structure. *Plant Biotechnol. J.* 9, 1109–1119. doi: 10.1111/j.1467-7652.2011.00636.x
- Song, W., Mentink, R. A., Henquet, M. G., Cordewener, J. H., van Dijk, A. D., Bosch, D., et al. (2013). N-glycan occupancy of *Arabidopsis* N-glycoproteins. *J. Proteomics* 93, 343–355. doi: 10.1016/j.jprot.2013.07.032
- Sourrouille, C., Marquet-Blouin, E., D'Aoust, M. A., Kiefer-Meyer, M. C., Seveno, M., Pagny-Salehabadi, S., et al. (2008). Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnol. J.* 6, 702–721. doi: 10.1111/j.1467-7652.2008.00353.x
- Soussilane, P., Soussilane, P., D'Alessio, C., Paccalet, T., Fitchette, A., Parodi, A., et al. (2009). N-glycan trimming by glucosidase II is essential for *Arabidopsis* development. *Glycoconj. J.* 26, 597–607. doi: 10.1007/s10719-008-9201-1
- Stanley, P., Narasimhan, S., Siminovitch, L., and Schachter, H. (1975). Chinese hamster ovary cells selected for resistance to the cytotoxicity of phytohemagglutinin are deficient in a UDP-N-acetylglucosamine-glycoprotein N-acetylglucosaminyltransferase activity. *Proc. Natl. Acad. Sci. U.S.A.* 72, 3323–3323. doi: 10.1073/pnas.72.9.3323
- Steklov, M. Y., Lomin, S. N., Osolodkin, D. I., and Romanov, G. A. (2013). Structural basis for cytokinin receptor signaling: an evolutionary approach. *Plant Cell Rep.* 32, 781–793. doi: 10.1007/s00299-013-1408-3
- Strasser, R., Altmann, E., Mach, L., Glössl, J., and Steinkellner, H. (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4
- Strasser, R., Bondili, J., Schoberer, J., Svoboda, B., Liebminger, E., Glössl, J., et al. (2007a). Enzymatic properties and subcellular localization of *Arabidopsis* N-acetylhexosaminidases. *Plant Physiol.* 145, 5–16. doi: 10.1104/pp.107.101162
- Strasser, R., Bondili, J., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007b). A unique beta1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* 19, 2278–2292. doi: 10.1105/tpc.107.052985
- Strasser, R., Mucha, J., Schwihla, H., Altmann, F., Glössl, J., and Steinkellner, H. (1999). Molecular cloning and characterization of cDNA coding for beta1,2N-acetylglucosaminyltransferase I (GlcNAc-TI) from *Nicotiana tabacum*. *Glycobiology* 9, 779–785. doi: 10.1093/glycob/9.8.779
- Strasser, R., Schoberer, J., Jin, C., Glössl, J., Mach, L., and Steinkellner, H. (2006). Molecular cloning and characterization of *Arabidopsis thaliana* Golgi alpha-mannosidase II, a key enzyme in the formation of complex N-glycans in plants. *Plant J.* 45, 789–803. doi: 10.1111/j.1365-313X.2005.02648.x
- Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., et al. (2008). Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* 6, 392–402. doi: 10.1111/j.1467-7652.2008.00330.x
- Strasser, R., Stadlmann, J., Svoboda, B., Altmann, F., Glössl, J., and Mach, L. (2005). Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem. J.* 387, 385–391. doi: 10.1042/BJ20041686
- Taylor, M., Ross, H., Mcrae, D., Stewart, D., Roberts, I., Duncan, G., et al. (2000). A potato alpha-glucosidase gene encodes a glycoprotein-processing alpha-glucosidase II-like activity. Demonstration of enzyme activity and effects of down-regulation in transgenic plants. *Plant J.* 24, 305–316. doi: 10.1046/j.1365-313X.2000.00873.x
- Tintor, N., and Saijo, Y. (2014). ER-mediated control for abundance, quality, and signaling of transmembrane immune receptors in plants. *Front. Plant Sci.* 5:65. doi: 10.3389/fpls.2014.00065
- von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M. (1993). Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol.* 102, 1109–1118. doi: 10.1104/pp.102.4.1109
- Wang, S., Xu, Y., Li, Z., Zhang, S., Lim, J. M., Lee, K. O., et al. (2014). OsMOGS is required for N-glycan formation and auxin-mediated root development in rice (*Oryza sativa* L.). *Plant J.* 78, 632–645. doi: 10.1111/tpj.12497
- Wilson, I., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C., Kamerling, J., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core alpha1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Wulfetange, K., Lomin, S. N., Romanov, G. A., Stolz, A., Heyl, A., and Schmülling, T. (2011). The cytokinin receptors of *Arabidopsis* are located mainly to the endoplasmic reticulum. *Plant Physiol.* 156, 1808–1818. doi: 10.1104/pp.111.180539
- Zhang, M., Henquet, M., Chen, Z., Zhang, H., Zhang, Y., Ren, X., et al. (2009). LEW3, encoding a putative alpha-1,2-mannosyltransferase (ALG11) in N-linked glycoprotein, plays vital roles in cell-wall biosynthesis and the abiotic stress response in *Arabidopsis thaliana*. *Plant J.* 60, 983–999. doi: 10.1111/j.1365-313X.2009.04013.x
- Zhang, Y., Giboulot, A., Zivy, M., Valot, B., Jamet, E., and Albenne, C. (2011). Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72, 1109–1123. doi: 10.1016/j.phytochem.2010.10.019
- Zielinska, D. F., Gnäd, F., Schropp, K., Wiśniewski, J. R., and Mann, M. (2012). Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell* 46, 542–548. doi: 10.1016/j.molcel.2012.04.031

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 May 2014; accepted: 08 July 2014; published online: 22 July 2014.

Citation: Strasser R (2014) Biological significance of complex N-glycans in plants and their impact on plant physiology. *Front. Plant Sci.* 5:363. doi: 10.3389/fpls.2014.00363

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Strasser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Sugars take a central position in plant growth, development and, stress responses. A focus on apical dominance

Wim Van den Ende*

Laboratory of Molecular Plant Biology, Biology, Katholieke Universiteit Leuven, Leuven, Belgium

*Correspondence: wim.vandenende@bio.kuleuven.be

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Sjef Smeekens, Utrecht University, Netherlands

Soulaiman Sakr, Agrocampus-ouest, France

Keywords: sugar signaling, apical dominance, auxin response, trehalose 6-phosphate, sucrose, phloem transport, hormone signaling

In plant glycobiology, free “metabolic” carbohydrates consist of small sugars (glucose, fructose, sucrose, trehalose), sucrose-derived oligosaccharides (fructans, Raffinose Family Oligosaccharides), starch and its breakdown products (Van den Ende, 2013). All these carbohydrates are directly or indirectly derived from photosynthesis. Plants typically accumulate higher carbohydrate levels as compared with other multicellular eukaryotes, especially when (milder) stresses compromise growth more than photosynthesis (Van den Ende and El-Esawe, in press).

In a historical perspective, the central role of sugars was already suggested many decades ago, with proposed roles in overall plant growth and development (Allsopp, 1954), disease susceptibility (Horsfall and Dimond, 1957), flowering (Kraus and Kraybill, 1918) and apical dominance (Loeb, 1924). The latter two processes were suggested to be regulated by the “nutrient diversion hypothesis” (Gregory and Veale, 1957; Corbesier et al., 1998). This theory, narrowed down to sugar nutrients, states that a minimal level of sugar assimilates needs to reach the apex (flowering) or lateral bud (removal of apical dominance) before floral transition or lateral bud outgrowth is initiated. The discovery that indole-3-acetic acid (IAA) repressed lateral bud outgrowth in decapitated shoots (Thimann, 1937) boosted plant hormone research at the expense of sugar-centered research. However, things changed when small sugars, similar to hormones, were considered as important signals in plants (Moore et al., 2003). Since then, a renewed

and strongly increasing worldwide interest in sugar signaling, sensing and metabolism was noticed (Ruan, 2014; Smeekens and Hellmann, 2014).

Mason et al. (2014) focused on the underlying mechanisms involved during apical dominance in pea. The authors challenge a long-held dogma in plant physiology, proposing that sugar signals, and not IAA, initiates lateral bud outgrowth after apex decapitation. For decades, textbooks declare that phloem-mediated IAA transport fluxes down the stem decrease after decapitation, relieving the IAA-mediated inhibition on bud outgrowth. However, the discovery that stem IAA cannot enter the bud, and the fact that IAA application on decapitated stems cannot always prevent bud burst, suggested that a positive signal could overrule IAA-mediated inhibitory effects (Mason et al., 2014).

Buds from the decapitated apex could be released starting from 2.5 h post-decapitation as shown by time-lapse photography (Mason et al., 2014). Importantly, this was not accompanied by a measurable IAA depletion in the adjacent stem. However, a much slower bud release was reported before (Wardlaw and Mortimer, 1970). Mason et al. (2014) designed an elegant set of physiological experiments demonstrating that leaves are the source of a rapid decapitation induced signal that promotes bud release. They reasoned that sucrose could be a candidate for this signaling role. Subsequently, [¹¹C] CO₂ was fed to leaves and the movement of [¹¹C] photo-assimilates was monitored

along the stem. They found a speed of 150 cm h⁻¹, which agreed with the timing of bud burst and phloem-mediated transport. Exogenous sucrose applications promoted bud burst. Moreover, in plants decapitated low on the stem, leaf removal caused a serious delay in bud release, and this could be rescued by feeding sucrose, but not sorbitol, via the petiole. Unfortunately, the sucrose-specific character of this response was not tested. Comparing glucose and sucrose responses could be informative to discriminate between glucose- or sucrose-mediated signaling events.

These data strongly suggest that a minimal threshold sucrose level is required to sustain lateral bud outgrowth. Initiation of bud outgrowth would not make much sense if not enough C would be available to sustain the later stages of bud growth. However, this leaves us with a remaining question: is there still room for hormone signaling events in this process?

I speculate that the answer on this question is “yes.” Internal sugar/IAA ratios within lateral buds or within the adjacent stem may be somehow integrated prior to lateral bud outgrowth initiation. In my opinion further studies should answer the following crucial questions (i) Are increased sugar levels associated with changed hormone levels in buds and how does this change over time? (ii) Do bud IAA levels depend on IAA biosynthesis from tryptophan within the bud, or is indol-3-aldehyde, a phloem-mobile lateral bud inhibitor (Nakajima et al., 2002), involved via a

mechanism that perhaps depends on sugar signaling? If so, can the presence of indol-3-aldehyde in the phloem sap be confirmed? (iii) Alternatively, could phloem residing Aux/IAA transcripts, involved in the regulation of auxin signaling, enter the bud followed by differential translation depending on the actual sugar status? Such transcripts may be important, since their upregulation in the phloem influences both root and shoot branching, as well as overall IAA sensitivity (Golan et al., 2013). Some of these views fit with emerging evidence that sugars can control auxin levels in plants (LeClere et al., 2010; Sairanen et al., 2012). Hexokinase (HXK) mediated sugar signaling may be central in such processes, since AtHXK1 overexpressors relieved their apical dominance (Kelly et al., 2012). This was associated with lowered expression levels of genes that encode crucial players in auxin signaling, suggesting that glucose signals control downstream auxin signaling in *Arabidopsis* (Kelly et al., 2012).

It cannot be excluded that light dependent signaling mechanisms, or any other positive factors, may have remained undetected in the pea apical dominance paper (Mason et al., 2014). Moreover, these experiments should be repeated in an array of other plant species, before any general conclusions can be derived that would apply to all higher plants. GA, sugars and light play crucial roles during bud outgrowth in roses through increasing sugar demand, by upregulating the expression and activity of vacuolar invertases (Choubane et al., 2012; Rabot et al., 2012).

Despite these critical notes, it should be recognized that the paper of Mason et al. (2014) boosts further research in both hormone and sugar signaling communities, by stimulating hormone workers to consider sugars and vice-versa. So far, attempts to integrate sugar and hormone signaling events are rather scarce (Bolouri Moghaddam and Van den Ende, 2013). The putative importance of sugar signaling events in apical dominance adds to a list of other physiological processes that are believed to be controlled by sugars or their phosphorylated derivatives such as trehalose 6-phosphate (T6P), which has been suggested as an important indicator of the carbohydrate status in plants and negative regulator of SnRK1 (Zhang et al.,

2009). The latter is a central player in overall energy homeostasis (Baena-González et al., 2007) together with TOR kinase (Robaglia et al., 2012; Xiong et al., 2013). The T6P/SnRK1 module is involved in sugar signaling processes (Baena-González et al., 2007) together with sucrose-specific DELLA-mediated processes (Li et al., 2014) controlling, for example, anthocyanin accumulation as part of the plants defense response (Nakabayashi et al., 2014).

In line with the earlier ideas of Allsopp (1954), recent molecular and biochemical evidences revealed that sugars regulate the juvenile-to-adult phase transition by modulating miR156 expression, with possible involvement of HXK-mediated signaling (Duarte et al., 2013; Yang et al., 2013; Yu et al., 2013). Moreover, it was demonstrated that rhythmic, endogenous sugar signals, independently of light signals, entrain circadian rhythms by regulating the expression of circadian clock components (Haydon et al., 2013). Recent data strongly suggest that leaf diurnal starch dynamics (Graf and Smith, 2011) intimately connect with T6P levels (Martins et al., 2013). In line with the nutrient diversion hypothesis, it is proposed that plants sense the T6P status prior to the transition to flowering (Wahl et al., 2013). Photoperiod modification of starch homeostasis by CONSTANS, a stimulator of the FLOWERING LOCUS T, may be crucial for increasing the sugar mobilization demanded by the floral transition (Ortiz-Marchena et al., 2014). These observations urge further research into T6P levels, starch dynamics and SnRK1 activities in the context of apical dominance.

Clearly, we are only at the beginning of our understanding of the complexity of cellular sugar homeostasis, and deciphering how this exactly connects to hormonal regulatory mechanisms is an important challenge.

ACKNOWLEDGMENT

Wim Van den Ende is supported by funding from the FSR, Flanders.

REFERENCES

- Allsopp, A. (1954). Juvenile stages of plants and the nutritional status of the shoot apex. *Nature* 173, 1032–1035. doi: 10.1038/1731032a0
- Baena-González, E., Rolland, F., Thevelein, J. M., and Sheen, J. (2007). A central integrator of transcription networks in plant stress and energy signaling. *Nature* 448, 938–942. doi: 10.1038/nature06069
- Bolouri Moghaddam, M., and Van den Ende, W. (2013). Sugars, the clock and transition to flowering. *Front. Plant Sci.* 4:22. doi: 10.3389/fpls.2013.00022
- Choubane, D., Rabot, A., Mortreau, E., Legourrierec, J., Peron, T., Foucher, F., et al. (2012). Photocontrol of bud burst involves gibberellin biosynthesis in *Rosa* sp. *J. Plant Physiol.* 169, 1271–1280. doi: 10.1016/j.jplph.2012.04.014
- Corbesier, L., Lejeune, P., and Bernier, G. (1998). The role of carbohydrates in the induction of flowering in *Arabidopsis thaliana*: comparison between the wild-type and a starchless mutant. *Planta* 206, 131–137. doi: 10.1007/s004250050383
- Duarte, G. T., Matioli, C. C., Pant, B. D., Schlereth, A., Scheible, W.-F., Stitt, M., et al. (2013). Involvement of microRNA-related regulatory pathways in the glucose-mediated control of *Arabidopsis* early seedling development. *J. Exp. Bot.* 64, 4301–4312. doi: 10.1093/jxb/ert239
- Golan, G., Betzer, R., and Wolf, S. (2013). Phloem-specific expression of a melon Aux/IAA in tomato plants alters auxin sensitivity and plant development. *Front. Plant Sci.* 4:329. doi: 10.3389/fpls.2013.00329
- Graf, A., and Smith, A. M. (2011). Starch and the clock: the dark side of plant productivity. *Trends Plant Sci.* 16, 169–175. doi: 10.1016/j.tplants.2010.12.003
- Gregory, F. G., and Veale, J. A. (1957). A reassessment of the problem of apical dominance. *Symp. Soc. Exp. Biol.* 11, 1–20.
- Haydon, M. J., Mielczarek, O., Robertson, F. C., Hubbard, K. E., and Webb, A. A. (2013). Photosynthetic entrainment of the *Arabidopsis thaliana* circadian clock. *Nature* 502, 689–692. doi: 10.1038/nature12603
- Horsfall, J. G., and Dimond, A. E. (1957). Interactions of tissue sugar, growth substances, and disease susceptibility. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz* 64, 415–421.
- Kelly, G., David-Schwartz, R., Sade, N., Moshelion, M., Levi, A., Alchanatis, V., et al. (2012). The pitfalls of transgenic selection and new roles of AtHXK1: a high level of AtHXK1 expression uncouples hexokinase1-dependent sugar signaling from exogenous sugar. *Plant Physiol.* 159, 47–51. doi: 10.1104/pp.112.196105
- Kraus, E. J., and Kraybill, H. R. (1918). *Vegetation and Reproduction with Special Reference to the Tomato*. Station bulletin (Oregon Agricultural Experiment Station). Corvallis, OR: Oregon Agricultural College Experiment Station.
- LeClere, S., Schmelz, E. A., and Chourey, P. S. (2010). Sugar levels regulate tryptophan-dependent auxin biosynthesis in developing maize kernels. *Plant Physiol.* 153, 306–318. doi: 10.1104/pp.110.155226
- Li, Y., Van den Ende, W., and Rolland, F. (2014). Sucrose induction of anthocyanin biosynthesis is mediated by DELLA. *Mol. Plant* 7, 570–572. doi: 10.1093/mp/sst161
- Loeb, J. (1924). *Regeneration from a Physico-Chemical Viewpoint*. New York, NY: McGraw-Hill.
- Martins, M. C., Hejazi, M., Fettke, J., Steup, M., Feil, R., Krause, U., et al. (2013). Feedback inhibition

- of starch degradation in *Arabidopsis* leaves mediated by trehalose 6-phosphate. *Plant Physiol.* 163, 1142–1163. doi: 10.1104/pp.113.226787
- Mason, M. G., Ross, J. J., Babst, B. A., Wienclaw, B. N., and Beveridge, C. A. (2014). Sugar demand, not auxin, is the initial regulator of apical dominance. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6092–6097. doi: 10.1073/pnas.1322045111
- Moore, B., Zhou, L., Rolland, F., Hall, Q., Cheng, W. H., Liu, Y. X., et al. (2003). Role of the *Arabidopsis* glucose sensor HXK1 in nutrient, light, and hormonal signalling. *Science* 300, 332–336. doi: 10.1126/science.1080585
- Nakabayashi, R., Yonekura Sakakibara, K., Urano, K., Suzuki, M., and Yamada, Y. (2014). Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. *Plant J.* 77, 367–379. doi: 10.1111/tpj.12388
- Nakajima, E., Nakano, H., Yamada, K., Shigemori, H., and Hasegawa, K. (2002). Isolation and identification of lateral bud growth inhibitor, indole-3-aldehyde, involved in apical dominance of pea seedlings. *Phytochemistry* 61, 863–865. doi: 10.1016/S0031-9422(02)00400-4
- Ortiz-Marchena, M. I., Albi, T., Lucas-Reina, E., Said, F. E., Romero-Campero, F. J., Cano, B., et al. (2014). Photoperiodic control of carbon distribution during the floral transition in *Arabidopsis*. *Plant Cell* 26, 565–584. doi: 10.1105/tpc.114.122721
- Rabot, A., Henry, C., Ben Baaziz, K., Mortreau, E., Azri, W., Lothier, J., et al. (2012). Insight into the role of sugars in bud burst under light in the rose. *Plant Cell Physiol.* 53, 1068–1082. doi: 10.1093/pcp/pcs051
- Robaglia, C., Thomas, M., and Meyer, C. (2012). Sensing nutrient and energy status by SnRK1 and TOR kinases. *Curr. Opin. Plant Biol.* 15, 301–307. doi: 10.1016/j.pbi.2012.01.012
- Ruan, Y.-L. (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu. Rev. Plant Biol.* 65, 26.1–26.35. doi: 10.1146/annurev-arplant-050213-040251
- Sairanen, I., Novák, O., Pencík, A., Ikeda, Y., Jones, B., Sandberg, G., et al. (2012). Soluble carbohydrates regulate auxin biosynthesis via PIF proteins in *Arabidopsis*. *Plant Cell* 24, 4907–4916. doi: 10.1105/tpc.112.104794
- Smeekens, J., and Hellmann, H. A. (2014). Sugar sensing and signaling in plants. *Front. Plant Sci.* 5:113. doi: 10.3389/fpls.2014.00113
- Thimann, K. V. (1937). On the nature of the inhibitions caused by auxin. *Am. J. Bot.* 24, 407–412. doi: 10.2307/2436422
- Van den Ende, W. (2013). Multifunctional fructans and raffinose family oligosaccharides. *Front. Plant Sci.* 4:247. doi: 10.3389/fpls.2013.00247
- Van den Ende, W., and El-Esawe, S. (in press). Sucrose signaling pathways leading to fructan and anthocyanin accumulation: a dual function in abiotic and biotic stress responses? *Environ. Exp. Bot.* doi: 10.1016/j.envexpbot.2013.09.017
- Wahl, V., Ponnus, J., Schlereth, A., Arrivault, S., Langencker, T., Franke, A., et al. (2013). Regulation of flowering by trehalose-6-phosphate signaling in *Arabidopsis thaliana*. *Science* 339, 704–707. doi: 10.1126/science.1230406
- Wardlaw, I. F., and Mortimer, D. C. (1970). Carbohydrate movement in pea plants in relation to axillary bud growth and vascular development. *Can. J. Bot.* 48, 229–237. doi: 10.1139/b70-035
- Xiong, Y., McCormack, M., Li, L., Hall, Q., Xiang, C., and Sheen, J. (2013). Glucose–TOR signalling reprograms the transcriptome and activates meristems. *Nature* 496, 181–186. doi: 10.1038/nature12030
- Yang, L., Xu, M., Koo, Y., He, J., and Poethig, R. S. (2013). Sugar promotes vegetative phase change in *Arabidopsis thaliana* by repressing the expression of MIR156A and MIR156C. *Elife* 2:e00260. doi: 10.7554/elife.00260
- Yu, S., Cao, L., Zhou, C.-M., Zhang, T.-Q., Lian, H., Sun, Y., et al. (2013). Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. *Elife* 2:e00269. doi: 10.7554/elife.00269
- Zhang, Y., Primavesi, L. F., Jhurrea, D., Andralojc, P. J., Mitchell, R. A. C., Powers, S. J., et al. (2009). Inhibition of SNF1-related protein kinase1 activity and regulation of metabolic pathways by trehalose-6-phosphate. *Plant Physiol.* 149, 1860–1871. doi: 10.1104/pp.108.133934

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 May 2014; accepted: 13 June 2014; published online: 30 June 2014.

Citation: Van den Ende W (2014) Sugars take a central position in plant growth, development and, stress responses. A focus on apical dominance. *Front. Plant Sci.* 5:313. doi: 10.3389/fpls.2014.00313

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Van den Ende. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Arabinogalactan proteins: focus on carbohydrate active enzymes

Eva Knoch, Adiphil Dilokpimol[†] and Naomi Geshi *

Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

Edited by:

Nausicaä Lanno, Ghent University, Belgium

Reviewed by:

Ian S. Wallace, University of Nevada, Reno, USA

Peter Ulvskov, Copenhagen University, Denmark

***Correspondence:**

Naomi Geshi, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Copenhagen, Denmark
e-mail: nge@plen.ku.dk

†Present address:

Adiphil Dilokpimol, Fungal Physiology, CBS-KNAW Fungal Biodiversity Center, Utrecht, Netherlands

INTRODUCTION

Arabinogalactan proteins (AGPs) are a family of proteoglycans found on the plasma membrane and in the cell walls of diverse species of plants. AGPs are synthesized by several post-translational modifications of proteins in the secretory pathway. The proteins generally contain repetitive dipeptide motifs, e.g., Ala-Pro, Ser-Pro, Thr-Pro, and Val-Pro, which are distinguished from the sequence motifs for extensin type glycosylation [e.g., Ser-(Pro)₂₋₃] known as another major class of O-glycosylation in plants (Kieliszewski, 2001). The Pro residues are hydroxylated by prolyl 4-hydroxylases and further O-glycosylated by glycosyltransferases (GTs). Moreover, many AGPs are attached by a glycosylphosphatidylinositol anchor, which attaches AGPs to the plasma membrane, but can be cleaved by phospholipases (Wang, 2001; Schultz et al., 2004). AGPs on the plasma membrane and cell wall may also be processed by proteolytic activities and glycosidases or transported by endocytotic multivesicular bodies to the vacuole where they are degraded (Herman and Lamb, 1992).

The glycan moiety of AGPs accounts for more than 90% of their total mass, which has been suggested to play an essential role in the function of AGPs, based on studies using synthetic phenylglycoside dyes (β -Yariv reagents) that specifically binds to the β -1,3-galactan moiety of AGPs (Kitazawa et al., 2013) as well as various monoclonal antibodies that recognize different AGP glycan epitopes (Seifert and Roberts, 2007). However, because of its complexity and heterogeneity, little is known about the structure-function relationship of AGP glycans. In fact, various structures have been reported for AGP glycans depending on samples and analytical methods. The common structural feature is a backbone of β -1,3-galactan, which is often substituted at O6 with side chains

Arabinogalactan proteins (AGPs) are a highly diverse class of cell surface proteoglycans that are commonly found in most plant species. AGPs play important roles in many cellular processes during plant development, such as reproduction, cell proliferation, pattern formation and growth, and in plant-microbe interaction. However, little is known about the molecular mechanisms of their function. Numerous studies using monoclonal antibodies that recognize different AGP glycan epitopes have shown the appearance of a slightly altered AGP glycan in a specific stage of development in plant cells. Therefore, it is anticipated that the biosynthesis and degradation of AGP glycan is tightly regulated during development. Until recently, however, little was known about the enzymes involved in the metabolism of AGP glycans. In this review, we summarize recent discoveries of carbohydrate active enzymes (CAZy; <http://www.cazy.org/>) involved in the biosynthesis and degradation of AGP glycans, and we discuss the biological role of these enzymes in plant development.

Keywords: arabinogalactan proteins, type II arabinogalactan, plant cell wall, carbohydrate active enzymes, glycosyltransferase, glycoside hydrolase

of β -1,6-galactan decorated further with arabinose, and less frequently also with fucose, rhamnose, and (methyl) glucuronic acid (Figures 1A,B). Tan et al. (2010) proposed that the backbone is composed of a repeat of a β -1,3-galactotriose unit with or without side chains, which is connected by β -1,6-linkages (kinks). This model is based on the AGPs synthesized onto synthetic peptides expressed in tobacco cells and analyzed by NMR (Tan et al., 2004, 2010). In this model, the side chains are rather short and composed of a single Gal decorated by 1–5 other sugars. However, longer β -1,6-galactan side chains have been reported for AGPs from radish root (Haque et al., 2005), wheat flour (Tryfona et al., 2010) and *Arabidopsis* leaf (Tryfona et al., 2012) based on the linkage and mass spectroscopy analysis.

Knowledge about each enzyme working on an individual step in the biosynthesis and degradation of AGP glycans is useful to understand the role of a particular sugar moiety of AGPs. This review outlines the recent findings for the carbohydrate active enzymes (CAZy; <http://www.cazy.org/>, Lombard et al., 2014) identified to be responsible for the biosynthesis and degradation of AGPs. The reader is referred to other excellent reviews for other topics with respect to structure, cell biological functions (Seifert and Roberts, 2007), localization, and commercial interests of AGPs (Nothnagel, 1997; Schultz et al., 1998; Majewska-sawka and Nothnagel, 2000; Gaspar et al., 2001; Showalter, 2001; Ellis et al., 2010; Tan et al., 2012).

GLYCOSYLTRANSFERASES INVOLVED IN AGP BIOSYNTHESIS

A large number of functionally distinct GTs are required for the biosynthesis of complex AGP glycans, e.g.,

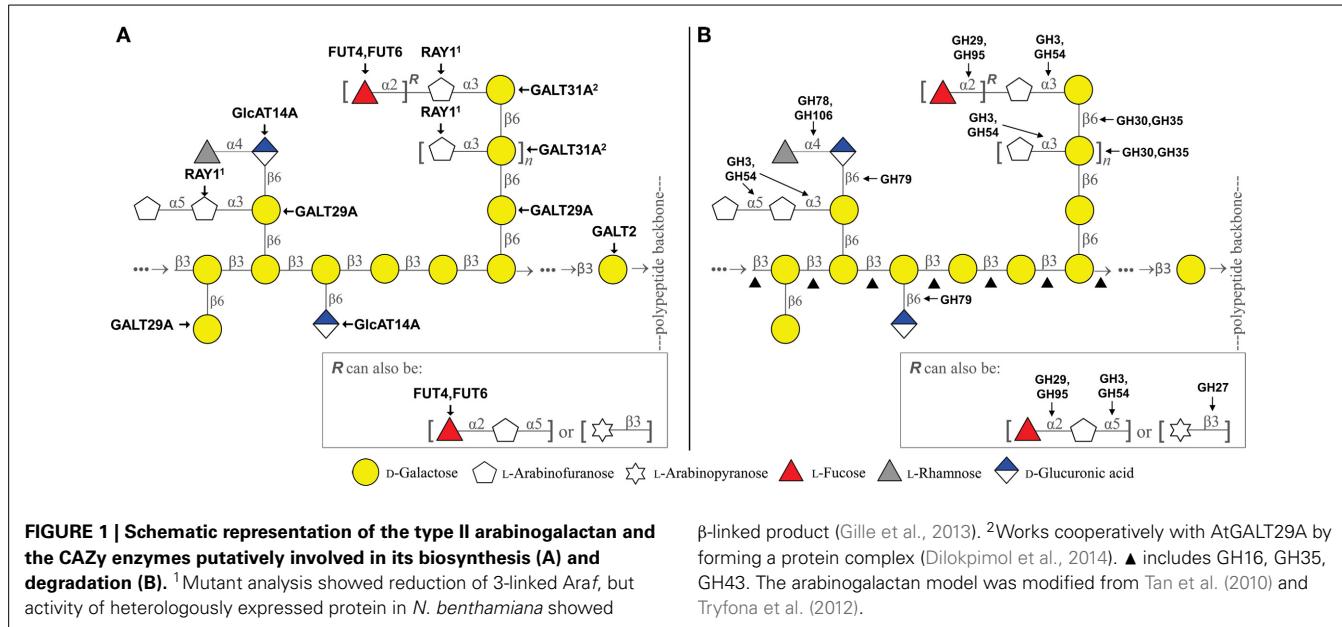


FIGURE 1 | Schematic representation of the type II arabinogalactan and the CAZy enzymes putatively involved in its biosynthesis (A) and degradation (B). ¹Mutant analysis showed reduction of 3-linked Araf, but activity of heterologously expressed protein in *N. benthamiana* showed

β -linked product (Gille et al., 2013). ²Works cooperatively with AtGALT29A by forming a protein complex (Dilokpimol et al., 2014). ▲ includes GH16, GH35, GH43. The arabinogalactan model was modified from Tan et al. (2010) and Tryfona et al. (2012).

β -1,3-galactosyltransferases (GalTs), β -1,6-GalTs, α -1,3- and α -1,5-arabinosyltransferases, fucosyltransferases, rhamnosyltransferases, glucuronosyltransferases, and glucuronic acid methyltransferases. Several GTs identified to date (Figure 1A and Table 1) are summarized below.

β -GALACTOSYLTTRANSFERASES

The first step in the glycosylation of AGPs is the transfer of Gal to hydroxyproline residues present in the peptide backbone. The Arabidopsis enzyme catalyzing this step was identified (At4g21060, AtGALT2, Basu et al., 2013). This enzyme belongs to the CAZy family GT31 and the recombinant protein expressed in *Pichia pastoris* demonstrated GalT activity transferring a Gal to hydroxyproline residues in the synthetic AGP peptides. Arabidopsis T-DNA knockout mutants contained reduced levels of Yariv precipitable AGPs and microsomes purified from mutants exhibited reduced levels of GalT activity compared to wild type. The mutant lines showed no detectable growth phenotype under normal growth conditions. Since the GalT activity was not completely abolished in the mutant microsomes, redundant activities encoded by other genes most likely exist. Nevertheless, based on this study, the quantity of AGP-glycans appears to not be crucial for plant development.

Another Arabidopsis GT from family GT31 encoded by At1g32930 was also characterized. Recombinant enzyme expressed in *Escherichia coli* and *Nicotiana benthamiana* demonstrated β -1,6-GalT activity elongating β -1,6-galactan side chains of AGP glycans in *in vitro* assays (AtGALT31A; Geshi et al., 2013). AtGALT31A is expressed specifically in the suspensor cells of the embryo proper and T-DNA insertion lines showed abnormal cell division in the hypophysis and arrested further development of embryos. Therefore, functional AtGALT31A is essential for normal plant embryogenesis. How AtGALT31A that is expressed in suspensor cells influences cell division in hypophysis remains unknown.

AtGALT29A (At1g08280) was identified as a gene co-expressed with AtGALT31A. Recombinant enzyme expressed in *Nicotiana benthamiana* demonstrated β -1,6-GalT activities elongating β -1,6-galactan and forming 6-Gal branches on β -1,3-galactan of AGP glycans (Dilokpimol et al., 2014). Moreover, Förster resonance energy transfer analysis revealed an interaction between AtGALT29A and AtGALT31A when both proteins are expressed as C-terminal fluorescent fusion proteins in *Nicotiana benthamiana* (Dilokpimol et al., 2014). The protein complex containing heterologously expressed AtGALT29A and AtGALT31A were purified and demonstrated increased levels of β -1,6-GalT activities by the AtGALT29A single enzyme. These results suggest cooperative action between AtGALT31A and AtGALT29A by forming an enzyme complex, which could be an important regulatory mechanism for producing β -1,6-galactan side chains of type II AG during plant development.

β -GLUCURONOSYLTTRANSFERASE

An Arabidopsis GT from family GT14 encoded by At5g39990 was identified as a glucuronosyltransferase involved in the biosynthesis of AGP glycans (AtGlcAT14A; Knoch et al., 2013). The enzyme expressed in *Pichia pastoris* demonstrated β -GlcAT activity by adding GlcA to both β -1,6- and β -1,3-galactan. Arabidopsis possesses 11 proteins in the GT14 family, of which two additional proteins encoded by At5g15050 and At2g37585 also demonstrated the same β -GlcAT activities and were named AtGlcAT14B and AtGlcAT14C, respectively (Dilokpimol and Geshi, 2014). The T-DNA insertion lines contained reduced levels of GlcA substitution of β -1,6-galactobiose and β -1,3-galactan in their AGPs compared to wild type. In addition to the altered levels of GlcA, a marked increase of Gal and a decrease of Ara were detected in the mutant AGPs. Mutant lines showed an increased cell elongation rate in dark grown hypocotyls and light grown roots during seedling growth compared to wild type. Since several sugars were altered in the mutant AGPs lines, it is unlikely that the observed phenotype

Table 1 | Characterized GTs and GHs, which process AGP-glycans.

CAZy Family	Activity	Protein name	Origin	No. of genes in Arabidopsis	Evidence ¹	Comments for enzyme activities and genetic manipulations	Selected references
GT14	β-glucuronosyltransferase	AtGlcAT14A	Arabidopsis thaliana	11	HE-P; MA	GlcAT activity to β-1,3 and β-1,6-galactan; [GlcA]↑, [Ara]↓ in AG and enhanced cell elongation in seedlings in atg/cat1/4a	Knoch et al., 2013
		AtGlcAT14B	Arabidopsis thaliana			GlcAT activity to β-1,3 and β-1,6-galactan	Dilokpimol and Geshi, 2014
		AtGlcAT14C	Arabidopsis thaliana				
GT29	β-1,6-galactosyltransferase	AtGALT29A	Arabidopsis thaliana	3	HE-N	β-1,6-GalT activity to β-1,3 and β-1,6-galactan; interaction with AtGALT31A enhances the activity	Dilokpimol et al., 2014
GT31	Hydroxyproline O-galactosyltransferase	AtGALT2	Arabidopsis thaliana	33	HE-P; MA	GalT activity to hydroxyproline; [Yariv-precipitable AG]↓ in gait2; no detectable growth phenotype under normal growth condition	Basu et al., 2013
	β-1,6-galactosyltransferase	AtGALT31A	Arabidopsis thaliana		HE-N, -E; MA	β-1,6-GalT activity elongating β-1,6-galactan; mutant is embryo-lethal	Geshi et al., 2013
GT37	α-1,2-fucosyltransferase	AtFUT4	Arabidopsis thaliana	10	HE-B; MA	Fuct activity to AGPs from BY2; [Fuc]↓ in AG in fut4 and fut6; no [Fuc] in AG from fut4/fut6; no detectable growth phenotype under normal growth condition, but reduced root growth under salt stress	Wu et al., 2010; Liang et al., 2013; Tryfona et al., 2014
GT77	Arabinofuranosyltransferase	AtRAY12	Arabidopsis thaliana	19	HE-N, MA	β-ArafT activity to methyl β-Gal; [3-linked Ara]↓ in AG of ray1, slower root growth ²	Gille et al., 2013
GH3	Exo-α-arabinofuranosidase	RsAraF1	Raphanus sativus L.	16 ³	HE-A	Cleaves α-linked Araf from AGPs, pectic α-1,5-arabinan, arabinoxylan. Overexpression in Arabidopsis resulted in [Ara]↓ in cell walls, but no growth phenotype	Kotake et al., 2006
GH16	Endo-β-1,3-galactanase	FvEn3GAL	Flammulina velutipes	33 ⁴	PUR, HE-P	Cleaves β-1,3-galactan in endo fashion	Kotake et al., 2011
GH27	Exo-β-arabinopyranosidase	SaArap27A	Streptomyces avermitillii	4	HE-S	Cleaves L-Arap from p-nitrophenyl-β-L-Arap and releases Ara	Fujimoto et al., 2009; Ichinose et al., 2009
GH30	Endo-β-1,6-galactanases	Tv6GAL	Trichoderma viride	0	HE-E	Cleaves β-1,6-galactan in endo fashion	Kotake et al., 2004
		FoGal1	Fusarium oxysporum		PUR, HE-E		Sakamoto et al., 2007
		Sa1,6Gal5A	Streptomyces avermitillii		HE-E		Ichinose et al., 2008
		Nc6GAL	Neurospora crassa		HE-P		Takata et al., 2010
GH35	Exo-β-1,3,1,6-galactosidase	RsBGAL1	Raphanus sativus	18	PUR, HE-P	Exo activity to β-1,3- and β-1,6-Gal, but not to β-1,4-Gal	Kotake et al., 2005
						Cooperative degradation of AG with arabinofuranosidase and glucuronidase	

(Continued)

Table 1 | Continued

CAZY Family	Activity	Protein name	Origin	No. of genes in <i>Arabidopsis</i>	Evidence ¹	Comments for enzyme activities and genetic manipulations	Selected references
		AtBGAL4	<i>Arabidopsis thaliana</i>	HE-E, HE-I	Preferred cleavage at β -1,3 and β -1,4-linked Gal rather than β -1,6-linked Gal	Ahn et al., 2007	
GH43	Exo- β -1,3-galactanase	Pc1,3Gal43A Ct1,3Gal43A Sa1,3Gal43A Il1,3Gal SGalase1, 2	<i>Phanerochaete chrysosporium</i> <i>Clostridium thermocellum</i> <i>Streptomyces avermitillii</i> <i>Lipex lacteus</i> <i>Streptomyces</i> sp.	2 HE-E HE-E HE-E HE-P HE-E	Cleaves β -1,3-linked Gal regardless the presence or absence of substituted side chains	Ishida et al., 2009 Ichinose et al., 2006b Ichinose et al., 2006a Kofake et al., 2009 Ling et al., 2012	
GH54	Exo- α -arabinofuranosidase	NcAraf1	<i>Neurospora crassa</i>	0	HE-P	Broad specificity to α -1,3 and α -1,5-Araf, which includes AGPs, pectic arabinan, arabinoxylan	Takata et al., 2010
GH78	Exo- α -rhamnosidase	SaRha78A	<i>Streptomyces avermitillii</i>	0	PUR, HE-E	Releases Rha from gum Arabic	Ichinose et al., 2013
GH79	Exo- β -glucuronidase	AnGlcAase NcGlcAase AtGUS2	<i>Aspergillus niger</i> <i>Neurospora crassa</i> <i>Arabidopsis thaliana</i>	3 PUR, HE-P PUR, HE-A, MA	Cleaves both GlcA and methyl GlcA from AG. Methyl GlcA from long β -1,6-galactan is cleaved, but not from short β -1,6-galactan Cleaves GlcA from AG Cleaves <i>p</i> -nitrophenyl- β -D-Glca. Mutant <i>atgus2</i> : [GlcA] \uparrow , [Ara] \downarrow , [Xyl] \downarrow in AG and reduced cell elongation in seedlings; overexpression AtGUS2: no [GlcA]; [Ara] \downarrow , [Xyl] \uparrow in AG and enhanced cell elongation in seedlings	Haque et al., 2005; Konishi et al., 2008 Konishi et al., 2008 Eudes et al., 2008	
GH95	Exo- α -1,2-fucosidase	AfcA	<i>Bifidobacterium bifidum</i>	1 ⁵	PUR, HE-E	Cleaves α -1,2-linked Fuc (linkage present in AG)	Nagae et al., 2007
GH106	Exo- α -rhamnosidase	Rham	<i>Sphingomonas paucimobilis</i>	0	PUR, HE-E	Broad specificity to α -Rha containing components. Involvement in AG degradation is unclear	Miyata et al., 2005

GTs are only from plants, while GHs are both from plants and microbial origins.

¹HE, activity demonstrated from heterologously expressed protein in: A. *thaliana*; B. tobacco BY2 cell; E. *E. coli*; N. *N. benthamiana*; P. *P. pastoris*; S. *Streptomyces cinnamonneus*. I. *baculovirus/insect cells*; MA, mutant analysis.

²Mutant analysis showed reduction of 3-linked Ara in AGPs, but heterologously expressed protein in *N. benthamiana* showed β -linked Ara to methyl β -Gal product, therefore a role of this protein in AGP glycosylation is not certain since arabinose exists as an α -linked sugar in AGPs (Gille et al., 2013).

³Characterized Arabidopsis enzymes demonstrated β -xylosidase activity toward xylan (Goujon et al., 2003; Minic et al., 2004).

⁴Characterized Arabidopsis enzymes demonstrated xyloglucan endo-transferase activity (Rose, 2002).

⁵Characterized Arabidopsis enzymes demonstrated α -1,2-fucosidase activity specifically toward xyloglucan (Léonard et al., 2008; Günl et al., 2011).

is solely a consequence of the reduced levels of GlcA, but most likely related to the dynamic conformational changes of AGP glycans caused in the mutants.

α -FUCOSYLTRANSFERASE

Two *Arabidopsis* GTs from family GT37 were identified as α -1,2-fucosyltransferases involved in the biosynthesis of AGP glycans (AtFUT4 and AtFUT6, encoded by *At2g15390* and *At1g14080*, respectively; Wu et al., 2010). AGPs from tobacco BY2 cells contain no fucose, but heterologous expression of AtFUT4 and AtFUT6 in BY2 cells resulted in fucosylated AGPs. The recombinant enzymes purified from BY2 cells demonstrated fucosyltransferase activity to endogenous AGPs. Single *Arabidopsis* T-DNA insertion lines, *atfut4* and *atfut6*, contained reduced levels of fucose in AGPs, and the double T-DNA insertion line *fut4/fut6* contained no detectable fucose in its AGPs (Liang et al., 2013); however, no obvious phenotype was observed in both types of mutants when they were grown under normal conditions. Differences between wild type and mutants were only seen in seedlings grown under salt stressed condition, where the mutant lines showed reduced root growth compared to wild type (Liang et al., 2013; Tryfona et al., 2014). This was somewhat surprising because the *Arabidopsis mur1* mutant, which is defective in a GDP-mannose-4,6-dehydratase (Bonin et al., 1997), contained a 40% reduction of fucose in root extracts (Reiter et al., 1997) and showed a 50% reduction of cell elongation rate in roots (Van Hengel and Roberts, 2002). The decrease of root cell elongation in *mur1* was previously attributed to the lack of fucose in AGPs (Van Hengel and Roberts, 2002), but the findings by Liang et al. (2013) and Basu et al. (2013) refute that hypothesis. The molecular changes behind the *mur1* phenotype are hard to pinpoint. Mutants affected in N-glycan fucosylation show reduced root growth under salt stress conditions (Kaulfürst-Soboll et al., 2011), and *mur2* plants, which lack only xyloglucan fucosylation, have no visible root phenotype (Van Hengel and Roberts, 2002; Vanzin et al., 2002). The *mur1* phenotype might be due to under-fucosylated rhamnogalaturonan II, or to the combination of several cell wall polysaccharides deficient in fucose.

α -ARABINOFRANOSYLTTRANSFERASE

An *Arabidopsis* GT encoded by *At1g70630* (named REDUCED ARABINOSE YARIV1, RAY1; Gille et al., 2013), was characterized as a putative arabinofuranosyltransferase since the mutation caused a reduced level of arabinofuranose (Araf) in its AGPs. This GT belongs to the GT77 family, which also contains *XEG113*, the mutation of which results in the reduction of β -linked arabinose in extensin (Gille et al., 2009). Therefore, Araf transferase activity that makes β -linkages was expected for RAY1, and indeed, microsomes isolated from *Nicotiana benthamiana* after expression of recombinant RAY1 demonstrated β -Araf transferase activity to methyl β -Gal. The T-DNA insertion lines contained reduced levels of 3-linked Ara in its AGP fractions compared to AGPs from wild type, and the mutant plants exhibited slower root growth as well as a reduced rosette size and inflorescence. However, β -1,3-Linked Araf has not been reported in AGPs, therefore the involvement of RAY1 in the biosynthesis of AGP glycans remains unclear.

GLYCOSIDE HYDROLASES

Glycoside hydrolases (GHs) acting on AGPs are potentially very important for the metabolism of these glycoproteins. AGPs from tobacco stylar transmitting tissue are degraded as the pollen tube grows and the released sugars are considered to be used as the carbohydrate resource necessary for the elongation of pollen tubes (Cheung et al., 1995). Similarly, rapid turnover of AGPs are observed in suspension cell culture and millet seedlings (Gibeaut and Carpita, 1991) and a substantial amount of AGPs are considered to be hydrolyzed to free sugars and recycled in the cytosol for the synthesis of new glycans (Gibeaut and Carpita, 1991) or degraded in the vacuole (Herman and Lamb, 1992). The appearance of distinct AGP epitopes in a developmentally regulated manner might be controlled by GHs in the cell walls. Additionally, the occurrence of free AG glycans detached from proteins observed in the cell walls may be a result of GH actions.

For the hydrolysis of AGP glycans, several GHs are required, e.g., β -galactosidases, β -galactanases, α -arabinofuranosidases, β -arabinopyranosidases, β -glucuronidases, α -fucosidases, and α -rhamnosidases. AGP degrading GHs from microbial origin have been relatively well characterized, while only a few plant GHs have been reported to degrade AGP glycans. Below is an overview for those GHs reported to possess hydrolase activity of AGP glycans from both microbial and plant origins (Figure 1B and Table 1).

β -GALACTOSIDASE AND β -GALACTANASE

A GH16 from the fungus *Flammulina velutipes* was characterized as an endo- β -1,3-galactanase degrading the AGP glycan β -1,3-galactan backbone (Kotake et al., 2011). The enzyme activity is distinct from other GH16 enzymes, which comprise β -1,3- and β -1,3:1,4-glucanases, xyloglucan endo-transglycosylase and β -agarase activities. *Arabidopsis* contains 33 proteins in the GH16 family and they are characterized as xyloglucan endo-transglycosylases and their homologs (Kaewthai et al., 2013).

Several enzymes of microbial origin in GH30 have been characterized as endo- β -1,6-galactanases that hydrolyze β -1,6-galactan side chains of AGP glycans (Kotake et al., 2004; Sakamoto et al., 2007; Ichinose et al., 2008; Takata et al., 2010). These β -1,6-galactanases were originally categorized as part of the GH5 family, but were moved to the GH30 family after additional bioinformatic analysis by (St John et al., 2010). *Tv6GAL* from *Trichoderma viride* was the first β -1,6-galactanase cloned and characterized (Kotake et al., 2004). This enzyme specifically recognizes β -1,6-galactan of AGPs and releases galactose and β -1,6-linked galactooligomers with a degree of polymerization from two to five. Efficiency of the hydrolysis of β -1,6-galactan is increased by pretreatment of the AGP substrate with α -L-arabinofuranosidase. The β -1,6-galactanases from *Fusarium oxysporum* (FoGAL1, Sakamoto et al., 2007), *Streptomyces avermitilis* (Sa1,6Gal5A, Ichinose et al., 2008), and *Neurospora crassa* (Nc6GAL, Takata et al., 2010) act in a similar manner, releasing galactose and β -1,6-galactobiose from β -1,6-galactan of AGP. All three enzymes show increased activity on de-arabinosylated AGP, similarly to *Tv5GAL*. Plants do not have any proteins classified in the GH16 CAZy family.

Several microbial β -1,3-galactosidases from GH43 have been characterized as exo- β -1,3-galactanase that degrades the β -1,3-galactan backbone of AGP glycans (Ichinose et al., 2005, 2006a,b; Kotake et al., 2009). All of these enzymes show similar substrate specificity and degrade β -1,3-linked Gal regardless of the substitution of side chains, which results in free Gal from unsubstituted β -1,3-galactan and side chains attached to β -1,3-linked Gal. Therefore, the GH43 enzymes have been used to release side chains from AGP glycans for structural analysis (Tryfona et al., 2010, 2012; Geshi et al., 2013; Knoch et al., 2013). The GH43 CAZy family contains two uncharacterized Arabidopsis proteins.

In plants, β -1,3-galactosidase has been purified from radish (*Raphanus sativus*) seed extracts (Kotake et al., 2005). Based on the deduced protein sequence, the enzyme RsBGAL1 was classified to the GH35 CAZy family. This enzyme was expressed heterologously in *Pichia pastoris* and demonstrated GH activity by degrading β -1,3- and β -1,6-galactan in an exo manner, but not β -1,4-galactan. The efficiency of degradation of AGP glycans by RsBGAL1 alone was limited, but co-treatment with arabinofuranosidase and glucuronidase resulted in the release of up to 90% of the bound sugars from AGPs, indicating the synergy of those GHs in the degradation of AGP glycans.

α -ARABINOFURANOSIDASE AND β -ARABINOPYRANOSIDASE

An α -arabinofuranosidase from the fungus *Neurospora crassa* with a broad substrate specificity toward AGPs, pectic arabinan, and arabinoxylan was identified and classified to the CAZy GH54 family (NcAraf1, Takata et al., 2010). This enzyme was heterologously expressed in *Pichia pastoris* and demonstrated α -arabinofuranosidase activity on both α -1,3- and α -1,5-linked Araf. NcAraf1 has been used extensively for the structural characterization of AGP glycans together with galactosidases, galactanases, and glucuronidases (Tsumuraya et al., 1990; Okemoto et al., 2003; Kotake et al., 2004, 2009; Konishi et al., 2008; Tryfona et al., 2010, 2012). Arabidopsis does not have proteins in the GH54 family.

Plant α -arabinofuranosidase acting on AGP glycans is classified to family GH3. Kotake et al. (2006) purified an α -arabinofuranosidase from radish seeds and named it RsAraf1. The recombinant enzyme expressed in Arabidopsis demonstrated hydrolytic activity on radish AGPs, pectic α -1,5-arabinan and arabinoxylan. Transgenic Arabidopsis overexpressing RsAraf1 showed decreased levels of Ara in the cell wall, but no obvious growth phenotype was observed compared to wild type plants.

β -Arabinopyranose generally represents only a minor part of Ara in AGP glycans, but has been reported from acacia, larch and wheat flour AGPs (Aspinall et al., 1958; Groman et al., 1994; Odonmazig et al., 1994; Tryfona et al., 2010). β -Arabinopyranosidase has been identified from *Streptomyces avermitilis* (Ichinose et al., 2009). The enzyme, named SaArap27A, belongs to family GH27, and the recombinant enzyme expressed in *Streptomyces* demonstrated the release of L-arabinopyranoside (Arap) from p-nitrophenyl- β -L-arabinopyranoside, as well as the release of L-arabinose from gum Arabic and larch AG. Arabidopsis contains four uncharacterized proteins in the GH27 family.

β -GLUCURONIDASE

Microbial β -glucuronidases are found in family GH79. Two fungal GH79 β -glucuronidases from *Neospora crassa* (NcGlcAase) and *Aspergillus niger* (AnGlcAase) have been cloned and recombinant proteins expressed in *Pichia pastoris* demonstrated β -glucuronidase activity (Konishi et al., 2008). AnGlcAase and NcGlcAase share high homology in their amino acid sequences, but possess slightly different substrate specificity. Both enzymes recognize unsubstituted and 4-methyl substituted β -GlcA on AGPs, but AnGlcAase cleaves both GlcA and 4-methyl GlcA with an equal efficiency, while NcGlcAase preferably cleaves GlcA and only small amounts of 4-methylGlcA. Arabidopsis contains three proteins in the GH79 family.

In plants, β -glucuronidases (GUS) are ubiquitously present and their activity is associated with cell elongation (Sudan et al., 2006). Eudes et al. (2008) partially purified a GUS from Arabidopsis stems, which cleaves *p*-nitrophenyl- β -D-glucuronide. The corresponding gene was identified and classified to the GH79 family (AtGUS2, At5g07830, Eudes et al., 2008). The T-DNA knockout insertion lines exhibited increased levels of GlcA, whereas plants overexpressing AtGUS2 lacked detectable levels of GlcA in their AGP fractions. The T-DNA insertion mutant of *AtGUS2* showed no clear changes in the elongation rate of plant organs, whereas the overexpression lines exhibited increased elongation of roots and stems. The increase of cell elongation observed in the overexpression lines of *AtGUS2* resembles similar observations of the *atglcat14a* T-DNA insertion lines. Although reduced levels of GlcA is observed in both types of plants, the altered profiles of other sugars present in AGP glycans are inconsistent. Therefore, it is unlikely that the increase of cell elongation is solely caused by the reduction of GlcA levels in AGPs.

α -FUCOSIDASE

α -Fucosidases from various prokaryotic and eukaryotic sources have been characterized, and several of them are commercially available. α -Fucosidases are classified into two GH families: GH29 and GH95 (Lombard et al., 2014). α -Fucosidases from GH29 are capable of hydrolyzing various types of linkages, mainly α -1,3/1,4-linked Fuc, whereas GH95 enzymes are active solely on α -1,2-linked Fuc. Only one GH from Arabidopsis is found in each of these GH families. *At2g28100* (*AtFUC1*; Zeleny et al., 2006) belongs to GH29 and the recombinant enzyme expressed in *Pichia pastoris* demonstrated α -1,3/1,4-fucosidase activity (Zeleny et al., 2006). *At4g34260* (Fuc95A, AXY8; Léonard et al., 2008; Günzl et al., 2011) belongs to GH95 and the enzyme heterologously expressed in *Nicotiana benthamiana* demonstrated α -1,2-fucosidase activity (Léonard et al., 2008). α -1,2-Fuc is present in both xyloglucan and AGP glycans, but Fuc95A (AXY8) acts specifically on α -1,2-Fuc on xyloglucan and not on AGPs (Günzl et al., 2011). α -1,2-Fucosidase purified from *Xanthomonas manihotis* apparently cleaves α -1,2-fucose on AGP glycans and has been used for the product analysis of AGP fucosyltransferases (*AtFUT4* and *AtFUT6*, Wu et al., 2010). Wu et al. (2010) also used α -1,3/4-fucosidase from almond meal for the characterization of the Fuc linkage. Both enzymes are commercially available, but are not classified to CAZy GH families.

α-RHAMNOSIDASE

Microbial α -rhamnosidases are classified into three GH families: GH28, GH78, and GH106 (Fujimoto et al., 2013; Lombard et al., 2014). An α -rhamnosidase from *Aspergillus niger* in GH28 was identified as specifically degrading pectic rhamnogalacturonan (RgxB; Martens-Uzunova et al., 2006). Several α -rhamnosidases from GH78 have been characterized and recently an α -rhamnosidase from *Streptomyces avermitilis* expressed in *Escherichia coli* demonstrated an α -rhamnosidase that releases rhamnose (Rha) from gum Arabic AGPs (SaRha78A; Fujimoto et al., 2013; Ichinose et al., 2013). GH106 exclusively contains bacterial α -rhamnosidases, of which only one has been characterized to date. This α -rhamnosidase was purified from *Sphingomonas paucimobilis* FP2001 (Rham; Miyata et al., 2005) and the enzyme expressed in *Escherichia coli* demonstrated α -rhamnosidase activity on a broad range of substrates containing α -Rha. One of those substrates is α -rhamnosyl-1,4-galactose, but whether the enzyme hydrolyzes α -rhamnosyl-1,4-GlcA, which is found as part of the side chains of AGP glycans, remains unknown. Arabidopsis contains 28 proteins in the GH28 family, but no plant proteins are present in GH78 and GH106. Among Arabidopsis GH28s, only pectin polygalacturonase has been characterized (Torki et al., 2000; Markovic and Janecek, 2001). Functions of other plant GHs in GH28 remain unknown.

CONCLUSIONS

Microbial GHs working on the degradation of plant AGPs have been reported in several studies, but little was known about the enzymes working on the biosynthesis and degradation of AGPs in plants. Recent discovery of plant GTs/GHs working on AGPs, together with the technical development of in-depth structural analysis of complex AGP glycans, has broadened our knowledge for AGP metabolism significantly. On the other hand, the attempt to elucidate the biological role of each sugar moiety or a particular part of AGP glycan structure by investigating knockout mutants or overexpressors of those enzymes did not result in straightforward answers. For instance, a mutation in *AtGlcAT14A* did not result in a sole reduction of GlcA but also exhibited an increase of Gal and a reduction of Ara in the AGP glycan as well as enhanced cell elongation in seedlings. Furthermore, the overexpression of *AtGUS2* resulted in a reduction of GlcA, Gal and Ara in AGP glycans and seedlings showed increased cell elongation, similarly to *atglcat14a*. The developmentally regulated appearance of different AGP glycan epitopes is well known, but the results available thus far are inconclusive concerning the molecular role of a particular part of AGP glycans in plant growth and development.

The carbohydrate active enzymes involved in the AGP metabolism have just begun to be identified and characterized. Further investigation of the remaining members in the AGP glycosylation pathway and their role *in vivo* is needed to understand the role of CAZy enzymes in relation to AGP glycans, the cell wall architecture, and in plant growth and development.

REFERENCES

- Ahn, Y. O., Zheng, M., Bevan, D. R., Esen, A., Shiu, S.-H., Benson, J., et al. (2007). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. *Phytochemistry* 68, 1510–1520. doi: 10.1016/j.phytochem.2007.03.021
- Aspinall, G. O., Hirst, E. L., Else, R., and Ramsta, E. (1958). The constitution of larch galactan. *J. Chem. Soc.* 593–601. doi: 10.1039/jr9580000593
- Basu, D., Liang, Y., Liu, X., Himmeldirk, K., Faik, A., Kieliszewski, M., et al. (2013). Functional identification of a hydroxyproline-o-galactosyltransferase specific for arabinogalactan protein biosynthesis in *Arabidopsis*. *J. Biol. Chem.* 288, 10132–10143. doi: 10.1074/jbc.M112.432609
- Bonin, C. P., Potter, I., Vanzin, G. F., and Reiter, W.-D. (1997). The MUR1 gene of *Arabidopsis thaliana* encodes an isoform of GDP-D-mannose-4,6-dehydratase, catalyzing the first step in the de novo synthesis of GDP-L-fucose. *Proc. Natl. Acad. Sci. U.S.A.* 94, 2085–2090. doi: 10.1073/pnas.94.5.2085
- Cheung, A. Y., Wang, H., and Wu, H. M. (1995). A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. *Cell* 82, 383–393. doi: 10.1016/0092-8674(95)90427-1
- Dilokpimol, A., and Geshi, N. (2014). *Arabidopsis thaliana* glucuronosyltransferase in family GT14. *Plant Signal. Behav.* 9:e28891. doi: 10.4161/psb.28891
- Dilokpimol, A., Poulsen, C. P., Vereb, G., Kaneko, S., Schulz, A., and Geshi, N. (2014). Galactosyltransferases from *Arabidopsis thaliana* in the biosynthesis of type II arabinogalactan: molecular interaction enhances enzyme activity. *BMC Plant Biol.* 14:90. doi: 10.1186/1471-2229-14-90
- Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-proteins (AGPs): key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000
- Eudes, A., Mouille, G., Thévenin, J., Goyallon, A., Minic, Z., and Jouanin, L. (2008). Purification, cloning and functional characterization of an endogenous beta-glucuronidase in *Arabidopsis thaliana*. *Plant Cell Physiol.* 49, 1331–1341. doi: 10.1093/pcp/pcn108
- Fujimoto, Z., Ichinose, H., Harazono, K., Honda, M., Uzura, A., and Kaneko, S. (2009). Crystallization and preliminary crystallographic analysis of beta-L-arabinopyranosidase from *Streptomyces avermitilis* NBRC14893. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 632–634. doi: 10.1107/S1744309109017230
- Fujimoto, Z., Jackson, A., Michikawa, M., Maehara, T., Momma, M., Henrissat, B., et al. (2013). The structure of a *Streptomyces avermitilis* α -L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement. *J. Biol. Chem.* 288, 12376–12385. doi: 10.1074/jbc.M113.460097
- Gaspar, Y., Johnson, K. L., McKenna, J. A., Bacic, A., and Schultz, C. J. (2001). The complex structures of arabinogalactan-proteins and the journey towards understanding function. *Plant Mol. Biol.* 47, 161–176. doi: 10.1023/A:1010683432529
- Geshi, N., Johansen, J. N., Dilokpimol, A., Rolland, A., Belcram, K., Verger, S., et al. (2013). A galactosyltransferase acting on arabinogalactan protein glycans is essential for embryo development in *Arabidopsis*. *Plant J.* 76, 128–137. doi: 10.1111/tpj.12281
- Gibeaut, D. M., and Carpita, N. C. (1991). Tracing cell wall biogenesis in intact cells and plants: selective turnover and alteration of soluble and cell wall polysaccharides in grasses. *Plant Physiol.* 97, 551–561. doi: 10.1104/pp.97.2.551
- Gille, S., Hänsel, U., Ziermann, M., and Pauly, M. (2009). Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14699–14704. doi: 10.1073/pnas.0905434106
- Gille, S., Sharma, V., Baidoo, E. E. K., Keasling, J. D., Scheller, H. V., and Pauly, M. (2013). Arabinosylation of a yariv-precipitable cell wall polymer impacts plant growth as exemplified by the *Arabidopsis* glycosyltransferase mutant ray1. *Mol. Plant* 6, 1369–1372. doi: 10.1093/mp/sts029
- Goujon, T., Minic, Z., El Amrani, A., Lerouxel, O., Aletti, E., Lapierre, C., et al. (2003). AtBXL1, a novel higher plant (*Arabidopsis thaliana*) putative beta-xylosidase gene, is involved in secondary cell wall metabolism and plant development. *Plant J.* 33, 677–690. doi: 10.1046/j.1365-313X.2003.01654.x
- Groman, E. V., Enriquez, P. M., Jung, C., and Josephson, L. (1994). Arabinogalactan for hepatic drug delivery. *Bioconjug. Chem.* 5, 547–556. doi: 10.1021/bc00030a010
- Günl, M., Neumetzler, L., Kraemer, F., de Souza, A., Schultink, A., Pena, M., et al. (2011). AXY8 encodes an α -fucosidase, underscoring the importance of apoplastic metabolism on the fine structure of *Arabidopsis* cell wall polysaccharides. *Plant Cell* 23, 4025–4040. doi: 10.1105/tpc.111.089193
- Haque, M. A., Kotake, T., and Tsumuraya, Y. (2005). Mode of action of beta-glucuronidase from *Aspergillus niger* on the sugar chains of arabinogalactan-protein. *Biosci. Biotechnol. Biochem.* 69, 2170–2177. doi: 10.1271/bbb.69.2170
- Herman, E. M., and Lamb, C. J. (1992). Arabinogalactan-rich glycoproteins are localized on the cell surface and in intravacuolar multivesicular bodies. *Plant Physiol.* 98, 264–272. doi: 10.1104/pp.98.1.264

- Ichinose, H., Fujimoto, Z., Honda, M., Harazono, K., Nishimoto, Y., Uzura, A., et al. (2009). A beta-L-Arabinopyranosidase from *Streptomyces avermitilis* is a novel member of glycoside hydrolase family 27. *J. Biol. Chem.* 284, 25097–25106. doi: 10.1074/jbc.M109.022723
- Ichinose, H., Fujimoto, Z., and Kaneko, S. (2013). Characterization of an α-L-Rhamnosidase from *Streptomyces avermitilis*. *Biosci. Biotechnol. Biochem.* 77, 213–216. doi: 10.1271/bbb.120735
- Ichinose, H., Kotake, T., Tsumuraya, Y., and Kaneko, S. (2006a). Characterization of an Exo-β-1,3-D-galactanase from *Streptomyces avermitilis* NBRC14893 acting on arabinogalactan-proteins. *Biosci. Biotechnol. Biochem.* 70, 2745–2750. doi: 10.1271/bbb.60365
- Ichinose, H., Kotake, T., Tsumuraya, Y., and Kaneko, S. (2008). Characterization of an endo-beta-1,6-Galactanase from *Streptomyces avermitilis* NBRC14893. *Appl. Environ. Microbiol.* 74, 2379–2383. doi: 10.1128/AEM.01733-07
- Ichinose, H., Kuno, A., Kotake, T., Yoshida, M., Sakka, K., Hirabayashi, J., et al. (2006b). Characterization of an exo-beta-1,3-galactanase from *Clostridium thermocellum*. *Appl. Environ. Microbiol.* 72, 3515–3523. doi: 10.1128/AEM.72.5.3515-3523.2006
- Ichinose, H., Yoshida, M., Kotake, T., Kuno, A., Igarashi, K., Tsumuraya, Y., et al. (2005). An exo-beta-1,3-galactanase having a novel beta-1,3-galactan-binding module from *Phanerochaete chrysosporium*. *J. Biol. Chem.* 280, 25820–25829. doi: 10.1074/jbc.M501024200
- Ishida, T., Fujimoto, Z., Ichinose, H., Igarashi, K., Kaneko, S., and Samejima, M. (2009). Crystallization of selenomethionyl exo-beta-1,3-galactanase from the basidiomycete *Phanerochaete chrysosporium*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 1274–1276. doi: 10.1107/S1744309109043395
- Kaewthai, N., Gendre, D., Eklöf, J. M., Ibatullin, F. M., Ezcurra, I., Bhalerao, R. P., et al. (2013). Group III-A XTH genes of *Arabidopsis* encode predominant xyloglucan endohydrolases that are dispensable for normal growth. *Plant Physiol.* 161, 440–454. doi: 10.1104/pp.112.207308
- Kaulfürst-Soboll, H., Rips, S., Koiwa, H., Kajiura, H., Fujiyama, K., and von Schaewen, A. (2011). Reduced immunogenicity of *Arabidopsis* hgl1 mutant N-glycans caused by altered accessibility of xylose and core fucose epitopes. *J. Biol. Chem.* 286, 22955–22964. doi: 10.1074/jbc.M110.196097
- Kieliszewski, M. J. (2001). The latest hype on Hyp-O-glycosylation codes. *Phytochemistry* 57, 319–323. doi: 10.1016/S0031-9422(01)00029-2
- Kitazawa, K., Tryfona, T., Yoshimi, Y., Hayashi, Y., Kawauchi, S., Antonov, L., et al. (2013). β-galactosyl Yariv reagent binds to the β-1,3-galactan of arabinogalactan proteins. *Plant Physiol.* 161, 1117–1126. doi: 10.1104/pp.112.211722
- Knoch, E., Dilokpimol, A., Tryfona, T., Poulsen, C. P., Xiong, G., Harholt, J., et al. (2013). A β-glucuronosyltransferase from *Arabidopsis thaliana* involved in biosynthesis of type II arabinogalactan has a role in cell elongation during seedling growth. *Plant J.* 76, 1016–1029. doi: 10.1111/tpj.12353
- Konishi, T., Kotake, T., Soraya, D., Matsuoka, K., Koyama, T., Kaneko, S., et al. (2008). Properties of family 79 beta-glucuronidases that hydrolyze beta-glucuronosyl and 4-O-methyl-beta-glucuronosyl residues of arabinogalactan-protein. *Carbohydr. Res.* 343, 1191–1201. doi: 10.1016/j.carres.2008.03.004
- Kotake, T., Dina, S., Konishi, T., Kaneko, S., Igarashi, K., Samejima, M., et al. (2005). Molecular cloning of a b-galactosidase from radish that residues of arabinogalactan protein 1. *Plant Physiol.* 138, 1563–1576. doi: 10.1104/pp.105.062562
- Kotake, T., Hirata, N., Degi, Y., Ishiguro, M., Kitazawa, K., Takata, R., et al. (2011). Endo-beta-1,3-galactanase from winter mushroom *Flammulina velutipes*. *J. Biol. Chem.* 286, 27848–27854. doi: 10.1074/jbc.M111.251736
- Kotake, T., Kaneko, S., Kubomoto, A., Haque, M. A., Kobayashi, H., and Tsumuraya, Y. (2004). Molecular cloning and expression in *Escherichia coli* of a *Trichoderma viride* endo-beta-(1→6)-galactanase gene. *Biochem. J.* 377, 749–755. doi: 10.1042/BJ20031145
- Kotake, T., Kitazawa, K., Takata, R., Okabe, K., Ichinose, H., Kaneko, S., et al. (2009). Molecular cloning and expression in *Pichia pastoris* of a *Irpea lacteum* Exo-β-(1→3)-galactanase Gene. *Biosci. Biotechnol. Biochem.* 73, 2303–2309. doi: 10.1271/bbb.90433
- Kotake, T., Tsuchiya, K., Aohara, T., Konishi, T., Kaneko, S., Igarashi, K., et al. (2006). An alpha-L-arabinofuranosidase/beta-D-xylosidase from immature seeds of radish (*Raphanus sativus* L.). *J. Exp. Bot.* 57, 2353–2362. doi: 10.1093/jxb/erj206
- Leonard, R., Pabst, M., Bondili, J. S., Chambat, G., Veit, C., Strasser, R., et al. (2008). Identification of an *Arabidopsis* gene encoding a GH95 alpha1,2-fucosidase active on xyloglucan oligo- and polysaccharides. *Phytochemistry* 69, 1983–1988. doi: 10.1016/j.phytochem.2008.03.024
- Liang, Y., Basu, D., Pattathil, S., Xu, W.-L., Venetos, A., Martin, S. L., et al. (2013). Biochemical and physiological characterization of fut4 and fut6 mutants defective in arabinogalactan-protein fucosylation in *Arabidopsis*. *J. Exp. Bot.* 64, 5537–5551. doi: 10.1093/jxb/ert321
- Ling, N. X.-Y., Lee, J., Ellis, M., Liao, M.-L., Mau, S.-L., Guest, D., et al. (2012). An exo-β-(1→3)-d-galactanase from *Streptomyces* sp. provides insights into type II arabinogalactan structure. *Carbohydr. Res.* 352, 70–81. doi: 10.1016/j.carres.2012.02.033
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Majewska-sawka, A., and Nothnagel, E. A. (2000). The multiple roles of arabinogalactan proteins in plant development. *Plant Physiol.* 122, 3–9. doi: 10.1104/pp.122.1.3
- Markovic, O., and Janecek, S. (2001). Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. *Protein Eng.* 14, 615–631. doi: 10.1093/protein/14.9.615
- Martens-Uzunova, E. S., Zandleven, J. S., Benen, J. A. E., Awad, H., Kools, H. J., Beldman, G., et al. (2006). A new group of exo-acting family 28 glycoside hydrolases of *Aspergillus niger* that are involved in pectin degradation. *Biochem. J.* 400, 43–52. doi: 10.1042/BJ20060703
- Minic, Z., Rihouey, C., Do, C. T., Lerouge, P., and Jouanin, L. (2004). Purification and characterization of enzymes exhibiting beta-D-xylosidase activities in stem tissues of *Arabidopsis*. *Plant Physiol.* 135, 867–878. doi: 10.1104/pp.104.041269
- Miyata, T., Kashige, N., Satho, T., Yamaguchi, T., Aso, Y., and Miake, F. (2005). Cloning, sequence analysis, and expression of the gene encoding *Sphingomonas paucimobilis* FP2001 alpha-L-rhamnosidase. *Curr. Microbiol.* 51, 105–109. doi: 10.1007/s00284-005-4487-8
- Nagae, M., Tsuchiya, A., Katayama, T., Yamamoto, K., Wakatsuki, S., and Kato, R. (2007). Structural basis of the catalytic reaction mechanism of novel 1,2-alpha-L-fucosidase from *Bifidobacterium bifidum*. *J. Biol. Chem.* 282, 18497–18509. doi: 10.1074/jbc.M702246200
- Nothnagel, E. A. (1997). Proteoglycans and related components in plant cells. *Int. Rev. Cytol.* 174, 195–291. doi: 10.1016/S0074-7696(08)62118-X
- Odonmazig, P., Ebringerová, A., Machová, E., and Alföldi, J. (1994). Structural and molecular properties of the arabinogalactan isolated from Mongolian larchwood (*Larix dahurica* L.). *Carbohydr. Res.* 252, 317–324. doi: 10.1016/0008-6215(94)90028-0
- Okemoto, K., Uekita, T., Tsumuraya, Y., Hashimoto, Y., and Kasama, T. (2003). Purification and characterization of an endo-beta-(1→6)-galactanase from *Trichoderma viride*. *Carbohydr. Res.* 338, 219–230. doi: 10.1016/S0008-6215(02)00405-6
- Reiter, W.-D., Chapple, C., and Somerville, C. R. (1997). Mutants of *Arabidopsis thaliana* with altered cell wall polysaccharide composition. *Plant J.* 12, 335–345. doi: 10.1046/j.1365-313X.1997.12020335.x
- Rose, J. K. C. (2002). The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant Cell Physiol.* 43, 1421–1435. doi: 10.1093/pcp/pcf171
- Sakamoto, T., Taniguchi, Y., Suzuki, S., Ihara, H., and Kawasaki, H. (2007). Characterization of *Fusarium oxysporum* beta-1,6-galactanase, an enzyme that hydrolyzes larch wood arabinogalactan. *Appl. Environ. Microbiol.* 73, 3109–3112. doi: 10.1128/AEM.02101-06
- Schultz, C., Gilson, P., Oxley, D., Youl, J., and Bacic, A. (1998). GPI-anchors on arabinogalactan-proteins: implications for signalling in plants. *Trends Plant Sci.* 3, 426–431. doi: 10.1016/S1360-1385(98)01328-4
- Schultz, C. J., Ferguson, K. L., Lahnestein, J., and Bacic, A. (2004). Post-translational modifications of arabinogalactan-peptides of *Arabidopsis thaliana*. Endoplasmic reticulum and glycosylphosphatidylinositol-anchor signal cleavage sites and hydroxylation of proline. *J. Biol. Chem.* 279, 45503–45511. doi: 10.1074/jbc.M407594200
- Seifert, G. J., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161. doi: 10.1146/annurev.arplant.58.032806.103801
- Showalter, A. M. (2001). Arabinogalactan-proteins: structure, expression and function. *Cell. Mol. Life Sci.* 58, 1399–1417. doi: 10.1007/PL00000784
- St John, F. J., Gonzalez, J. M., and Pozharski, E. (2010). Consolidation of glycosyl hydrolase family 30: a dual domain 4/7 hydrolase family consisting of two structurally distinct groups. *FEBS Lett.* 584, 4435–4441. doi: 10.1016/j.febslet.2010.09.051

- Sudan, C., Prakash, S., Bhomkar, P., Jain, S., and Bhalla-Sarin, N. (2006). Ubiquitous presence of beta-glucuronidase (GUS) in plants and its regulation in some model plants. *Planta* 224, 853–864. doi: 10.1007/s00425-006-0276-2
- Takata, R., Tokita, K., Mori, S., Shimoda, R., Harada, N., Ichinose, H., et al. (2010). Degradation of carbohydrate moieties of arabinogalactan-proteins by glycoside hydrolases from *Neurospora crassa*. *Carbohydr. Res.* 345, 2516–2522. doi: 10.1016/j.carres.2010.09.006
- Tan, L., Qiu, F., Lampert, D. T. A., and Kieliszewski, M. J. (2004). Structure of a hydroxyproline (Hyp)-arabinogalactan polysaccharide from repetitive Ala-Hyp expressed in transgenic *Nicotiana tabacum*. *J. Biol. Chem.* 279, 13156–13165. doi: 10.1074/jbc.M311864200
- Tan, L., Showalter, A. M., Egelund, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci.* 3:140. doi: 10.3389/fpls.2012.00140
- Tan, L., Varnai, P., Lampert, D. T. A., Yuan, C., Xu, J., Qui, F., et al. (2010). Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated sidechains. *J. Biol. Chem.* 285, 24575–24583. doi: 10.1074/jbc.M109.4200497-7
- Torki, M., Mandaron, P., Mache, R., and Falconet, D. (2000). Characterization of a ubiquitous expressed gene family encoding polygalacturonase in *Arabidopsis thaliana*. *Gene* 242, 427–436. doi: 10.1016/S0378-1119(99)00497-7
- Tryfona, T., Liang, H.-C., Kotake, T., Kaneko, S., Marsh, J., Ichinose, H., et al. (2010). Carbohydrate structural analysis of wheat flour arabinogalactan protein. *Carbohydr. Res.* 345, 2648–2656. doi: 10.1016/j.carres.2010.09.018
- Tryfona, T., Liang, H.-C., Kotake, T., Tsumuraya, Y., Stephens, E., and Dupree, P. (2012). Structural characterization of *Arabidopsis* leaf arabinogalactan polysaccharides. *Plant Physiol.* 160, 653–666. doi: 10.1104/pp.112.202309
- Tryfona, T., Theys, T. E., Wagner, T., Stott, K., Keegstra, K., and Dupree, P. (2014). Characterisation of FUT4 and FUT6 α -(1→2)-Fucosyltransferases reveals that absence of root arabinogalactan fucosylation increases *Arabidopsis* root growth salt sensitivity. *PLoS ONE* 9:e93291. doi: 10.1371/journal.pone.0093291
- Tsumuraya, Y., Mochizuki, N., Hashimoto, Y., and Kovac, P. (1990). Purification of an exo-beta-(1→3)-D-galactanase of *Irpex lacteus* (*Polyporus tulipiferae*) and its action on arabinogalactan-proteins. *J. Biol. Chem.* 265, 7207–7215.
- Van Hengel, A. J., and Roberts, K. (2002). Fucosylated arabinogalactan-proteins are required for full root cell elongation in *Arabidopsis*. *Plant J.* 32, 105–113. doi: 10.1046/j.1365-313X.2002.01406.x
- Vanzin, G. F., Madson, M., Carpita, N. C., Raikhel, N. V., Keegstra, K., and Reiter, W.-D. (2002). The mur2 mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proc. Natl. Acad. Sci. U.S.A.* 99, 3340–3345. doi: 10.1073/pnas.052450699
- Wang, X. (2001). PLANT PHOSPHOLIPASES. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 211–231. doi: 10.1146/annurev.arplant.52.1.211
- Wu, Y., Williams, M., Bernard, S., Driouch, A., Showalter, A. M., and Faik, A. (2010). Functional identification of two nonredundant *Arabidopsis* alpha(1,2)fucosyltransferases specific to arabinogalactan proteins. *J. Biol. Chem.* 285, 13638–13645. doi: 10.1074/jbc.M110.102715
- Zeleny, R., Leonard, R., Dorfner, G., Dalik, T., Kolarich, D., and Altmann, F. (2006). Molecular cloning and characterization of a plant α 1,3/4-fucosidase based on sequence tags from almond fucosidase I. *Phytochemistry* 67, 641–648. doi: 10.1016/j.phytochem.2006.01.021

Conflict of Interest Statement: The Review Editor, Dr. Ulvskov, declares that, despite being affiliated to the same institution and having collaborated with author(s) Eva Knoch, Adiphon Dilokpimol, Naomi Geshi, the review process was handled objectively. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 January 2014; accepted: 24 April 2014; published online: 11 June 2014. Citation: Knoch E, Dilokpimol A, and Geshi N (2014) Arabinogalactan proteins: focus on carbohydrate active enzymes. *Front. Plant Sci.* 5:198. doi: 10.3389/fpls.2014.00198 This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Knoch, Dilokpimol and Geshi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



An update on post-translational modifications of hydroxyproline-rich glycoproteins: toward a model highlighting their contribution to plant cell wall architecture

May Hijazi^{1,2†}, Silvia M. Velasquez^{3†}, Elisabeth Jamet^{1,2}, José M. Estevez^{3*} and Cécile Albenne^{1,2*}

¹ Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, UPS, UMR 5546, Castanet-Tolosan, France

² CNRS, UMR 5546, Castanet-Tolosan, France

³ Facultad de Ciencias Exactas y Naturales, Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE-CONICET), Universidad de Buenos Aires, Buenos Aires, Argentina

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Georgios Merkouropoulos, Centre for Research & Technology, Hellas, Greece

Maura C. Cannon, University of Massachusetts Amherst, USA

***Correspondence:**

José M. Estevez, Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE-CONICET), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires C1428EGA, Intendente Guiraldes 2160, Pabellon 2, Buenos Aires, Argentina

e-mail: jestevez@fbmc.fcen.uba.ar;
Cécile Albenne, Laboratoire de Recherche en Sciences Végétales, UMR 5546 UPS/CNRS, 24 chemin de Borderouge, BP 42617 Auzeville, F-31326 Castanet-Tolosan, France

e-mail: albenne@lrsv.ups-tls.fr

[†]Co-first authors.

Plant cell walls are composite structures mainly composed of polysaccharides, also containing a large set of proteins involved in diverse functions such as growth, environmental sensing, signaling, and defense. Research on cell wall proteins (CWPs) is a challenging field since present knowledge of their role into the structure and function of cell walls is very incomplete. Among CWPs, hydroxyproline (Hyp)-rich O-glycoproteins (HRGPs) were classified into three categories: (i) moderately glycosylated extensins (EXTs) able to form covalent scaffolds; (ii) hyperglycosylated arabinogalactan proteins (AGPs); and (iii) Hyp/proline (Pro)-Rich proteins (H/PRPs) that may be non-, weakly- or highly-glycosylated. In this review, we provide a description of the main features of their post-translational modifications (PTMs), biosynthesis, structure, and function. We propose a new model integrating HRGPs and their partners in cell walls. Altogether, they could form a continuous glyco-network with non-cellulosic polysaccharides via covalent bonds or non-covalent interactions, thus strongly contributing to cell wall architecture.

Keywords: arabinogalactan protein, extensin, hydroxyproline, O-glycosylation, proline-rich protein

INTRODUCTION

Plant cell walls are composite structures mainly composed of polysaccharides, namely cellulose, hemicelluloses and pectins, containing also a large set of proteins involved in the cell dynamics through diverse functions such as growth, environmental sensing, signaling, and defense (Fry, 2004). Research on cell wall components emerged in the nineteen sixties (Lampert and Northcote, 1960; Rees and Wight, 1969) and is still a very active field with continuous advances on the nature, structure and functions of polysaccharides (Carpita and Gibeaut, 1993; Willats et al., 2006; Scheller and Ulvskov, 2010) and of proteins (Rose and Lee, 2010; Albenne et al., 2013). However, the question of how these components are connected to make a functional matrix is still a matter of debate (Keegstra et al., 1973; Park and Cosgrove, 2012; Wang et al., 2012).

Among cell wall proteins (CWPs), hydroxyproline (Hyp)-rich O-glycoproteins (HRGPs) are complex macromolecules with various structures and functions. Identified several decades ago, HRGPs were classified into three categories: (i) moderately glycosylated extensins (EXTs); (ii) hyperglycosylated

arabinogalactan proteins (AGPs); and (iii) Hyp/Pro-rich proteins (H/PRPs) that may be non-, weakly- or highly-glycosylated. Each HRGP sub-family is characterized by repetitive consensus sequences which determine the way they are glycosylated according to the so-called Hyp-O-glycosylation code (Kieliszewski, 2001; Tan et al., 2004; Estevez et al., 2006). From a functional point of view, HRGPs are also very diverse. AGPs are implicated in a variety of physiological processes including cell expansion, reproductive development, embryogenesis, signaling, and defense (Seifert and Roberts, 2007). EXTs are mostly described as structural proteins able to form covalent scaffolds (Qi et al., 1995; Brady et al., 1996; Cannon et al., 2008; Velasquez et al., 2012). Finally, H/PRPs are the less documented HRGPs and little is known about their structure and function. They seem to be associated to development and defense against biotic and abiotic stresses (Bradley et al., 1992; Bernhardt and Tierney, 2000; Battaglia et al., 2007). Hybrid and chimeric HRGPs also exist, enlarging the diversity of this superfamily. As previously defined, hybrid HRGPs are composed of HRGP modules from different families, and chimeric HRGPs are composed of one or

more HRGP modules within a non-HRGP protein (Showalter et al., 2010). An expert bioinformatics analysis of the *Arabidopsis thaliana* genome identified 166 HRGPs classified in 85 AGPs, 59 EXTs, 18 H/PRPs, and 4 AGP/EXT hybrid proteins (Showalter et al., 2010). Besides, related to HRGPs but not classified in any of its three sub-families, some allergen proteins containing Hyp residues substituted by arabinogalactans (AGs) were identified in *Artemisia vulgaris* and *Ambrosia artemisiifolia* (Léonard et al., 2005, 2010).

Despite the great interest that plant biologist have had in HRGPs for more than 50 years, many questions about their mode of action in cell walls are still unanswered and HRGP research is still very challenging. In this review, we provide an update on (i) their post-translational modifications (PTMs) which consist in Pro-hydroxylation and O-glycosylation on serine (Ser) and Hyp residues and (ii) their roles in cell walls. We also focus on new insights into HRGP supramolecular assembly and propose a model including most recent data on covalent and non-covalent networks connecting HRGPs and polysaccharides.

EXTENSINS (EXTs)

EXTs AS STRUCTURAL MOLECULES IN PLANT CELL WALLS

EXTs are modular, highly repetitive HRGPs showing similar features as collagen that contain Tyr cross-linking motifs. Unlike collagen, EXTs also undergo plant specific post-translational O-glycosylation on Ser-(Hyp)_{n≥2} motifs. EXTs are represented in the *A. thaliana* genome by 59 members, some are classical EXTs while others are EXT-like chimeras and hybrid-EXTs that also contain other domains. Despite the high number of proteins with EXT domains in plant cell walls (Lamport et al., 2011), we know little about their exact functions and how this protein diversity is coordinated during plant development. There are several reasons that may explain our current lack of understanding of the EXT biology: (i) a high similarity in their protein sequences that make their characterization at the molecular level very difficult; (ii) the highly repetitive nature of their sequences since they are modular proteins, large in size and with complex chemical structures that carry several PTMs. Consequently, the biochemical characterization of a single EXT protein is still today very challenging; (iii) large number of EXTs and EXTs-related proteins encoded in known plant genomes; and (iv) several EXT genes are expressed at the same time in the same plant tissues (see Genevestigator database, <https://www.genevestigator.com>). In addition, most of the available EXT mutants analyzed until now show no clear phenotype. Few exceptions are the mutants *atext3* (embryo lethal), *atext6, 7, 10, 12* (shorter root hairs) and *lrx1, 2* (root hair morphogenesis) that showed clear phenotypes (see Table 1).

PTMs OF EXTs AND THE ENZYMES INVOLVED

Structural O-glycoproteins containing EXT domains that are ultimately secreted into plant cell walls are first shaped by several and complex PTMs that include: (i) signal peptide processing (in the ER), (ii) hydroxylation of Pro into Hyp residues, (iii) O-glycosylation on Hyp and Ser residues (in the ER-Golgi apparatus) and finally, (iv) Tyr cross-linking to promote the formation of a covalent network (in the cell wall). In the last few years,

great progress has been made in our knowledge of the molecular players that act on the EXT biosynthetic pathway with the identification of several enzymes involved in their PTMs (summarized in Table A1). Hydroxylation of peptidyl-Pro is catalyzed by prolyl 4-hydroxylases (P4Hs) providing reactive hydroxyl groups for further O-glycosylation. Plant P4Hs are membrane-bound enzymes that belong to a family of 2-oxoglutarate-dependent dioxygenases (Hieta and Myllyharju, 2002; Koski et al., 2007, 2009). Partial *in vitro* and *in vivo* characterization of plant P4Hs (see Table A1) has been carried out in several plant model systems (Hieta and Myllyharju, 2002; Taininen et al., 2005; Yuasa et al., 2005; Keskiaho et al., 2007; Vlad et al., 2007, 2010; Asif et al., 2009; Velasquez et al., 2011, 2012, *in revision*; Parsons et al., 2013). Most P4Hs are able to hydroxylate with different affinities several types of substrates containing collagen-like, polyproline EXT-type as well as AGP-like sequences. On the other hand, structural information on plant P4Hs is scarce since only one P4H from *Chlamydomonas reinhardtii* (CrP4H1) has been crystallized (Koski et al., 2007, 2009) and few P4Hs were characterized *in vivo* (Velasquez et al., *in revision*). Recent evidence showed that in *A. thaliana*, P4H5 forms homo-/hetero-dimers with P4H2 and P4H13 in the Golgi, suggesting the existence of P4H complexes required for proper Pro hydroxylation. It is plausible that more than one type of P4H complex would be formed in the ER-Golgi compartment, and in the case of the hetero-complexes, the presence of specific P4Hs (e.g., AtP4H5) may be required either for the correct recruitment or the scaffolding of the other P4Hs (e.g., AtP4H2) (Velasquez et al., *in revision*).

Hydroxylated EXTs are usually O-glycosylated with chains of up to four linear Ara residues on each Hyp (Velasquez et al., 2011; Ogawa-Ohnishi et al., 2013). The usual arabinoside structure linked to each Hyp unit is composed of β-L-Araf-(1,2)-β-L-Araf-(1,2)-β-L-Araf-(1,3)-α-L-tAraf. A fifth arabinose unit was reported in some tissues (Lamport, 1973). Specifically, three groups of arabinosyltransferases (AraTs) HPAT1-HPAT3 (GT8 CAZy family), RRA1-RRA3 (GT77 family), and XEG113 (GT77 family) have recently been implicated in the sequential addition of the innermost three L-Ara residues (Egelund et al., 2007; Gille et al., 2009; Velasquez et al., 2011; Ogawa-Ohnishi et al., 2013). The AraT that would transfer the fourth (1,3)-α-L-Araf moiety was identified very recently as Extensin Arabinose Deficient transferase (ExAD) within the GT47 family (Petersen et al., *in preparation*). In addition, one novel peptidyl-Ser galactosyltransferase named as SGT1 has been reported to add a single α-Galp residue to each Ser residue in Ser-(Hyp)₄ motifs of EXTs and it would belong to a new family of CAZy (Saito et al., 2014). Glycosylated EXTs are cross-linked, at least *in vitro*, by putative type-III peroxidases (PERs) at the Tyr residues (Schnabelrauch et al., 1996; Jackson et al., 2001; Price et al., 2003) forming *intra-* and *inter-*EXT linkages (Cannon et al., 2008; Lamport et al., 2011). Thus, EXTs are able to form a three-dimensional glycoprotein network that possibly interacts with other cell wall components like pectins (Nuñez et al., 2009; Dick-Perez et al., 2011). Although the *in vivo* molecular mechanism of the covalent cross-link is unknown, there is evidence of PER-catalyzed oxidative coupling of Tyr residues *in vitro* that mediates the insolubilization of the

Table 1 | Examples of EXTs and EXT-related proteins characterized in the last years.

Protein/Gene name	Tissue or sub-cellular localization	Assumed function /Phenotype of mutants	References
EXTs			
AtEXT1 (At1g76930)	Roots and Inflorescences	Cell wall formation/Induction in response to mechanical wounding, pathogen infection, senescence and at abscission zones, and treatment with hormones (methyl jasmonate, salicylic acid, auxin, brassinosteroids)	Merkouropoulos and Shirsat, 2003
AtEXT3 (At1g21310)	Embryo	Cell wall formation/Embryo-lethal mutant. Incomplete cross wall assembly	Hall and Cannon, 2002; Cannon et al., 2008
AtEXT6 (At2g24980)	Root hairs	Cell wall formation/Short root hair	Velasquez et al., 2011
AtEXT7 (At4g08400)			
AtEXT10 (At5g06640)			
AtEXT11 (At5g49080)			
AtEXT12 (At4g13390)			
AtMOP10 (At5g05500)	Root hairs	Cell wall formation/Short root hair	Velasquez et al., 2011
AtEXT-LIKE (At4g26750)	Root hairs	Cell wall formation/Short root hair	Velasquez et al., 2011
SIEXT1	Trichoblasts	-/Induced by ethylene	Bucher et al., 2002
BnExtA	External and internal phloem of the main stem	-/Greatest expression in regions where a maximum tensile stress is exerted	Shirsat et al., 1996
NtEXT1.4	Stems, Roots and Carpels	-/Cells under mechanical stress: emergence of lateral roots, junction stem/petiole, fusion of carpels. Induction by mechanical stress in roots and stems	Hirsinger et al., 1999; Salvà and Jamet, 2001
NsEXT1.2A	Stems and Roots	-/Expression in the root transition zone, in stem inner and outer phloem and in cortical cells at the stem/petiole junction. Induced by wounding	Guzzardi et al., 2004
LRXs			
AtLRR-EXT (At4g29240)	Root hairs	Cell wall formation /Short root hair	Velasquez et al., 2011
AtLRX1 (At1g12040)	Root hairs	Cell wall formation/Morphogenesis of root hair	Baumberger et al., 2001
AtLRX2 (At1g62440)	Root hairs	<i>atlrx2</i> acts synergistically with <i>atlrx1</i> . <i>atlrx1/atrx2</i> show osmophilic aggregates and local disintegration of the cell wall	Baumberger et al., 2003
VcISG (Inversion-Specific Glycoprotein)	Extracellular matrix	-	Ertl et al., 1992
ZmPex1/ZmPex2/SIPEX (Pollen extensin-like)	Callose portion of the pollen tube cell wall	-	Rubinstein et al., 1995; Stratford et al., 2001

Dc, *Daucus carota*; Dca, *Dianthus caryophyllus*; La, *Lupines albus*; Ns, *Nicotiana sylvestris*; Nt, *N. tabacum*; Sl, *Solanum lycopersicon*; Vc, *Volvox carteri*; Zm, *Zea mays*.

proteins (Schnabelrauch et al., 1996; Jackson et al., 2001; Price et al., 2003). Recently, six apoplastic type-III PERs were identified as putative candidates for the cross-linking of EXTs specifically in the root hairs of *A. thaliana* (Velasquez et al., in revision). Structural proteins with polyproline sequences like collagen can also be Tyr-cross-linked by the action of a PER not only *in vitro* but also *in vivo* (Edens et al., 2001) suggesting that EXTs and

collagen, as extracellular building blocks, would share structural features and functions.

Root hair as models to study EXT functions and related GTs

Root hairs have been used as a single-cell model to study cell wall biosynthesis in general and specifically EXTs during tip-growth (Park et al., 2011; Velasquez et al., 2011). Mutants deficient in

the synthesis of a single wall polymer specifically in the root hair are generally impaired in growth because their cell wall structure is severely compromised (Diet et al., 2006; Cavalier et al., 2008; Ringli, 2010; Park et al., 2011; Pena et al., 2012; Velasquez et al., 2012). In this framework of interconnected wall polymers (Cosgrove, 2005; Dick-Perez et al., 2011), cross-linked EXTs have a key role during cell expansion and growth (Cannon et al., 2008; Ringli, 2010; Lampert et al., 2011; Velasquez et al., 2011). EXT domains seem to be important during polarized cell expansion since several EXT-related mutants have shorter root hairs such as classical *ext6*, *ext7*, *ext10* and *ext12* (Velasquez et al., 2011, 2012) and *lrx1* and *lrx2* mutants (Baumberger et al., 2001, 2003; Ringli, 2010).

Impact of O-glycosylation on EXT function

It is accepted that O-glycans increase HRGP solubility, resistance to proteolytic degradation and thermal stability (Kieliszewski et al., 1989; Ferris et al., 2001; Shpak et al., 2001; Kieliszewski et al., 2011; Lampert et al., 2011). Most of the mutants that correspond to glycosyltransferases (GTs) known to glycosylate EXTs (Table A1) have been related to root hair drastic phenotypes, highlighting that even minor changes in the O-glycosylation status of EXTs affect EXT function during polarized cell expansion (Velasquez et al., in revision). In addition, it was found that both O-glycosylation types present in EXTs (Hyp-O-arabinosylation and Ser-O-galactosylation) are required and have additive effects for correct EXT function in root hair growth (Velasquez et al., in revision). The known roles of EXTs in cell wall assembly, cell shape and growth raise the question about the function of each individual EXT molecule (Hall and Cannon, 2002; Cannon et al., 2008; Velasquez et al., 2011). Some examples of already characterized EXT or EXT-related genes are presented in **Table 1**. Recently, it was reported that EXTs can form, at least *in vitro*, a tridimensional covalent network through Tyr-linkages mediated by EXT PERs between individual EXT molecules and also via self-recognition and alignment of hydrophilic O-glycosylated Ser-(Hyp)_{3–4} repeats and hydrophobic peptide-cross-linking modules (Cannon et al., 2008). Thus, the ordered EXT monomer assembly in plant cell walls would

involve a zipper-like endwise association via cross-linking at the ends of the molecules (Kieliszewski et al., 2011; Lampert et al., 2011). Recently, molecular dynamics and homology modeling experiments suggested that classical EXTs would be able to form a putative triple helix structure by lateral staggered alignment (Cannon et al., 2008) and Tyr cross-linking, similar to that present in collagen (Velasquez et al., in revision). It is also proposed that EXTs interact with pectins by a simple acid-base reaction forming a supramolecular ionic structure in the nascent cell wall (Valentin et al., 2010), which would serve as a template for further cell wall deposition (Cannon et al., 2008; Lampert et al., 2011). In addition, covalent EXT-pectin cross-links were also suggested (Qi et al., 1995; Nuñez et al., 2009). However, it is unclear how EXT monomers are secreted and assembled into the glyco-network and how EXT-pectin interactions are controlled in a coordinated way during new cell wall formation. In addition, pectin methyl esterases de-esterify galacturonic acid residues in homogalacturonans and liberate acidic charges for ionic interactions (Micheli, 2001) with positively-charged domains in molecules like EXTs.

ARABINOGALACTAN PROTEINS (AGPs)

Many articles reporting the state of the art concerning AGP structure, function and biosynthesis have been published recently (Seifert and Roberts, 2007; Ellis et al., 2010; Tan et al., 2012; Lampert and Várnai, 2013; Nguema-Ona et al., 2013; Knoch et al., 2014). AGPs are HRGPs containing a high proportion of sugars, up to 90%. They are characterized by repetitive X(Pro)_n motifs in which X is mostly alanine (Ala) or Ser. In this review, we focus on specific aspects concerning (i) the characterization of their O-glycan moiety and (ii) their interactions with cell wall polysaccharides.

STRUCTURE OF O-GLYCANS OF AGPs

A remarkable work performed on proteins deriving from synthetic genes and produced in cell suspension cultures has allowed to characterize AGP O-glycans covalently linked to [Ser(Hyp)]_n and [Ala(Hyp)]_n motifs (Tan et al., 2010). It has been possible to precisely define the structure of type II AGs by combining monosaccharide and linkage analyses to mass spectrometry and

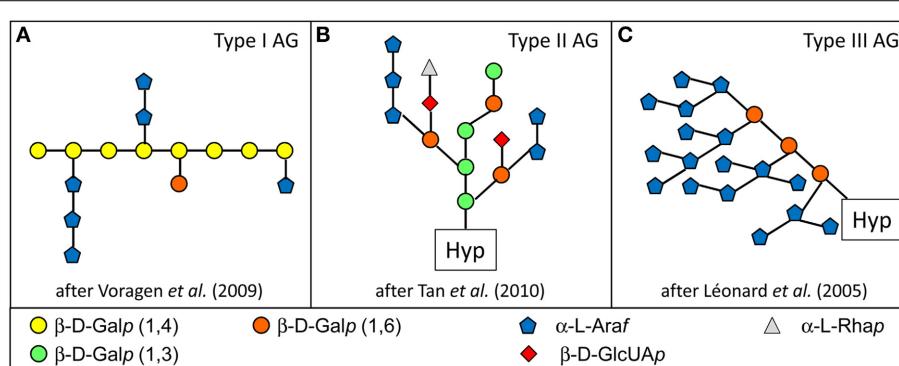


FIGURE 1 | The three main types of AGs. One of the main differences between these AG types consist in the type of linkages between Gal residues of the main chain: β -1,4 in type I AG (A); β -1,3 and β -1,6 in type II AG

(B); β -1,6 in type III AG (C). These differences have been highlighted on the figure by using different colors for Gal residues. Other differences are described in the text.

NMR. An example of type II AG is given in **Figure 1B**. Type II AGs contain a β -D-Galp backbone formed by a succession of three β -1,3 linked D-Galp interrupted by a β -1,6 linkage causing a reverse turn. Gal residues of side chains can be substituted with α -L-Araf, α -L-Rhap or Me-GlcP α (Tan et al., 2010; Tryfona et al., 2012). The chelation of Ca^{2+} ions could occur at the level of GlcA residues located in close proximity. It should be noted that different variants of this basic structure exist, for example β -1,6 side chains can vary in length from 1 over 20 Gal residues (Tryfona et al., 2012). Type II AGs react positively with the β -glucosyl and β -galactosyl Yariv reagents, but not with the α -glucosyl and α -galactosyl Yariv reagents (Kitazawa et al., 2013). The β -galactosyl Yariv reagent has been shown to recognize the β -1,3-Gal main chains of type II AGs. They are different from previously described type I AGs which constitute lateral branches of RGI (Voragen et al., 2009). Type I AGs are formed by a linear chain of β -D-Galp (1,4) with lateral chains of α -L-Araf (1,5 attached to Gal O-3) and β -D-Galp (attached to Gal-O-6) (**Figure 1A**). Type II AGs also differ from type III AGs found on allergens like the *Artemisia vulgaris* Art v 1 (Léonard et al., 2005). The structure of type III AGs has been determined by combining the results of linkage analysis, NMR and enzymatic degradation. They are formed by a short linear chain of β -D-Galp (1,6). They only contain Gal and Ara residues, and they have large branched Ara chains. The linkage analysis indicates the presence of terminal Araf, 5-Araf, 3,5-Araf, 2,5-Araf, 2,3,5-Araf and 3,6-Galp. This Hyp O-glycan was shown to consist of Hyp₁Gal₃Ara_{5–28} series by MALDI-TOF MS. Type III AGs react with the β -glucosyl Yariv reagent suggesting that Art v 1 is an AGP. As for type II AGs, type III AGs probably exist in various forms and only a consensus model can be proposed (**Figure 1C**). Another kind of type III AGs has been later described for Amb a 4, an allergen of *Ambrosia artemisiifolia* (Léonard et al., 2010). It differs from the that of Art v 1 by the presence of different Hyp₁Gal₁Ara_{5–20}series with a lower amount of Gal, the presence of more α -Araf (1,5) and less α -Araf (1,3).

The existence of different types of AGs linked to AGPs (types II and III) raises the questions of (i) the role of the amino acid sequence and (ii) the presence of different types of GTs in plants to ensure the appropriate O-glycosylation of HRGPs (Léonard et al., 2005).

INTERACTIONS OF AGPs WITH POLYSACCHARIDES

The question of how AGPs are connected to other cell wall components and the involvement of their carbohydrate moieties in the interactions is of great importance, but still poorly documented. It has been assumed that AGPs could act as covalent cross-linkers in polysaccharide networks. Several lines of evidence suggested associations between AGPs and pectins. More than 40 years ago, it was hypothesized that Rha residues on type II AG side chains could be attached to RGI (Keegstra et al., 1973). Since then, several studies have reported the existence of strong associations between AGPs and pectins from different plant tissues, including grape (Pellerin et al., 1995), carrot (Immerzeel et al., 2006) or sugar beet (McKenna et al., 2006). Pectins were shown to co-localize with AGPs in pollen tubes (Li et al., 1995; Jauh and Lord, 1996; Mollet et al., 2002). Besides, enzymatic treatment of cell wall

fractions with pectin-degrading enzymes allowed for an increased release of AGPs (Immerzeel et al., 2006; Lampert et al., 2006). One study also suggested the existence of AGP/xylan complexes (Kwan and Morvan, 1995). However, all these AGP/polysaccharide complex analyses were either indirect or achieved on preparations containing a mixture of AGPs, thus preventing a detailed characterization of the interactions. The first in depth structural study of an AGP polysaccharide complex involving pure AGP was only recently reported (Tan et al., 2013). It was shown that two isoforms of a purified *A. thaliana* AGP, At3g45230, are covalently attached to pectins and hemicelluloses. Linkages have been demonstrated between: (i) RGI/homogalacturonan and the Rha residue in the AGP type II AG domain and (ii) arabinoxylan and either a Rha residue of RGI or an Ara residue in the type II AG domain. A model was proposed for this complex called Arabinoxylan Pectin Arabinogalactan Protein1 (APAP1). The *apap1* mutant showed an increased extractability of pectin and xylan, suggesting a structural role for APAP1 (Tan et al., 2013). However, since APAP1 was isolated from suspension culture media, it could correspond to a simplified structure with pectin and xylan domains smaller than expected in plant cell walls. Larger APAP1 complexes may exist in cell walls, but their extraction is undoubtedly the bottleneck preventing their characterization. Large AGP/pectin/xylan complexes should also be found in other plants, corroborating all prior studies indirectly suggesting their existence (Tan et al., 2013).

Present knowledge on AGP/polysaccharide interactions indicates that some AGPs may serve as cross-linker in cell walls and act as polysaccharide plasticizers as previously assumed (Lampert, 2001; Lampert et al., 2006). Chimeric proteins containing AGP domains were also suggested to interact with polysaccharides. In particular, SOS5 (SALT-OVERLY SENSITIVE 5), a Fasciclin-AGP, was assumed to interact with pectins, thus mediating mucilage adherence (Griffiths et al., 2014). SOS5 interacting partners were not identified. Further efforts will be necessary to highlight the contribution of AGPs to cell wall architecture and to give more insight into its molecular basis.

HYP/PRO-RICH PROTEINS

Like EXTs and AGPs, H/PRPs belong to the HRGP superfamily and some of them are chimeric proteins. As mentioned above, little is known about the O-glycosylation of H/PRPs and their interactions with polysaccharides. With regard to O-glycosylation, information is only available for H/PRPs having X(Pro/Hyp)_{n≥2}X motifs. This type of domain can be associated with a short N-terminal AGP domain, a histidine (His)-stretch and a C-terminal PAC (Proline-rich protein and AGP, containing Cys) domain like in the *A. thaliana* AtAGP31 (Liu and Mehdy, 2007; Hijazi et al., 2012). Up to now, twelve such proteins have been identified in *A. thaliana*, *Daucus carota*, *Gossypium hirsutum*, *Nicotiana alata*, *N. tabacum*, *Phaseolus vulgaris*, *Capsicum annuum* and *Petunia hybrida* (Hijazi et al., 2014).

STRUCTURE OF THE O-GLYCANS OF H/PRPs OF THE ATAGP31 TYPE

O-glycosylated amino acid motifs of the H/PRP domain of AtAGP31 have been characterized by mass spectrometry: Lys(Ala/Ser)HypVal, Lys(Pro/Hyp)(Hyp/Pro)(Thr/Val), Thr(Pro/Hy

p)(Hyp/Pro)Val, and Tyr(Pro/Hyp)(Hyp/Pro)Thr (Hijazi et al., 2012). The monosaccharide linked to Hyp is an hexose which is most probably a Gal based on the monosaccharide analysis of the purified protein (53.2% Gal, 39.5% Ara, 2.2% Xyl, 1.9% Fuc, 1.8% Glc, 1.3% Man, 0.3% GlcUA). It should be noted that this global analysis includes O-glycans linked to the AGP domain of AtAGP31 and N-glycans linked to its PAC domain. The O-glycan linked to the H/PRP domain of AtAGP31 is not recognized by the β -D-glucosyl Yariv reagent, but it interacts with the Peanut Agglutinin (PNA), a lectin having a high affinity for Gal residues (Hijazi et al., 2012). It was called Gal/Ara-rich motif (Hijazi et al., 2012). *Nicotiana alata* NaPRP4 shares the same type of H/PRP domain and a PAC domain with AtAGP31 (Sommer-Knudsen et al., 1996). The predominant monosaccharide of this O-glycoprotein is Gal (83%) whereas Ara, GlcNac, Man, Xyl are in minor amounts (7, 4, 4, 1% respectively). The

linkage analysis has shown the presence of terminal Araf (6%), terminal Galp (48%), 1,3-Galp (4%), 1,6-Galp (14%), 1,3,6 Galp (25%), 1,2-Manp (1%) and Xylp (1%). Altogether, H/PRPs with $X(\text{Pro/Hyp})_{n \geq 2}X$ motifs are O-glycosylated with Gal-Ara-rich glycans which seems to be slightly different from the previously described type I, II and III AGs. Further characterization, especially by NMR will be required to fully describe these structures.

INTERACTIONS OF H/PRPs WITH POLYSACCHARIDES

H/PRPs are assumed to be cross-linked in cell walls, but direct evidence is still lacking (Bradley et al., 1992; Brisson et al., 1994; Frueauf et al., 2000). Nothing is known about the possible roles of O-glycosylations. AtAGP31 was recently proposed to be involved in non-covalent interactions networks (Hijazi et al., 2014). Consistently and unlike HRGPs which are covalently insolubilized in cell walls, AtAGP31 is easily extracted

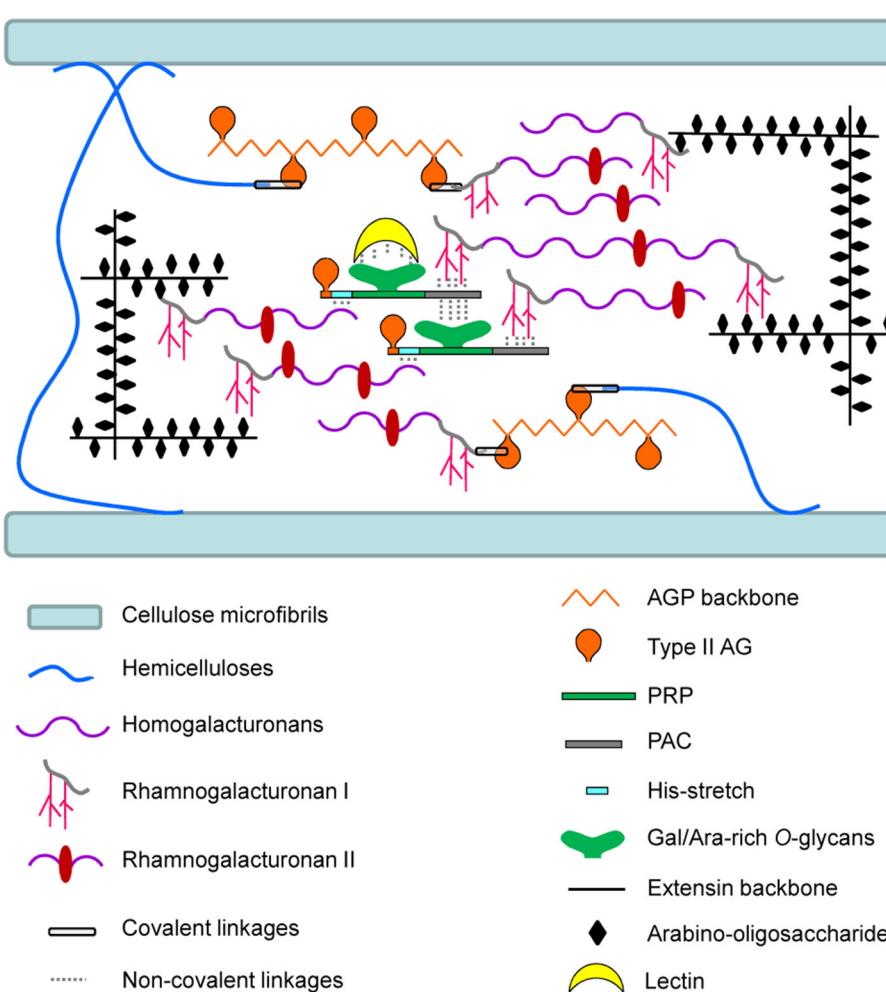


FIGURE 2 | Schematic representation of interactions between HRGPs and cell wall polysaccharides *in muro*. This model proposes an overview of the interactions assumed or demonstrated between HRGPs and polysaccharides according to most relevant publications in this field. For clarity, the model does not represent the whole complexity of the polysaccharide networks. AGPs are represented with covalent linkages with

pectins and hemicelluloses, as proposed by Tan et al. (2013) for the so-called APAP1 complex. EXTs are drawn attached covalently with pectins as proposed by Qi et al. (1995). Finally, non-covalent networks between chimeric HRGPs and polysaccharides are represented according to Hijazi et al. (2014) for AtAGP31. Lectins assumed to bind to Gal/Ara-rich O-glycans of AtAGP31 are also integrated into the model.

from cell walls of etiolated hypocotyls (Hijazi et al., 2012). It should be noted that NaPRP4 is not insolubilized in cell walls as well (Sommer-Knudsen et al., 1996). AtAGP31 was shown to interact *in vitro* with RGI type I AG branches through its PAC domain and with methyl-esterified polygalacturonic acid, probably through its His-stretch. Protein/protein interactions were also assumed for AtAGP31, with (i) self-recognition between its PAC domain and its H/PRP domain O-glycans, and (ii) interaction with cell wall lectins. It was proposed that the AtAGP31 multi-domain organization results in complex supra-molecular scaffolds with different cell wall components, thus contributing to the strengthening of cell walls of quickly growing organs like etiolated hypocotyls. Such non-covalent networks have not been described before for HRGPs. A similar behavior may exist for proteins sharing features with AtAGP31 (Hijazi et al., 2014). However, as mentioned above, except NaPRP4 whose glycosylation has been characterized (Sommer-Knudsen et al., 1996), these proteins were not described at the molecular level and their interactions with cell wall polysaccharides were not investigated. TTS-1 and TTS-2 (Transmitting Tissue-Specific) from *N. tabacum*, and DcAGP1 from *D. carota* were shown to display an ellipsoidal shape and to self-assemble into higher-order structures using microscopy techniques (Baldwin et al., 2000, 2001; Wu et al., 2001). Interestingly, the deglycosylation of TTS disrupts its ability to aggregate, suggesting a regulation of self-association by its level of O-glycosylation (Wu et al., 1995). Self-assembly in a head-to-tail fashion through interactions between the O-glycans of H/PRP domain and the PAC domain can be proposed for proteins like TTS and DcAGP1, similarly to AtAGP31.

CONCLUDING REMARKS AND FUTURE DEVELOPMENTS

In this review, we have focused on some structural features of HRGP O-glycans and we have highlighted their possible interactions *in muro* through covalent glycosidic linkages or non-covalent interactions. As proposed in the model shown in **Figure 2**, HRGPs could serve as cross-linkers in cell walls, connecting non-cellulosic polysaccharides, thus forming a continuous network. Large covalent complexes connecting AGP, hemicelluloses and pectins, as proposed in APAP1, are represented (Tan et al., 2013). However, the relevance of such covalent complexes in cell walls need to be confirmed. EXTs appear to form covalent linkages with pectins as reported (Qi et al., 1995; Nuñez et al., 2009). The precise moieties involved in these linkages have not been identified so far. Finally, chimeric HRGPs with H/PRP and PAC domains like AtAGP31 may form non-covalent networks with a set of cell wall components, including polysaccharides and lectins (Hijazi et al., 2014). It can be speculated that these protein/polysaccharide networks contribute to the cell wall architecture, by reinforcing the polysaccharide scaffold and by controlling its porosity. A recent high-resolution solid-state NMR study elucidating the 3D-architecture of the polysaccharides and proteins *in muro* revealed that the structural proteins in the primary cell wall are separated from the polysaccharides by more than one nanometer (Wang et al., 2012). This corroborates the assumption that O-glycans acts as spacers between HRGP backbones and cell wall polysaccharides.

These new features render even more complex the cell wall architecture. Plant cell walls contain a variety of complex macromolecules, possibly interconnected, resulting from a sophisticated metabolism. A tremendous set of carbohydrate active enzymes is required to achieve (i) polysaccharide synthesis and assembly, (ii) protein glycosylation, and (iii) possible polysaccharide/protein linkages. Non-cellulosic polymer synthesis occurs in the Golgi (Mohnen, 2008; Brown et al., 2011), and HRGP synthesis starts in the ER and continues in the Golgi (Basu et al., 2013; Knoch et al., 2014). An important issue is now to determine in which sub-cellular compartment covalent HRGP/polysaccharide complexes are formed and by which mechanism. Is there a code for establishing these links or are they occurring randomly? Which enzymes are involved? Answering these questions constitutes a real challenge toward a better understanding of cell wall biosynthesis and architecture. Further studies will also be necessary to elucidate the molecular basis of HRGP functions in cell walls and their involvement in physiological processes like cell plate formation or root hair cell expansion (Cannon et al., 2008; Velasquez et al., 2011).

ACKNOWLEDGMENTS

May Hijazi, Cécile Albenne, and Elisabeth Jamet are thankful to Université Paul Sabatier (Toulouse, France), CNRS for supporting their research work and to the Lebanon Ecological Association (grant to May Hijazi). Part of this work has been done at LRSV which belongs to the *Laboratoire d'Excellence* (LabEx) entitled TULIP (ANR-10-LABX-41; ANR-11-IDEX-0002-02). This work was also supported by PICT FONCyT 2011-0054, Argentina (JME) and Mizutani Foundation Glycoscience Grant 2013, Japan (JME). The authors wish to apologize for the papers not quoted in this review, due to lack of space.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00395/abstract>

REFERENCES

- Albenne, C., Canut, H., and Jamet, E. (2013). Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*. *Front. Plant Sci.* 4:111. doi: 10.3389/fpls.2013.00111
- Asif, M. H., Trivedi, P. K., Misra, P., and Nath, P. (2009). Prolyl-4-hydroxylase (AtP4H1) mediates and mimics low oxygen response in *Arabidopsis thaliana*. *Funct. Integr. Genomics* 9, 525–535. doi: 10.1007/s10142-009-0118-y
- Baldwin, T. C., Domingo, C., Schindler, T., Seetharaman, G., Stacey, N., and Roberts, K. (2001). DcAGP1, a secreted arabinogalactan protein, is related to a family of basic proline-rich proteins. *Plant Mol. Biol.* 45, 421–435. doi: 10.1023/A:1010637426934
- Baldwin, T. C., Hengel, A., and Roberts, K. (2000). “The C-terminal PAC domain of a secreted arabinogalactan protein from carrot defines a family of basic proline-rich proteins,” in *Cell and Developmental Biology of Arabinogalactan Proteins*, eds E. A. Nothnagel, A. Bacic, and A. E. Clarke (New York, NY: Kluwer Academic Publishers), 43–50. doi: 10.1007/978-1-4615-4207-0_4
- Basu, D., Liang, Y., Liu, X., Himmeldirk, K., Faik, A., Kieliszewski, M., et al. (2013). Functional identification of a hydroxyproline-O-galactosyltransferase specific for arabinogalactan protein biosynthesis in *Arabidopsis*. *J. Biol. Chem.* 288, 10132–10143. doi: 10.1074/jbc.M112.432609
- Battaglia, M., Solorzano, R. M., Hernandez, M., Cuellar-Ortiz, S., Garcia-Gomez, B., Marquez, J., et al. (2007). Proline-rich cell wall proteins accumulate in growing regions and phloem tissue in response to water deficit in common bean seedlings. *Planta* 225, 1121–1133. doi: 10.1007/s00425-006-0423-9

- Baumberger, N., Ringli, C., and Keller, B. (2001). The chimeric leucine-rich repeat/extensin cell wall protein LRx1 is required for root hair morphogenesis in *Arabidopsis thaliana*. *Genes Dev.* 15, 1128–1139. doi: 10.1101/gad.200201
- Baumberger, N., Steiner, M., Ryser, U., Keller, B., and Ringli, C. (2003). Synergistic interaction of the two paralogous *Arabidopsis* genes *LRX1* and *LRX2* in cell wall formation during root hair development. *Plant J.* 35, 71–81. doi: 10.1046/j.1365-313X.2003.01784.x
- Bernhardt, C., and Tierney, M. L. (2000). Expression of AtPRP3, a proline-rich structural cell wall protein from arabidopsis, is regulated by cell-type-specific developmental pathways involved in root hair formation. *Plant Physiol.* 122, 705–714. doi: 10.1104/pp.122.3.705
- Bradley, D. J., Kjellbom, P., and Lamb, C. J. (1992). Elicitor-induced and wound induced oxidative cross-linking of a proline rich plant cell wall protein: a novel, rapid defense response. *Cell* 70, 21–30. doi: 10.1016/0092-8674(92)90530-P
- Brady, J. D., Sadler, J. H., and Fry, S. C. (1996). Di-isodityrosine, a novel tetrameric derivative of tyrosine in plant cell wall proteins: a new potential cross-link. *Biochem. J.* 315, 323–327.
- Brisson, L. F., Tenhaken, R., and Lamb, C. (1994). Function of oxidative cross linking of cell wall structural proteins in plant disease resistance. *Plant Cell* 6, 1703–1712. doi: 10.1105/tpc.6.12.1703
- Brown, D., Wightman, R., Zhang, Z., Gomez, L., Atanassov, I., Bukowski, J. P., et al. (2011). Arabidopsis genes *IRREGULAR XYLEM* (*IRX15*) and *IRX15L* encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall. *Plant J.* 66, 401–403. doi: 10.1111/j.1365-313X.2011.04501.x
- Bucher, M., Brunner, S., Zimmermann, P., Zardi, G. I., Amrhein, N., Willmitzer, L., et al. (2002). The expression of an extensin-like protein correlates with cellular tip growth in tomato. *Plant Physiol.* 128, 911–923. doi: 10.1104/pp.010998
- Cannon, M. C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B. L., et al. (2008). Self-assembly of the plant cell wall requires an extensin scaffold. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2226–2231. doi: 10.1073/pnas.0711980105
- Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants, consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x
- Cavalier, D. M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., et al. (2008). Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* 20, 1519–1537. doi: 10.1105/tpc.108.059873
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi: 10.1038/nrm1746
- Dick-Perez, M., Zhang, Y., Hayes, J., Salazar, A., Zabotina, O. A., and Hong, M. (2011). Structure and interactions of plant cell-wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry* 50, 989–1000. doi: 10.1021/bi101795q
- Diet, A., Link, B., Seifert, G. J., Schellenberg, B., Wagner, U., Pauly, M., et al. (2006). The *Arabidopsis* root hair cell wall formation mutant *lrx1* is suppressed by mutations in the *RHM1* gene encoding a UDP-L-rhamnose synthase. *Plant Cell* 18, 1630–1641. doi: 10.1105/tpc.105.038653
- Edens, W. A., Sharling, L., Cheng, G., Shapira, R., Kinkade, J. M., Lee, T., et al. (2001). Tyrosine cross-linking of extracellular matrix is catalyzed by Duox, a multidomain oxidase/peroxidase with homology to the phagocyte oxidase subunit gp91phox. *J. Cell Biol.* 154, 879–891. doi: 10.1083/jcb.200103132
- Egelund, J., Obel, N., Ulvskov, P., Geshi, N., Pauly, M., Bacic, A., et al. (2007). Molecular characterization of two *Arabidopsis thaliana* glycosyltransferase mutants, *rra1* and *rra2*, which have a reduced residual arabinose content in a polymer tightly associated with the cellulosic wall residue. *Plant Mol. Biol.* 64, 439–451. doi: 10.1007/s11003-007-9162-y
- Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000
- Ertl, H., Hallmann, A., Wenzl, S., and Sumper, M. (1992). A novel extensin that may organize extracellular matrix biogenesis in *Volvox carteri*. *EMBO J.* 11, 2055–2062.
- Estevez, J. M., Kieliszewski, M. J., Khitrov, N., and Somerville, C. (2006). Characterization of synthetic hydroxyproline-rich proteoglycans with arabinogalactan protein and extensin motifs in *Arabidopsis*. *Plant Physiol.* 142, 458–470. doi: 10.1104/pp.106.084244
- Ferris, P. J., Woessner, J. P., Waffenschmidt, S., Kilz, S., Drees, J., and Goodenough, U. W. (2001). Glycosylated polyproline II rods with kinks as a structural motif in plant hydroxyproline-rich glycoproteins. *Biochemistry* 40, 2978–2987. doi: 10.1021/bi0023605
- Frueauf, J. B., Dolata, M., Leykam, J. F., Lloyd, E., Gonzales, M., Vandebosch, K., et al. (2000). Peptides isolated from cell walls of *Medicago truncatula* nodules and uninfected root. *Phytochemistry* 55, 429–438. doi: 10.1016/S0031-9422(00)00336-8
- Fry, S. C. (2004). Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. *New Phytol.* 161, 641–675. doi: 10.1111/j.1469-8137.2004.00980.x
- Gille, S., Hansel, U., Ziemann, M., and Pauly, M. (2009). Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14699–14704. doi: 10.1073/pnas.0905434106
- Griffiths, J. S., Tsai, A. Y., Xue, H., Voiniciuc, C., Sola, K., Seifert, G., et al. (2014). SALT-OVERLY SENSITIVE5 mediates *Arabidopsis* seed coat mucilage adherence and organization through pectins. *Plant Physiol.* 165, 991–1004. doi: 10.1104/pp.114.239400
- Guzzardi, P., Genot, G., and Jamet, E. (2004). The *Nicotiana sylvestris* extensin gene, *Ext 1.2A*, is expressed in the root transition zone and upon wounding. *Biochim. Biophys. Acta* 1680, 83–92. doi: 10.1016/j.bbexp.2004.08.012
- Hall, Q., and Cannon, M. C. (2002). The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in *Arabidopsis*. *Plant Cell* 14, 1161–1172. doi: 10.1105/tpc.010477
- Hieta, R., and Myllyharju, J. (2002). Cloning and characterization of a low molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*. Effective hydroxylation of proline-rich, collagen-like, and hypoxia-inducible transcription factor alpha-like peptides. *J. Biol. Chem.* 277, 23965–23971. doi: 10.1074/jbc.M201865200
- Hijazi, M., Durand, J., Pichereaux, C., Pont, F., Jamet, E., and Albenne, C. (2012). Characterization of the arabinogalactan protein 31 (AGP31) of *Arabidopsis thaliana*: new advances on the Hyp-O-glycosylation of the Pro-rich domain. *J. Biol. Chem.* 287, 9623–9632. doi: 10.1074/jbc.M111.247874
- Hijazi, M., Roujol, D., Nguyen-Kim, H., Del Rocio Cisneros Castillo, L., Saland, E., Jamet, E., et al. (2014). Arabinogalactan protein 31 (AGP31), a putative network-forming protein in *Arabidopsis thaliana* cell walls? *Ann Bot.* doi: 10.1093/aob/mcu038. [Epub ahead of print].
- Hirsinger, C., Salva, I., Marbach, J., Durr, A., Fleck, J., and Jamet, E. (1999). The tobacco extensin gene *Ext 1.4* is expressed in cells submitted to mechanical constraints and in cells proliferating under hormone control. *J. Exp. Bot.* 50, 343–355.
- Immerzeel, P., Eppink, M. M., De Vriesb, S. C., Schols, H. A., and Voragen, A. G. J. (2006). Carrot arabinogalactan proteins are interlinked with pectins. *Physiol. Plant.* 128, 18–28. doi: 10.1111/j.1399-3054.2006.00712.x
- Jackson, P. A., Galinha, C. I., Pereira, C. S., Fortunato, A., Soares, N. C., Amancio, S. B., et al. (2001). Rapid deposition of extensin during the elicitation of grapevine callus cultures is specifically catalyzed by a 40-kilodalton peroxidase. *Plant Physiol.* 127, 1065–1076. doi: 10.1104/pp.010192
- Jauh, G. Y., and Lord, E. M. (1996). Localization of pectins and arabinogalactan-proteins in lily (*Lilium longitubum* L.) pollen tube and style, and their possible roles in pollination. *Planta* 199, 251–261. doi: 10.1007/BF00196566
- Keegstra, K., Talmadge, K. W., Bauer, W. D., and Albersheim, P. (1973). The structure of plant cell walls: III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant Physiol.* 51, 188–197. doi: 10.1104/pp.51.1.188
- Keskitalo, K., Hieta, R., Sormunen, R., and Myllyharju, J. (2007). *Chlamydomonas reinhardtii* has multiple prolyl 4-hydroxylases, one of which is essential for proper cell wall assembly. *Plant Cell* 19, 256–269. doi: 10.1105/tpc.106.042739
- Kieliszewski, M. J. (2001). The latest hype on Hyp-O-glycosylation codes. *Phytochemistry* 57, 319–323. doi: 10.1016/S0031-9422(01)00029-2
- Kieliszewski, M. J., Lamport, D. T., Tan, L., and Cannon, M. C. (2011). “Hydroxyproline-rich glycoproteins: form and function,” in *Plant Polysaccharides: Biosynthesis and Bioengineering*, ed. P. Ulvskov. (Oxford: Wiley-Blackwell), 321–342.
- Kieliszewski, M. J., Leykam, J. F., and Lamport, D. T. (1989). Trypsin cleaves lysyl-proline in a hydroxyproline-rich glycoprotein from *Zea mays*. *Pept. Res.* 2, 246–248.

- Kitazawa, K., Tryfona, T., Yoshimi, Y., Hayashi, Y., Kawauchi, S., Antonov, L., et al. (2013). β -galactosyl Yariv reagent binds to the β -1,3-galactan of arabinogalactan proteins. *Plant Physiol.* 161, 1117–1126. doi: 10.1104/pp.112.211722
- Knoch, E., Dilokpimol, A., and Geshi, N. (2014). Arabinogalactan proteins: focus on carbohydrate active enzymes. *Front. Plant Sci.* 5:198. doi: 10.3389/fpls.2014.00198
- Koski, M. K., Hieta, R., Bollner, C., Kivirikko, K. I., Myllyharju, J., and Wierenga, R. K. (2007). The active site of an algal prolyl 4-hydroxylase has a large structural plasticity. *J. Biol. Chem.* 282, 37112–37123. doi: 10.1074/jbc.M706554200
- Koski, M. K., Hieta, R., Hirsila, M., Ronka, A., Myllyharju, J., and Wierenga, R. K. (2009). The crystal structure of an algal prolyl 4-hydroxylase complexed with a proline-rich peptide reveals a novel buried tripeptide binding motif. *J. Biol. Chem.* 284, 25290–25301. doi: 10.1074/jbc.M109.014050
- Kwan, J. S., and Morvan, H. (1995). Characterization of extracellular beta-(1,4)-xylan backbone O-substituted by arabinogalactans type-II in a plant-cell suspension. *Carbohydr. Polym.* 26, 99–107. doi: 10.1016/0144-8617(94)00098-E
- Lampert, D. T. (1973). Galactosylserine in extensin. *Biochem. J.* 133, 125–131.
- Lampert, D. T. (2001). Life behind cell walls: paradigm lost, paradigm regained. *Cell. Mol. Life Sci.* 58, 1363–1385. doi: 10.1007/PL00000782
- Lampert, D. T., Kieliszewski, M. J., Chen, Y., and Cannon, M. C. (2011). Role of the extensin superfamily in primary cell wall architecture. *Plant Physiol.* 156, 11–19. doi: 10.1104/pp.110.169011
- Lampert, D. T., Kieliszewski, M. J., and Showalter, A. M. (2006). Salt stress upregulates periplasmic arabinogalactan proteins: using salt stress to analyse AGP function. *New Phytol.* 169, 479–492. doi: 10.1111/j.1469-8137.2005.01591.x
- Lampert, D. T., and Northcote, D. H. (1960). Hydroxyproline in primary cell walls of higher plants. *Nature* 188, 665–666. doi: 10.1038/188665b0
- Lampert, D. T., and Várnai, P. (2013). Periplasmic arabinogalactan glycoproteins act as a calcium capacitor that regulates plant growth and development. *New Phytol.* 197, 58–64. doi: 10.1111/nph.12005
- Leonard, R., Petersen, B. O., Himly, M., Kaar, W., Wopfner, N., Kolarich, D., et al. (2005). Two novel types of O-glycans on the mugwort pollen allergen Art v 1 and their role in antibody binding. *J. Biol. Chem.* 280, 7932–7940. doi: 10.1074/jbc.M410407200
- Leonard, R., Wopfner, N., Pabst, M., Stadlmann, J., Petersen, B. O., Duus, J. O., et al. (2010). A new allergen from ragweed (*Ambrosia artemisiifolia*) with homology to Art v 1 from mugwort. *J. Biol. Chem.* 285, 27192–27200. doi: 10.1074/jbc.M110.127118
- Li, Y.Q., Faleri, C., Geitmann, A., Zhang, H. Q., and Cresti, M. (1995). Immunogold localization of arabinogalactan proteins, unesterified and esterified pectins in pollen grains and pollen tubes of *Nicotiana tabacum* L. *Protoplasma* 189, 26–36. doi: 10.1007/BF01280289
- Liu, C., and Mehdy, M. C. (2007). A nonclassical arabinogalactan protein gene highly expressed in vascular tissues, AGP31, is transcriptionally repressed by methyl jasmonic acid in *Arabidopsis*. *Plant Physiol.* 145, 863–874. doi: 10.1104/pp.107.102657
- McKenna, C., Al-Assaf, S., Phillips, G. O., and Funami, T. (2006). The protein component in pectin - is it a AGP? *Foods Food Ingred. J. Jpn.* 211, 264–271.
- Merkouropoulos, G., and Shirsat, A. H. (2003). The unusual *Arabidopsis* extensin gene *atExt1* is expressed throughout plant development and is induced by a variety of biotic and abiotic stresses. *Planta* 217, 356–366. doi: 10.1007/s00425-003-1002-y
- Micheli, F. (2001). Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends Plant Sci.* 6, 414–419. doi: 10.1016/S1360-1385(01)02045-3
- Mohnen, D. (2008). Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* 11, 266–277. doi: 10.1016/j.pbi.2008.03.006
- Mollet, J. C., Kim, S., Jauh, G. Y., and Lord, E. M. (2002). Arabinogalactan proteins, pollen tube growth, and the reversible effects of Yariv phenylglycoside. *Protoplasma* 219, 89–98. doi: 10.1007/s007090200009
- Nguema-Ona, E., Vicré-Gibouin, M., Cannesan, M. A., and Driouch, A. (2013). Arabinogalactan proteins in root-microbe interactions. *Trends Plant Sci.* 18, 440–449. doi: 10.1016/j.tplants.2013.03.006
- Núñez, A., Fishman, M. L., Fortis, L. L., Cooke, P. H., and Hotchkiss, A. T. J. (2009). Identification of extensin protein associated with sugar beet pectin. *J. Agric. Food Chem.* 57, 10951–10958. doi: 10.1021/jf902162t
- Ogawa-Ohnishi, M., Matsushita, W., and Matsubayashi, Y. (2013). Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*. *Nat. Chem. Biol.* 9, 726–730. doi: 10.1038/nchembio.1351
- Park, S., Szumlanski, A. L., Gu, F., Guo, F., and Nielsen, E. (2011). A role for CSLD3 during cell-wall synthesis in apical plasma membranes of tip-growing root-hair cells. *Nat. Cell Biol.* 13, 973–980. doi: 10.1038/ncb2294
- Park, Y. B., and Cosgrove, D. J. (2012). A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiol.* 158, 1933–1943. doi: 10.1104/pp.111.192880
- Parsons, J., Altmann, F., Graf, M., Stadlmann, J., Reski, R., and Decker, E. L. (2013). A gene responsible for prolyl-hydroxylation of moss-produced recombinant human erythropoietin. *Sci. Rep.* 3:319. doi: 10.1038/srep03019
- Pellerin, P., Pellerin, P., Vidal, S., Williams, P., and Brillouet, J. M. (1995). Characterization of five type II arabinogalactan-protein fractions from red wine of increasing uronic acid content. *Carbohydr. Res.* 277, 135–143. doi: 10.1016/0008-6215(95)00206-9
- Pena, M. J., Kong, Y., York, W. S., and O'Neill, M. A. (2012). A galacturonic acid-containing xyloglucan is involved in *Arabidopsis* root hair tip growth. *Plant Cell* 24, 4511–4524. doi: 10.1105/tpc.112.103390
- Price, N. J., Pinheiro, C., Soares, C. M., Ashford, D. A., Ricardo, C. P., and Jackson, P. A. (2003). A biochemical and molecular characterization of LEP1, an extensin peroxidase from lupin. *J. Biol. Chem.* 278, 41389–41399. doi: 10.1074/jbc.M304519200
- Qi, X. Y., Behrens, B. X., West, P. R., and Mort, A. J. (1995). Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension-cultures, evidence for a covalent cross-link between extensin and pectin. *Plant Physiol.* 108, 1691–1701. doi: 10.1104/pp.108.4.1691
- Rees, D. A., and Wight, N. J. (1969). Molecular cohesion in plant cell walls. Methylation analysis of pectic polysaccharides from the cotyledons of white mustard. *Biochem. J.* 115, 431–439.
- Ringli, C. (2010). The hydroxyproline-rich glycoprotein domain of the *Arabidopsis* LRX1 requires Tyr for function but not for insolubilization in the cell wall. *Plant J.* 63, 662–669. doi: 10.1111/j.1365-313X.2010.04270.x
- Rose, J. K. C., and Lee, S.-J. (2010). Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome. *Plant Physiol. Biochem.* 153, 433–436. doi: 10.1104/pp.110.154872
- Rubinstein, A. L., Marquez, J., Suarez-Cervera, M., and Bedinger, P. A. (1995). Extensin-like glycoproteins in the maize pollen tube wall. *Plant Cell* 7, 2211–2225. doi: 10.1105/tpc.7.12.2211
- Saito, F., Suyama, A., Oka, T., Yoko-O, T., Matsuoka, K., Jigami, Y., et al. (2014). Identification of novel peptidyl serine O-galactosyltransferase gene family in plants. *J. Biol. Chem.* 289, 20405–20420. doi: 10.1074/jbc.M114.553933
- Salvà, I., and Jamet, E. (2001). Expression of the tobacco *Ext1.4* extensin gene upon mechanical constraints and localization of regulatory regions. *Plant Biol.* 3, 1–10. doi: 10.1055/s-2001-11746
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Schnabelrauch, L. S., Kieliszewski, M. J., Upham, B. L., Alizedeh, H., and Lampert, D. T. (1996). Isolation of pI 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val-Tyr-Lys as putative intermolecular cross-link site. *Plant J.* 9, 477–489. doi: 10.1046/j.1365-313X.1996.09040477.x
- Seifert, G. J., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161. doi: 10.1146/annurev.arplant.58.032806.103801
- Shirsat, A. H., Wieczorek, D., and Kozbial, P. (1996). A gene for *Brassica napus* extensin is differentially expressed on wounding. *Plant Mol. Biol.* 30, 1291–1300. doi: 10.1007/BF00019559
- Showalter, A. M., Keppler, B., Lichtenberg, J., Gu, D., and Welch, L. R. (2010). A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol.* 153, 485–513. doi: 10.1104/pp.110.156554
- Shpak, E., Barbar, E., Leykam, J. F., and Kieliszewski, M. J. (2001). Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in *Nicotiana tabacum*. *J. Biol. Chem.* 276, 11272–11278. doi: 10.1074/jbc.M011323200
- Sommer-Knudsen, J., Clarke, A. E., and Bacic, A. (1996). A galactose-rich, cell-wall glycoprotein from styles of *Nicotiana alata*. *Plant J.* 9, 71–83. doi: 10.1046/j.1365-313X.1996.09010071.x
- Stratford, S., Barne, W., Hohorst, D. L., Sagert, J. G., Cotter, R., Golubiewski, A., et al. (2001). A leucine-rich repeat region is conserved in pollen extensin-like (Pex) proteins in monocots and dicots. *Plant Mol. Biol.* 46, 43–56. doi: 10.1023/A:1010659425399

- Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., et al. (2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25, 270–287. doi: 10.1105/tpc.112.107334
- Tan, L., Qiu, F., Lampert, D. T., and Kieliszewski, M. J. (2004). Structure of a hydroxyproline (Hyp)-arabinogalactan polysaccharide from repetitive Ala-Hyp expressed in transgenic *Nicotiana tabacum*. *J. Biol. Chem.* 279, 13156–13165. doi: 10.1074/jbc.M311864200
- Tan, L., Showalter, A. M., Egelund, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci.* 3:140. doi: 10.3389/fpls.2012.00140
- Tan, L., Varnai, P., Lampert, D. T., Yuan, C., Xu, J., Qiu, F., et al. (2010). Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated side chains. *J. Biol. Chem.* 285, 24575–24583. doi: 10.1074/jbc.M109.100149
- Tiainen, P., Myllyharju, J., and Koivunen, P. (2005). Characterization of a second *Arabidopsis thaliana* prolyl 4-hydroxylase with distinct substrate specificity. *J. Biol. Chem.* 280, 1142–1148. doi: 10.1074/jbc.M411109200
- Tryfona, T., Liang, H. C., Kotake, T., Tsunuraya, Y., Stephens, E., and Dupree, P. (2012). Structural characterization of *Arabidopsis* leaf arabinogalactan polysaccharides. *Plant Physiol.* 160, 563–666. doi: 10.1104/pp.112.202309
- Valentin, R., Cerclier, C., Geneix, N., Aguirre-Beghin, V., Gaillard, C., Ralet, M. C., et al. (2010). Elaboration of extensin-pectin thin film model of primary plant cell wall. *Langmuir* 26, 9891–9898. doi: 10.1021/la100265d
- Velasquez, S. M., Ricardi, M. M., Dorosz, J. G., Fernandez, P. V., Nadra, A. D., Pol-Fachin, L., et al. (2011). O-glycosylated cell wall proteins are essential in root hair growth. *Science* 332, 1401–1403. doi: 10.1126/science.1206657
- Velasquez, S. M., Salgado Salter, J., Petersen, B. L., and Estevez, J. M. (2012). Recent advances on the post-translational modifications of EXTs and their roles in plant cell walls. *Front. Plant Sci.* 3:93. doi: 10.3389/fpls.2012.00093
- Vlad, F., Spano, T., Vlad, D., Daher, F. B., Ouelhadj, A., Fragkostefanakis, S., et al. (2007). Involvement of *Arabidopsis* prolyl 4 hydroxylases in hypoxia, anoxia and mechanical wounding. *Plant Signal. Behav.* 2, 368–369. doi: 10.4161/psb.2.5.4462
- Vlad, F., Tiainen, P., Owen, C., Spano, T., Daher, F. B., Oualid, F., et al. (2010). Characterization of two carnation petal prolyl 4 hydroxylases. *Physiol. Plant.* 140, 199–207. doi: 10.1111/j.1399-3054.2010.01390.x
- Voragen, A. G. J., Coenen, G. J., Verhoef, R. P., and Schols, H. A. (2009). Pectin, a versatile polysaccharide present in plant cell walls. *Struct. Chem.* 20, 263–275. doi: 10.1007/s11224-009-9442-z
- Wang, T., Zabotina, O., and Hong, M. (2012). Pectin-cellulose interactions in the *Arabidopsis* primary cell wall from two-dimensional magic-angle-spinning solid-state nuclear magnetic resonance. *Biochemistry* 51, 9846–9856. doi: 10.1021/bi3015532
- Willats, W. G. T., Knox, P., and Mikkelsen, J. D. (2006). Pectin: new insights into an old polymer are starting to gel. *Trends Food Sci. Technol.* 17, 97–104. doi: 10.1016/j.tifs.2005.10.008
- Wu, H., De Graaf, B., Mariani, C., and Cheung, A. Y. (2001). Hydroxyproline-rich glycoproteins in plant reproductive tissues: structure, functions and regulation. *Cell. Mol. Life Sci.* 58, 1418–1429. doi: 10.1007/PL00000785
- Wu, H. M., Wang, H., and Cheung, A. Y. (1995). A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell* 82, 395–403. doi: 10.1016/0092-8674(95)90428-X
- Yuasa, K., Toyooka, K., Fukuda, H., and Matsuoka, K. (2005). Membrane-anchored prolyl hydroxylase with an export signal from the endoplasmic reticulum. *Plant J.* 41, 81–94. doi: 10.1111/j.1365-313X.2004.02279.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 July 2014; accepted: 24 July 2014; published online: 14 August 2014.

Citation: Hijazi M, Velasquez SM, Jamet E, Estevez JM and Albenne C (2014) An update on post-translational modifications of hydroxyproline-rich glycoproteins: toward a model highlighting their contribution to plant cell wall architecture. *Front. Plant Sci.* 5:395. doi: 10.3389/fpls.2014.00395

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Hijazi, Velasquez, Jamet, Estevez and Albenne. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Loss of *Arabidopsis GAUT12/IRX8* causes anther indehiscence and leads to reduced G lignin associated with altered matrix polysaccharide deposition

Zhangying Hao^{1,2,3}, Utku Avci^{2,3}, Li Tan^{2,3}, Xiang Zhu^{2,4}, John Glushka², Sivakumar Pattathil^{2,3}, Stefan Eberhard², Tipton Sholes², Grace E. Rothstein^{2,5}, Wolfgang Lukowitz¹, Ron Orlando^{2,4}, Michael G. Hahn^{1,2,3} and Debra Mohnen^{2,3,6*}

¹ Department of Plant Biology, University of Georgia, Athens, GA, USA

² Complex Carbohydrate Research Center, University of Georgia, Athens, GA, USA

³ BioEnergy Science Center (BESC), Oak Ridge National Laboratory, Oak Ridge, TN, USA

⁴ Department of Chemistry, University of Georgia, Athens, GA, USA

⁵ Department of Biology, Lawrence University, Appleton, WI, USA

⁶ Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA

Edited by:

Cécile Albenne, Université Paul Sabatier Toulouse (UPS), France

Reviewed by:

Henrik Scheller, Lawrence Berkeley National Laboratory, USA

Kian Hématy, Institut National de la Recherche Agronomique, France

Rui Shi, North Carolina State University, USA

*Correspondence:

Debra Mohnen, Department of Biochemistry and Molecular Biology, Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Rd., Athens, GA 30602, USA
e-mail: dmohnen@ccrc.uga.edu

GAlectUronosylTransferase12 (GAUT12)/IRregular Xylem8 (IRX8) is a putative glycosyltransferase involved in *Arabidopsis* secondary cell wall biosynthesis. Previous work showed that *Arabidopsis irregular xylem8* (*irx8*) mutants have collapsed xylem due to a reduction in xylan and a lesser reduction in a subfraction of homogalacturonan (HG). We now show that male sterility in the *irx8* mutant is due to indehiscent anthers caused by reduced deposition of xylan and lignin in the endothecium cell layer. The reduced lignin content was demonstrated by histochemical lignin staining and pyrolysis Molecular Beam Mass Spectrometry (pyMBMS) and is associated with reduced lignin biosynthesis in *irx8* stems. Examination of sequential chemical extracts of stem walls using 2D ¹³C-¹H Heteronuclear Single-Quantum Correlation (HSQC) NMR spectroscopy and antibody-based glycome profiling revealed a reduction in G lignin in the 1 M KOH extract and a concomitant loss of xylan, arabinogalactan and pectin epitopes in the ammonium oxalate, sodium carbonate, and 1 M KOH extracts from the *irx8* walls compared with wild-type walls. Immunolabeling of stem sections using the monoclonal antibody CCRC-M138 reactive against an unsubstituted xylopentaose epitope revealed a bi-lamellate pattern in wild-type fiber cells and a collapsed bi-layer in *irx8* cells, suggesting that at least in fiber cells, GAUT12 participates in the synthesis of a specific layer or type of xylan or helps to provide an architecture framework required for the native xylan deposition pattern. The results support the hypothesis that GAUT12 functions in the synthesis of a structure required for xylan and lignin deposition during secondary cell wall formation.

Keywords: secondary cell walls, xylan, lignin, pectin, wall glycan epitopes, anther dehiscence

INTRODUCTION

The plant extracellular matrix (i.e., cell wall) consists of a variety of complex carbohydrate polymers with distinct chemical and physical properties. The covalent and noncovalent interactions between these polymers in the final composite determine many of the characteristics of the cell wall. Accordingly, mutations in individual glycosyltransferases (GTs), each of which presumably participates in the biogenesis of a single cell wall component or domain, often affect multiple classes of cell wall polymers and sometimes result in dwarf plants. For example, the *Arabidopsis* mutant *irx7* (Zhong et al., 2005) is defective in both xylan and cellulose deposition, whereas *qua1* (Bouton et al., 2002; Leboeuf et al., 2005; Orfila et al., 2005), *parvus-3/gat1* (Lao et al., 2003; Shao et al., 2004; Brown et al., 2007; Lee et al., 2007b; Kong et al.,

2009), and *irx8/gaut12* mutants (Peña et al., 2007; Persson et al., 2007) are affected in pectin and xylan biosynthesis. These complex effects make it difficult to infer primary gene function on the basis of mutant phenotypes alone.

The gene mutated in the xylan- and pectin-deficient mutant *irregular xylem8* (*irx8*) is *GAUT12/IRX8* (*At5g54690*), a member of CAZy family GT8 that contains GTs related to homogalacturonan (HG): α -1,4-galacturonosyltransferase (GalAT) GAUT1 (Sterling et al., 2006). In this paper we refer to the gene affected in the *irx8* mutant as *GAUT12* and its protein as GAlectUronosylTransferase12 (GAUT12). GAUT12 is predicted to be a type II transmembrane protein with its C-terminal catalytic domain facing the Golgi lumen. Transient expression of YFP-tagged GAUT12 protein showed that it co-localizes with

CFP-tagged MUR4, consistent with the localization of GAUT12 in the Golgi apparatus (Peña et al., 2007). Transcription of *GAUT12* is strongest in xylem vessels and interfascicular fiber cells, and *irx8* mutant cell walls show a substantial reduction in glucuronoxyylan (Peña et al., 2007; Persson et al., 2007) as well as a modest reduction in α -1,4-linked GalA (Persson et al., 2007). Xylan is one of the major components of the secondary wall, and pectin is a major matrix polysaccharide in primary walls, but is also found in low abundance in walls prepared from cells synthesizing secondary walls. Additionally, *irx8* mutant xylan is nearly devoid of a xylan reducing-end glycosyl sequence [XRES; β -D-Xylp-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 2)- α -D-GalpA-(1 \rightarrow 4)-D-Xylp] (Peña et al., 2007). XRES has been proposed to act either as a primer to initiate xylan biosynthesis or a terminator to control xylan chain length (York and O'Neill, 2008). However, it remains unknown how XRES synthesis is initiated. Since stem microsomes isolated from *irx8* plants contain comparable amounts of xylan:xylosyltransferase and xylan:glucuronosyltransferase activity as their wild-type counterparts (Brown et al., 2007; Lee et al., 2007a), it seems unlikely that GAUT12 is involved in the elongation or branching of the xylan backbone (York and O'Neill, 2008; Scheller and Ulvskov, 2010). Based on analyses of *irx8* cell walls and GAUT12 protein homology to GAUT1, it has been hypothesized that GAUT12 is a GalAT that either synthesizes a subfraction of HG (Persson et al., 2007) or catalyzes the addition of GalA into the nascent XRES (Peña et al., 2007). The biochemical function of GAUT12, however, remains unresolved to date.

In addition to being severely dwarfed and slow growing, Arabidopsis *irx8* mutants are sterile (Persson et al., 2007). Consistent with a role in secondary wall formation and reproduction, GAUT12 expression is regulated by transcription factors that regulate vessel and fiber formation, such as MYB46 (Ko et al., 2009), MYB83 (McCarthy et al., 2009), VND6, and VND7 (Yamaguchi et al., 2010), as well as by transcription factors that act in anther development, such as MYB26/MALE STERILE35 (Steiner-Lange et al., 2003; Yang et al., 2007), NST1/NST2 (Mitsuda et al., 2005), and AHP4 (Jung et al., 2008). Within anthers, secondary wall thickenings in the endothelium cell layer provide part of the biophysical force that enables dehiscence, the programmed rupture of the anthers to release mature pollen (Wilson et al., 2011). Several lignin-defective mutants have recently been shown to be indehiscent and to generate defective pollen grains (Schilmiller et al., 2009; Weng et al., 2010; Thevenin et al., 2011).

The phenotypes of the Arabidopsis *irx8* mutant include not only a loss of \sim 60% xylan and \sim 13% pectin, but also \sim 25% cellulose. This is a significantly smaller reduction in cellulose than observed in the cellulose defective mutants *irx1* and *irx3* (Brown et al., 2005; Persson et al., 2005). Based on phloroglucinol-HCl staining, a reduction of lignin in *irx8* vessels and fiber cells was also recently reported that was suggested to be associated with reduced xylan biosynthesis (Petersen et al., 2012). Lignin, a resin-like molecular network generated by oxidative polymerization of phenolic subunits within the extracellular space of many terminally differentiated cells constitutes up to 30% of most secondary cell walls (Boerjan et al., 2003). The three subunits of lignin, namely *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) are

laid down in specific spatio-temporal patterns. For example, in wood formation H and G lignins are deposited at early stages of lignification in the middle lamella and tricellular junctions, while G lignin is deposited earlier than S lignin in vessels and fibers and S lignin is mainly deposited in fibers (Donaldson, 2001). Lignin subunits also become covalently linked to hemicelluloses and pectins (Jeffries, 1990), and the carbohydrate polymers are thought to guide expansion of the lignin lamellae during lignification (Donaldson, 2001; Donaldson and Knox, 2012).

The goal of this study was to elucidate the biological function(s) of GAUT12 in Arabidopsis. We identified the cause of sterility in the *irx8* mutant and a reduction in lignin in this mutant. We attempted to solve the biochemical function of GAUT12 and found that GAUT12 does not have HG:GalAT activity comparable to that of GAUT1. In addition, an increased expression of an RG-I epitope in *irx8* fiber cell walls was revealed. Our results suggest a connection between an RG-I-containing structure and the GAUT12-dependent wall product. We propose that GAUT12 participates in the synthesis of a structure required for xylan and lignin deposition during the formation of the secondary cell wall and that pectin is associated with this structure.

MATERIALS AND METHODS

PLANT MATERIALS

Arabidopsis wild type (*Col-0*), *irx8-5* (SALK_044387), *irx8-2* (SAIL_603_G02), *parvus-3* (SALK_045368), and *irx9-1* (SALK_058238) plants were grown on soil in a controlled-environment chamber (Conviron, Pembina, ND) under a 14-h-light/10-h-dark cycle at 19 and 15°C, respectively. Light intensity was 150 μ Em $^{-2}$ s $^{-1}$ and relative humidity was maintained at 50%. Plants were harvested after 7–8 weeks. T-DNA insertions were confirmed using primers from genomic regions flanking the T-DNA and the general T-DNA left border primer (Supplemental Table S1). Arabidopsis *Col-0* and *irx8* heterozygote plants were transformed via the floral dip method (Clough and Bent, 1998) and transgenic plants selected on $\frac{1}{2}$ MS media plates containing 15 mg/L hygromycin. Transgenic plants harboring the construct in *Col-0* and *irx8* homozygote mutant backgrounds were genotyped using PCR (Primers listed in Supplemental Table S1).

GENERATION OF THE GAUT12-EGFP CONSTRUCT

GAUT12 coding sequence (CDS) was amplified from total RNA (0.5 μ g) isolated from 7-week-old Arabidopsis *Col-0* stem by RT-PCR using the SuperScriptTM III One-Step RT-PCR System with Platinum Tag High Fidelity (Invitrogen 12574-030) and cloned into pGEM®-T Easy vector (Promega) (primers listed in Supplemental Table S1). The amplified *GAUT12* CDS was sequence-verified and cloned into the over-expression construct pCambia35t:egfps2#4 (Pattathil et al., 2005) between the NcoI and KpnI restriction sites to produce the *GAUT12-EGFP* construct driven by the CaMV 35S promoter. The plasmid was electroporated into *Agrobacterium tumefaciens* strain GV3101 competent cells and the transformed cells used to transform both wild-type (WT) and heterozygous *irx8* Arabidopsis plants.

HISTOCHEMICAL STAINING

Mäule reagent was prepared as described (Chapple et al., 1992) with slight modifications and used to detect S lignin. Arabidopsis open flowers and hand-cut stem transverse sections were treated with 0.5% (w/v) KMnO₄ solution for 10 min and rinsed with water. For flower samples, the solution was supplemented with 0.01% (v/v) 7X detergent (Linbro, Flow Laboratories) to break surface tension. Samples were treated with 10% (v/v) HCl for 5 min, rinsed with water and mounted in concentrated ammonia for microscopic observation.

Phloroglucinol-HCl stain was prepared freshly as described (Guo et al., 2001). Two parts of 2% (w/v) phloroglucinol in 95% (v/v) ethanol were mixed with one part concentrated HCl. Pictures were taken 10 min after applying the stain. Stained flowers were viewed using a dissecting scope (Olympus SZH) under dark field. Stained stem transverse sections were viewed using a Nikon Eclipse80i microscope under bright field. Images were captured using a Nikon DS-Ri1 camera head (Nikon, Melville, NY).

SCANNING ELECTRON MICROSCOPY

Using a dissecting scope, anthers from WT and mutant open flowers were removed with dissecting forceps (Sigma-Aldrich T4537). Pollen was released onto specimen stubs topped with double-sided sticky carbon tabs by gently tapping the forceps, or lightly tapping the anthers onto the stubs. The *irxa* mutant anthers were first manually dissected to open them and pollen was gently scooped out using the forceps tip and transferred onto the stub surface. Samples were dehydrated and coated with gold particles for 120 s in a Sputter Coater, and imaged using either a JEOL JSM-5800 (SEM/EDAX) scanning electron microscope or a Topcon Aquila—Hybrid SEM.

IN VITRO POLLEN TUBE GROWTH AND RNA PREPARATION

Pollen tubes were grown *in vitro* as previously described (Dardelle et al., 2010). In brief, WT pollen grains were collected from 40 open flowers by vortexing for 3 min in a micro-centrifuge tube containing 1 ml pollen germination medium (PGM) composed of 5 mM CaCl₂·2H₂O, 0.01% (w/v) H₃BO₃, 5 mM KCl, 1 mM MgSO₄·7H₂O, and 10% (w/v) sucrose (pH adjusted to 7.5 using KOH). The flowers were carefully removed and pollen grains pelleted by 3200 g centrifugation for 6 min. The old media was removed, the pollen pellet gently re-suspended in 250 µL of fresh (PGM), and the pollen grains transferred into a 13 × 100 mm glass tube, covered with 3M micropore tape and set in the dark at 22°C for 6 or 24 h.

For RNA isolation, pollen grains from 200 open WT flowers were collected in PGM in five tubes. Pollen from 200 flowers was either directly harvested as hydrated pollen grains (0.5 h) or grown as pollen tubes for 6 and 24 h. Hydrated pollen grains (0.5 h) were combined and ground in liquid nitrogen using a plastic pestle and a microcentrifuge tube. Pollen tubes grown for 6 h and 24 h for RNA isolation were collected by centrifugation for 6 min at 3200 g and ground in microcentrifuge tubes. RNA isolation was repeated using three batches of independently collected tissues.

TISSUE FIXATION AND IMMUNOLABELING

Freshly cut plant tissues were fixed in 25 mM sodium phosphate buffer (pH 7.1) with 1.6% (w/v) paraformaldehyde and 0.2% (w/v) glutaraldehyde overnight at 4°C. Using a lab-grade microwave (PELCO BioWave Pro, Ted Pella, CA) set at 250 Watt, tissues were washed three times for 1 min each with 25 mM sodium phosphate buffer (pH 7.1) followed by two washes with water. Samples then underwent a series of 40-s ethanol gradient incubations (35, 50, 75, 95, 100, 100, and 100% [v/v]) to dehydrate the tissue. Samples were infiltrated with cold LR White embedding resin (Ted Pella) in a gradient (1:3, 1:1, 3:1 resin:ethanol [v/v]) and finally three times with 100% resin. Each step was conducted under vacuum (20" Hg) for 2.5 min. After the last resin change, samples were kept at 4°C for 24 h, transferred into gelatin capsules filled with resin, and polymerized under 365 nm UV light at 4°C for 48 h. Tissue cross sections (250 nm) were cut with a Leica EM UC6 ultramicrotome (Leica Microsystems), mounted on pre-coated slides (Colorfrost/Plus, Fisher Scientific) and used for immunolabeling or stained with 0.05% (w/v) toluidine blue for light microscopy.

Immunolabeling for fluorescent microscopy was done as described (Avci et al., 2012). For LM series and JIM series antibodies, a wash buffer containing 10 mM KPBS (pH 7.2) and 100 mM NaCl was used because we found that the use of 500 mM NaCl adversely affected consistent binding of these antibodies. The secondary antibody used for the LM and JIM series was Alexa fluor 488 goat anti-rat IgG (Cat#A11006, Invitrogen) which was applied to the sections in the same manner as described above, but diluted in the low salt wash buffer. Images were captured using a Nikon DS-Ri1 camera head (Nikon, Melville, NY). All data shown depict representative images out of three images viewed for each type of tissue section stained with each antibody.

TRANSMISSION ELECTRON MICROSCOPY

Ultrathin sections (80 nm) were prepared using an ultramicrotome (Leica EM UC6, Austria) and collected on Formvar-coated nickel grids. Grids were stained with 2% (w/v) uranyl acetate for 4 min followed by 10 dips in three changes of deionized water and dried by wicking. Micrographs were recorded on film in a JEOL 100S transmission electron microscope. The negatives were developed and scanned in Adobe Photoshop.

For immunogold labeling, ultrathin sections were blocked in TBS (10 mM Tris buffer, 150 mM NaCl, pH 7.5) containing 0.06% (w/v) bovine serum albumin for 30 min at room temperature in a petri-dish with a folded piece of water-soaked Kimwipe set to one side of the dish. Sections were transferred onto 10 µL drops of primary antibody diluted (1:5) in TBS for 1 h. After washing three times by dipping the grids (10 times) in TBS, secondary antibody (goat anti-rat IgG coupled to 15 nm gold,) diluted 1:10 in TB (10 mM Tris buffer) containing 0.06% (w/v) bovine serum albumin was applied to the sections. The sections were dipped in TB and distilled water for washing and dried by wicking. Images shown are representative of four images viewed for each sample stained with each antibody.

STEM RNA ISOLATION AND QUANTITATIVE PCR ANALYSES

The bottom half of stem tissues from wild type and *irxa8*, in which secondary cell walls are actively synthesized, were flash frozen in liquid nitrogen and ground to a fine powder in liquid nitrogen in pre-chilled mortars with pestles. Total RNA was isolated from ~100 mg of frozen powder from three individual tissue samples using RNeasy Plant Mini Kit (Qiagen, 74904). First-strand cDNA was synthesized from 1 µg of total RNA using SuperScript III First-Strand Synthesis Super mix (Invitrogen, 18080-400) followed by quantitative PCR analysis using iQ™ SYBR Green Supermix (Bio-Rad 170-8882) on a CFX96™ Real-Time PCR Detection System (Bio-Rad) following the manufacturer's instructions. Melt curve analyses were performed after each run to ensure single size amplicon production. Data are the average ± standard deviation of three biological samples. The data were analyzed as described (Livak and Schmittgen, 2001). Primer sequences are provided in Supplemental Table S1.

CELL WALL (AIR) PREPARATION

Whole stem tissues were harvested from 7-week-old *Col-0*, *irxa8-5*, *irxa8-5+GAUT12*, *irxa8-2+GAUT12*, and *WT+GAUT12*, ground with a mortar and pestle to a fine powder in liquid nitrogen, re-suspended in 80% (v/v) ethanol and rotated end-to-end for 12 h. The pellet obtained upon 4000 rpm centrifugation was sequentially washed in 80% (v/v) ethanol, 100% ethanol, chloroform: methanol (1:1, v/v), and acetone by re-suspension, rotation for 12 h, and re-centrifugation. The resulting alcohol insoluble residue (AIR) was dried for 24 h at room temperature.

GLYCOSYL RESIDUE COMPOSITION ANALYSES

Neutral sugar analysis was performed as described (Albersheim et al., 1967) with slight modifications. Each AIR sample (0.4 mg with 20 µg *myo*-inositol as an internal standard) was hydrolyzed with 30 drops of 2 N trifluoroacetic acid (TFA) for 2 h at 120°C. Samples were cooled to room temperature and dried under an air stream, washed twice with isopropanol to remove TFA, and reduced by incubation for 1 h at room temperature in 10 drops of sodium borohydride (10 mg/ml) dissolved in 1 M ammonium hydroxide solution. The reaction was quenched with 30–40 drops of acetone and dried down with air. Once the volume was reduced to half, isopropanol (1 ml) was added to each sample to facilitate drying. O-acetylation was performed by adding 250 µl of acetic anhydride followed by 230 µl of concentrated TFA and the samples were incubated for 10 min at 50°C. Samples were then washed with isopropanol and dried down with air. H₂O (1 ml) and dichloromethane (DCM, 1 ml) were added and the samples vortexed and centrifuged to allow phase separation. The top aqueous layer was discarded and the bottom DCM layer containing the alditol acetate derivatives was dried down. Ten drops of DCM were added to each sample before analysis by gas-liquid chromatography-flame ionization detection using an Agilent 7890A GC system.

For GalA and GlcA measurements, AIR samples (0.5 mg) were hydrolyzed in 2N TFA at 105°C for 1 h. The hydrolysates were dried down with an air stream, re-hydrolyzed in 3N methanolic HCl (Thermo Scientific, Rockford, IL) overnight at 80°C, dried down individually and dissolved in 100 µl distilled

H₂O. The samples were centrifuged and 12.5 µl of hydrolsate supernatant was loaded onto a CarboPac PA20 analytical column (3 × 150 mm, Dionex, Sunnyvale, CA). The loaded column was washed with a solution of 49 mM NaOH, 20 mM NaOAc and eluted with a linear gradient from 49 mM NaOH, 20 mM NaOAc to 40 mM NaOH, 200 mM NaOAc over 25 min. The column was eluted at 30°C and a flow rate of 0.4 ml/min and effluent was monitored with an ECD detector. The amount of GalA and GlcA was determined by comparison of peak areas to standards separated under the same conditions.

Each analysis was repeated five times for each sample, and the bar represents the average mol% of each sugar residue ± standard error.

SEQUENTIAL EXTRACTION AND GLYCOME PROFILING

Sequential extractions of cell wall samples and glycome profiling were carried out as described previously (Demartini et al., 2011; Pattathil et al., 2012). Briefly, AIR samples were sequentially extracted with 50 mM ammonium oxalate, pH 5; 50 mM sodium carbonate [containing 0.5% (w/v) sodium borohydride], pH 10; 1 M KOH with 1% (w/v) sodium borohydride; 4 M KOH with 1% (w/v) sodium borohydride; 100 mM acidified sodium chlorite; and finally with 4 M KOH with 1% (w/v) sodium borohydride for the post-chlorite extraction. The wall extracts were used for NMR analyses (see below) and were screened by ELISA using plant glycan-directed monoclonal antibodies (CCRC, JIM, and MAC series) from Complex Carbohydrate Research Center stocks available through CarboSource Services (<http://www.carbosource.net>). Detailed description of each mAbs used in this study can be found in the Supporting Information (Supplemental Table S2) that includes links to a web database named WallMAbDB (<http://www.wallmabdb.net>).

PYROLYSIS MOLECULAR BEAM MASS SPECTROMETRY (pyMBMS)

About 4 mg of sample was weighed and transferred into 80-µl stainless steel sample cups of an auto sampler of a double shot pyrolyzer (PY-2020iD, Frontier Ltd.). The samples were pyrolyzed at 500°C and the residues analyzed using a custom built Super Sonic Molecular Beam Mass Spectrometer (Extrel Model MAX-1000).

Mass spectral data from m/z 30–450 were acquired on a Merlin Automation Data System version 3.3. Multivariate analysis was performed using Unscrambler software version 10.1 (CAMO). The intensities of the lignin peaks were summed and averaged to estimate the lignin content in the sample (Evans and Milne, 1987). Total lignin peaks corresponded to m/z 120, 124, 137, 138, 150, 152, 154, 164, 167, 168, 178, 180, 181, 182, 194, 208, and 210. The syringyl peaks corresponded to m/z 154, 167, 168, 182, 194, 208, and 210, the guaiacol peaks corresponded to m/z 124, 137, 138, 150, 164, and 178, and phenol peaks corresponded to m/z 120 and 122. Syringyl to Guaiacol (S/G) ratios were determined by summing syringyl peaks and dividing by the sum of guaiacol peaks. The lignin values thus generated and calculated were compared with the WT. Because the NIST (National Institute Standards and Technology) standard for *Arabidopsis* is not available, the lignin percentage was corrected using a standard of eastern cottonwood (NIST 8492) to obtain the lignin

percentage value equivalent to the Klason lignin. This analysis was conducted at the Complex Carbohydrate Research Center Analytical Services, University of Georgia (<http://www.ccrc.uga.edu/services/crcanalyticalservices/index.html>).

DETERMINATION OF LIGNIN MONOMER COMPOSITION BY HSQC NMR SPECTROSCOPY

Perdeuterated pyridinium molten salt (ionic liquid) was synthesized as described (Jiang et al., 2009). Approximately 2 mg of cell wall extracts were weighed and dissolved in 180 μ L of the ionic liquid [DMSO-*d*₆/pyridine-*d*₅ (2:1, v/v)] at 65°C. Data were collected at 60°C on Agilent 600 MHz Direct Drive spectrometers equipped with either a 5 or 3 mm cold probe. A standard Agilent pulse program (“HSQCAD”) was used to acquire the ¹³C-¹H heteronuclear correlated spectra. The proton dimension of 1200 complex data points covered 20 ppm centered at 6 ppm, and the carbon dimension of 48 or 64 complex points was centered at 92 ppm with a width of either 100 or 184 ppm, respectively. In the former case, some resonances in the alkyl or aromatic regions were folded along the F1 axis. Total data acquisition times ranged from 7 to 16 h, with the number of transients between 240 and 400 per *t*₁ increment. Data were processed with NMRPipe (NIH) and visualized with NMRViewJ (One Moon Scientific) or with MNova (Mestrelab Research). Typically, squared cosine window functions were applied in both dimensions after zero filling and linear prediction in *t*₁. Chemical shifts were referenced to DMSO at 2.50 ppm in proton and 39.51 ppm in carbon. Heteronuclear Single-Quantum Correlation (HSQC) cross-peak assignments were determined and referenced as described (Kim and Ralph, 2010).

GENERATION OF ARABIDOPSIS SUSPENSION CULTURES AND GAUT12-EGFP TRANSGENIC CELLS

A wild-type Arabidopsis suspension culture was generated from seed callus as previously described (Doelling and Pikaard, 1993). Cell were subcultured every 10 days by transfer of 7 ml of packed cells into 100 ml fresh callus inducing medium (CIM). CIM contained 3.2 g/L Gamborg's B-5 basal medium with minimal organics (Sigma-Aldrich G5893 or PhytoTechnology Laboratories G398), 2 mg/L 2,4-D, 0.05 mg/L kinetin, and 20 g/L sucrose (pH 5.7). The GAUT12-EGFP construct was stably transformed into wild-type suspension culture cells via *Agrobacterium tumefaciens* strain GV3101. An aliquot of 1 ml packed cells from a four-day-old culture was co-cultured in 8 ml of fresh CIM with 100 μ L of *Agrobacterium* cells harboring the GAUT12-EGFP vector re-suspended in CIM to an OD₆₀₀ of 0.8. The *Agrobacteria* were previously seed-cultured overnight in 3 ml YEP medium supplemented with rifampicin 50 mg/L, kanamycin 50 mg/L, and gentamycin 50 mg/L and re-cultured in 25 ml of the above YEP medium. The co-culture was done in a 1-inch-deep petri-dish in the dark on a gyrotory shaker at 130 rpm for 24 h. Cells were removed from the co-cultivation medium and washed thrice with CIM and then with CIM supplemented with 500 mg/L cefotaxime in 50 ml falcon tubes. In each wash, the old medium was completely removed and 20 ml fresh medium introduced and vortexed for 30 s. The washed cells were plated onto CIM 0.6% (w/v) agar plates containing 300 mg/L cefotaxime and 15 mg/L

hygromycin. Hygromycin-resistant calli emerged in 2–3 weeks and were transferred onto fresh media plates and grown for up to 4 weeks. After two to three transfer cycles, the calli were free of *Agrobacteria* and used to initiate suspension cultures. Surviving calli were genotyped, and RT-PCR was used to determine the expression level of the GAUT12-EGFP transcript. The GAUT12-EGFP transcript was highest on day 5 compared to day 2 and 8. Therefore, cells were harvested on day 6 for preparation of microsomes.

MICROSOMAL MEMBRANE PREPARATION

Microsomal membranes used for enzyme activity assays were prepared at 4°C as described (Orfila et al., 2005) with modification. Arabidopsis stems (10 grams) were cut into small pieces, flash frozen in liquid nitrogen, and homogenized on ice in 20 ml of pre-chilled homogenization buffer containing 50 mM Hepes (pH 7.3), 0.4 M sucrose, 0.1 M sodium ascorbate, 0.25 mM MnCl₂, 25 mM KCl, 1% (w/v) polyvinylpyrrolidone (PVPP), and EDTA-free protease inhibitor cocktail (Roche) until the tissues were pureed. The homogenate was filtered through three layers of miracloth and the filtrate centrifuged at 4°C for 30 min at 4000 g to remove cell debris and intact cells. The supernatant was ultra-centrifuged at 110,000 g for 1 h, yielding the microsome pellet which was re-suspended on ice in pre-chilled storage buffer (homogenization buffer without PVPP, 30 μ L buffer/gram fresh weight) using a glass homogenizer. Total protein was measured using the Bradford assay (Bio-Rad Protein Assay 500-0006) with BSA as a standard. Aliquots of the microsomes were used directly or flash frozen in liquid nitrogen and stored at -80°C for later use in enzyme assays and immunoprecipitation experiments.

GENERATION OF anti-GAUT12 POLYCLONAL ANTIBODY AND WESTERN BLOTTING

The anti-GAUT12 polyclonal antibody was generated against synthetic peptides corresponding to GAUT12 amino acid residues 101–114 (EQPLSEQELKGRSD) and antigen-purified over a column packed with the antigenic-peptide (service via New England Peptide, <http://www.newenglandpeptide.com/>). The purified anti-GAUT12 antibody did not cross-react with GAUT1 or GAUT7 (Supplemental Figure S11). Pre-immune serum did not contain anti-GAUT12 antibody. Anti-GAUT1 and anti-GAUT7 antibodies were generated previously (Sterling et al., 2006; Atmodjo et al., 2011). For western blotting, a dilution factor of 1:5000, 1:10000, 1:3000, and 1:2000 was applied for anti-GAUT12, anti-GAUT7, anti-GAUT1, and anti-GFP (abcam, ab6556), respectively. Either horseradish peroxidase (HRP)- or alkaline phosphate (AP)-conjugated goat anti-rabbit secondary antibody (Sigma) was used followed by a reaction with corresponding substrates to yield a blue or purple color precipitant on the target protein band.

IMMUNOPRECIPITATION OF GAUT12 FOR HG:GALAT ENZYME ASSAYS

The purified anti-GAUT12 antibody (1.14 mg/ml) was incubated with Dynabeads M-280 Sheep anti-Rabbit IgG (Invitrogen, 112-04D) in a ratio of 1:20 (v/v) for 2 h at 4°C on a tube rotator and the beads collected on a magnet stand and washed thrice with isotonic PBS (pH 7.4, 139 mM NaCl, 5.5 mM Na₂HPO₄, 1.2 mM

NaH_2PO_4). The beads were then washed once with storage buffer (see microsome preparation section). Each reaction contained 30 μl anti-GAUT12-conjugated beads incubated with ~500 μg Triton X-100-treated microsomes at 4°C for 2 h on the tube rotator. The nonionic detergent Triton X-100 (TX-100) has been used to solubilize GalAT activity during GAUT1 and GAUT7 purification (Doong and Mohnen, 1998; Sterling et al., 2006; Atmodjo et al., 2011). A concentration of 4% (v/v) TX-100 was used to homogenize WT stem microsomes on ice and they were immediately diluted with storage buffer to a final detergent concentration of 0.5% (v/v) for incubation with anti-GAUT12-conjugated magnetic beads.

After a 2 h end-to-end incubation at 4°C, the beads were washed with pre-chilled storage buffer on ice thrice and the original 30 μl beads were re-suspended in 13 μl pre-chilled storage buffer for each enzyme reaction. Immunoabsorbed-GAUT1 was prepared by incubating Dynabeads M-280 Sheep anti-Rabbit IgG with anti-GAUT7 anti-serum in a ratio of 1:3 (v/v) in parallel as described (Atmodjo et al., 2011). Anti-GAUT7 antibody was previously shown to immunoabsorb the HG:GalAT activity-containing the GAUT1-GAUT7 core complex (Atmodjo et al., 2011).

HG:GALAT ENZYME ACTIVITY ASSAY

UDP-D-[¹⁴C]GalpA (specific activity 180.3 mCi/mmol; 1 Ci = 37 GBq) was synthesized enzymatically from UDP-D-[¹⁴C]GlcP-A using UDP-D-GlcP-A 4-epimerase as described (Atmodjo et al., 2011). HG:GalAT activity was assayed in 30- μL reactions containing either enzyme (15 μl total microsomes, ~100 μg total protein) or 13 μl of immunoabsorbed beads, 50 mM Hepes (pH 7.3), 0.2 M sucrose, 0.05% (w/v) BSA, 25 mM KCl, 1.9 mM MnCl₂, 1 mM HG oligosaccharides (referred to as oligogalacturonides, OGA) of a degree of polymerization (DP) of 7–23, and 6.9 μM UDP-[¹⁴C]GalpA (specific activity 180.3 mCi/mmol; 1 Ci = 37 GBq). The reactions were incubated for 3 h at 29°C in a water-bath and terminated by the addition of 5 μl of 400 mM NaOH with vortexing. The entire reaction (~35 μl) was spotted onto 1-inch² filter paper squares that had been pre-treated with cetylpyridinium chloride (CPC) as described (Sterling et al., 2005). The spotted filters were dried, washed 3× in 150 mM NaCl each for 15 min to remove access UDP-[¹⁴C]GalpA, dried, and added to 4 ml ScintiVerse™ BD cocktail for scintillation counting.

LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY (LC-MS/MS)

A large-scale GAUT12 immunoprecipitation was conducted for LC-MS/MS analysis. An aliquot of 500 μl of the anti-GAUT12-conjugated magnetic beads was incubated at 4°C overnight on a tube rotator with WT Arabidopsis stem microsomes (5 mg total protein). The microsomes on ice were homogenized with 4% (v/v) TX-100 supplemented with 200 mM NaCl, 100 mM NaOAc, and 2 mM EDTA and immediately diluted with PBS to a final detergent concentration of 0.5% (v/v) for incubation with anti-GAUT12-conjugated magnetic beads. The beads were washed 5× with PBS (pH 7.4) and the recovered beads denatured in 3% (w/v) SDS and reduced in 25 mM DTT. The material bound

to the beads was released by magnetic separation and separated by electrophoresis on a 10% SDS-PAGE gel. A gel piece corresponding to the size of GAUT12 (between 55 and 70 KDa protein marker) was cut out and in-gel trypsin-digested as described (Atmodjo et al., 2011). The peptide samples from the proteolytic digestions were analyzed on an Agilent 1100 capillary LC (Palo Alto, CA) interfaced directly to a LTQ linear ion trap mass spectrometer (Thermo Fisher, San Jose, CA). Mobile phases A and B were H₂O-0.1% (v/v) formic acid and acetonitrile-0.1% (v/v) formic acid, respectively. Peptides were eluted from the C18 column into the mass spectrometer during an 80 min linear gradient from 5 to 55% (v/v) mobile phase B at a flow rate of 4 $\mu\text{l}/\text{min}$. The instrument was set to acquire MS/MS spectra on the nine most abundant precursor ions. Generated raw tandem mass spectra were converted into the mzXML format and then into peak lists using ReAdW software followed by mzML2Other software (Pedrioli et al., 2004). The peak lists were searched using Mascot 2.2 (Matrix Science, Boston, MA).

DATABASE SEARCHING AND PROTEIN IDENTIFICATION

A target database was created using the *Arabidopsis* annotated sequences obtained from the TAIR10_pep_20101214 protein database (ftp://ftp.arabidopsis.org/home/tair/Genes/TAIR10_genome_release/TAIR10_blastsets/TAIR10_pep_20101214_updated). A decoy database (decoy) was constructed by reversing the sequences in the normal database. Searches were performed against the normal and decoy databases using the following parameters: full tryptic enzymatic cleavage with two possible missed cleavages, peptide tolerance of 1000 ppm, fragment ion tolerance of 0.6 Da. Fixed modification was set as carbamidomethyl due to carboxyamidomethylation of cysteine residues (+57 Da) and variable modifications were chosen as oxidation of methionine residues (+16 Da) and deamidation of asparagine residues (+1 Da). Statistically significant proteins from both searches were determined at a ≤1% protein false discovery rate (FDR) using the ProValT algorithm, as implemented in ProteoIQ (BioInquire, LLC, Athens, GA) (Weatherly et al., 2005).

RESULTS

THE EFFECTS OF *irx8* MUTATION ON GROWTH HABIT AND SECONDARY WALL SYNTHESIS IN XYLEM AND FIBERS ARE COMPLEMENTED BY A GAUT12-OVEREXPRESSION CONSTRUCT

All reported *irx8* mutant alleles are dwarf and show a collapsed xylem phenotype (Peña et al., 2007; Persson et al., 2007). We selected two of these GAUT12 T-DNA insertion mutants, *irx8-2* and *irx8-5*, for this study (Supplemental Figure S1A). The T-DNA insertions in *irx8-2* and *irx8-5* are in the promoter region and fourth intron of GAUT12, respectively (Supplemental Figure S1A). Both alleles are associated with phenotypes similar to those previously described (Peña et al., 2007; Persson et al., 2007). Specifically, both mutants have a dwarfed growth habit (Supplemental Figure S1B), reduced secondary cell wall thickness in xylem and fiber cells, and collapsed xylem vessels (Supplemental Figures S1D,H). Semi-quantitative RT-PCR analyses revealed that *irx8-2* plants contain trace amounts of full-length GAUT12 transcript, whereas *irx8-5* plants contain transcripts that

are truncated at the insertion site (Supplemental Figure S1L). Since both *irxa-2* and *irxa-5* mutants express similar overall phenotypes, they were used interchangeably in the present study.

The *irxa* mutant phenotypes in both alleles are complemented (Supplemental Figures S1E,I,K) by the constitutive expression of an EGFP-tagged *GAUT12* construct (*GAUT12-EGFP*). The EGFP (Pattathil et al., 2005) was connected to the C-terminus of *GAUT12* via a Val-Pro linker to facilitate structural flexibility between the two parts of the fusion protein and access of substrates to the predicted C-terminal catalytic domain of *GAUT12* (Supplemental Figure S1A). The *GAUT12-EGFP* construct restored the phenotype of *irxa* mutants as evidenced by the normal stature (Supplemental Figure S2B) and cell wall thickness (Supplemental Figures S3C,G) in the complemented transgenic plants. Cell wall sugar composition analyses revealed that *irxa-5+GAUT12* and *irxa-2+GAUT12* plants had a xylose content that was closer to normal (65 and 74 mol% of wild-type level, respectively) compared to the *irxa-5* mutant (33 mol% of wild-type xylose content, Supplemental Figure S2C).

Finally, the 20 and 63 mol% reduction of GalA and GlcA content, respectively, associated with the *irxa-5* mutation was also complemented by the *GAUT12-EGFP* construct (Supplemental Figure S2D).

REDUCTION IN LIGNIN AND XYLAN LEADS TO INDEHISCENT ANTERS IN *irxa*

Prior studies reported that multiple alleles of homozygous *irxa* mutants were “semi-sterile” (Brown et al., 2005; Persson et al., 2007). Both the *irxa-2* and *irxa-5* mutants used for this study produced small and empty siliques with almost no seeds (Supplemental Figure S1M).

We investigated the cause of *irxa* sterility. First, we used a dissecting microscope to observe open flowers of *irxa* and wild type (WT) at stages 13 and 14 (anthesis and fertilization) (Bowman, 1994), respectively (Figure 1). An open flower of *irxa-5* contains reproductive organs of the correct shape but smaller in size than wild type (Figure 1A) and has shorter stamens (as previously described; Persson et al., 2007). At this stage, wild-type anthers

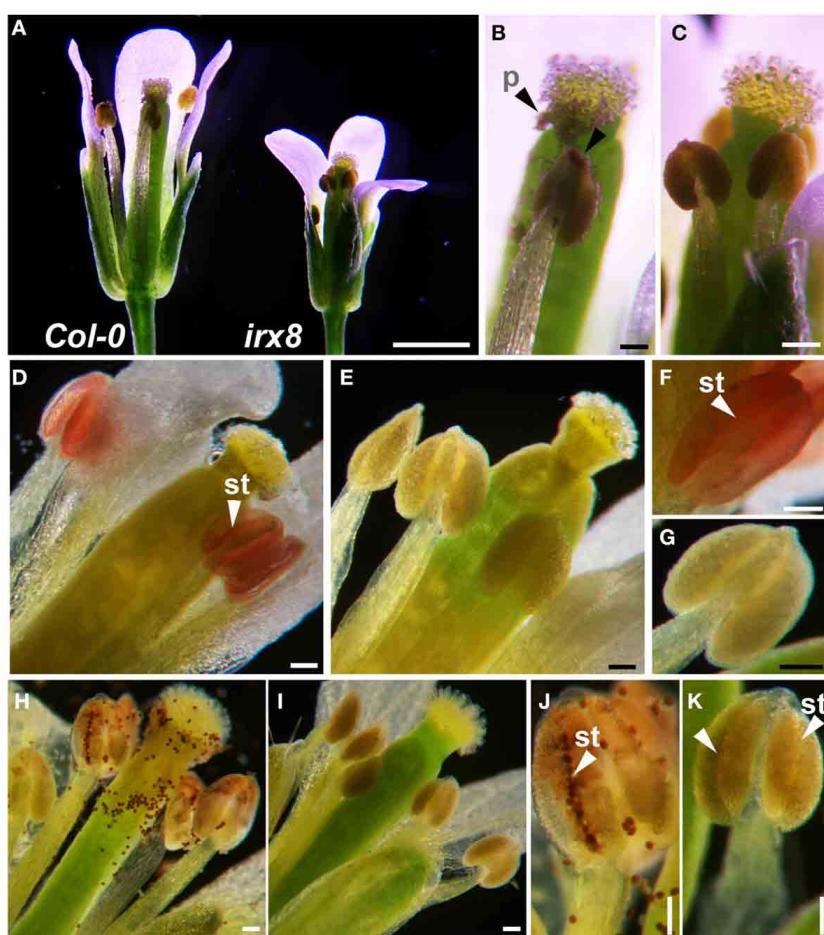


FIGURE 1 | Loss of lignin in anther endothecium cells results in failure of anther dehiscence in *irxa* mutants. (A) Light microscope images of wild-type and *irxa* open flowers (stage 13, anthesis). (B) Close-up of *Col-0* pistil and anthers. (C) Close-up of *irxa-5* pistil and anthers. (D–G) Red phloroglucinol-HCl staining of flower anthers indicates presence of lignin. Endothecium layer stains red in wild type (D), the lack of staining in *irxa-5* (E), close-up of a wild-type

anther (stomium indicated by arrow, F), and close-up of an *irxa* anther showing no staining (G). (H–K) Mäule staining of flowers of wild type: pollen and dehisced anther walls stained red (H); lack of staining in *irxa-5* (I); close-up of wild-type anther stomium (arrowhead), released pollen stained dark red (J); close-up of *irxa-5* stomium (arrowhead), no lignin staining and no pollen released (K). Bar in A = 1 mm; bar in B–K = 0.1 mm. p, pollen; st, stomium.

have already dehisced and pollen grains are released onto the stigma and style (**Figure 1B**, arrowheads). The *irx8-5* mutant, however, has smooth intact anthers and shows no pollen release (**Figure 1C**). The same phenotype was also observed in the *irx8-2* mutant (Supplemental Figure S4).

Secondary wall thickening in endothecium cells of pollen sacs provides the mechanical force for anther dehiscence. After anther dehydration and pollen swelling, the stomium breaks open to release pollen at anthesis, followed by anther filament extension to achieve fertilization (Wilson et al., 2011). Phloroglucinol-HCl staining (Wiesner test), which colorimetrically identifies coniferaldehyde end-groups in G lignin, and Mäule reagent, which reacts with syringylpropane moieties of S lignin to produce a rose red color (Lewis and Yamamoto, 1990) were used to compare lignification in the wild type and mutant. The lignin of the secondary wall thickening along the stomium furrow in wild-type mature anther stained red using phloroglucinol-HCl (**Figures 1D,F**, arrowheads). In contrast, the stomium furrow was not stained at all in *irx8* anthers at the same stage (**Figures 1E,G**), suggesting a reduction of G lignin in the *irx8* endothecium cell layer. In addition, released pollen from the wild type stained brownish-red with Mäule reagent at the stomium opening (**Figure 1H**). The wild-type pollen sacs were also partially stained since the stain was able to access the endothecium

layer through the open stomium (**Figure 1J**). However, at the same developmental stage (anthesis), such staining was absent in *irx8* anthers (**Figures 1I,K**). Despite no pollen release in *irx8* flowers, the anthers appeared to be turgid (**Figure 1C**; Supplemental Figure S4), suggesting that the pollen inside was swollen. The enzymatic lysis of the *irx8* stomium and septum breakage also occurred in the anther, as seen in transverse sections (**Figure 2E**, black arrowheads). The indehiscent anther phenotype was confirmed in both the *irx8-2* and *irx8-5* mutants using SEM. This phenotype was complemented by the GAUT12-EGFP construct (**Figure 3A**; Supplemental Figure S4).

The secondary wall thickening of the endothecium cells (**Figure 2A**) contains not only lignin, but also xylan and cellulose (Wilson et al., 2011). Since *irx8* has a known xylan defect, we fixed, embedded, and sectioned open flowers of *irx8* and analyzed anther cell walls in the sectioned tissues using plant cell wall glycan-directed monoclonal antibodies against xylan and other major polysaccharides (Pattathil et al., 2010). In wild-type endothecium cell walls, LM10 (**Figure 2B**) which binds low-substituted xylan and LM11 (**Figure 2C**) which binds both low- and high-substituted xylan (McCartney et al., 2005) showed almost identical punctate labeling patterns, demonstrating a common location of the two xylan epitopes recognized by these two antibodies. Both epitopes are reduced in *irx8* (**Figures 2F,G**),

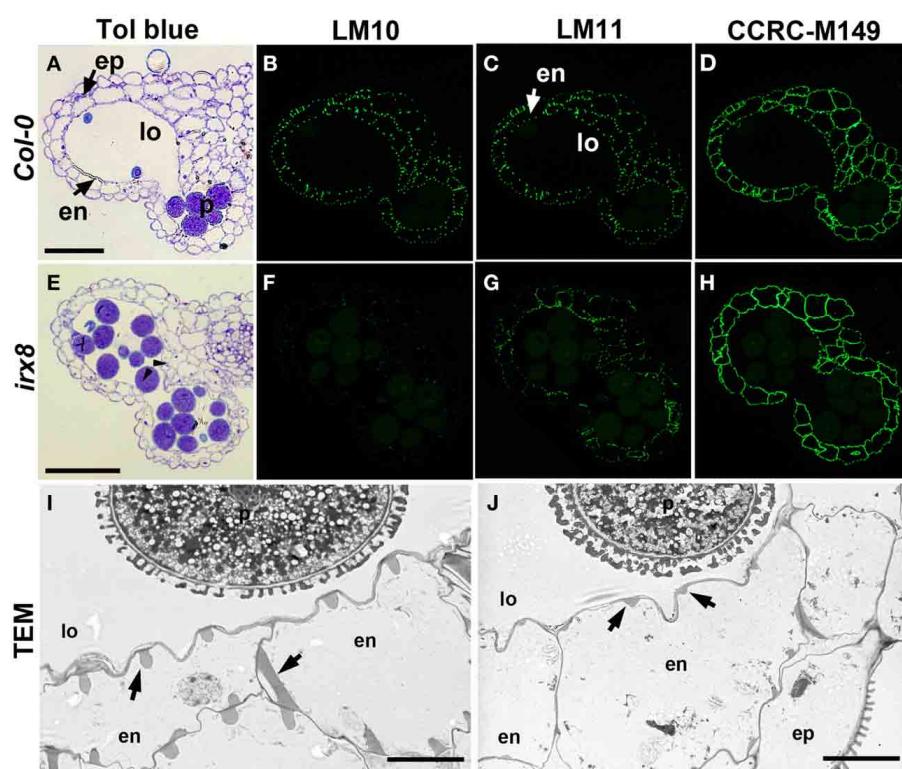


FIGURE 2 | Reduction in xylan deposition in anther endothecium cells in *irx8* mutants. LR White-embedded WT (**A**) and *irx8-5* (**E**) anther transverse sections (250 nm-thick) stained with toluidine blue O (Tol blue), arrowheads point to septum breakage. WT (**B–D**) and *irx8* (**F–H**) anthers immunolabeled with xylan-reactive antibodies LM10

(**B,F**), LM11 (**C,G**), and CCRC-M149 (**D,H**). Bar = 50 μm for (**A–H**). TEM of WT (**I**) and *irx8-5* (**J**) anther. Arrows point to the secondary wall thickenings in the endothecium cell layer, which is reduced in *irx8-5*. ep, epidermis; en, endothecium; p, pollen; lo, locule. Bar = 5 μm for (**I**) and (**J**).

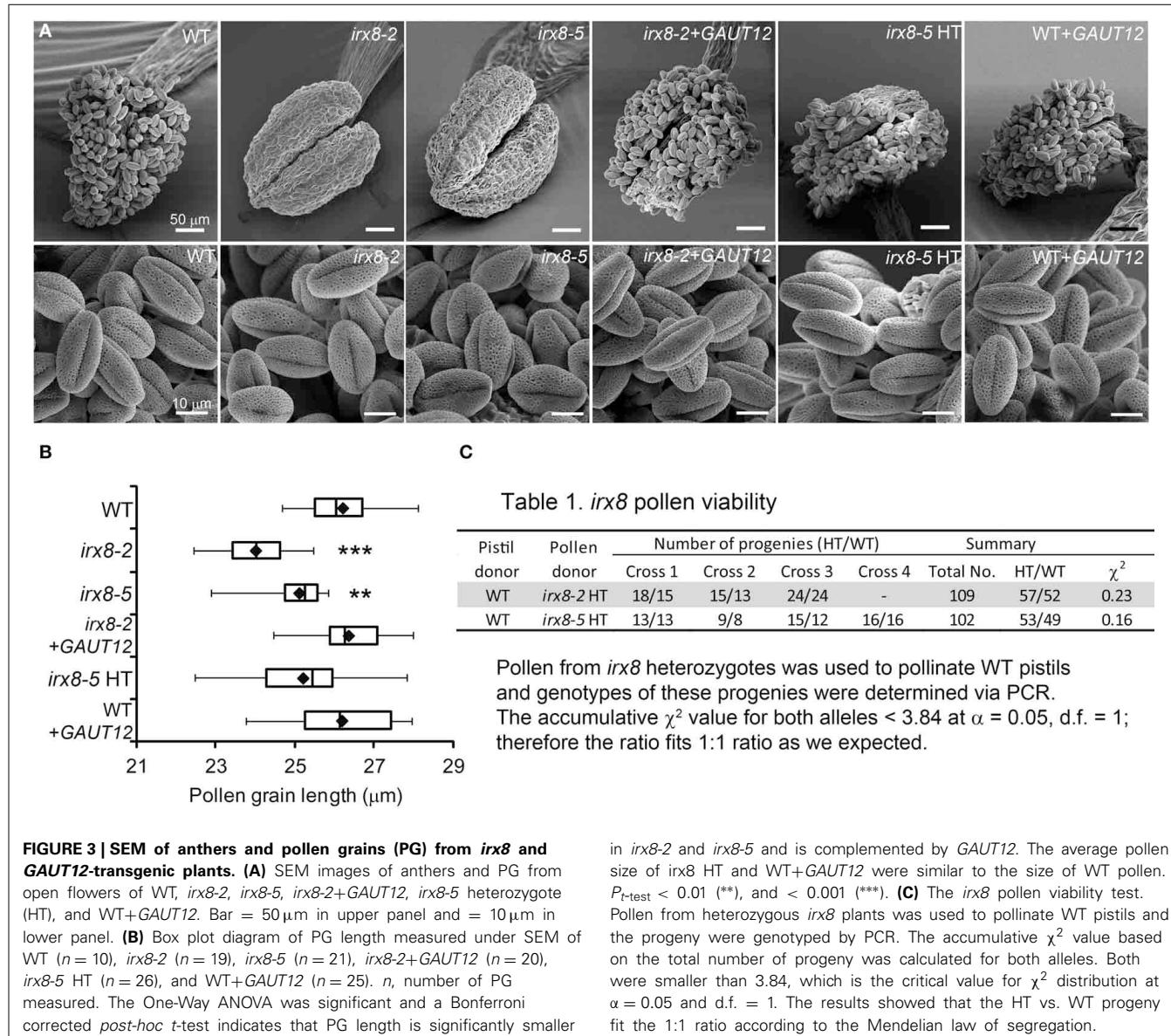


FIGURE 3 | SEM of anthers and pollen grains (PG) from *irx8* and *GAUT12*-transgenic plants. (A) SEM images of anthers and PG from open flowers of WT, *irx8-2*, *irx8-5*, *irx8-2+GAUT12*, *irx8-5* heterozygote (HT), and WT+*GAUT12*. Bar = 50 μ m in upper panel and = 10 μ m in lower panel. **(B)** Box plot diagram of PG length measured under SEM of WT ($n = 10$), *irx8-2* ($n = 19$), *irx8-5* ($n = 21$), *irx8-2+GAUT12* ($n = 20$), *irx8-5* HT ($n = 26$), and WT+*GAUT12* ($n = 25$). n , number of PG measured. The One-Way ANOVA was significant and a Bonferroni corrected post-hoc *t*-test indicates that PG length is significantly smaller

in *irx8-2* and *irx8-5* and is complemented by *GAUT12*. The average pollen size of *irx8* HT and WT+*GAUT12* were similar to the size of WT pollen. $P_{t\text{-test}} < 0.01$ (**), and < 0.001 (***)�. **(C)** The *irx8* pollen viability test. Pollen from heterozygous *irx8* plants was used to pollinate WT pistils and the progeny were genotyped by PCR. The accumulative χ^2 value based on the total number of progeny was calculated for both alleles. Both were smaller than 3.84, which is the critical value for χ^2 distribution at $\alpha = 0.05$ and d.f. = 1. The results showed that the HT vs. WT progeny fit the 1:1 ratio according to the Mendelian law of segregation.

with LM10 labeling almost completely absent in *irx8* (Figure 2F), indicating a pronounced reduction in low-substituted xylan recognized by LM10. Transmission electron microscopy (TEM) showed that the secondary wall thickening in wild-type endothecium cells, in the shape of teeth- and ribbon-like structures (Figure 2I, arrows), is significantly reduced in *irx8* endothecium cells (Figure 2J, arrows). Xylan present in the *irx8* endothecium was detected using CCRC-M149 (Figure 2H), CCRC-M137, CCRC-M138, and CCRC-M160 antibodies (Supplemental Figures S5I,M,N), suggesting that *GAUT12* affects some xylan synthesis in endothecium cells and that some xylan epitopes remain unchanged in *irx8* anthers compared to wild type (Figure 2D, Supplemental Figures S5B,F,G). No difference was seen in *irx8* endothecium walls using anti-peptin antibodies JIM5, JIM7, and CCRC-M38 (Supplemental Figures S5U–W) that recognize HG epitopes with various degrees of methylesterification,

and little change was observed in this tissue using antibodies CCRC-M14, JIM13, and CCRC-M1, which recognize RG-I backbone, arabinogalactan protein (AGP), and fucosylated xyloglucan epitopes, respectively (Supplemental Figures S5X–Z).

TEM revealed smaller pollen size in *irx8-5* (Figure 2J), which was confirmed by SEM (Figure 3A). While the wild-type pollen grains are uniform in size and shape, both mutant alleles of *irx8* have smaller pollen grains when manually released from pollen sacs (Figure 3B). Occasionally we observed defective *irx8* pollen grains under both SEM and TEM. However, since pollen development is easily affected by growth conditions, and pollen formation involves both sporophytic and gametophytic factors (Ariizumi and Toriyama, 2011), we wanted to test directly whether *GAUT12* affects pollen fertility. We found that manually released *irx8* pollen was able to pollinate both wild-type and heterozygous *irx8* pistils, and wild-type pollen was able to

pollinate the *irx8* pistil, both producing viable seeds. Interestingly, however, manual fertilization of the homozygous *irx8* pistil with its own pollen was not successful. We observed that the *irx8* inflorescence often dried soon after the stem stopped elongating, which may be due to the poor water conduction in this mutant. This may explain why manual fertilization of the *irx8* pistil was unsuccessful.

We used quantitative PCR to determine if *GAUT12* is expressed in pollen. Small amounts of *GAUT12* transcript was detected in hydrated pollen grains and pollen tubes (Supplemental Figure S6). However, no expression of *CESA4* or *IRX9* transcript, genes known to encode secondary wall cellulose and xylan biosynthetic proteins, respectively (Taylor et al., 2003; Peña et al., 2007), was detected in either tissue. In contrast, relatively high levels of expression were observed for *GAUT1* and *CESA1* (Supplemental Figure S6B), genes known to be involved in primary cell wall pectin and cellulose biosynthesis, respectively (Arioli et al., 1998; Sterling et al., 2006; Atmodjo et al., 2011).

To further analyze whether *GAUT12* affects pollen tube formation and viability, we manually fertilized wild-type pistils with pollen from the *irx8* heterozygote, a pollen population consisting of both *irx8* and wild-type pollen produced in equal amounts. The number of heterozygote vs. WT progeny from these crosses were 57:52 and 53:49 for *irx8-2* and *irx8-5* heterozygotes, respectively, fitting the 1:1 ratio expected for Mendelian segregation according to a χ^2 test (Figure 3C). This result demonstrated that *irx8* pollen has normal fertility. From these results we conclude that the sterility of *irx8* mutants is due to indehiscent anthers and to the resulting inability of the mutant to release pollen for fertilization.

THE *irx8* MUTANT HAS LOW LIGNIN CONTENT IN BASAL STEMS AND REDUCED EXPRESSION OF MAJOR LIGNIN BIOSYNTHETIC GENES

Reduction of lignin in the *irx8* endothecium cell layer led us to examine the lignin content in *irx8* stems by histochemical staining. We found a reduction in total lignin compared to wild type (Figure 4), an observation recently confirmed by Petersen et al. (2012). Wild-type xylem vessels and fibers had thick secondary walls with evenly distributed lignin (Figures 4A,D). In contrast, interfascicular fiber cells of *irx8* mutants stained very weakly with phloroglucinol-HCl (Figure 4B), indicating a loss of G lignin. Staining with Mäule reagent revealed a brown staining in *irx8* xylem cells and reduced red staining in the interfascicular fiber cells (Figure 4E), indicative of a possible but lesser reduction in S lignin deposition. The lignin content was recovered in the *GAUT12*-complemented *irx8* (*irx8+GAUT12*) plants, as shown by staining using both stains (Figures 4C,F).

To determine whether the reduced lignin content was due to reduced biosynthesis of lignin monomers, we quantified the steady state levels of transcripts for 10 enzymes of the lignin biosynthetic pathway (Raes et al., 2003) in the basal half of *irx8* stems (Figure 4G). Significantly reduced expression was observed for several key enzymes including trans-cinnamate 4-hydroxylase (C4H), 4-coumaroyl shikimate 3'-hydroxylase (C3'H), caffeoyl CoA O-methyltransferase (CCoAOMT1), and caffeic acid/5-hydroxyferulic acid O-methyltransferase (COMT1). We conclude that the generation of lignin precursors is likely down-regulated in the *Arabidopsis* *irx8* mutant.

To further evaluate lignin structure and lignin composition in *irx8* stems, AIR was analyzed using pyrolysis molecular beam mass spectrometry (pyMBMS) (Sykes et al., 2008). There was a moderate but significant reduction (9%) in total lignin which was complemented by the *GAUT12-EGFP* transgene (Figure 4H). The reduction of total lignin in *irx8* stems was largely due to a lower amount (18% reduction) of guaiacyl (G) subunits, whereas there was a relatively normal amount of *p*-hydroxyphenyl (H) and syringyl (S) subunits in *irx8* (Figure 4I). The reduction in G lignin resulted in an increased S/G ratio (*irx8-2*: 0.9 ± 0.09; WT: 0.7 ± 0.06; $P_{t\text{-test}} = 0.0006$). The *irx8+GAUT12* plants had an S/G ratio of 0.6 ± 0.04 ($P_{t\text{-test}} = 0.011$), slightly lower than that of the wild type.

We further characterized the lignification in *irx8* and wild-type stems by sequentially extracting stem AIR with 50 mM ammonium oxalate, 50 mM sodium carbonate, 1 M and 4 M KOH, acidified sodium chlorite, and post-chlorite 4 M KOH. We first used 2D ^{13}C - ^1H HSQC NMR spectroscopy to analyze the cell wall fraction extracted with acidified sodium chlorite, which delignifies the biomass. The aromatic region in the HSQC spectrum of the *irx8* chlorite extract revealed a dramatic loss of G lignin C/H-2, 5, 6 signals (Figure 5B) compared to that of the wild-type chlorite extract (Figure 5A). This result is consistent with the reduction in G lignin identified by phloroglucinol-HCl staining (Figure 4B) and pyMBMS (Figure 4I). Consistent with the complementation result (Supplemental Figures S1, S2), G lignin cross peaks were present in the HSQC spectrum in the chlorite extract of *irx8+GAUT12* complemented plants (Figure 5C). Only trace amounts of H and S lignin signals were identified in the chlorite extract across all three samples, indicating that either these monomers were lost during sample preparation or located in other wall extracts.

We thus used 2D ^{13}C - ^1H HSQC NMR spectroscopy to examine all other wall extracts from both wild-type and *irx8* stems, including ammonium oxalate-, sodium carbonate-, 1 M KOH-, 4 M KOH-, post-chlorite 4 M KOH (PC4MKOH) fractions, and residual pellets. Surprisingly, we found that both the 1 M and 4 M KOH extracts of wild type and *irx8* contained most of the H lignin, with the major H lignin signals located in the 1 M KOH extract (Figures 6A–D). The G lignin in the *irx8* mutant, although significantly reduced in amount, was found almost exclusively in the 1 M KOH extract (Figure 6B), while the G lignin in the wild type was located in both the chlorite- (Figure 5A) and PC4MKOH extracts (Figure 6E). These results reveal that the G lignin in the *irx8* mutant is more easily extracted than in the wild type. Only trace amounts of S lignin were found across all five wall extracts, indicating that S lignin was either composed of small molecules or degraded to small molecules during sequential extractions, and thus lost during sample dialysis. The aromatic regions of wild-type and *irx8* pectin-enriched fractions, i.e., ammonium oxalate- and sodium carbonate-extracts, showed comparable amounts of H lignin. The cellulose-enriched pellets did not show recognizable signals for major lignin structures (Supplemental Figure S7). Compared to wild type, the lignin aliphatic (side-chain) regions in the HSQC NMR spectra of the *irx8* chlorite extract suggested a reduction in the signals for β -O-4, β -5, and β - β linked lignin (Supplemental Figure S8B), which were calculated by density functional theory to be the major

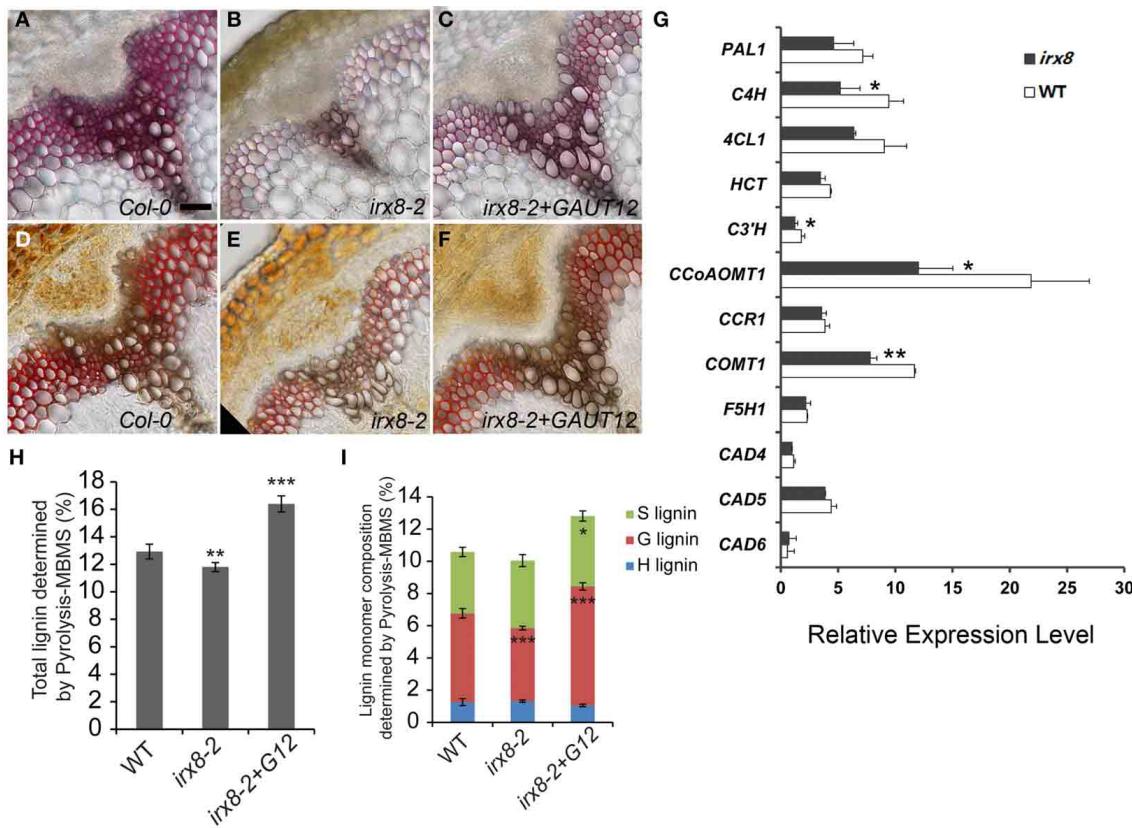


FIGURE 4 | Measurement of lignin and lignin biosynthetic gene expression in the *Irxa* mutant. Phloroglucinol-HCl staining of free-hand basal stem cross-sections of WT (A), *Irxa*-2 (B), and *Irxa*-2+GAUT12 (C). Mäule staining of free-hand basal stem cross-sections of WT (D), *Irxa*-2 (E), and *Irxa*-2+GAUT12 (F). Bar = 50 μ m for A–F. (G) Expression analyses in Arabidopsis WT and *Irxa* lower stems of lignin biosynthetic genes using Real-Time PCR. Genes labeled with asterisks have significantly lower expression in *Irxa*-2 compared to WT (* $p < 0.05$; ** $p < 0.01$). The relative expression level of each gene was normalized using *Actin2* as the reference gene and the expression of *C3'H* in wild-type basal stem was set to 1. Values are mean \pm standard deviation ($n = 3$). Lower-stem refers to the lower half of the inflorescence. PAL, phenylalanine ammonia lyase; C4H, trans-cinnamate

4-hydroxylase; 4CL, 4-coumarate: CoA ligase; HCT, hydroxycinnamoyl-CoA: shikimate/quinate hydroxycinnamoyltransferase; C3'H, 4-coumaroyl shikimate 3'-hydroxylase; CCoAOMT1, caffeoyl-CoA 3-O-methyltransferase; CCR, cinnamyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid/5-hydroxyferulic acid O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase; PER/LAC, peroxidases/laccases. (H) Total lignin (%) determined by pyrolysis molecular beam mass spectrometry (pyMBMS) of stem alcohol insoluble residues. (I) Lignin monomer composition determined by pyMBMS. Lignin content percentage (%) was corrected by equivalence to Klason lignin. The *Irxa*-2 and *Irxa*-2+GAUT12 values were compared to WT as determined by ANOVA and post-hoc t-tests with Bonferroni correction (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; $n = 6$) for (H,I).

and thermodynamically favorable linkages within lignin polymers (Sangha et al., 2012), and hence the most stable during acidic chlorite extractions.

GLYCOME PROFILING INDICATES ALTERED EXTRACTABILITY OF XYLAN, PECTIN, AND AG EPITOPE IN WALLS OF THE *Irxa* MUTANT

To examine cell wall glycan epitope changes associated with the loss of GAUT12 function and to correlate *Irxa* wall polysaccharide epitope changes with the lignin signal alterations observed by NMR, we performed glycome profiling analyses on wall extracts from wild type, *Irxa*, and GAUT12-complemented *Irxa* (*Irxa*+GAUT12) stems. The analyses correlated ELISA signals of different monoclonal antibody (mAb) groups with carbohydrates released in each fraction. Cell walls (i.e., AIR) prepared from *Irxa* showed major differences in their glycome profiles when compared with WT walls. These differences are highlighted by dotted blocks in Figure 7.

A pronounced difference was noted in the extractability of xylan epitopes in *Irxa*, compared to the wild type. Ammonium oxalate-, sodium carbonate-, and 1 M KOH-extracts prepared from *Irxa* walls contained significantly less xylan epitopes recognized by the xylan 4 through 7 groups of xylan-directed mAbs. Since less or similar amounts of carbohydrate mass was isolated in these extracts from *Irxa* walls compared to wild-type walls, there was significantly less xylan in the oxalate-, carbonate-, and 1 M KOH extracts in *Irxa* (Figure 7). The results indicate a loss of easily extractable xylan that is potentially associated with pectin in *Irxa*, since these wall extracts contain large amounts of pectin. Both the 4 M KOH- and PC4MKOH extracts of *Irxa*, however, displayed a marginal increase in binding of xylan mAbs compared to the corresponding WT extracts. The result that more xylan in the *Irxa* mutant was extracted under harsher conditions (i.e., with 4 M KOH and post-chlorite 4 M KOH) correlates with the previous finding that the 4 M KOH fractions from *Irxa* stems contain

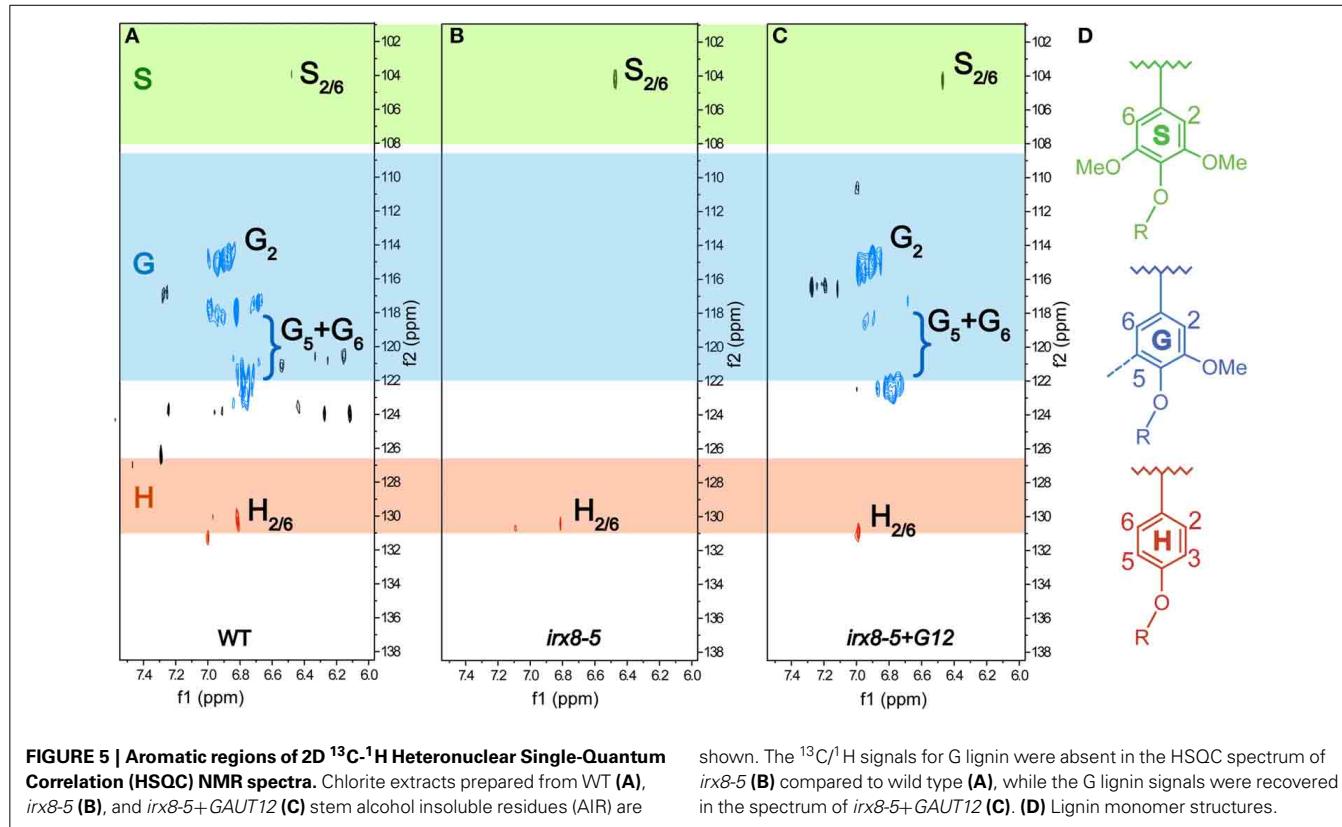


FIGURE 5 | Aromatic regions of 2D ¹³C-¹H Heteronuclear Single-Quantum Correlation (HSQC) NMR spectra. Chlorite extracts prepared from WT (A), *irx8-5* (B), and *irx8-5+GAUT12* (C) stem alcohol insoluble residues (AIR) are

shown. The ¹³C/¹H signals for G lignin were absent in the HSQC spectrum of *irx8-5* (B) compared to wild type (A), while the G lignin signals were recovered in the spectrum of *irx8-5+GAUT12* (C). (D) Lignin monomer structures.

some xylan of higher molecular weight than the WT counterparts (Brown et al., 2007; Peña et al., 2007).

Other differences in the glycome profile of the *irx8* mutant include a reduced presence of HG backbone epitopes (those recognized by HG Backbone-I group of mAbs, Figure 7) and of arabinogalactan (AG) epitopes in the oxalate and carbonate extracts (those recognized by AG-3 and AG-4 groups of mAbs, Figure 7), as well as a reduction in pectic arabinogalactan epitopes (recognized by RG-I/AG antibodies) and AG (those recognized by the AG-4 group of mAbs) in the 1 M KOH extract (Figure 7). These results indicate a concomitant loss of pectin, AG, and xylan in these wall extracts. The complemented line *irx8+GAUT12* shows a partial reversion of the glycome profile pattern to that of the WT, including an enhanced extractability of xylan, HG, and AG epitopes in the oxalate and carbonate extracts. The chlorite extracts of *irx8* showed significantly enhanced levels of hemicellulose epitopes including xylan and xyloglucan (Non-Fuc XG-1 to XG-6, Fuc-XG, and xylan-1/XG) compared to those of wild type and *irx8+GAUT12*. There was also a marginal increase in binding of antibodies against pectic arabinogalactan (RG-I/AG and AG2) in the *irx8* chlorite extract. This wall extract contained more mass in the *irx8* extract compared to the wild type, suggesting an increased weight ratio of carbohydrates to lignin in this extract (possibly caused by the reduction of lignin content described above; Figure 5B). An increase in pectic arabinogalactan (RG-I/AG and AG2) epitope content was also observed in the *irx8* PC4MKOH extract, which contained less mass compared to the wild type and *irx8+GAUT12*. These results indicate a shift in

the extractability of RG-I/AG epitopes in the *irx8* mutant that may compensate for the slight reduction of these epitopes in the oxalate- and carbonate fractions. Together, these results show that *irx8* has significant changes in the extractability of glycan epitopes, particularly of xylan, pectin, and some AG.

THE *irx8* MUTANT EXHIBITS INCREASED RG-I LABELING AND ALTERED XYLAN LOCALIZATION PATTERNS IN FIBER CELL WALLS

In addition to the significant reduction in xylan in *irx8* fiber cell walls based on reduced immunolabeling by anti-xylan antibodies LM10 and LM11 (Figure 8; Supplemental Figures S3F, S8), there was also an unexpected increase in immunolabeling by the antibody CCRC-M14, which recognizes an RG-I backbone epitope (Figure 8). CCRC-M14 binds to an empty triangle-shaped region in tricellular junctions of WT fiber cells (Figure 8). However, in 6–7-week-old *irx8* basal stems, this antibody also labels the inner layer of fiber cell walls in a dotted and lamellate pattern, suggesting possible increases in the amount or accessibility of RG-I-associated cell wall epitopes in *irx8* fiber cells. Consistent with this result, a slight increase in CCRC-M14 labeling was also observed in the glycome profiles in the 4 M KOH-, chlorite-, and PC4MKOH extracts of *irx8* compared to the WT and *irx8+GAUT12* counterparts (Figure 7, green arrows). It is worth mentioning that we have also observed the loss of the CCRC-M14 labeling pattern in fiber cells in 8-week-old (or older) WT and *irx8* basal stem sections, indicating that this phenotype may relate to a specific developmental stage. The GAUT12-EGFP construct complemented the CCRC-M14 phenotype in

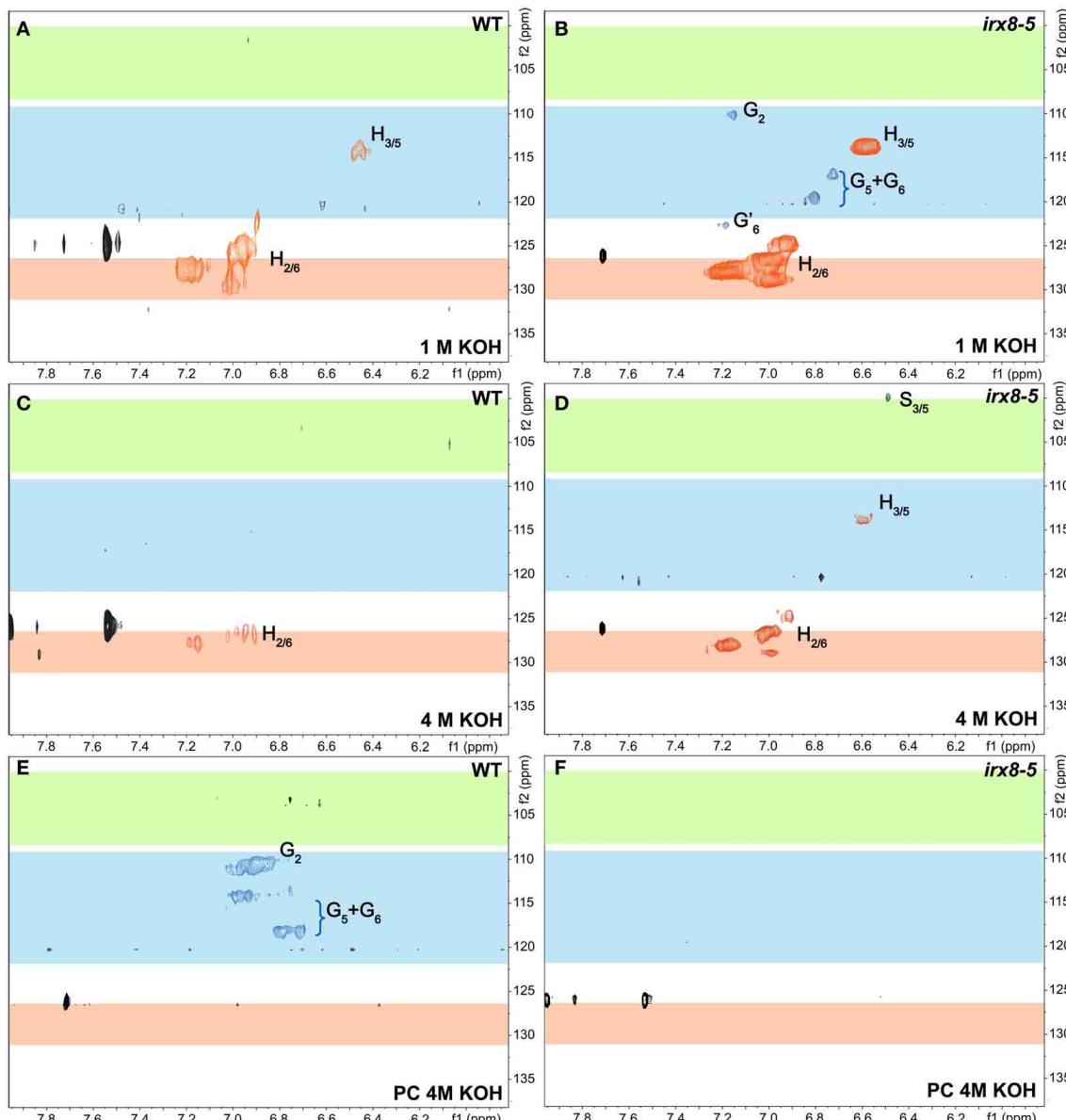


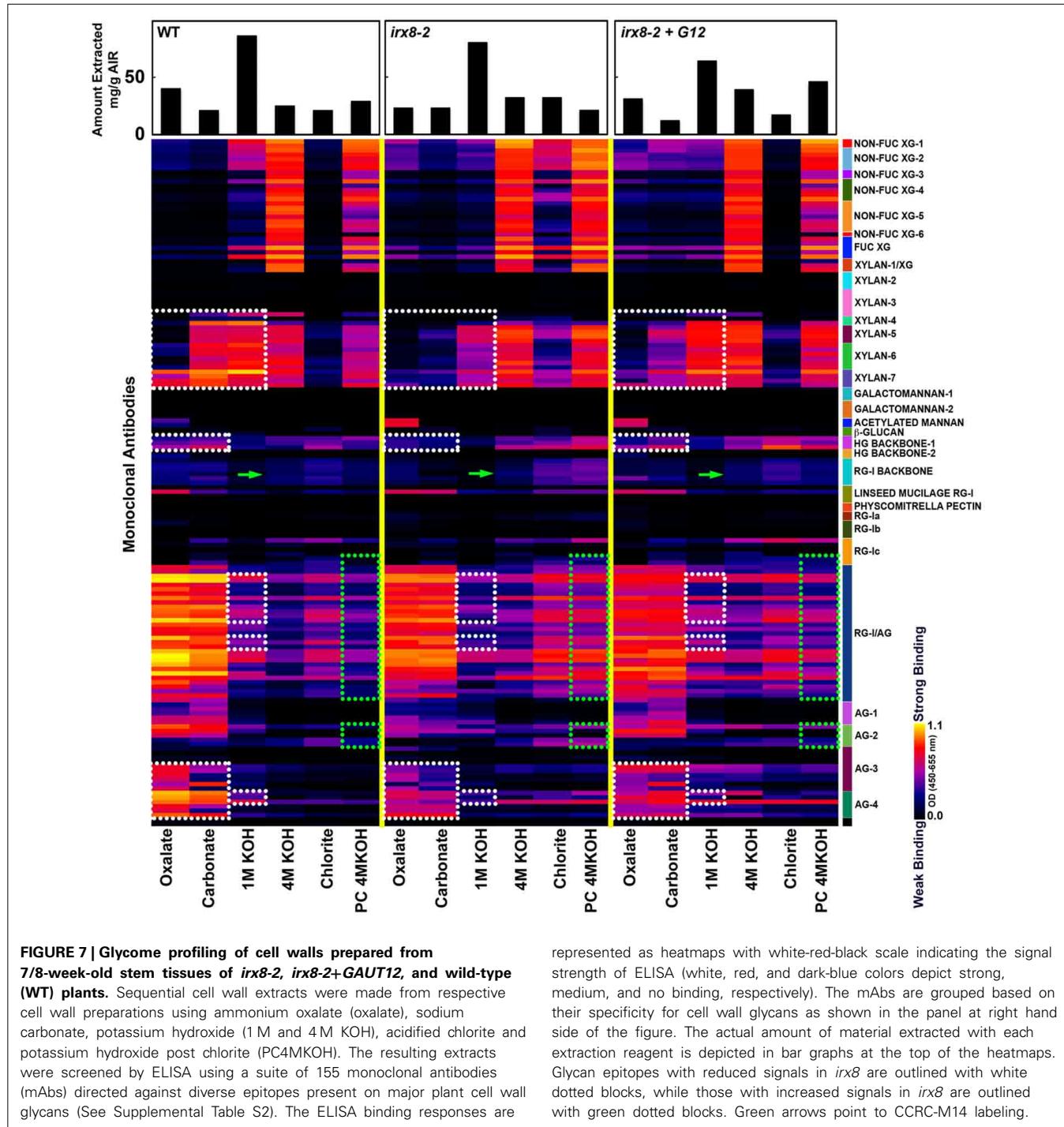
FIGURE 6 | Aromatic regions of 2D ^{13}C - ^1H Heteronuclear Single-Quantum Correlation (HSQC) NMR spectra of 1 M KOH, 4 M KOH, and post-chlorite-4 M KOH extracts of WT and *irx8-5* mutant walls. WT

(A) and *irx8-5* (B) 1 M KOH extracts. WT (C) and *irx8-5* (D) 4 M KOH extracts. WT (E) and *irx8-5* (F) post-chlorite 4 M KOH extracts. The signals of H, S, and G lignin monomers are as labeled.

irx8 fiber cells (Figure 8). Over-expression of *GAUT12* in WT (WT+*GAUT12*) plants, on the other hand, led to an occasional accumulation of CCRC-M14 reactive material in areas of the wall outside the triangular cell corner regions observed in the WT (Figure 8). There was, however, no obvious growth phenotype or altered sugar composition in WT vs. *GAUT12* over-expression lines (Supplemental Figures S1G, S3D,H, S4, S10).

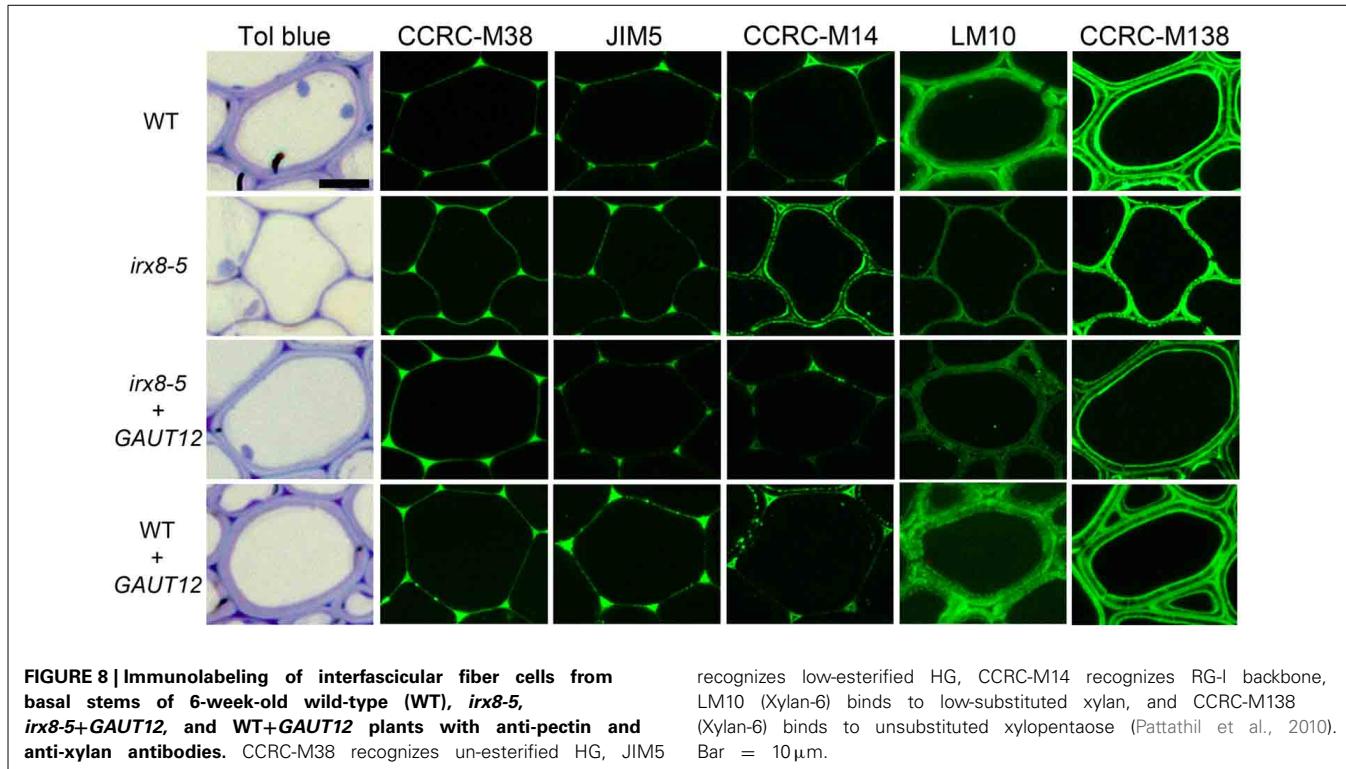
We applied a selection of six xylan-directed antibodies to basal stem sections of wild type, *irx8*, and *GAUT12*-transgenic lines (Supplemental Figure S9). Compared to WT fiber cells, there is a reduction in the labeling intensity in *irx8* fibers with LM10, CCRC-M137, and CCRC-M160. LM10 binds to low-substituted

xylan (McCartney et al., 2005), while both CCRC-M138 and CCRC-M160 bind to unsubstituted xylopentaose in ELISAs. The fiber walls of the *irx8+GAUT12* plant are much thicker than those of *irx8*, but are still slightly thinner than WT fiber walls and show lower labeling intensities with LM10, LM11, and CCRC-M137, a pattern resembling those of the *irx8* mutant labeled using these three antibodies (Supplemental Figure S9). Interestingly, we noticed a double-ring labeling pattern in WT fiber cell walls labeled with CCRC-M138, a monoclonal antibody that recognizes unsubstituted xylopentaose (Figure 8). In WT fiber cells, both an inner, plasma membrane-proximal wall domain and an outer, middle lamella-proximal wall domain



are solidly and continuously labeled with CCRC-M138, whereas xylan between these two rings, the middle layer, is mostly not labeled with this antibody. In *irxa* fibers the CCRC-M138-labeled double-ring, while still present as can be observed readily in cell corners, has collapsed and displays a discontinuous (dotted) pattern, which is clearly different from that of WT fiber walls. The reduced thickness or loss of the middle layer in the *irxa* mutant results in much thinner fiber cell walls. Overall, the GAUT12 construct complemented the CCRC-M138

phenotype in *irxa* fiber cells (Figure 8), although the labeling intensity in the labeled wall domains appears slightly uneven (Supplemental Figure S9). CCRC-M160 shows a very similar labeling pattern to CCRC-M138 in the GAUT12-complemented (*irxa*+GAUT12) line, albeit its double-ring pattern is less manifest in WT and WT+GAUT12 fibers (Supplemental Figure S9). The xylan deposition changes were also reflected in the altered xylan extraction patterns identified in *irxa* by glycome profiling (Figure 7).



IMMUNOABSORBED GAUT12 IS NOT AN HG:GALAT WITH CHARACTERISTICS COMPARABLE TO GAUT1

The increased CCRC-M14 labeling in *irxa* stems which suggested a change in RG-I, along with the previously reported reduction in a subfraction of HG (Persson et al., 2007), is consistent with the hypothesis that GAUT12 functions as an HG:GalAT required for secondary wall and xylan formation. To test this possibility, we measured HG:GalAT activity in microsomes from *irxa* stems which were shown to contain no GAUT12 protein (Figure 9B). The *irxa* microsomes contained 55% of the GAUT1-like HG:GalAT activity present in wild-type microsomes (Figure 9A). However, it remained unclear whether this reduction was directly due to the loss of GAUT12 activity or due to indirect effects on other enzymes, such as the HG:GalAT GAUT1, which is known to be expressed in Arabidopsis stems (Atmodjo et al., 2011). To directly test if GAUT12 has HG:GalAT activity, we generated a polyclonal anti-GAUT12 antibody that specifically recognizes GAUT12 and does not cross-react with GAUT1 or GAUT7 (Figures 9D,E; Supplemental Figure S11). LC-MS/MS was used to verify the specificity of the anti-GAUT12 antibody because the GAUT protein family contains 15 members with high sequence identity and similarity (Sterling et al., 2006). LC-MS/MS showed that peptides recovered from the protein immunoprecipitated by anti-GAUT12 are GAUT12-specific sequences and not those belonging to other GAUT proteins (Figure 9E). The antigen-purified anti-GAUT12 antibody was used to immunoabsorb-GAUT12 from detergent-permeabilized microsomes from wild-type stems (Figure 9D), and the immunoabsorbed GAUT12 was assayed for HG:GalAT activity. Neither the immunoabsorbed-GAUT12 from wild-type

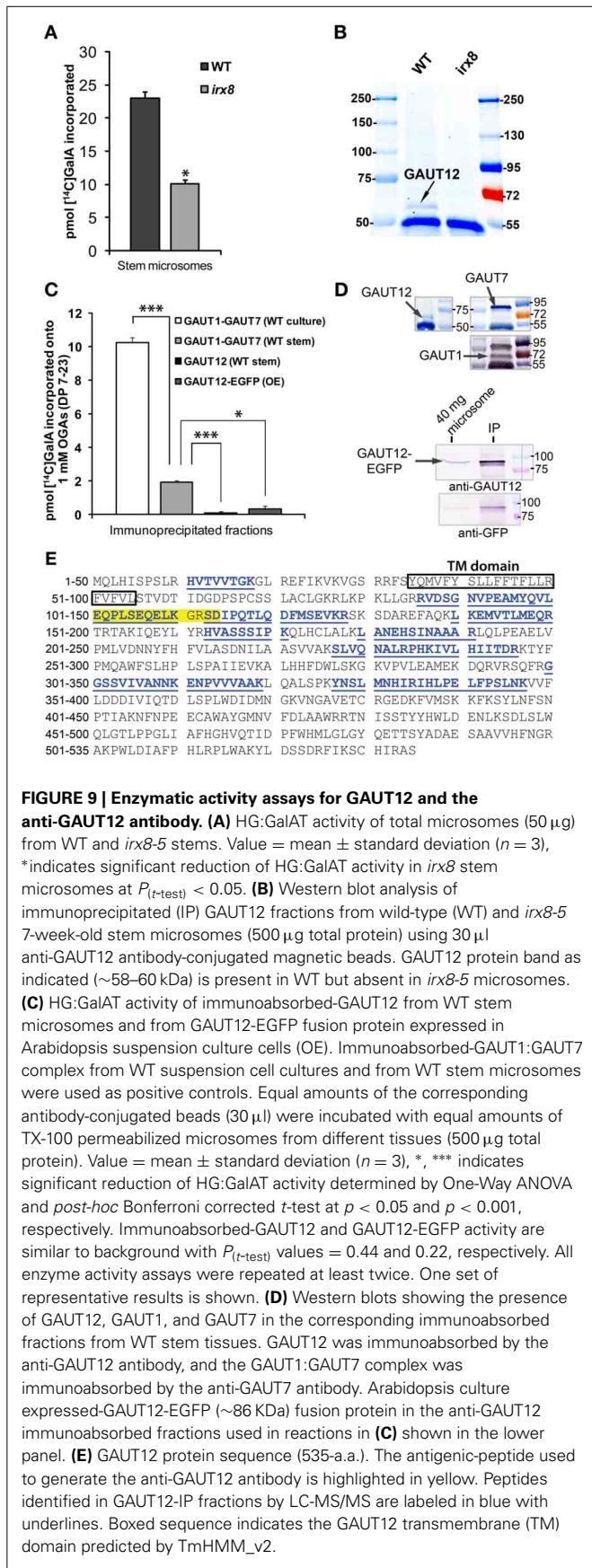
(WT) stem microsomes nor the immunoabsorbed-GAUT12-EGFP from Arabidopsis suspension culture cells over-expressing the GAUT12-EGFP fusion protein showed significant HG:GalAT activity (Figure 9C), although both proteins were confirmed to be present in these fractions by western blotting (Figure 9D). Thus, GAUT12 either does not have HG:GalAT activity, or its HG:GalAT activity is biochemically distinct from that of GAUT1 and cannot be assayed under the same reaction conditions.

DISCUSSION

GAUT12 IS REQUIRED FOR ANTER DEHISCENCE

Persson et al. (2007) have described the *irxa* mutants as semi-sterile (*irxa*-1 and *irxa*-2) having shorter anther filaments, less pollen than the wild type, and no seeds (Persson et al., 2007). Indeed, in our hands we were not able to recover any seeds from either *irxa*-2 or *irxa*-5 plants. Other dwarf mutants with collapsed xylem phenotypes, particularly *irx9* and *parvus*-3, have dehiscent anthers that release pollen (Supplemental Figure S4) and produce seeds under the same growth conditions. It is believed that *IRX9* is involved in xylan backbone elongation and GAUT12 together with *PARVUS/GATL1* are involved in XRES biosynthesis (Brown et al., 2007; Lee et al., 2007a,b; Peña et al., 2007). Our results suggest that the function of GAUT12 may, at least in the endothecium cell layer that is critical for anther dehiscence, be distinct from that of *IRX9* and *PARVUS/GATL1* in regards to xylan and lignin synthesis and deposition. Alternatively, anther dehiscence in both the *irx9* and *parvus*-3 mutants may be due to expression of functionally redundant genes in the endothecium cell layer.

The reduction of lignin in *irxa* anther cell walls (Figures 1E,I), together with the reduction in xylan as recognized by LM10 and



LM11 in endothecium cell walls (Figures 2F,G), contributes to a lack of secondary wall thickening in the *irx8* endothecium layer (Figure 2J). Consequently, relatively low tension during anther wall dehydration leads to indehiscent anthers in *irx8* mutants. The pollen grains produced by *irx8*, albeit smaller in size (Figure 3B), are viable upon manual release and able to fertilize both wild-type and *irx8* heterozygote pistils. Furthermore, manual fertilization of wild-type pistils with *irx8* heterozygote pollen demonstrated that *irx8* pollen has similar viability to wild-type pollen (Figure 3C). Thus, the function of GAUT12 is not essential for pollen viability and fertilization. Prior published transcriptomic and proteomic analyses of pollen and pollen tubes have not detected GAUT12 transcript or protein in these tissues (Wang et al., 2008; Zou et al., 2009). However, using quantitative RT-PCR we detected low GAUT12 expression in hydrated pollen grains and elongating pollen tubes (Supplemental Figure S6), suggesting a potential role of GAUT12 in pollen, perhaps associated with pollen size (Figure 3B).

LACK OF GAUT12 FUNCTION RESULTS IN REDUCED AMOUNTS AND ALTERED EXTRACTABILITY OF G LIGNIN IN ARABIDOPSIS STEM

We identified a reduction of lignin in the *irx8* mutant and used immunohistochemical staining (Figures 4B,E), pyMBMS (Figures 4H,I), and 2D $^{13}\text{C}-^1\text{H}$ HSQC NMR spectroscopy (Figure 5) to characterize this mutant phenotype in detail as well as study the possible connections between lignin and xylan deposition. In the semi-quantitative 2D $^{13}\text{C}-^1\text{H}$ HSQC NMR analyses, only trace amounts of H and S lignin were found in chlorite extracts from the WT, *irx8*-5, and *irx8*-5+GAUT12 (Figure 5). H lignin signals, however, were identified in the oxalate-, carbonate-, 1 M and 4 M KOH extracts of both WT and *irx8*, with the major signals located in the 1 M KOH extracts (Figure 6; Supplemental Figure S7). These results suggest that H lignin is present in the pectin and hemicellulose-enriched wall fractions. The presence of H lignin in pectin fractions is consistent with the observation that H lignin is deposited in middle lamella and cell wall corners (Donaldson, 2001) and hence co-extracted with pectin. The observation that a major portion of the H lignin signals were found in the 1 M KOH fraction suggests either that this portion of H lignin is directly or indirectly (e.g., via pectin) connected to the KOH-solubilized xylan or xyloglucan in these fractions, or that H lignin is connected through alkaline-labile ester linkages (Balakshin et al., 2011), and thus, extracted in alkaline buffers. Specific linkages between the hemicellulose/pectin/H lignin, however, remain to be determined. H lignin signals in *irx8* appeared to be slightly more prominent than in the WT, possibly due to the lower amount of xylan in the 1 M KOH extract of *irx8* (Figure 7), resulting in an increased lignin to xylan weight ratio. The total amount of H lignin in *irx8*, however, was similar to that of WT as determined by pyMBMS (Figure 4I). Thus, the results do not support a role for GAUT12 in producing a structure required for H lignin deposition.

It is estimated that the chlorite extractions conducted in this study could remove 50–80% of total lignin based on studies of the efficiency of acidic sodium chlorite treatment in the removal of lignin in black spruce, switchgrass, and poplar (Ahlgren and Goring, 1971; Kumar et al., 2013). Therefore, the lignin signals

observed in the chlorite- and PC4MKOH extracts by HSQC NMR were likely portions of the residual lignin (20–50%) recovered from each extraction at, or after, the sodium chlorite treatment. This may explain why we did not identify S lignin signals across all fractions, although it is unclear whether S lignin is composed of smaller molecules and depleted during sequential extractions and dialyses. The G lignin in WT was identified mostly in the chlorite- and PC4MKOH extracts (**Figures 5A, 6E**), indicating that a portion of G lignin is linked to wall polysaccharides via alkaline-resistant linkages, such as benzyl ethers and phenyl glucosides, and that only harsh conditions like acidic sodium chlorite are able to degrade these linkages. In *irx8*, however, the chlorite fraction is nearly depleted of G lignin signals (**Figure 5B**). Surprisingly, the bulk of the *irx8* G lignin signals, albeit much reduced in amount compared to the WT, was found in the 1 M KOH extract (**Figure 6B**), a wall fraction released prior to the acidic chlorite treatment, clearly indicating an altered G lignin extractability in the *irx8* mutant. It remains unknown, however, whether there is a portion of G lignin in the *irx8* mutant that does not withstand the acidic chlorite treatment and hence is removed during subsequent dialysis. Overall the results are consistent with the reduced phloroglucinol-HCl staining of *irx8* stem cross-sections (**Figure 4B**) and the reduced G lignin monomer content identified by pyMBMS (**Figure 4I**). Our results suggest that in the *irx8* mutant either (i) the G lignin linkages themselves or (ii) the polymers to which the G lignin is connected are partially alkaline-labile, and hence, more easily extracted from the walls in this mutant than in the WT.

INTERDEPENDENCE OF LIGNIN, PECTIN AND XYLAN IN WALL BIOGENESIS

The pleiotropic effects of *irx8* on xylan, lignin, and pectin during secondary wall formation confirms that normal wall biogenesis is dependent on an interaction between different cell wall polymers and suggests that there may be a requisite order in which they are deposited *in muro*. Lignification, which creates a relatively rigid and impermeable resin within the polysaccharide network of the wall, has been proposed to occur in two distinct stages. In the primary wall it appears that lignin deposition starts as early as during the formation of the middle lamella at the cell plate, followed by deposition at the cell wall corners—a sequence that has been taken to imply the existence of possible pectin initiation/nucleation sites for primary wall lignin deposition (Donaldson, 2001). Our data demonstrate that *Arabidopsis* interfascicular fiber tricellular junctions are filled with un-esterified to low-esterified HG, as recognized in a solid triangle shape by anti-HG antibodies CCRC-M38 and JIM5 (**Figure 8**). This tricellular junction triangle also contains a small amount of high-esterified HG as labeled by JIM7 (Supplemental Figure S9). RG-I backbone, as recognized by CCRC-M14, appears to be present toward the outer layer of this HG triangle, where labeling with this antibody is observed as an empty triangle at the tricellular junction (**Figure 8**). The lignin in this tricellular junction area was solidly stained for both G and S lignin in *irx8* fibers (**Figures 4B,E**), suggesting that this process is unlikely to have been affected by the mutation in *GAUT12*. In a second stage of lignification, secondary cell wall lignin is deposited in specific, terminally differentiated

cell types (Donaldson, 2001). We showed by colorimetric staining that both G and S lignin are reduced in the *irx8* xylem vessels and fiber cells, apparently due to the reduction in wall thicknesses in these cells. The G lignin in *irx8* is significantly reduced (**Figures 4I, 5B**), and released in the 1 M KOH extract (**Figure 6B**) rather than in the chlorite- and PC4MKOH extracts as occurs in the WT (**Figures 5A, 6E**), suggesting a possible correlation between the xylan reduction and lignin alteration in *irx8*. We also found a significant reduction in expression of four key lignin biosynthetic enzymes (C4H, C3'H, CCoAOMT1, and COMT1) that may lead to reduced lignin precursor production in *irx8* (**Figure 4G**). These results are consistent with a reduction in total lignin in *irx8* walls, in particular through a reduction in G lignin. Our data support the proposition that lignin formation is down-regulated in a xylan defective mutant, and that reduced *GAUT12* function affects both xylan and lignin deposition.

It has been reported that loss of the primary cellulose synthase subunit CES3 in the *eli1* mutant is associated with ectopic lignin deposition and increased defense responses (Cano-Delgado et al., 2003), which was interpreted as an effort by the cells to maintain cell wall integrity. By contrast, a decrease in lignin was observed in the *irx8* mutant, and reportedly also in the *irx7* and *irx9* mutants (Petersen et al., 2012). These observations show that loss of matrix polysaccharides causes a reduction in lignin in such *irx* mutants and suggest that matrix polysaccharides provide a different structural function than cellulose during lignification. The reduction of lignin in the *irx8* mutant may be a secondary effect due to the ~60% loss of xylan in *irx8*, since linkages between xylan and lignin have been reported. For example, 4-O-methylglucuronoxylan is the major carbohydrate linked to lignin in wood (Yuan et al., 2011) and xylan and lignin are linked via ferulate esters in maize (Grabber et al., 2000). The exact lignin-carbohydrate linkages in the *irx8* mutant are currently under investigation.

GAUT12 FUNCTION IS REQUIRED TO ESTABLISH A LAMELLATE STRUCTURE IN THE SECONDARY CELL WALL

We observed interesting lamellate-like patterns of xylan labeling in wild-type fiber cell walls upon immunolabeling with selected xylan-directed antibodies (**Figure 8**; Supplemental Figure S9). These results suggest that the deposition of different xylan structure may be spatially controlled. LM10 and LM11 bind to xylo-oligomers as small as a disaccharide, while CCRC-M149 binding requires an unsubstituted xylotriose structure, and CCRC-M138 and CCRC-M160 require unsubstituted xylopentaose for recognition (S. Pattathil, U. Avci, and M.G. Hahn, unpublished results). In wild-type fiber cells, LM10 labeling is more intense on the side of the fiber wall adjacent to adjoining cells, and the labeling decreases in intensity toward the lumen side of the wall. LM11 has an overall dotted labeling pattern that covers the entire secondary wall in fiber cells. CCRC-M137 and CCRC-M149 show comparable and relatively even labeling throughout fiber walls. All four antibodies had reduced labeling in *irx8* fiber cell walls due either to a reduced number of epitopes recognized by these antibodies or to reduced thickness of the wall. The fiber cells in *GAUT12*-complemented (*irx8+GAUT12*) plants, however, show a partial restoration of the WT labeling patterns with LM10, LM11, and

CCRC-M137. Although *irx8+GAUT12* fibers have thicker walls than those of *irx8*, the lower labeling intensity (i.e., epitope density) with LM10, LM11, and CCRC-M137 in *irx8+GAUT12* fibers resembles that of the *irx8* mutant (Supplemental Figure S9). In contrast, CCRC-M149 labels a thinner wall in *irx8* fibers with similar intensity (i.e., epitope density) in both *irx8* and *irx8+GAUT12* plants compared to WT plants, suggesting that the CCRC-M149-reactive xylan was perhaps indirectly affected by the *irx8* mutation. In other words, the reduction in CCRC-149 labeling in *irx8* fiber walls is likely due to reduced wall thickness and not reduced epitope density.

The most interesting labeling pattern was observed using antibodies CCRC-M138 and CCRC-M160, which showed double-ring labeling patterns in wild-type fiber cell walls (**Figure 8**; Supplemental Figure S9). The pattern suggests that the middle layer of wild-type fiber secondary walls between the labeled double rings contains xylan (as recognized by CCRC-M149) with fewer regions of unsubstituted xylan as recognized by CCRC-M138. The double-ring still exists in *irx8* fiber cell walls, as can be clearly seen at cell corners, but it has collapsed in *irx8* fiber cell walls due to reduced thickness of the middle layer material of the wall. In addition, the texture of the CCRC-M138 labeling in *irx8* fiber walls is different than in the WT, indicating a re-organized secondary wall with an altered configuration of xylan deposition in the *irx8* mutant. This was also reflected in the altered xylan extraction patterns in *irx8* observed by glycome profiling, revealing that significantly less xylan was extracted in the ammonium oxalate-, sodium carbonate-, and 1 M KOH extracts in the *irx8* mutant walls compared to WT walls (**Figure 7**). It is unclear why CCRC-M138 labeling is slightly uneven in the *GAUT12*-complemented line (Supplemental Figure S9). A possible explanation is that the 35S promoter drives an inconsistent expression of *GAUT12* leading to the uneven production of corresponding xylan epitopes.

The cell wall immunolabeling patterns in fiber cell cross sections obtained using the six monoclonal antibodies reactive against diverse xylan epitopes indicates that *GAUT12* is required for the formation of a middle xylan-containing layer located between an outer (middle lamella-proximal) and an inner (plasma membrane-proximal) xylan layer. These three layers are clearly seen in cross sections of WT fiber cells immunolabeled using antibodies LM11, CCRC-M138 and CCRC-M160. Immunolabeling of fiber cell cross sections of the *irx8* mutant reveals a loss of the middle xylan-containing layer and a collapse of the inner and outer xylan layers onto each other. Since several of the xylan-directed antibodies label the middle layer, as well as the outer and inner xylan layers (e.g., CCRC-M137 and CCRC-M149), it is clear that the middle layer contains xylan. Thus, the lack of the middle layer, and the retention of the outer and inner xylan layers, albeit in a discontinuous pattern in the *irx8* mutants, indicate that *GAUT12* is required to produce a WT xylan architecture and that, in the absence of *GAUT12*, this architecture is either not made or collapses. The current results do not clarify whether the reduction in xylan in the *irx8* mutant is due to the inability of the plant to make a subfraction of xylan that requires *GAUT12* for synthesis, or rather, whether the absence of *GAUT12* results in an altered xylan architecture that leads to an accumulation of

surplus xylan which acts as a negative signal to down-regulate xylan synthesis.

GAUT12 FUNCTION

How may the complex phenotype of *irx8* mutants be explained? The reduction in α -1,4-linked GalA in the endopolygalacturonase/pectin methylesterase (EPG/PME)-accessible fraction of *irx8* mutant walls raises the possibility that *GAUT12* is involved in the synthesis of an as-yet-to-be-determined HG species. Such a *GAUT12*-dependent HG may be tightly associated with xylan and, when missing, disrupt xylan biogenesis and/or deposition (Persson et al., 2007). Such *GAUT12*-dependent pectin and/or xylan may further be a foundation upon which G lignin accumulates.

Recently, Tan et al. (2013) discovered a novel cell wall structure isolated from *Arabidopsis* suspension cell cultures named ARABINOXYLAN PECTIN ARABINOGLACTAN PROTEIN1 (APAP1), which contains a unique proteoglycan with xylan connected to stretches of RG-I that are flanked by short HG oligomers (Tan et al., 2013). Although we show in the present study that *GAUT12* does not appear to have *GAUT1*-like HG:GalAT activity, i.e., *GAUT12* does not catalyze the addition of GalA from UDP- α -GalA onto oligogalacturonide (DP 7–23) acceptors (**Figure 9C**), a possible function for *GAUT12* is the synthesis of such an HG sub-domain using the RG-I sub-domain (oligomer) as a primer for synthesis of an APAP1-like structure in fiber walls. This hypothesis is consistent with the significant decreases in Xyl and GalA in *irx8* stem cell walls (Supplemental Figure S2C) and with the increased immunolabeling of fiber walls with the RG-I backbone-reactive antibody CCRC-M14 (**Figure 8**). Although there is no obvious change in the wall glycosyl residue composition when *GAUT12* is over-expressed in the WT background (Supplemental Figure S10E), *GAUT12* over-expression may lead to increased production of RG-I/HG sub-domains which may overload the capacity of the xylan biosynthesis machinery, resulting in accumulation of the CCRC-M14 reactive material at the plasma membrane as observed in the WT+*GAUT12* fiber cell walls (**Figure 8**).

Alternatively, the reduction in the *irx8* mutant of the xylan reducing end sequence [XRES, β -D-Xylp-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 2)- α -D-GalpA-(1 \rightarrow 4)-D-Xylp] led to speculations that *GAUT12* may catalyze the addition of GalA into the nascent XRES (Peña et al., 2007). However, it remains unknown how XRES synthesis is initiated or whether it acts as a primer or terminator during xylan biosynthesis (York and O'Neill, 2008). *GAUT12* appears not to have any functional homologs in graminaceous monocots (Caffall et al., 2009; Yin et al., 2010). The absence of XRES in grass species examined to date (Kulkarni et al., 2012) would be consistent with a function of *GAUT12* in the synthesis of this structure. So far, however, we were not able to demonstrate that *GAUT12* adds the GalA into the XRES. Further study is required to identify *GAUT12* enzymatic function, as well as its role in the carbohydrate-lignin connection.

In summary, we demonstrate that *GAUT12* has a role in anther dehiscence and affects the amount of G lignin and its connectivity to xylan in *Arabidopsis*. Our work shows that a

mutation in a single glycosyltransferase leads to alterations in xylan, pectin, and lignin, thereby providing further evidence for possible associations or connections between different wall polymers. Although the catalytic activity of GAUT12 remains to be determined, we have shown that GAUT12 is not an HG:GalAT with substrate specificities comparable to GAUT1 and that the product of GAUT12 may connect to a structure that contains RG-I and that is required for native xylan architecture in the secondary cell wall.

ACKNOWLEDGMENTS

We thank Maor Bar-Peled for the 35tlegfps2#4 plasmid as the base for the *GAUT12-EGFP* (*GAUT12*) constructs. This research was supported by the Department of Energy Center Grant DOE DE-FG02-09ER20097 and BioEnergy Science Center Grant DE-AC05-00OR22725. The BioEnergy Science Center is a US Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the Department of Energy Office of Science. The generation of the CCRC series of glycan-directed monoclonal antibodies used in this work was supported by the National Science Foundation Plant Genome Program (Awards DBI-0421683 and IOS-0923992 to Michael G. Hahn).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00357/abstract>

REFERENCES

- Ahlgren, P. A., and Goring, D. A. I. (1971). Removal of wood components during chlorite delignification of black spruce. *Can. J. Chem.* 49, 1272–1275. doi: 10.1139/V71-207
- Albersheim, P., Nevis, D. J., English, P. D., and Karr, A. (1967). A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. *Carbohydr. Res.* 5, 340–345.
- Ariizumi, T., and Toriyama, K. (2011). Genetic regulation of sporopollenin synthesis and pollen exine development. *Annu. Rev. Plant Biol.* 62, 437–460. doi: 10.1146/annurev-aplant-042809-112312
- Arioli, T., Peng, L., Betzner, A. S., Burn, J., Wittke, W., Herth, W., et al. (1998). Molecular analysis of cellulose biosynthesis in *Arabidopsis*. *Science* 279, 717–720. doi: 10.1126/science.279.5351.717
- Atmodjo, M. A., Sakuragi, Y., Zhu, X., Burrell, A. J., Mohanty, S. S., Atwood, J. A. 3rd, et al. (2011). Galacturonosyltransferase (GAUT)1 and GAUT are the core of a plant cell wall pectin biosynthetic homogalacturonan:galacturonosyltransferase complex. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20225–20230. doi: 10.1073/pnas.1112816108
- Avci, U., Pattathil, S., and Hahn, M. G. (2012). Immunological approaches to plant cell wall and biomass characterization: immunolocalization of glycan epitopes. *Methods Mol. Biol.* 908, 73–82. doi: 10.1007/978-1-61779-956-3_7
- Balakshin, M., Capanema, E., Gracz, H., Chang, H. M., and Jameel, H. (2011). Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta* 233, 1097–1110. doi: 10.1007/s00425-011-1359-2
- Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.aplant.54.031902.134938
- Bouton, S., Leboeuf, E., Mouille, G., Leydecker, M. T., Talbotec, J., Granier, F., et al. (2002). QUASIMODO1 encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in *Arabidopsis*. *Plant Cell* 14, 2577–2590. doi: 10.1105/tpc.004259
- Bowman, J. (1994). *Arabidopsis: an Atlas of Morphology and Development*. London: Springer, Limited, 2011.
- Brown, D. M., Goubet, F., Wong, V. W., Goodacre, R., Stephens, E., Dupree, P., et al. (2007). Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J.* 52, 1154–1168. doi: 10.1111/j.1365-313X.2007.03307.x
- Brown, D. M., Zeef, L. A., Ellis, J., Goodacre, R., and Turner, S. R. (2005). Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17, 2281–2295. doi: 10.1105/tpc.105.031542
- Caffall, K. H., Pattathil, S., Phillips, S. E., Hahn, M. G., and Mohnen, D. (2009). *Arabidopsis thaliana* T-DNA mutants implicate GAUT genes in the biosynthesis of pectin and xylan in cell walls and seed testa. *Mol. Plant* 2, 1000–1014. doi: 10.1093/mp/ssp062
- Cano-Delgado, A., Penfield, S., Smith, C., Catley, M., and Bevan, M. (2003). Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. *Plant J.* 34, 351–362. doi: 10.1046/j.1365-313X.2003.01729.x
- Chapple, C. C., Vogt, T., Ellis, B. E., and Somerville, C. R. (1992). An *Arabidopsis* mutant defective in the general phenylpropanoid pathway. *Plant Cell* 4, 1413–1424. doi: 10.1105/tpc.4.11.1413
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313X.1998.00343.x
- Dardelle, F., Lehner, A., Ramdani, Y., Bardor, M., Lerouge, P., Driouich, A., et al. (2010). Biochemical and immunocytological characterizations of *Arabidopsis* pollen tube cell wall. *Plant Physiol.* 153, 1563–1576. doi: 10.1104/pp.110.158881
- Demartini, J. D., Pattathil, S., Avci, U., Szekalski, K., Mazumder, K., Hahn, M. G., et al. (2011). Application of monoclonal antibodies to investigate plant cell wall deconstruction for biofuels production. *Energy Environ. Sci.* 4, 4332–4339. doi: 10.1039/c1ee02112e
- Doelling, J. H., and Pikaard, C. S. (1993). Transient expression in *Arabidopsis thaliana* protoplasts derived from rapidly established cell-suspension cultures. *Plant Cell Rep.* 12, 241–244. doi: 10.1007/BF00237127
- Donaldson, L. A. (2001). Lignification and lignin topochemistry—an ultrastructural view. *Phytochemistry* 57, 859–873. doi: 10.1016/S0031-9422(01)00049-8
- Donaldson, L. A., and Knox, J. P. (2012). Localization of cell wall polysaccharides in normal and compression wood of radiata pine: relationships with lignification and microfibril orientation. *Plant Physiol.* 158, 642–653. doi: 10.1104/pp.111.184036
- Doong, R. L., and Mohnen, D. (1998). Solubilization and characterization of a galacturonosyltransferase that synthesizes the pectic polysaccharide homogalacturonan. *Plant J.* 13, 363–374. doi: 10.1046/j.1365-313X.1998.00042.x
- Evans, R. J., and Milne, T. A. (1987). Molecular characterization of the pyrolysis of biomass I. fundamentals. *Energy Fuels* 1, 123–137. doi: 10.1021/Ef0002a001
- Grabber, J. H., Ralph, J., and Hatfield, R. D. (2000). Cross-linking of maize walls by ferulate dimerization and incorporation into lignin. *J. Agric. Food Chem.* 48, 6106–6113. doi: 10.1021/jf0006978
- Guo, D., Chen, F., Inoue, K., Blount, J. W., and Dixon, R. A. (2001). Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* 13, 73–88. doi: 10.1105/tpc.13.1.73
- Jeffries, T. W. (1990). Biodegradation of lignin-carbohydrate complexes. *Biodegradation* 1, 163–176. doi: 10.1007/bf00058834
- Jiang, N., Pu, Y. Q., Samuel, R., and Ragauskas, A. J. (2009). Perdeuterated pyridinium molten salt (ionic liquid) for direct dissolution and NMR analysis of plant cell walls. *Green Chem.* 11, 1762–1766. doi: 10.1039/b913609f
- Jung, K. W., Oh, S. I., Kim, Y. Y., Yoo, K. S., Cui, M. H., and Shin, J. S. (2008). *Arabidopsis* histidine-containing phosphotransfer factor 4 (AHP4) negatively regulates secondary wall thickening of the anther endothecium during flowering. *Mol. Cells* 25, 294–300.
- Kim, H., and Ralph, J. (2010). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d₆/pyridine-d₅. *Org. Biomol. Chem.* 8, 576–591. doi: 10.1039/b916070a
- Ko, J. H., Kim, W. C., and Han, K. H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in *Arabidopsis*. *Plant J.* 60, 649–665. doi: 10.1111/j.1365-313X.2009.03989.x
- Kong, Y., Zhou, G., Avci, U., Gu, X., Jones, C., Yin, Y., et al. (2009). Two poplar glycosyltransferase genes, *PdGATL1.1* and *PdGATL1.2*, are functional

- orthologs to *PARVUS/AtGATL1* in *Arabidopsis*. *Mol. Plant* 2, 1040–1050. doi: 10.1093/mp/ssp068
- Kulkarni, A. R., Peña, M. J., Avci, U., Mazumder, K., Urbanowicz, B. R., Pattathil, S., et al. (2012). The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. *Glycobiology* 22, 439–451. doi: 10.1093/glycob/cwr117
- Kumar, R., Hu, F., Hubbell, C. A., Ragauskas, A. J., and Wyman, C. E. (2013). Comparison of laboratory delignification methods, their selectivity, and impacts on physiochemical characteristics of cellulosic biomass. *Bioresour. Technol.* 130, 372–381. doi: 10.1016/j.biortech.2012.12.028
- Lao, N. T., Long, D., Kiang, S., Coupland, G., Shoue, D. A., Carpita, N. C., et al. (2003). Mutation of a family 8 glycosyltransferase gene alters cell wall carbohydrate composition and causes a humidity-sensitive semi-sterile dwarf phenotype in *Arabidopsis*. *Plant Mol. Biol.* 53, 647–661. doi: 10.1023/B:PLAN.0000019074.60542.6c
- Leboeuf, E., Guillou, F., Thoiron, S., and Lahaye, M. (2005). Biochemical and immunohistochemical analysis of pectic polysaccharides in the cell walls of *Arabidopsis* mutant *QUASIMODO 1* suspension-cultured cells: implications for cell adhesion. *J. Exp. Bot.* 56, 3171–3182. doi: 10.1093/jxb/eri314
- Lee, C., O'Neill, M. A., Tsumuraya, Y., Darvill, A. G., and Ye, Z. H. (2007a). The *irregular xylem9* mutant is deficient in xylan xylosyltransferase activity. *Plant Cell Physiol.* 48, 1624–1634. doi: 10.1093/pcp/pcm135
- Lee, C., Zhong, R., Richardson, E. A., Himmelsbach, D. S., McPhail, B. T., and Ye, Z. H. (2007b). The *PARVUS* gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant Cell Physiol.* 48, 1659–1672. doi: 10.1093/pcp/pcm155
- Lewis, N. G., and Yamamoto, E. (1990). Lignin: occurrence, biogenesis and biodegradation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41, 455–496. doi: 10.1146/annurev.pp.41.060190.002323
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- McCarthy, R. L., Zhong, R., and Ye, Z. H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell Physiol.* 50, 1950–1964. doi: 10.1093/pcp/pcp139
- McCartney, L., Marcus, S. E., and Knox, J. P. (2005). Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *J. Histochem. Cytochem.* 53, 543–546. doi: 10.1369/jhc.4B6578.2005
- Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17, 2993–3006. doi: 10.1105/tpc.105.036004
- Orfila, C., Sorensen, S. O., Harholt, J., Geshi, N., Crombie, H., Truong, H. N., et al. (2005). *QUASIMODO1* is expressed in vascular tissue of *Arabidopsis thaliana* inflorescence stems, and affects homogalacturonan and xylan biosynthesis. *Planta* 222, 613–622. doi: 10.1007/s00425-005-0008-z
- Pattathil, S., Avci, U., Baldwin, D., Swennes, A. G., McGill, J. A., Popper, Z., et al. (2010). A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* 153, 514–525. doi: 10.1104/pp.109.151985
- Pattathil, S., Avci, U., Miller, J. S., and Hahn, M. G. (2012). Immunological approaches to plant cell wall and biomass characterization: glycane profiling. *Methods Mol. Biol.* 908, 61–72. doi: 10.1007/978-1-61779-956-3_6
- Pattathil, S., Harper, A. D., and Bar-Peled, M. (2005). Biosynthesis of UDP-xylene: characterization of membrane-bound *AtUxs2*. *Planta* 221, 538–548. doi: 10.1007/s00425-004-1471-7
- Pedrioli, P. G., Eng, J. K., Hubley, R., Vogelzang, M., Deutsch, E. W., Raught, B., et al. (2004). A common open representation of mass spectrometry data and its application to proteomics research. *Nat. Biotechnol.* 22, 1459–1466. doi: 10.1038/nbt1031
- Peña, M. J., Zhong, R., Zhou, G. K., Richardson, E. A., O'neill, M. A., Darvill, A. G., et al. (2007). *Arabidopsis irregular xylem8* and *irregular xylem9*: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19, 549–563. doi: 10.1105/tpc.106.049320
- Persson, S., Caffall, K. H., Freshour, G., Hilley, M. T., Bauer, S., Poindexter, P., et al. (2007). The *Arabidopsis irregular xylem8* mutant is deficient in glucuronoxylan and homogalacturonan, which are essential for secondary cell wall integrity. *Plant Cell* 19, 237–255. doi: 10.1105/tpc.106.047720
- Persson, S., Wei, H., Milne, J., Page, G. P., and Somerville, C. R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8633–8638. doi: 10.1073/pnas.0503392102
- Petersen, P. D., Lau, J., Ebert, B., Yang, F., Verhertbruggen, Y., Kim, J. S., et al. (2012). Engineering of plants with improved properties as biofuels feedstocks by vessel-specific complementation of xylan biosynthesis mutants. *Biotechnol. Biofuels* 5:84. doi: 10.1186/1754-6834-5-84
- Raes, J., Rohde, A., Christensen, J. H., Van De Peer, Y., and Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol.* 133, 1051–1071. doi: 10.1104/pp.103.026484
- Sangha, A. K., Parks, J. M., Standaert, R. F., Ziebell, A., Davis, M., and Smith, J. C. (2012). Radical coupling reactions in lignin synthesis: a density functional theory study. *J. Phys. Chem. B* 116, 4760–4768. doi: 10.1021/JP2122449
- Scheller, H. V., and Ulvsborg, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-aplant-042809-112315
- Schilmiller, A. L., Stout, J., Weng, J. K., Humphreys, J., Ruegger, M. O., and Chapple, C. (2009). Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in *Arabidopsis*. *Plant J.* 60, 771–782. doi: 10.1111/j.1365-313X.2009.03996.x
- Shao, M., Zheng, H., Hu, Y., Liu, D., Jang, J. C., Ma, H., et al. (2004). The *GAOLAOZHUANGREN1* gene encodes a putative glycosyltransferase that is critical for normal development and carbohydrate metabolism. *Plant Cell Physiol.* 45, 1453–1460. doi: 10.1093/pcp/pch168
- Steiner-Lange, S., Unte, U. S., Eckstein, L., Yang, C., Wilson, Z. A., Schmelzer, E., et al. (2003). Disruption of *Arabidopsis thaliana* *MYB26* results in male sterility due to non-dehiscent anthers. *Plant J.* 34, 519–528. doi: 10.1046/j.1365-313X.2003.01745.x
- Sterling, J. D., Atmodjo, M. A., Inwood, S. E., Kumar Kolli, V. S., Quigley, H. F., Hahn, M. G., et al. (2006). Functional identification of an *Arabidopsis* pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5236–5241. doi: 10.1073/pnas.0600120103
- Sterling, J. D., Lemons, J. A., Forkner, I. F., and Mohnen, D. (2005). Development of a filter assay for measuring homogalacturonan: α -(1,4)-galacturonosyltransferase activity. *Anal. Biochem.* 343, 231–236. doi: 10.1016/j.ab.2005.05.037
- Sykes, R., Kodrzycki, B., Tuskan, G., Foutz, K., and Davis, M. (2008). Within tree variability of lignin composition in *Populus*. *Wood Sci. Technol.* 42, 649–661. doi: 10.1007/s00226-008-0199-0
- Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., et al. (2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25, 270–287. doi: 10.1105/tpc.112.107334
- Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., and Turner, S. R. (2003). Interactions among three distinct CesA proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1450–1455. doi: 10.1073/pnas.0337628100
- Thevenin, J., Pollet, B., Letarne, B., Saulnier, L., Gissot, L., Maia-Grondard, A., et al. (2011). The simultaneous repression of CCR and CAD, two enzymes of the lignin biosynthetic pathway, results in sterility and dwarfism in *Arabidopsis thaliana*. *Mol. Plant* 4, 70–82. doi: 10.1093/mp/qqq045
- Wang, Y., Zhang, W. Z., Song, L. F., Zou, J. J., Su, Z., and Wu, W. H. (2008). Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in *Arabidopsis*. *Plant Physiol.* 148, 1201–1211. doi: 10.1104/pp.108.126375
- Weatherly, D. B., Atwood, J. A. 3rd, Minning, T. A., Cavola, C., Tarleton, R. L., and Orlando, R. (2005). A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results. *Mol. Cell. Proteomics* 4, 762–772. doi: 10.1074/mcp.M400215-MCP200
- Weng, J. K., Mo, H., and Chapple, C. (2010). Over-expression of F5H in COMT-deficient *Arabidopsis* leads to enrichment of an unusual lignin and disruption of pollen wall formation. *Plant J.* 64, 898–911. doi: 10.1111/j.1365-313X.2010.04391.x
- Wilson, Z. A., Song, J., Taylor, B., and Yang, C. (2011). The final split: the regulation of anther dehiscence. *J. Exp. Bot.* 62, 1633–1649. doi: 10.1093/jxb/err014
- Yamaguchi, M., Goue, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J. C., et al. (2010). VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem

- vessel elements under control of an induction system. *Plant Physiol.* 153, 906–914. doi: 10.1104/pp.110.154013
- Yang, C., Xu, Z., Song, J., Conner, K., Vizcay Barrena, G., and Wilson, Z. A. (2007). *Arabidopsis MYB26/MALE STERILE35* regulates secondary thickening in the endothecium and is essential for anther dehiscence. *Plant Cell* 19, 534–548. doi: 10.1105/tpc.106.046391
- Yin, Y., Chen, H., Hahn, M. G., Mohnen, D., and Xu, Y. (2010). Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. *Plant Physiol.* 153, 1729–1746. doi: 10.1104/pp.110.154229
- York, W. S., and O'Neill, M. A. (2008). Biochemical control of xylan biosynthesis—which end is up? *Curr. Opin. Plant Biol.* 11, 258–265. doi: 10.1016/j.pbi.2008.02.007
- Yuan, T. Q., Sun, S. N., Xu, F., and Sun, R. C. (2011). Characterization of lignin structures and lignin-carbohydrate complex (LCC) linkages by quantitative ¹³C and 2D HSQC NMR spectroscopy. *J. Agric. Food Chem.* 59, 10604–10614. doi: 10.1021/jf2031549
- Zhong, R., Peña, M. J., Zhou, G. K., Nairn, C. J., Wood-Jones, A., Richardson, E. A., et al. (2005). *Arabidopsis Fragile Fiber8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. *Plant Cell* 17, 3390–3408. doi: 10.1105/tpc.105.035501
- Zou, J., Song, L., Zhang, W., Wang, Y., Ruan, S., and Wu, W. H. (2009). Comparative proteomic analysis of *Arabidopsis* mature pollen and germinated pollen. *J. Integr. Plant Biol.* 51, 438–455. doi: 10.1111/j.1744-7909.2009.00823.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 May 2014; paper pending published: 20 June 2014; accepted: 03 July 2014; published online: 28 July 2014.

Citation: Hao Z, Avci U, Tan L, Zhu X, Glushka J, Pattathil S, Eberhard S, Sholes T, Rothstein GE, Lukowitz W, Orlando R, Hahn MG and Mohnen D (2014) Loss of *Arabidopsis GAUT12/IRX8* causes anther indehiscence and leads to reduced G lignin associated with altered matrix polysaccharide deposition. *Front. Plant Sci.* 5:357. doi: 10.3389/fpls.2014.00357

This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Hao, Avci, Tan, Zhu, Glushka, Pattathil, Eberhard, Sholes, Rothstein, Lukowitz, Orlando, Hahn and Mohnen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Lectin domains at the frontiers of plant defense

Nausicaä Lannoo and Els J. M. Van Damme*

Laboratory of Biochemistry and Glycobiology, Department of Molecular Biotechnology, Ghent University, Ghent, Belgium

Edited by:

Elisabeth Jamet, Laboratoire de Recherche en Sciences Végétales, France

Reviewed by:

Raimund Tenhaken, University of Salzburg, Austria

Thomas Sebastian Nuhse, The University of Manchester, UK

***Correspondence:**

Els J. M. Van Damme, Laboratory of Biochemistry and Glycobiology, Department of Molecular Biotechnology, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium

e-mail: elsjm.vandamme@ugent.be

Plants are under constant attack from pathogens and herbivorous insects. To protect and defend themselves, plants evolved a multi-layered surveillance system, known as the innate immune system. Plants sense their encounters upon perception of conserved microbial structures and damage-associated patterns using cell-surface and intracellular immune receptors. Plant lectins and proteins with one or more lectin domains represent a major part of these receptors. The whole group of plant lectins comprises an elaborate collection of proteins capable of recognizing and interacting with specific carbohydrate structures, either originating from the invading organisms or from damaged plant cell wall structures. Due to the vast diversity in protein structures, carbohydrate recognition domains and glycan binding specificities, plant lectins constitute a very diverse protein superfamily. In the last decade, new types of nucleocytoplasmic plant lectins have been identified and characterized, in particular lectins expressed inside the nucleus and the cytoplasm of plant cells often as part of a specific plant response upon exposure to different stress factors or changing environmental conditions. In this review, we provide an overview on plant lectin motifs used in the constant battle against pathogens and predators during plant defenses.

Keywords: carbohydrate, innate immunity, lectin, protein–carbohydrate interaction, PRR

INTRODUCTION

In nature, plants are constantly exposed to a plethora of different pathogens including bacteria, viruses, and fungi. Whereas the interaction with some of these organisms can be beneficial, most microbial infection is harmful for the plant (Dangl et al., 2013). In order to resist pathogen colonization, plants developed a highly sophisticated, multilayered system enabling the plant to recognize invading pathogens and mount rapidly efficient defense responses (Muthamilarasan and Prasad, 2013; Wirthmueller et al., 2013).

Microbial entry into the host tissue is a critical step in causing infection. Pathogens can enter plants through natural openings such as stomata, hydathodes, lateral roots, or through accidental wounds, but can also form specialized structures such as haustoria to penetrate directly into the plant surface (Melotto et al., 2006; Gudesblat et al., 2009). Many phytopathogens also produce lytic enzymes to damage the plant cell wall in favor of pathogen invasion. Perception of the invading pathogen is the first step in the plant's defense and is governed by cell surface transmembrane receptors. These pattern recognition receptors or PRRs are able to recognize two types of molecules.

The first group encompasses the damage-associated molecular patterns (DAMPs) which are produced in the plant apoplast as a consequence of pathogen entry. Examples include cell wall fragments such as oligogalacturonides and cellulose fragments, cutin monomers, and peptides such as systemin, defensins, and phytosulfokines (Ryan, 2000; Nühse, 2012; Albert, 2013). PRRs are also able to recognize conserved microbial structures, known as pathogen- or microbe-associated molecular patterns (PAMPs/MAMPs), which are essential for the microbial physiology and the pathogen's fitness (Newman et al., 2013; Wirthmueller et al., 2013). Examples of PAMPs/MAMPs include

lipopolysaccharides (LPS) of Gram-negative bacteria, peptidoglycan (PGN) of Gram-positive bacteria, bacterial flagellins, eubacterial elongation factors (EF-Tu), and fungal cell wall derived glucans, chitins, and proteins.

Upon PAMP/MAMP and DAMP perception by the PRRs, the so-called PAMP/MAMP-triggered immunity (PTI/MTI) response is activated which gives rise to downstream intracellular signaling events such as activation of mitogen-activated protein kinases, production of reactive oxygen species and transcriptional reprogramming ultimately leading to a complex output response of the plant that limits microbial growth (Wirthmueller et al., 2013).

However, successful pathogens have elaborated a counter defense response to overcome PTI by means of expression of specific elicitors or effector proteins [also known as avirulence (Avr) proteins; Grant et al., 2006]. Pathogenic bacteria typically inject these effectors directly into the cytoplasm of the plant host cell through type III secretion mechanisms to suppress and/or block PRR-dependent signaling, to facilitate nutrient acquisition and to contribute to the pathogen's dispersal which can lead to effector-triggered susceptibility (ETS; Block et al., 2013; Cui et al., 2013).

As a counter move, plants have co-evolved a second layer of defense, known as effector-triggered immunity (ETI) which, in contrast to PTI/MTI acts mostly inside the plant cell. In ETI, specific resistance (*R*) genes become expressed upon recognition of an effector to produce defense proteins. The majority of the *R* proteins include nucleotide-binding leucine-rich repeat (NB-LRR)-containing proteins. The outcome of PTI/MTI and ETI can lead to programmed cell death of the host cell via (local) activation of a hypersensitive response (HR), but can also initiate systemic acquired resistance (SAR) to activate defenses in distal,

non-infected parts of plants in order to establish a heightened state of immunity throughout the plant (Thomma et al., 2011).

PATHOGEN RECOGNITION RECEPTORS (PRRs)

The cell wall confers the first tier of the plant's immunity (Malinovsky et al., 2014). The extracellular PRRs are able to detect pathogen determinants (the so-called PAMPs/MAMPs), DAMPs and effectors at the surface of the plant cell and are used to translocate the extracellular message of 'danger' to the intracellular environment to trigger appropriate defense mechanisms (**Figure 1A**; Dubery et al., 2012). The PRR family encompasses two groups of plasma membrane-localized proteins: the receptor-like kinases (RLKs) and the receptor-like proteins (RLPs). RLKs are single-pass transmembrane proteins with an extracellular domain that is responsible for the perception of the P/M/DAMPs and an intracellular serine/threonine kinase domain that activates the downstream signaling responses. RLPs possess a similar structure but, because they only have a short cytosolic domain without an obvious signaling module, they depend on the association with kinases for signaling. However, there is emerging evidence that upon ligand binding RLKs also form homodimers or heterodimers with other kinases and RLPs and as such function in multiprotein complexes to initiate plant immunity (Boller and Felix, 2009; Böhm et al., 2014; Han et al., 2014; Macho and Zipfel, 2014).

Thus far, only a limited number of RLKs and RLPs that may function in plant immunity have been functionally characterized. Matching these proteins to their ligands is still a challenging study. The majority of the known PRR ectodomains contains LRRs for direct/indirect recognition of pathogenic effector proteins (**Table 1**). In addition, a large diversity of membrane-bound and soluble PRRs have been described to carry lectin domains that are implicated in the recognition of carbohydrate structures from microbial organisms or derived from plant cell wall damage (**Tables 2 and 3**).

PATHOGEN RECOGNITION BASED ON PROTEIN–PROTEIN INTERACTIONS

The study of plant–pathogen interactions has focused on those PRRs which use protein–protein interactions to recognize invading pathogens. Phytopathogens are recognized upon perception of characteristic epitopes present on their surface and essential for the pathogen's survival. These epitopes are mostly recognized by plant cell surface receptors carrying LRR motifs in their ectodomain structures and a kinase domain in their intracellular domain, collectively named the LRR-RLKs. Since these protein–protein interactions have been the subject of several recent overview papers, we only briefly summarize some plant LRR-RLKs and LRR-RLPs (**Figure 1B** and **Table 1**).

Amongst the plant PRRs of the LRR-RLK type, the *Arabidopsis* LRR-RLK *AtFLS2* (Flagellin Sensing 2) is the best-studied protein, containing 28 extracellular LRRs. This FLS2 recognizes bacterial flagellin via perception of the conserved 22-amino acid epitope flg22. *AtFLS2* directly interacts with flg22 resulting in phosphorylation of *AtFLS2* and immediate dimerization with its co-receptor BAK1/SERK, another LRR-RLK. Transphosphorylation of the kinase domain of BAK1 enables conformational

changes and subsequent release of phosphorylated BAK1 leading to activation of downstream MAPK defense signaling (Gómez-Gómez et al., 2001; Chinchilla et al., 2006, 2007; Schulze et al., 2010). In the absence of PAMP recognition, BAK1 itself interacts with the pseudokinase BIR2 (also LRR-RLK-type) to prevent FLS2-BAK1 heterodimerization (Halter et al., 2014). After flg22 perception, *AtFLS2* is subject to endocytosis and degradation by the E3 ubiquitin ligase PUB12/13 to prevent continuous defense signaling. Newly synthesized *AtFLS2* is incorporated in the plasma membrane at later times (Smith et al., 2014). In turn, virulent *Pseudomonas syringae* pathovars produce effector proteins, such as AvrPto, AvrPtoB, and AvrPphB to destabilize *AtFLS2* and thus compromise host immunity (Block and Alfano, 2011).

The transmembrane protein *AtEFR* represents another *Arabidopsis* LRR-RLK-type receptor involved in bacterial PAMP signaling (Zipfel et al., 2006). The ectodomain of *AtEFR* consists of 24 LRRs and is involved in the perception of the elf18 peptide, a conserved N-terminal fragment of bacterial elongation factor Tu. Many of the signaling compounds downstream of *AtEFR* are shared with *AtFLS2*, and *AtEFR* also requires dimerization with *AtBAK1* for signaling. However, the action of *AtEFR* is independent of flagellin perception and unlike *AtFLS2*, *AtEFR* requires N-glycosylation to become functional. Indeed, a single N-glycan is crucial for receptor abundance and ligand recognition between the pathogen and the plant cell surface (Häwcker et al., 2010).

Rice plants use the transmembrane XA21 receptor kinase to confer immunity toward a number of *Xanthomonas oryzae* pv *oryzae* (*Xoo*) isolates, which cause leaf blight in rice. The XA21 receptor recognizes Ax21, a sulfated 17-amino acid peptide derived from the *Xoo* type I secreted protein (Lee et al., 2009). Also here, the action of XA21 is tightly regulated. Without PAMP, XA21 is kept in an inactive state through binding with and autoprophosphorylation by the ATPase XB24. Upon binding of Ax21 to XA21, the XB24/XA21 dimer dissociates and the XA21 kinase domain is released and translocated to the cell nucleus for subsequent immune signaling (Park and Ronald, 2012). Chen et al. (2014) recently reported that XA21 can also be found in a constitutive heteromeric complex with a BAK1 ortholog, named OsSERK2, and undergoes bidirectional transphosphorylation to confer resistance to the *Xanthomonas* bacterium.

Tomato plants encode several cell-surface LRR-RLPs such as LeEIX1/EIX2 and Ve1 which confer resistance toward *Trichoderma* and race 1 strains of *Verticillium* pathogens, respectively (Ron and Avni, 2004; Fradin et al., 2009). The ethylene-inducing xylanase EIX is a fungal β-1–4-endoxylanase that is used by *Trichoderma viride* to enter tomato and tobacco plants. The epitope that is recognized by the plants to elicit defense responses constitutes five amino acids that are not involved in the enzymatic activity (Rotblat et al., 2002). Both LeEIX1/EIX2 can bind EIX, but only LeEIX2 transmits the signal to activate immune responses (Ron and Avni, 2004). The ligand of the Ve1 receptor is the Ave1 peptide, a peptide conserved in several fungi and phytopathogenic bacteria. BAK1 signaling is involved in induced defense responses for both LeEIX1 and Ve1 (Fradin et al., 2009; Bar et al., 2010).

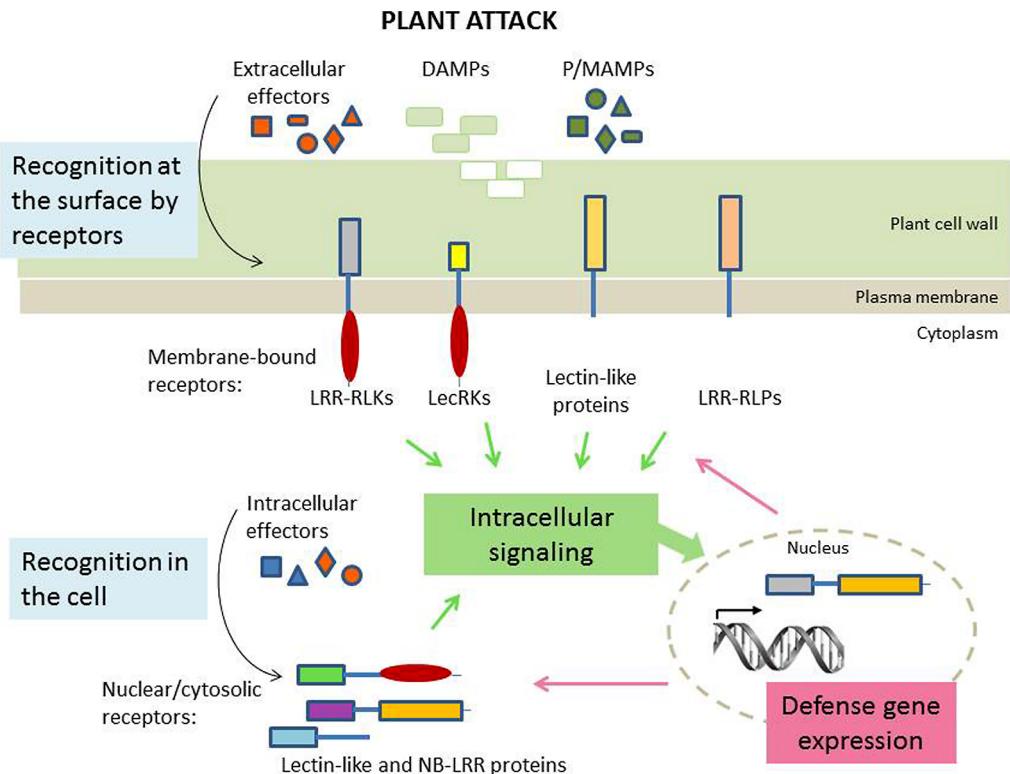
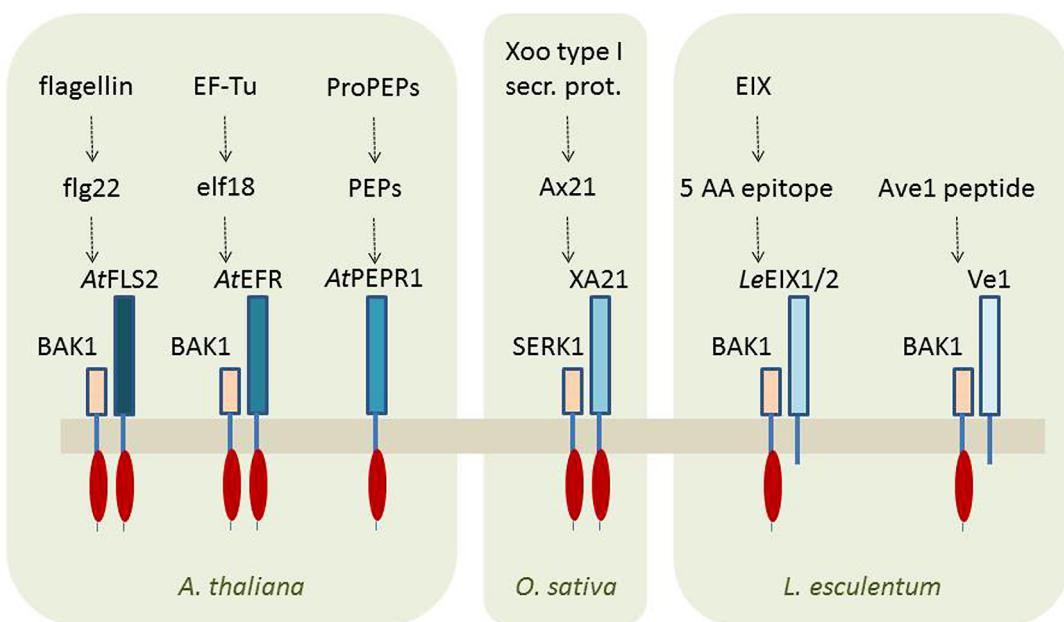
A**B**

FIGURE 1 | Plant innate immunity. (A) Perception of pathogen/microbe-associated molecular patterns (P/MAMPs), damage-associated molecular patterns (DAMPs) and pathogen-derived effector proteins. Plants sense P/MAMPs, DAMPs, and effectors through membrane-bound and intracellular (soluble) receptors. Four types of membrane-bound receptors can be distinguished: the LRR-type receptor kinases (LRR-RLKs) and proteins (LRR-RLPs), and the receptor kinases and proteins with lectin domains (called LecRKs and lectin-like proteins, respectively). Soluble receptors known thus

far include NB-LRR proteins as well as nucleocytoplasmic lectins. Upon ‘danger’ perception, these receptors trigger intracellular signals which ultimately will result in altered expression of defense-related genes. Legend: ellipses represent kinase domains; bars represent other protein motifs, including LRRs and lectin domains. **(B)** Transmembrane PRRs detect P/MAMPs through protein–protein interactions. Bars represent LRR domains, red ellipses indicate functional kinase domains.

(Continued)

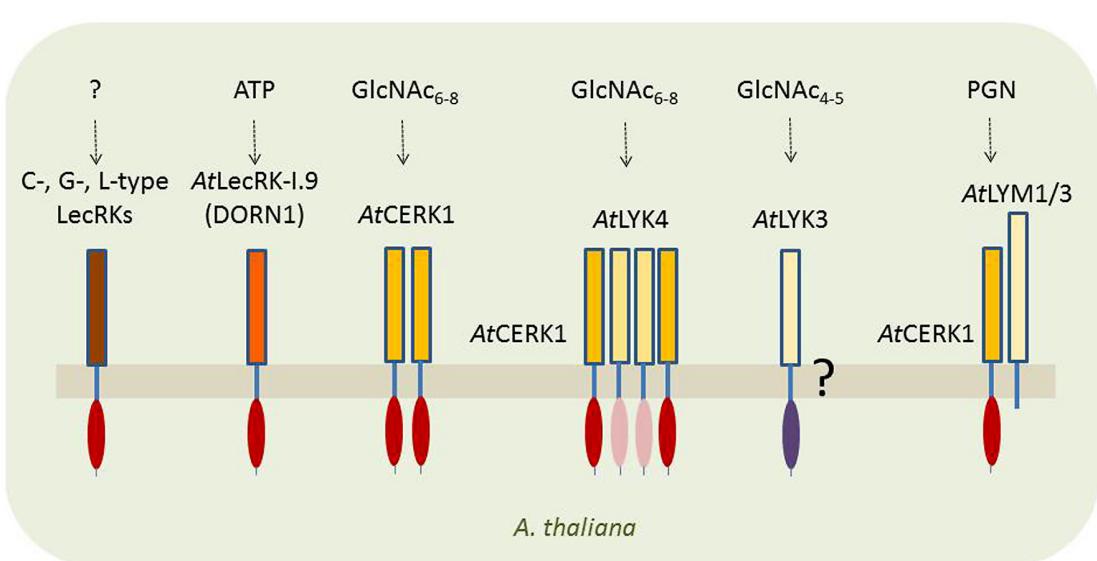
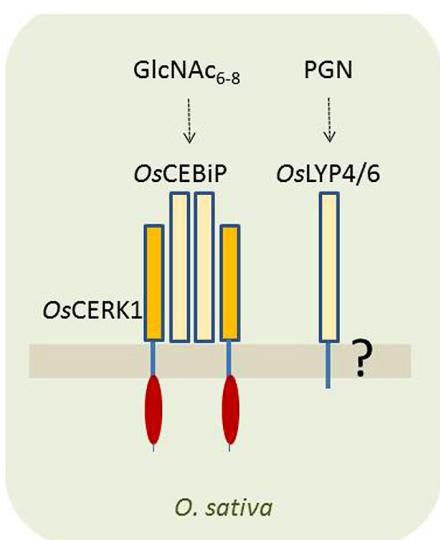
C**D**

FIGURE 1 | (C) Transmembrane PRRs with lectin domains identified in *Arabidopsis thaliana*. In the case of the LysM domain evidence supports protein–carbohydrate interactions to detect P/M/DAMPs. Bars represent lectin domains, including C-, G-, and L-type (brown) and LysM (yellow) domains. Ellipses represent kinase domains (red = functional,

pink = non-functional, purple = putative kinase domain). **(D)** Transmembrane PRRs with lectin domains identified in *Oryza sativa*. The LysM domain recognizes P/M/DAMPs through specific binding of chitin fragments. Bars represent the LysM lectin domains; ellipses represent functional kinase domains.

PATHOGEN RECOGNITION BASED ON PROTEIN–CARBOHYDRATE INTERACTIONS THE CARBOHYDRATES

Major part of the P/M/DAMPs that are perceived in the plant as ‘danger’ molecules include carbohydrate structures which are either present at the cell surface of the invading pathogen or originate from the plant itself, when released from cell wall degradation caused by pathogen entry. These structures comprise bacterial LPS and PGNs and fungal chitin molecules as well as plant-derived oligogalacturonides and cellulose fragments. Also arabinogalactan

proteins residing in the plant cell wall have been reported to be involved in plant immune responses (Newman et al., 2013).

Lipopolysaccharides are large outer membrane glycoconjugates found in Gram-negative bacteria that are composed of a lipid, a core oligosaccharide and an O-antigen polysaccharide chain. The lipid, called Lipid A, is embedded in the bacterial membrane and is linked to the core oligosaccharide by the KDO sugar (3-deoxy-D-mannose-2-octulosonate). The core sugar ends in the O-antigen which is composed of oligorhamnans in many phytopathogens.

Table 1 | LRR-type PRRs involved in plant defense signaling.

PRR	Plant species	Ligand	Reference
LRR-RLK type			
AtFLS2	<i>Arabidopsis thaliana</i>	Flagellin (Flg22)	Chinchilla et al. (2006)
AtEFR	<i>A. thaliana</i>	Ef-TU	Zipfel et al. (2006)
XA21	<i>Oryza sativa</i>	Activator of XA21 (Ax21)	Lee et al. (2009)
XA3/XA26		Epitope derived from <i>Xanthomonas oryzae</i> pv <i>oryzae</i>	Sun et al. (2004)
SR160	<i>Lycopersicon peruvianum</i>	(pro)systemin	Scheer and Ryan (2002)
PEPR1	<i>A. thaliana</i>	PEPR1	Krol et al. (2010)
NORK	<i>Medicago truncatula</i>	?	Endre et al. (2002)
SYMRK	<i>Lotus japonicus</i>	?	Stracke et al. (2002)
LRR-RLP type			
LeEIX2	<i>Lycopersicon esculentum</i>	Xylanase (EIX)	Ron and Avni (2004)
Ve1	<i>L. esculentum</i>	Ave1 peptide	Fradin et al. (2009)

Table 2 | Membrane-bound lectin-type PRRs involved in plant defense signaling and symbiosis.

PRR	Plant species	Ligand	Reference
LysM-RLK type			
AtCERK1	<i>Arabidopsis thaliana</i>	Chitin	Miya et al. (2007), Brotman et al. (2012)
AtLYK3	<i>A. thaliana</i>	Chitin	Paparella et al. (2014)
AtLYK4	<i>A. thaliana</i>	Chitin	Wan et al. (2012)
OsCERK1	<i>Oryza sativa</i>	Chitin (when in combination with OsCEBiP)	Kaku et al. (2006)
NFR1	<i>Lotus japonicus</i>	Lipochitoooligosaccharide Nod factors	Radutoiu et al. (2003)
LYK3	<i>Medicago truncatula</i>	Lipochitoooligosaccharide Nod factors	Knogge and Scheel (2006)
NFR5	<i>L. japonicus</i>	Lipochitoooligosaccharide Nod factors	Radutoiu et al. (2003), Madsen et al. (2003)
SYM10	<i>Pisum sativum</i>	Lipochitoooligosaccharide Nod factors?	Madsen et al. (2003)
LYK4	<i>M. truncatula</i>	Lipochitoooligosaccharide Nod factors	Limpens et al. (2003)
NFP	<i>M. truncatula</i>	Lipochitoooligosaccharide Nod factors	Mulder et al. (2006)
LysM-RLP type			
LYM1/AtLYP2, LYM3/AtLYP3	<i>A. thaliana</i>	Peptidoglycan	Willmann et al. (2011), Tanaka et al. (2013)
OsCEBiP	<i>O. sativa</i>	Chitin	Shimizu et al. (2010)
OsLYP4, OsLYP6	<i>O. sativa</i>	Chitin + Peptidoglycan	Liu et al. (2012a)

Table 3 | Nucleocytoplasmic lectin domains involved in plant defense signaling.

Lectin domain	Carbohydrate Specificity	Subcellular localisation	Examples
Amaranthin domain	GalNAc, T-antigen	nucleus, cytosol	Amaranthin, Hfr-2
EUL domain	Galactosides, high-mannose <i>N</i> -glycans	nucleus, cytosol	EEA, ArathEULS3
Jacalin domain	Mannose-specific subgroup / galactose-specific subgroup	nucleus, cytosol / vacuole	Orysata, TaVER2, TaHfr-1, TaJA-1
Nictaba domain	(GlcNAc) _n , high-mannose <i>N</i> -glycans, complex <i>N</i> -glycans	nucleus, cytosol	Nictaba, PPL
Ricin-B domain	Gal/GalNAc, Sia α 2,6Gal/GalNAc	Vacuole, nucleus, cytosol	Ricin, abrin, SNA-I, SNA-V

Peptidoglycans are essential cell wall components of Gram-positive and Gram-negative bacteria, and comprise alternating $\beta(1\rightarrow 4)$ linked *N*-acetylmuramic acid (MurNAc) and *N*-acetylglucosamine (GlcNAc) residues, with a short peptide chain attached to MurNAc.

Chitin is a long-chain polymer of $\beta(1\rightarrow 4)$ linked GlcNAc residues and is the main component of the fungal cell wall and the exoskeleton of insects. In case of fungi, chitin is cross-linked to β -glucan.

Oligogalacturonides are oligomers of $\alpha(1\rightarrow 4)$ linked galacturonosyl residues that are released from plant cell walls upon partial degradation of homogalacturonan (i.e., the major component of pectin) by pathogen attack and also upon mechanical damage.

Cellulose is an important component of the plant cell wall, built up of hundreds of $\beta(1\rightarrow 4)$ linked glucose residues which form long polymer chains. These chains are packed into microfibrils which give strength and flexibility to the plant cell wall.

Arabinogalactan proteins are a distinct class of complex, extensively glycosylated hydroxyproline-rich proteins (the so-called proteoglycans), widely distributed among plant species. They consist of a rather small core protein backbone which is *O*-glycosylated by type II arabinogalactan glycans and often contain an N-terminal GPI anchor. AGPs are located near the cell surface, including the plasma membrane, the apoplast, the cell wall, and the intercellular matrix, and have been implicated in many aspects of plant growth and development, such as cell expansion, proliferation, and differentiation. These AGPs are not only involved in establishing a connection between the cell wall and the plasma membrane, but would also extend to the cytoplasm, establishing a continuum between intracellular and extracellular compartments.

LECTIN DOMAINS INVOLVED IN PLANT IMMUNITY

Lectins are proteins that contain at least one non-catalytic domain which enables them to selectively recognize and bind in a reversible way to specific glycans that are either present in a free form or are part of glycoproteins and glycolipids. Plants express a huge number of highly diverse lectins, exhibiting different molecular structures and binding specificities toward endogenous (plant) glycans as well as to glycans from exogenous (non-plant) origin (Van Damme et al., 2008, 2011).

A lot of plant lectins are constitutively expressed in high amounts in seeds and vegetative storage tissues where they have been shown to play a role in plant defense (Peumans and Van Damme, 1995). In addition, plants also express minute amounts of specific lectins as particular responses toward environmental stresses and pathogen attack. In the absence of plant stress, the inducible lectins are not expressed at detectable levels. Most of the constitutively expressed lectins are synthesized with a signal peptide, and are sequestered in the vacuole or secreted to the extracellular space. In contrast, most of the inducible plant lectins reside in the nucleus and the cytoplasm of a plant cell (Lanno and Van Damme, 2010).

The majority of the known plant lectins are built up of one or more lectin-like domains coupled to un-related domains such as aerolysin, AIG1, chitinase, dirigent, F-box, Kelch, kinase, LRR, NB-ARC, PAG, or TIR domains (Van Damme et al., 2008). Up

till now, most attention of the scientific community dealing with plant innate immunity has been given to transmembrane receptor proteins containing one or more ectopic lectin domains coupled to an intracellular kinase domain. Amongst these lectin receptor kinases (LecRKs), those comprising LysM-type lectin domains are the most studied ones (Table 2; Singh and Zimmerli, 2013). However, plants use a broad variety of lectin domains to counteract pathogen attack (Table 3).

Membrane-bound proteins with a lectin domain

Lectin receptor kinases (LecRKs). Typically, LecRKs are two-domain proteins composed of an N-terminal extracellular lectin domain and a C-terminal cytosolic Ser/Thr kinase domain, separated by a transmembrane region (Figure 1C). Based on their lectin domain LecRKs are classified into 4 types; G-, C-, L-, and LysM-type (Singh and Zimmerli, 2013; Vaid et al., 2013). Although these LecRKs consist of at least one domain that shows striking sequence similarity with a lectin motif, very little information is available with respect to the ability of this domain to recognize and interact with specific carbohydrate structures.

G-type LecRKs contain an extracellular lectin domain which resembles the *Galanthus nivalis* agglutinin (GNA). However, it remains to be shown whether this sugar binding domain is indeed involved in ligand binding. Based on genome-wide analyses, 32 G-type LecRKs have been identified in *Arabidopsis thaliana* and 100 G-type LecRKs in rice (Vaid et al., 2012). G-type LecRKs function in self-incompatibility reactions in flowering plants (the so-called SRKs) and in plant defense to biotic stress as well as to abiotic stress (Sherman-Broyles et al., 2007; Kim et al., 2009; Sun et al., 2013).

C-type (calcium-dependent) LecRKs are mostly found in mammalian proteins that mediate immune responses and play a role in pathogen recognition. In plants, C-type RLKs are rather rare. At present only one C-type LecRK encoding gene has been identified in *A. thaliana* (At1g52310), though its function has not been elucidated yet (Cambi et al., 2005; Bouwmeester and Govers, 2009).

L-type (legume-like) LecRKs represent a more abundant group of LecRKs. Thus far, 45 L-type LecRKs have been identified in *A. thaliana*. Based on phylogenetic relationships the genes encoding *Arabidopsis* L-type LecRK can be classified into nine clusters and nine clades (designated with the Roman numerals I to IX). These genes showed variable expression patterns in different tissues and developmental stages in response to stimuli (Bouwmeester and Govers, 2009). Some LecRKs were indeed reported to be involved in plant resistance to pathogens, e.g., AtLecRK-I.9 is involved in sensing cell wall integrity and defense response to *Phytophthora infestans* (Bouwmeester et al., 2011). AtLecRK-VI.2 is critical for resistance against *Pseudomonas syringae* and *Pectobacterium carotovorum* (Singh et al., 2012; Huang et al., 2014) while AtLecRK-IV.3 induces resistance against *Botrytis cinerea* (Huang et al., 2013). Some AtLecRKs have also been reported to act in hormone signaling (ABA) and stomatal immunity (e.g., AtLecRK-VI.2 and AtLecRK-V.5; Singh et al., 2012). L-type LecRKs have also been identified in other plants. For instance, tobacco plants express L-type LecRKs

with a role in plant immunity against pathogens and insects (Kanzaki et al., 2008; Gilardoni et al., 2011). In turn, *Medicago* plants contain L-type LecRKs that are involved in symbiosis (Navarro-Gochicoa et al., 2003).

At present, it is not yet clear whether the L-type LecRKs possess lectin activity since the amino acids important for interaction of the legume lectin domain with its specific carbohydrate ligand are poorly conserved. In contrast the hydrophobic site present in the legume-type lectin domain is preserved, suggesting that LecRKs may act in the recognition of small hydrophobic ligands (Huang et al., 2013; Choi et al., 2014). Recently, evidence was obtained that the plasma membrane localized DORN1, encoded by the *AtLecRK-I.9* gene, plays an important role as a receptor for extracellular ATP (**Figure 1C**; Choi et al., 2014). DORN1 lacks the conserved Ca^{2+} and Mn^{2+} binding residues that are critical for carbohydrate binding activity of legume lectins. Early studies also suggested the ability of the legume lectin domain to bind adenine, a component of ATP (Roberts and Goldstein, 1983). However, since adenine was unable to compete with ATP for binding to DORN1, the exact ATP binding site in DORN1 remains to be determined. These data are in good agreement with the fact that extracellular ATP is now perceived as a central signaling molecule in plant stress responses (Cao et al., 2014; Choi et al., 2014).

LysM LecRKs are the most studied LecRKs (**Figures 1C,D** and **Table 2**). They contain ectopic lysin motifs, referred to as LysM domains, which are considered to mediate binding to various types of bacterial PGN and fungal chitin, upon recognition of the GlcNAc moieties (Buist et al., 2008; Gust et al., 2012). The lysine motif, approximately 40 amino acids in length, is a ubiquitous protein domain found in most living organisms except the Archaea. It can be used as a single domain, but is also present in the form of two or occasionally three repeats in a large number of proteins. In most cases LysM motifs are coupled to other protein domains exhibiting some enzymatic activity, such as GlcNAc modification in the case of microbial hydrolases or intracellular signaling for plant kinases.

The *Arabidopsis* chitin elicitor receptor kinase 1 (*AtCERK1*, also known as LYK1/RLK1) is the major chitin receptor found in *A. thaliana* (Miya et al., 2007; Petutschnig et al., 2010; Tanaka et al., 2013). It is a membrane-anchored protein with an extracellular domain containing three LysM motifs coupled to an intracellular kinase domain. This kinase domain contains a canonical RD (Arginine-Aspartate) motif in its catalytic loop and possesses autophosphorylation activity, unlike the non-RD kinase domain of typical PRRs. *AtCERK1* was reported to be involved in fungal resistance. The protein directly binds to fungal chitooligosaccharides (GlcNAc_n with $n > 2$) through its LysM domains, but only longer oligomers ($n > 4$) trigger immune responses. Liu et al. (2012b) demonstrated that binding of chitin oligomers ($n = 8$) to *AtCERK1* induces homodimerization of the receptor, which is essential for the activation of downstream intracellular signaling, most likely by phosphorylation of both CERK1 cytoplasmic kinase domains.

AtCERK1 can also mediate perception of PGN when part of a complex with *AtLYM1* and *AtLYM3*, two other transmembrane LysM containing proteins lacking an intracellular kinase domain (the so-called LYP proteins; Willmann et al., 2011). Next to *AtCERK1* *A. thaliana* contains four additional LysM RLKs,

designated *AtLYK2–5*. Since *AtLYK4* and *AtLYK5* can also recognize and bind to chitin molecules but have a non-functional pseudokinase domain it has been suggested they should form a complex with *AtCERK1* to compose a functional chitin receptor (Wan et al., 2012). *AtLYK3* possesses a functional kinase domain, but was suggested be involved in ABA signaling rather than in chitin recognition (Paparella et al., 2014).

Unlike *AtCERK1*, the rice ortholog *OsCERK1* cannot bind directly to chitooligosaccharides. *OsCERK1*, a LysM LecRK with a single extracellular LysM domain, requires heterodimerization with its co-receptor *OsCEBiP* (chitin elicitor binding protein) for chitin binding and subsequent activation of innate immunity (**Figure 1D**). *OsCEBiP* is a transmembrane LysM receptor protein containing three LysM domains and lacking a kinase domain, resembling the *Arabidopsis* proteins *AtLYM1* and *AtLYM3*. However, whereas *AtLYM1* and *AtLYM3* are involved in PGN binding through receptor formation with *AtCERK1*, *OsCEBiP* seems to play a major role in fungal chitin perception. Upon binding to the fungal GlcNAc_{6-8} oligomers, *OsCEBiP* homo-dimerizes at the plasma membrane of the plant cell. After this ligand-induced dimerization, the *OsCEBiP* sandwich-like structure forms a heteromeric complex with two *OsCERK1* proteins, to activate intracellular defenses, including ROS signaling, callose deposition, and defense gene expression (Shimizu et al., 2010; Shinya et al., 2012; Hayafune et al., 2014; Kouzai et al., 2014). *OsLYP4* and *OsLYP6* are two additional rice LysM proteins lacking a kinase domain which presumably mediate perception of both PGN and chitin, but their action and transmembrane signal transfer remain unclear (Liu et al., 2012a).

Plants not only use LysM LecRKs to recognize pathogenic organisms, they also use them to perceive beneficial organisms such as mycorrhizal fungi and rhizobacteria implicating a dual role of LysM in both innate immunity and symbiosis (**Table 2**; Gust et al., 2012). Examples include the *Lotus japonicus* NRF1 and NRF5 (Radutoiu et al., 2003) and the *Medicago truncatula* LYK3 and LYK4 (Knogge and Scheel, 2006) which can recognize rhizobial lipochitin-oligosaccharide signals or Nod factors.

Soluble proteins with a lectin domain

An overview of soluble proteins with a lectin domain involved in plant defense signaling is given in **Table 3**.

Amaranthins. The Amaranthin family groups all proteins related to amaranthin, a lectin present in the seeds of *Amaranthus caudatus*. Native amaranthin is a homodimeric protein built of two 33 kDa subunits, each comprising two tandem-arrayed homologous amaranthin domains. Amaranthin specifically recognizes the T-antigen disaccharide $\text{Gal}\beta(1,3)\text{GalNAc}$ but also interacts with GalNAc. Interestingly, the amaranthin domain itself possesses no sugar binding site(s), but the specific head-to-tail arrangement of two amaranthin subunits is necessary to establish the T-antigen disaccharide binding site (Van Damme et al., 2008). Up till now, only amaranthins originating from *Amaranthus* species have been purified and biochemically characterized. This nucleocytoplasmic lectin was reported to enhance the plant's resistance against aphids

when ectopically expressed in transgenic tobacco, potato, and cotton by affecting growth and development of the invading aphids (Wu et al., 2006; Xin et al., 2011).

Screening of the publicly available genome databases revealed that the amaranthin domain is widespread throughout the plant kingdom. Several chimeric proteins containing N-terminal amaranthin domain(s) coupled to unrelated protein domains have been identified (Van Damme et al., 2011). Columbine plants (*Aquilegia formosa* × *Aquilegia pubescens*) encode a protein in which two amaranthin domains are coupled to a kinase domain. Since this protein does not have a transmembrane domain, it is suggested to reside inside the cell. Cucumber, maize, and wheat plants were found to contain genes that encode proteins with amaranthin domain(s) coupled to an aerolysin domain. Aerolysins are cytolytic toxins that are mostly produced by the bacterium *Aeromonas* and can kill host cells upon pore formation in the plasma membrane (Degiacomi et al., 2013). In wheat, a chimeric protein called Hfr-2 (Hessian fly responsive-2) is up-regulated in the leaf sheaths after feeding of virulent Hessian fly larvae and armyworms, and enhances wheat tolerance against Hessian fly larvae (Puthoff et al., 2005).

Calreticulin/calnexin. Calreticulin (CRT) and calnexin (CNX) are glucose binding lectins residing in the endoplasmic reticulum (ER) of eukaryotic cells. Both CRT and CNX act as molecular chaperones and are essential ER components ensuring proper folding and quality control of newly synthesized secretory and membrane-bound glycoproteins before ER release (Ellgaard and Frickel, 2003; Kapoor et al., 2004; Williams, 2006). While CNX is a type-I integral membrane protein, CRT is a soluble protein. Both CRT and CNX act together with their co-chaperones ERp57 and PDI, two soluble thiol-disulfide oxidoreductases. Whereas the classical chaperones associate with the peptide moiety of their substrates, CNX and CRT bind to their glycoprotein substrates primarily through specific recognition and binding to the oligosaccharide intermediates Glc₁Man₇–₉GlcNAc₂ present on nascent glycoproteins.

In the ER quality control system, a growing polypeptide initially gets N-glycosylated with the core glycan Glc₃Man₉GlcNAc₂. By successive action of glucosidase I and II, the outer glucoses are trimmed resulting in a monoglycosylated glycoprotein which then serves as the substrate of CNX/CRT for proper folding. Once the glycoprotein is correctly folded, the terminal glucose of its oligosaccharide is cleaved by glucosidase II and the glycoprotein is released from the CNX/CRT/ERp57 complex for further processing. If the glycoprotein is not correctly folded, it is recognized by the UDP-glucose:glycoprotein glucosyltransferase enzyme, which acts as a folding sensor, and gets re-glucosylated to promote its renewed association with CNX/CRT. As such, de- and re-glucosylation of glycoproteins facilitates their correct folding. When folding ultimately fails, the misfolded glycoproteins are sorted out of the ER and are degraded by the proteasome, a system known as ER associated degradation or ERAD. Under adverse environmental conditions, the demand for protein folding exceeds the capacity of the system resulting in the accumulation of misfolded proteins in the

ER, giving rise to so-called ER stress (Howell, 2013; Liu and Li, 2014).

The correct folding of membrane-bound PRRs is a critical step in plant immunity. The ER quality control system not only regulates the abundance and quality of transmembrane receptors, it also affects downstream signaling of the receptor (Tintor and Saijo, 2014). The LRR-RLKs *AtEFR* and *AtFLS2* from *A. thaliana* as well as the LysM LecRK NFP from *M. truncatula* require proper N-glycosylation for accurate functioning. Nevertheless, *AtEFR* and *AtFLS2* production are coordinated by different ER components (Li et al., 2009; Häweker et al., 2010). *Arabidopsis* plants contain two types of CRTs: CRT1/2 and CRT3 isoforms (Thelin et al., 2011). Whereas CRT1 is a key chaperone in plant ER stress, CRT3 is typically involved in the quality control of *AtEFR* and the brassinosteroid receptor BRI1, but is not essential for *AtFLS2* biogenesis (Jin et al., 2009). CRT2 appears to have a dual regulatory role in plant defense against the biotrophic pathogen *P. syringae* pv tomato DC3000 (Qiu et al., 2012). Upon pathogen invasion, CRT2 is involved in the up-regulation of SA-dependent immune signaling through its Ca²⁺ buffering capacity. However, CRT2 negatively influences these SA-dependent responses through its chaperone activity, resulting in the overall suppression of plant resistance toward *P. syringae* pv tomato DC3000.

EUL-related lectins. The family of EUL-related lectins groups all nucleocytoplasmic proteins that comprise at least one *Euonymus* lectin (EUL) domain. The prototype of this family is the so-called *Euonymus europaeus* agglutinin (EEA) which is present at very high concentrations in the arillus tissue of the spindle tree (*E. europaeus*). EEA is a non-glycosylated homodimeric protein composed of 17 kDa subunits, and recognizes two structurally different classes of glycans. Glycans with carbohydrate epitopes containing galactose, such as Galα1–3Gal and Galα1–3Galβ1–4GlcNAc, blood group B [Galα1–3(Fucα1–2)Gal-], and O (Fucα1–2Gal-) epitopes are bound with a higher affinity compared to high-mannose N-glycans. Based on inhibition studies, it was suggested that the EUL domain might contain two different binding sites (Fouquaert et al., 2008).

Sequences with an EUL domain are present in almost all sequenced plant genomes from Embryophyta, ranging from liverworts to flowering plants. The sequence of the EUL domain is well conserved suggesting that the corresponding EUL proteins fulfill an essential role in plants. Based on the overall protein domain architecture, the EUL family can be divided into two groups containing proteins either composed of a single EUL domain (S-type EUL proteins) or of two tandemly arrayed EUL domains separated by a spacer sequence (D-type EUL proteins; Fouquaert and Van Damme, 2012). The majority of the EUL sequences encode chimeric proteins, in which the EUL domain is linked to other unknown domains. Whereas most dicot species encode one or two EUL S-type proteins, monocot, and lower plant species contain a whole set of S- and D-type EUL sequences.

In contrast to EEA, which is expressed at high concentrations in the arilli of spindle tree seeds, the EUL proteins from *Oryza sativa* and *A. thaliana* are very low abundant proteins. In both plants the amount of EUL transcripts is enhanced after the plant

was subjected to different abiotic (such as dehydration, salinity, osmotic stress, and ABA treatment) and biotic (such as bacterial and fungal infection) stresses (Fouquaert and Van Damme, 2012; Al Atalah et al., 2014a).

Next to the differential regulation of gene expression, the EUL proteins show different carbohydrate binding properties. ArathEULS3 from *A. thaliana* preferentially interacts with N-glycans containing galactosylated structures such as Lewis X [Gal β 1–4(Fuc α 1–3)GlcNAc], Lewis Y [Fuc α 1–2Gal β 1–4(Fuc α 1–3)GlcNAc] and lactosamine (Gal β 1–4GlcNAc) motifs (Van Hove et al., 2011). Similarly, both EUL domains composing the two-domain EUL protein from rice, OrysaEULD1A, show carbohydrate specificity toward galactose containing glycans (Al Atalah et al., 2014b). In contrast, the rice protein OrysaEULS2 preferably binds mannosylated N-glycan structures (Al Atalah et al., 2012). All these EUL proteins are located in the nucleus and the cytoplasm of the plant cell. In search for interacting proteins, Li et al. (2014) recently reported that ArathEULS3 interacts with the nuclear/cytosolic ABA receptor RCAR1. Furthermore, Berendzen et al. (2012) demonstrated interaction of ArathEULS3 with CPK3, a Ca $^{2+}$ dependent kinase involved in the ABA response in guard cells, supporting a role for ArathEULS3 in ABA signaling and stomatal closure.

Jacalin-related lectins (JRL). The family of jacalin-related lectins is named after jacalin, a 18 kDa T-antigen disaccharide-binding lectin domain first isolated from the seeds of jackfruit (*Artocarpus integrifolia*). Based on differences in molecular structure, subcellular localization, and carbohydrate binding properties, the large group of jacalins can be subdivided into two subgroups, further referred to as the galactose binding and mannose binding JRLs, residing in the vacuolar and nucleocytoplasmic compartment of the plant cell, respectively.

Galactose-specific JRLs have been reported mainly within the family Moraceae, whereas the mannose-specific JRLs are widespread in Viridiplantae. Furthermore, recent studies have shown that the jacalin domain is not restricted to plant proteins, since a similar domain has been reported in eukaryotes outside the plant kingdom as well as in some prokaryotes (Van Damme et al., 2008; Kanagawa et al., 2014). In plants, chimeric proteins comprising one or more jacalin domains linked to an unrelated domain are widely distributed. For instance, in *A. thaliana*, sequences composed of one or two jacalin domains C-terminally linked to five in tandem arranged Kelch domains are present. In addition, multiple *Arabidopsis* genes encode a putative F-box protein with a C-terminal jacalin domain (Nagano et al., 2008). Several Poaceae species (wheat, rice, maize) express proteins consisting of an N-terminal dirigent domain (also called disease-response domain) C-terminally fused to a jacalin domain. In rice, additional types of chimerolectins were identified, either composed of an N-terminal tyrosine kinase domain coupled to two or three jacalin domains or an N-terminal NB-ARC motif coupled to a LRR and a C-terminal jacalin domain (Van Damme et al., 2008).

Many jacalin-related lectin genes have been shown to be associated with disease resistance, abiotic stress signaling, wounding, insect damage or multiple stresses (Song et al., 2014). Especially the

jacalin proteins with a dirigent domain are functionally involved in plant defense. To our knowledge, these chimeric proteins have only been reported in Poaceae species. In wheat, nearly half of the jacalin-related lectin genes encode dirigent domain-containing jacalin-related proteins. Several of these proteins have been studied in some detail, amongst them TaVER2, TaHfr-1, and TaJA-1 (Song et al., 2014). Interestingly all these mannose binding lectins are expressed as a response toward plant stress. TaVER2 is specifically expressed during vernalization in wheat (Yong et al., 2003) but is also up-regulated upon jasmonate and ABA treatment (Feng et al., 2009). TaHfr-1 is up-regulated after herbivory of Hessian fly larvae (Williams et al., 2002) and TaJA-1 is specifically accumulating after jasmonate treatment (Ma et al., 2013). The mannose specific TaHfr-1 was shown to effectively inhibit Hessian fly larval feeding resulting in the delay of larval development and premature death of the pest insects (Subramanyam et al., 2008). Transgenic tobacco plants overexpressing Ta-JA1 revealed increased resistance to bacterial, fungal, and viral pathogens (Ma et al., 2010). TaVER2 homologs have been found in maize and sorghum (β -glucosidase aggregating factor) and in rice (OsJAC1; Esen and Blanchard, 2000; Jiang et al., 2006; Kittur et al., 2009). Transgenic rice plants overexpressing OsJAC1 indicated the importance of OsJAC1 for rice growth and development (Jiang et al., 2007).

Similar to the chimerolectins also jacalin-related proteins composed only of jacalin domains are up-regulated in plant tissues subjected to certain stress treatments. For instance, OrySata was first reported as a salt inducible mannose binding lectin in the leaves of *O. sativa* (Zhang et al., 2000). Later it was shown that OrySata is also expressed upon JA and ABA treatment, after infection with an incompatible *Magnaporthe grisea* strain as well as during senescence (Lee et al., 2001; de Souza Filho et al., 2003; Qin et al., 2003). Glycan array analyses revealed that OrySata preferentially interacts with high-mannose and some more complex N-glycans (Al Atalah et al., 2011). In recent years, several orthologs of OrySata have been identified in Gramineae species but also in other plants, such as *Helianthus tuberosus* and *Ipomoea batatas*. The mannose specific wheat protein TaJRL1 consisting of two jacalin-like domains is considered a component of SA and JA dependent plant defense signaling mechanisms and is activated upon fungal infection (*Fusarium graminearum* and *Blumeria graminis*; Xiang et al., 2011). However, not all jacalin-related defense proteins depend on hormone signaling. For example, the *Arabidopsis* jacalin-related JAX1 confers broad but specific resistance to potex viruses by inhibition of viral RNA accumulation, independent of hormone signaling (Yamaji et al., 2012). Other *Arabidopsis* jacalin-related proteins interact with proteins of ER bodies, i.e., ER-derived organelles presumably involved in defense against herbivores and/or pathogens (Nagano et al., 2008). Recently, the jacalin-related protein from sunflower seedlings named Helja was reported as a lectin with antifungal properties toward some pathogenic fungi of the genus *Candida*. Helja induces morphological changes as well as ROS production in yeast cells. Furthermore, lectin treatment also altered the membrane permeability of the cells (Regente et al., 2014).

Nictaba-related lectins. The family of Nictaba-related lectins was named after the *Nicotiana tabacum* agglutinin, abbreviated as Nictaba, a 19 kDa lectin domain originally discovered in tobacco leaves (Chen et al., 2002a) after jasmonate treatment. Though the lectin was first reported as a chito-oligosaccharide binding protein, glycan array analyses revealed that Nictaba also reacts with the inner core structure ($\text{Man}_3\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4\text{GlcNAc}\beta\text{-N-Asn}$ of high-mannose and complex *N*-glycans. Biochemical assays confirmed that Nictaba can interact in a sugar-inhibitable way with many *N*-glycosylated proteins (Lannoo et al., 2006, 2007). A nuclear proteomics approach revealed the interaction of Nictaba with the core histone proteins H2A, H2B, and H4 through their O-GlcNAc modification (Schouppé et al., 2011), which was later confirmed at the microscopical level (Delporte et al., 2014a).

An extensive survey of the genome/transcriptome databases indicated that Nictaba-like domains are widespread among the Embryophyta but are absent from other eukaryotes and prokaryotes. Few proteins belonging to the Nictaba-like family consist of a single Nictaba domain. Furthermore, numerous sequences were identified encoding chimeric proteins comprising the Nictaba domain fused to unrelated *N*-terminal domains (e.g., F-box domain) or a Nictaba domain C-terminally fused to an *N*-terminal TIR (toll/interleukin-1 receptor) domain (found in *Arabidopsis* and tomato), an AIG1 (avrRpt2-induced gene) domain (identified in *Arabidopsis*) or a kinase domain (found in rice; Delporte et al., 2014b).

At present, only few Nictaba-related proteins have been studied for their biological properties and physiological role. A comparative analysis of the carbohydrate binding properties of Nictaba from *N. tabacum*, the Cucurbitaceae phloem lectin PPL, the *A. thaliana* homolog PP2-A1 and the *A. thaliana* F-box-Nictaba protein encoded by the gene *At2g02360* revealed that despite the sequence similarity and the presence of conserved amino acids in the carbohydrate binding site, different Nictaba domains can interact with different glycan motifs (Delporte et al., 2014b), suggesting different biological roles.

Since insect herbivory by Lepidopteran insects also triggers the JA pathway the tobacco lectin also accumulates after caterpillar attack. Furthermore, feeding assays with (transgenic) tobacco lines demonstrated the entomotoxic activity of Nictaba. It was suggested that the entomotoxic effect of Nictaba is caused by interaction of the lectin with glycoconjugates present in the digestive tract of the insect (Vandenborre et al., 2010, 2011). Within the plant cell, insect herbivory results in enhanced Nictaba expression in the cytoplasm, followed by partial translocation of the lectin to the nucleus, where it can interact with core histone proteins through their O-GlcNAc modification. It is hypothesized that Nictaba binding to chromatin results in enhanced transcription of defense related genes (Lannoo and Van Damme, 2010).

The Cucurbitaceae phloem lectins are a group of Nictaba-related lectins that are found in phloem exudates of different Cucurbitaceae species. Unlike Nictaba, the PP2 lectins are exclusively and continuously expressed in the companion cells of the phloem and then translocated into the phloem sap. The pumpkin lectin PPL and the PP2-like protein from *A. thaliana*, PP2-A1, show high binding affinity for chitin oligomers ($\text{GlcNAc}_3\text{-}_6$). Similar to Nictaba, PP2-A1 also binds with the $\text{Man}_3\text{GlcNAc}_2$

core of high-mannose *N*-glycans (Beneteau et al., 2010). Interestingly, the expression of PP2-A1 was enhanced by ethylene treatment and *Pseudomonas* infection. Furthermore, PP2-A1 represses phloem feeding of the green peach aphid *Myzus persicae* (Zhang et al., 2011), displayed antifungal activity against various fungal strains (Lee et al., 2014), and negatively affects Cucurbit aphid borne yellow virus transmission (Bencharki et al., 2010), strongly supporting a role in plant defense.

Several F-box Nictaba proteins are encoded in the *A. thaliana* genome. Glycan array analyses revealed the F-box protein encoded by *At2g02360* exhibits carbohydrate binding activity toward *N*- and *O*-glycans with *N*-acetyllactosamine (LacNAc; $\text{Gal}\beta 1\text{-}3\text{GlcNAc}$ and $\text{Gal}\beta 1\text{-}4\text{GlcNAc}$) and poly-LacNAc structures as well as with Lewis A ($\text{Gal}\beta 1\text{-}3(\text{Fuc}\alpha 1\text{-}4)\text{GlcNAc}$), Lewis X ($\text{Gal}\beta 1\text{-}4(\text{Fuc}\alpha 1\text{-}3)\text{GlcNAc}$), Lewis Y ($\text{Fuc}\alpha 1\text{-}2\text{Gal}\beta 1\text{-}4(\text{Fuc}\alpha 1\text{-}3)\text{GlcNAc}$), and type-1 B motifs ($\text{Gal}\alpha 1\text{-}3(\text{Fuc}\alpha 1\text{-}2)\text{Gal}\beta 1\text{-}3\text{GlcNAc}$). Since these glycan structures have been reported in bacteria, viruses, and animals rather than in plants the physiological importance of this glycan interaction remains enigmatic (Lannoo et al., 2008; Stefanowicz et al., 2012). Furthermore, the same *Arabidopsis* F-box Nictaba protein was shown to interact with core members of an E3-type ubiquitin ligase complex which resulted in the hypothesis assuming a role of the Nictaba domain in glycoprotein degradation (Takahashi et al., 2004; Arabidopsis Interactome Mapping Consortium, 2011).

Ricin-B lectins. The ricin-B lectin family is one of the most widespread families of carbohydrate binding proteins in nature. The most famous member of this family is ricin, a toxic protein from castor bean (*Ricinus communis* L.) seeds, which was the very first lectin discovered in plants by Peter Hermann Stillmark in 1888. Ricin is a chimeric protein consisting of an A chain with enzymatic activity linked through a disulfide bridge with a B chain with lectin activity. This B chain consists of a duplicated ricin-B domain, responsible for the carbohydrate binding activity of the protein toward galactosylated structures. The enzymatic activity of ricin involves RNA *N*-glycosidase activity and is responsible for the removal of a highly conserved adenine residue from the sarcin/ricin loop of the 28S ribosomal RNA. As a result, the ribosomes are no longer able to bind elongation factor 2 and protein synthesis is blocked. Because of their catalytic activity these chimeric proteins are also referred to as type 2 ribosome-inactivating proteins (RIPs), and are considered as toxic proteins if they succeed in entering the cell. The uptake of the protein by the host cell is aided by their lectinic B-chain which can specifically interact with glycoconjugate structures on the cell surface. Except for ricin and abrin (from the jequirity bean *Abrus precatorius*), most type 2 RIPs are only moderately or even weakly toxic (Van Damme et al., 2001; Stirpe and Battelli, 2006).

The family of ricin-B related lectins is widespread in the plant kingdom and has been characterized in detail for what concerns its biological activity and toxicity in several plant species, especially *R. communis* (castor bean), *Abrus precatorius* (jequirity bean), *Viscum album* (mistletoe), and *Sambucus nigra* (elderberry). Unlike the other soluble lectins described above, most ricin-B related proteins accumulate in the plant vacuole or are secreted to the extracellular space (Van Damme et al., 2008). Over the years the ricin-B domain

was identified in numerous plants, animals, fungi, and bacteria. All these proteins with ricin-B domains are also classified as the R-type lectins (Cummings and Etzler, 2009).

Within the genus *Sambucus* (elderberry) an extended family of ricin-B related proteins, including several chimolectins and hololectins has been identified (Shang and Van Damme, 2014). Detailed hapten inhibition assays and glycan array studies revealed that all these *S. nigra* proteins exhibit different carbohydrate binding properties, and allowed to classify the *Sambucus* lectins into three groups. A first group covers the lectins SNA-II and SNA-IV as well as the type 2 RIP SNA-V with specificity toward Gal/GalNAc and Gal/GalNAc-containing glycan structures. The second group comprises only the type 2 RIP SNA-I, which specifically interacts with terminal sialic acid residues (Neu5Ac; α 2–6) linked to Gal/GalNAc. Finally, the type 2 RIP SNLRP represents a third specificity group with strong interaction with GlcNAc oligomers (Shang and Van Damme, 2014).

Several lines of evidence support the idea that ricin-B related lectins play a role in plant defense against pathogens (Chen et al., 2002b; Vandenbussche et al., 2004a,b) and insects (Wei et al., 2004; Shahidi-Noghabi et al., 2009). Over-expressing SNA-I' or SNA-V from *S. nigra* in transgenic tobacco (Samsun NN) plants enhances the tobacco plant's resistance against infection with tobacco mosaic virus. Though the antiviral effect is clearly related to the amount of protein expressed it cannot be related to an increased expression of pathogenesis-related proteins (Chen et al., 2002b; Vandenbussche et al., 2004a,b). Furthermore, no clear correlation was observed between *in planta* antiviral activity of the transgenic tobacco lines and the *in vitro* N-glycosidase activity of the proteins toward genomic RNA of the tobacco mosaic virus, suggesting that the *in planta* antiviral activity of these RIPs may rely on a direct interaction with the virus (Vandenbussche et al., 2004a). At present the importance of the lectin activity in the antiviral activity of the proteins remains unclear.

The first evidence for the insecticidal activity of ricin-B related proteins came from feeding assays with ricin and cinnamomin (from *Cinnamomum camphora* tree). Ricin showed strong toxicity to several insects such as cowpea weevil (*Callosobruchus maculatus*), cotton boll weevil (*Anthonomus grandis*), house fly (*Musca domestica*), and larvae of the silkworm *Bombyx mori* (Wei et al., 2004). The differences in activity between ricin and cinnamomin could not be related to the enzymatic activity but rather were attributed to differences in the activity of the lectin chain of the proteins (Wei et al., 2004). Shahidi-Noghabi et al. (2009) reported the enhanced resistance of transgenic tobacco plants overexpressing SNA-I or its isoform SNA-I' toward different pest insect species including aphids and caterpillars. Since mutation of the carbohydrate binding site can abolish or reduce the toxic effect, the entomotoxic properties of the proteins can be linked to their carbohydrate binding activity (Shahidi-Noghabi et al., 2008). In addition, the cytotoxic effects of *S. nigra* RIPs toward insect cells was accompanied by caspase 3-like protease-induced apoptosis (Shahidi-Noghabi et al., 2008, 2011). More research is needed to identify the interacting proteins for the *Sambucus* lectins on the cell surface.

CONCLUSION

Plant genomes encode a plethora of RLKs, RLPs, and lectins to protect themselves against the vast array of pathogenic bacteria, viruses, fungi, oomycetes, and pest insects. A key feature of the plant's innate immunity is the ability to recognize D/P/MAMPs of potential pathogens through PRRs, and subsequently respond in a highly sensitive and specific manner. Many advances have been made in the understanding how different proteins function in plant innate immunity. As more structural and biochemical data become available, common themes are emerging on receptor organization, ligand perception and binding, receptor activation, and intracellular defense signaling. It is clear now that PRRs are ultra-dynamic multiprotein structures which often use phosphorylation to activate downstream signaling.

A first interaction between the pathogen and the plant occurs at the level of the cell wall and the plasma membrane where extracellular effectors, DAMPs, and P/MAMPs will be recognized by membrane-bound receptors (Figure 1; Tables 1 and 2), among which a large group of RLPs and RLKs some of which carry an extracellular lectin domain. Though different lectin motifs have been recognized, only the LysM domain was unambiguously shown to be dependent on carbohydrate interactions for recognition of fungal and bacterial components and subsequent signal transmission into the plant cell. Interestingly, the LysM motif shows high specificity for chito-oligosaccharides, an abundant component in different pathogens but absent from plants. Other (lectin) receptor kinases/proteins probably depend on protein–protein interactions to recognize specific ligands. Though our knowledge on receptor kinase function and signaling has greatly improved, several issues still remain with respect to the potential ligands for pathogen recognition and the downstream signaling events, especially for those receptors lacking the kinase domain.

In addition to the lectin motifs present at the level of the cell wall/plasma membrane, plants synthesize well-defined soluble lectins upon exposure to multiple abiotic and biotic stresses (Table 3). Most of these inducible lectins reside in the nucleus and the cytosol of the plant cell and evidence is emerging for their role in signal transduction inside the plant cell as part of multiple plant defense pathways. Hence protein–carbohydrate interactions should not only be envisaged at the level of the interaction between the pathogen and the plant cell, but also play an important role inside the plant cell as part of the intracellular signaling resulting from the recognition of plant attackers. At least for some cytoplasmic lectins (Nictaba-related proteins, EUL-related proteins, amaranthins) it was shown that they are also translocated inside the plant nucleus. The Nictaba-related proteins in particular have been shown to interact with glycosylated histones, and therefore are suggested to act as chromatin remodelers, enabling altered gene expression as a result of stress signaling. Surveying the plant genome sequences also revealed that most lectin domains are part of larger proteins, consisting of one or more lectin domains linked to un-related protein domains, most often with unknown functions. Future challenges include the characterization of the ligands for these soluble lectins or lectin domains present as part of a larger protein to elucidate the biological relevance of these interactions.

Large-scale experiments integrating genomics, biochemistry, cell biology, structural biology, and bioinformatics will enable to elucidate the physiological importance of the lectin motifs in protein–carbohydrate interactions in signal transduction and plant defense.

ACKNOWLEDGMENTS

This work was supported by the Fund for Scientific Research – Flanders and the Research Council of Ghent University. N. Lannoo acknowledges the receipt of a Postdoctoral fellowship from the Fund for Scientific Research – Flanders.

REFERENCES

- Al Atalah, B., De Vleesschauwer, D., Xu, J., Fouquaert, E., Höfte, M., and Van Damme, E. J. (2014a). Transcriptional behavior of EUL-related rice lectins toward important abiotic and biotic stresses. *J. Plant Physiol.* 171, 986–992. doi: 10.1016/j.jplph.2014.04.004
- Al Atalah, B., Vanderschaeghe, D., Bloch, Y., Proost, P., Plas, K., Callewaert, N., et al. (2014b). Characterization of a type D1A EUL-related lectin from rice expressed in *Pichia pastoris*. *Biol. Chem.* 395, 413–424. doi: 10.1515/hsz-2013-0267
- Al Atalah, B., Fouquaert, E., Vanderschaeghe, D., Proost, P., Balzarini, J., Smith, D. F., et al. (2011). Expression analysis of the nucleocytoplasmic lectin “Oryysata” from rice in *Pichia pastoris*. *FEBS J.* 278, 2064–2079. doi: 10.1111/j.1742-4658.2011.08122.x
- Al Atalah, B., Rougé, P., Smith, D. F., Proost, P., Lasanajak, Y., and Van Damme, E. J. M. (2012). Expression analysis of a type S2 EUL-related lectin from rice in *Pichia pastoris*. *Glycoconj. J.* 29, 467–479. doi: 10.1007/s10719-012-9405-2
- Albert, M. (2013). Peptides as triggers of plant defence. *J. Exp. Bot.* 64, 5269–5279. doi: 10.1093/jxb/ert275
- Arabidopsis Interactome Mapping Consortium. (2011). Evidence for network evolution in an *Arabidopsis* interactome map. *Science* 333, 601–607. doi: 10.1126/science.1203877
- Bar, M., Sharfman, M., Ron, M., and Avni, A. (2010). BAK1 is required for the attenuation of ethylene-inducing xylanase (Eix)-induced defense responses by the decoy receptor LeEix1. *Plant J.* 63, 791–800. doi: 10.1111/j.1365-313X.2010.04282.x
- Benchmarki, B., Boissinot, S., Revillon, S., Ziegler-Graff, V., Erdinger, M., Wiss, L., et al. (2010). Phloem protein partners of Cucurbit aphid borne yellows virus: possible involvement of phloem proteins in virus transmission by aphids. *Mol. Plant Microbe Interact.* 23, 799–810. doi: 10.1094/MPMI-23-6-0799
- Beneteau, J., Renard, D., Marché, L., Douville, E., Lavenant, L., Rahbé, Y., et al. (2010). Binding properties of the N-acetylglucosamine and high-mannose N-glycan PP2-A1 phloem lectin in *Arabidopsis*. *Plant Physiol.* 153, 1345–1361. doi: 10.1104/pp.110.153882
- Berendzen, K. W., Böhmer, M., Wallmeroth, N., Peter, S., Vesić, M., Zhou, Y., et al. (2012). Screening for in planta protein–protein interactions combining bimolecular fluorescence complementation with flow cytometry. *Plant Methods* 8:25. doi: 10.1186/1746-4811-8-25
- Block, A., and Alfano, J. R. (2011). Plant targets for *Pseudomonas syringae* type III effectors: virulence targets or guarded decoys? *Curr. Opin. Microbiol.* 14, 39–46. doi: 10.1016/j.mib.2010.12.011
- Block, A., Toruño, T. Y., Elowsky, C. G., Zhang, C., Steinbrenner, J., Beynon, J., et al. (2013). The *Pseudomonas syringae* type III effector HopD1 suppresses effector-triggered immunity, localizes to the endoplasmic reticulum, and targets the *Arabidopsis* transcription factor NTL9. *New Phytol.* 201, 1358–1370. doi: 10.1111/nph.12626
- Böhm, H., Albert, I., Fan, L., Reinhard, A., and Nürnberger, T. (2014). Immune receptor complexes at the plant cell surface. *Curr. Opin. Plant Biol.* 20, 47–54. doi: 10.1016/j.pbi.2014.04.007
- Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* 60, 379–406. doi: 10.1146/annurev.arplant.57.032905.105346
- Bouwmeester, K., de Sain, M., Weide, R., Gouget, A., Klamer, S., Canut, H., et al. (2011). The lectin receptor kinase LecRK-I.9 is a novel Phytophthora resistance component and a potential host target for a RXLR effector. *PLoS Pathog.* 7:e1001327. doi: 10.1371/journal.ppat.1001327
- Bouwmeester, K., and Govers, F. (2009). *Arabidopsis* L-type lectin receptor kinases: phylogeny, classification, and expression profiles. *J. Exp. Bot.* 60, 4383–4396. doi: 10.1093/jxb/erp277
- Brotman, Y., Landau, U., Pnini, S., Liseck, J., Balazadeh, S., Mueller-Roeber, B., et al. (2012). The LysM receptor-like kinase LysM RLK1 is required to activate defense and abiotic-stress responses induced by overexpression of fungal chitinases in *Arabidopsis* plants. *Mol. Plant* 5, 1113–1124. doi: 10.1093/mp/sss021
- Buist, G., Steen, A., Kok, J., and Kuipers, O. P. (2008). LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.* 68, 838–847. doi: 10.1111/j.1365-2958.2008.06211.x
- Cambi, A., Koopman, M., and Figdor, C. G. (2005). How C-type lectins detect pathogens. *Cell. Microbiol.* 7, 481–488. doi: 10.1111/j.1462-5822.2005.00506.x
- Cao, Y., Tanaka, K., Nguyen, C. T., and Stacey, G. (2014). Extracellular ATP is a central signaling molecule in plant stress responses. *Curr. Opin. Plant Biol.* 20, 82–87. doi: 10.1016/j.pbi.2014.04.009
- Chen, X., Zuo, S., Schwessinger, B., Chern, M., Canlas, P. E., Ruan, D., et al. (2014). An XA21-associated kinase (OsSERK2) regulates immunity mediated by the XA21 and XA3 immune receptors. *Mol. Plant* 7, 874–892. doi: 10.1093/mp/ssu003
- Chen, Y., Peumans, W. J., Hause, B., Bras, J., Kumar, M., Proost, P., et al. (2002a). Jasmonic acid methyl ester induces the synthesis of a cytoplasmic/nuclear chitooligosaccharide-binding lectin in tobacco leaves. *FASEB J.* 16, 905–907.
- Chen, Y., Peumans, W. J., and Van Damme, E. J. M. (2002b). The *Sambucus nigra* type-2 ribosome-inactivating protein SNA-I' exhibits in planta antiviral activity in transgenic tobacco. *FEBS Lett.* 516, 27–30. doi: 10.1016/S0014-5793(02)02455-9
- Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell* 18, 465–476. doi: 10.1105/tpc.105.036574
- Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J. D., et al. (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* 448, 497–500. doi: 10.1038/nature05999
- Choi, J., Tanaka, K., Cao, Y., Qi, Y., Qiu, J., Liang, Y., et al. (2014). Identification of a plant receptor for extracellular ATP. *Science* 343, 290–294. doi: 10.1126/science.343.6168.290
- Cui, F., Wu, S., Sun, W., Coaker, G., Kunkel, B., He, P., et al. (2013). The *Pseudomonas syringae* type III effector AvrRpt2 promotes pathogen virulence via stimulating *Arabidopsis* auxin/indole acetic acid protein turnover. *Plant Physiol.* 162, 1018–1029. doi: 10.1104/pp.113.219659
- Cummings, R. D., and Etzler, M. E. (2009). “R-type lectins,” in *Essentials of Glycobiology*, 2nd Edn, Chap. 28, eds E. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, et al. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Dangl, J. L., Horvath, D. M., and Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science* 341, 746–751. doi: 10.1126/science.1236011
- Degiocomi, M. T., Iacovache, I., Pernot, L., Chami, M., Kudryashev, M., Stahlberg, H., et al. (2013). Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. *Nat. Chem. Biol.* 9, 623–629. doi: 10.1038/nchembio.1312
- Delporte, A., De Vos, W. H., and Van Damme, E. J. M. (2014a). In vivo interaction between the tobacco lectin and the core histone proteins. *J. Plant Physiol.* 171, 986–992. doi: 10.1016/j.jplph.2014.04.008
- Delporte, A., Van Holle, S., Lannoo, N., and Van Damme, E. J. M. (2014b). The tobacco lectin, prototype of the family of Nictaba-related proteins. *Curr. Prot. Pept. Sci.* (in press).
- de Souza Filho, G. A., Ferreira, B. S., Dias, J. M. R., Queiroz, K. S., Branco, A. T., Bressan-Smith, R. E., et al. (2003). Accumulation of SALT protein in rice plants as a response to environmental stresses. *Plant Sci.* 164, 623–628. doi: 10.1016/S0168-9452(03)00014-1
- Dubery, I. A., Sanabria, N. M., and Huang, J.-C. (2012). “Chapter 6: nonself perception in plant innate immunity,” in *Self and Nonself*, ed. C. López-Larrea (Austin, TX: Landes Bioscience; New York, NY: Springer Science+Business Media).

- Ellgaard, L., and Frickel, E. M. (2003). Calnexin, calreticulin, and ERp57: teammates in glycoprotein folding. *Cell Biochem. Biophys.* 39, 223–247. doi: 10.1385/CBB:39:3:223
- Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., and Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature* 417, 962–966. doi: 10.1038/nature00842
- Esen, A., and Blanchard, D. J. (2000). A specific β -glucosidase-aggregating factor is responsible for the β -glucosidase null phenotype in maize. *Plant Physiol.* 122, 563–572. doi: 10.1104/pp.122.2.563
- Feng, H., Xu, W.-Z., Lin, H.-H., and Chong, K. (2009). Transcriptional regulation of wheat VERT2 promoter in rice in response to abscisic acid, jasmonate, and light. *J. Genet. Genomics* 36, 371–377. doi: 10.1016/S1673-8527(08)60126-5
- Fouquaert, E., Peumans, W. J., Smith, D. F., Proost, P., Savvides, S. N., and Van Damme, E. J. M. (2008). The “old” *Euonymus europaeus* agglutinin represents a novel family of ubiquitous plant proteins. *Plant Physiol.* 147, 1316–1324. doi: 10.1104/pp.108.116764
- Fouquaert, E., and Van Damme, E. J. M. (2012). Promiscuity of the euonymus carbohydrate-binding domain. *Biomolecules* 2, 415–434. doi: 10.3390/biom2040415
- Fradin, E. F., Zhang, Z., Juarez Ayala, J. C., Castroverde, C. D., Nazar, R. N., Robb, J., et al. (2009). Genetic dissection of *Verticillium* wilt resistance mediated by tomato Ve1. *Plant Physiol.* 150, 320–332. doi: 10.1104/pp.109.136762
- Gilardoni, P. A., Hettenhausen, C., Baldwin, I. T., and Bonaventure, G. (2011). *Nicotiana attenuata* LECTIN RECEPTOR KINASE 1 suppresses the insect-mediated inhibition of induced defense responses during *Manduca sexta* herbivory. *Plant Cell* 23, 3512–3532. doi: 10.1105/tpc.111.088229
- Gómez-Gómez, L., Bauer, Z., and Boller, T. (2001). Both the extracellular leucine-rich repeat domain and the kinase activity of FLS2 are required for flagellin binding and signaling in *Arabidopsis*. *Plant Cell* 13, 1155–1163. doi: 10.1105/tpc.13.5.1155
- Grant, S. R., Fisher, E. J., Chang, J. H., Mole, B. M., and Dangl, J. L. (2006). Suberfuge and manipulation: type III effector proteins of phytopathogenic bacteria. *Annu. Rev. Microbiol.* 60, 425–449. doi: 10.1146/annurev.micro.60.080805.142251
- Gudesblat, G. E., Torres, P. S., and Vojnov, A. A. (2009). *Xanthomonas campestris* overcomes *Arabidopsis* stomatal innate immunity through a DSF cell-to-cell signal-regulated virulence factor. *Plant Physiol.* 149, 1017–1027. doi: 10.1104/pp.108.126870
- Gust, A. A., Willmann, R., Desaki, Y., Grabherr, H. M., and Nürnberg, T. (2012). Plant LysM proteins: modules mediating symbiosis and immunity. *Trends Plant Sci.* 17, 495–502. doi: 10.1016/j.tplants.2012.04.003
- Halter, T., Imkampe, J., Mazzotta, S., Wierzba, M., Postel, S., Bücherl, C., et al. (2014). The leucine-rich repeat receptor kinase BIR2 is a negative regulator of BAK1 in plant immunity. *Curr. Biol.* 24, 134–143. doi: 10.1016/j.cub.2013.11.047
- Han, Z., Sun, Y., and Chai, J. (2014). Structural insight into the activation of plant receptor kinases. *Curr. Opin. Plant Biol.* 20, 55–63. doi: 10.1016/j.pbi.2014.04.008
- Häweker, H., Rips, S., Koiwa, H., Salomon, S., Saito, Y., Chinchilla, D., et al. (2010). Pattern recognition receptors require N-glycosylation to mediate plant immunity. *J. Biol. Chem.* 285, 4629–4636. doi: 10.1074/jbc.M109.063073
- Hayafune, M., Berisio, R., Marchetti, R., Silipo, A., Kayama, M., Desaki, Y., et al. (2014). Chitin-induced activation of immune signaling by the rice receptor CEBiP relies on a unique sandwich-type dimerization. *Proc. Natl. Acad. Sci. U.S.A.* 111, E404–E413. doi: 10.1073/pnas.1312099111
- Howell, S. H. (2013). Endoplasmic reticulum stress responses in plants. *Annu. Rev. Plant Biol.* 64, 477–499. doi: 10.1146/annurev-plant-050312-120053
- Huang, P., Ju, H.-W., Min, J.-H., Zhang, X., Kim, S.-H., Yang, K.-Y., et al. (2013). Overexpression of L-type lectin-like protein kinase 1 confers pathogen resistance and regulates salinity response in *Arabidopsis thaliana*. *Plant Sci.* 203–204, 98–106. doi: 10.1016/j.plantsci.2012.12.019
- Huang, P.-Y., Yeh, Y.-H., Liu, A.-C., Cheng, C.-P., and Zimmerli, L. (2014). The *Arabidopsis* LeCRK-VL2 associates with the pattern-recognition receptor FLS2 and primes *Nicotiana benthamiana* pattern-triggered immunity. *Plant J.* 79, 243–255. doi: 10.1111/tpj.12557
- Jiang, J. F., Han, Y., Xing, L. J., Xu, Y. Y., Xu, Z. H., and Chong, K. (2006). Cloning and expression of a novel cDNA encoding a mannose-specific jacalin-related lectin from *Oryza sativa*. *Toxicon* 47, 133–139. doi: 10.1016/j.toxicon.2005.10.010
- Jiang, J. F., Xu, Y. Y., and Chong, K. (2007). Overexpression of OsJAC1, a lectin gene, suppresses the coleoptiles and stem elongation in rice. *J. Integr. Plant Biol.* 49, 230–237. doi: 10.1111/j.1744-7909.2007.00428.x
- Jin, H., Hong, Z., Su, W., and Li, J. (2009). A plant-specific calreticulin is a key retention factor for a defective brassinosteroid receptor in the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13612–13617. doi: 10.1073/pnas.0906144106
- Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiya, C., Dohmae, N., Takio, K., et al. (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11086–11091. doi: 10.1073/pnas.0508882103
- Kanagawa, M., Liu, Y., Hanashima, S., Ikeda, A., Chai, W., Nakano, Y., et al. (2014). Structural basis for multiple sugar recognition of jacalin-related human ZG16p lectin. *J. Biol. Chem.* 289, 16954–16965. doi: 10.1074/jbc.M113.539114
- Kanzaki, H., Saitoh, H., Takahashi, Y., Berberich, T., Ito, A., Kamoun, S., et al. (2008). NbLRK1, a lectin-like receptor kinase protein of *Nicotiana benthamiana*, interacts with *Phytophthora infestans* INF1 elicitor and mediates INF1-induced cell death. *Planta* 228, 977–987. doi: 10.1007/s00425-008-0797-y
- Kapoor, M., Ellgaard, L., Gopalakrishnapai, J., Schirra, C., Gemma, E., Oscarsson, S., et al. (2004). Mutational analysis provides molecular insight into the carbohydrate-binding region of calreticulin: pivotal roles of tyrosine-109 and aspartate-135 in carbohydrate recognition. *Biochemistry* 43, 97–106. doi: 10.1021/bi0355286
- Kim, Y. T., Oh, J., Kim, K. H., Uhm, J. Y., and Lee, B. M. (2009). Isolation and characterization of NgRLK1, a receptor-like kinase of *Nicotiana glutinosa* that interacts with the elicitor of *Phytophthora capsici*. *Mol. Biol. Rep.* 37, 717–727. doi: 10.1007/s11033-009-9570-y
- Kittur, F. S., Yu, H. Y., Bevan, D. R., and Esen, A. (2009). Homolog of the maize β -glucosidase aggregating factor from sorghum is a jacalin-related GalNAc-specific lectin but lacks protein aggregating activity. *Glycobiology* 19, 277–287. doi: 10.1093/glycob/cwn132
- Knogge, W., and Scheel, D. (2006). LysM receptors recognize friend and foe. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10829–10830. doi: 10.1073/pnas.0604601103
- Kouzai, Y., Nakajima, K., Hayafune, M., Ozawa, K., Kaku, H., Shibuya, N., et al. (2014). CEBiP is the major chitin oligomers-binding protein in rice and plays a main role in the perception of chitin oligomers. *Plant Mol. Biol.* 84, 519–528. doi: 10.1007/s11103-013-0149-6
- Krol, E., Mentzel, T., Chinchilla, D., Boller, T., Felix, G., Kemmerling, B., et al. (2010). Perception of the *Arabidopsis* danger signal peptide 1 involves the pattern recognition receptor AtPEPR1 and its close homologue AtPEPR2. *J. Biol. Chem.* 285, 13471–13479. doi: 10.1074/jbc.M109.07394
- Lannoo, N., Peumans, W. J., and Van Damme, E. J. M. (2008). Do F-box proteins with a C-terminal domain homologous with the tobacco lectin play a role in protein degradation in plants? *Biochem. Soc. Trans.* 36, 843–847. doi: 10.1042/BST0360843
- Lannoo, N., Peumans, W. J., Van Pamel, E., Alvarez, R., Xiong, T., Hause, G., et al. (2006). Localization and *in vitro* binding studies suggest that the cytoplasmic/nuclear tobacco lectin can interact *in situ* with high-mannose and complex N-glycans. *FEBS Lett.* 580, 6329–6337. doi: 10.1016/j.febslet.2006.10.044
- Lannoo, N., and Van Damme, E. J. M. (2010). Nucleocytoplasmic plant lectins. *Biochem. Biophys. Acta* 1800, 190–201. doi: 10.1016/j.bbagen.2009.07.021
- Lannoo, N., Vandeborre, G., Miersch, O., Smagghe, G., Wasternack, C., Peumans, W. J., et al. (2007). The jasmonate-induced expression of the *Nicotiana tabacum* leaf lectin. *Plant Cell Physiol.* 48, 1207–1218. doi: 10.1093/pcp/pcm090
- Lee, J. R., Boltz, K. A., and Lee, S. Y. (2014). Molecular chaperone function of *Arabidopsis thaliana* phloem protein 2-A1 encodes a protein similar to phloem lectin. *Biochem. Biophys. Res. Commun.* 443, 18–21. doi: 10.1016/j.bbrc.2013.11.034
- Lee, R. H., Wang, C. H., Huang, L. T., and Chen, S. C. (2001). Leaf senescence in rice plants: cloning and characterization of senescence up-regulated genes. *J. Exp. Bot.* 52, 1117–1121. doi: 10.1093/jexbot/52.358.1117
- Lee, S. W., Han, S. W., Sririanum, M., Park, C. J., Seo, Y. S., and Ronald, P. C. (2009). A type I-secreted, sulfated peptide triggers XA21-mediated innate immunity. *Science* 326, 850–853. doi: 10.1126/science.1173438

- Li, D., Wang, X., Yuan, D., Zhang, L., Jiang, X., Tao, Z., et al. (2014). Over-expression of ArathEULS3 confers ABA sensitivity and drought tolerance in *Arabidopsis*. *Plant Cell Tissue Organ Cult.* 117, 431–442. doi: 10.1007/s11240-014-0453-0
- Li, J., Zhao-Hui, C., Batoux, M., Nekrasov, V., Roux, M., Chinchilla, D., et al. (2009). Specific ER quality control components required for biogenesis of the plant innate immune receptor EFR. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15973–15978. doi: 10.1073/pnas.0905532106
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* 302, 630–633. doi: 10.1126/science.1090074
- Liu, B., Li, J. F., Ao, Y., Qu, J., Li, Z., Su, J., et al. (2012a). Lysin motif-containing proteins LYP4 and LYP6 play dual roles in peptidoglycan and chitin perception in rice innate immunity. *Plant Cell* 24, 3406–3419. doi: 10.1105/tpc.112.102475
- Liu, T., Liu, Z., Song, C., Hu, Y., Han, Z., She, J., et al. (2012b). Chitin-induced dimerization activates a plant immune receptor. *Science* 336, 1160–1164. doi: 10.1126/science.1218867
- Liu, Y., and Li, J. (2014). Endoplasmic reticulum-mediated protein quality control in *Arabidopsis*. *Front. Plant Sci.* 5:162. doi: 10.3389/fpls.2014.00162
- Ma, Q. H., Tian, B., and Li, Y. L. (2010). Overexpression of a wheat jasmonate-regulated lectin increases pathogen resistance. *Biochimie* 92, 187–193. doi: 10.1016/j.biochi.2009.11.008
- Ma, Q. H., Zhen, W. B., and Liu, Y. C. (2013). Jacalin domain in wheat jasmonate-regulated protein Ta-JA1 confers agglutinating activity and pathogen resistance. *Biochimie* 95, 359–365. doi: 10.1016/j.biochi.2012.10.014
- Macho, A. P., and Zipfel, C. (2014). Plant PRRs and the activation of innate immune signaling. *Mol. Cell* 54, 263–272. doi: 10.1016/j.molcel.2014.03.028
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczylowski, K., et al. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* 425, 637–640. doi: 10.1038/nature02045
- Malinovsky, F. G., Fangel, J. U., and Willats, W. G. T. (2014). The role of the cell wall in plant immunity. *Front. Plant Sci.* 5:178. doi: 10.3389/fpls.2014.00178
- Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S. Y. (2006). Plant stomata function in innate immunity against bacterial invasion. *Cell* 126, 969–980. doi: 10.1016/j.cell.2006.06.054
- Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., et al. (2007). CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19613–19618. doi: 10.1073/pnas.0705147104
- Mulder, L., Lefebvre, B., Cullimore, J., and Imbert, A. (2006). LysM domains of *Medicago truncatula* NFP protein involved in Nod factor perception. Glycosylation state, molecular modeling and docking of chitoooligosaccharides and Nod factors. *Glycobiology* 16, 801–809. doi: 10.1093/glycob/cwl006
- Muthamilarasan, M., and Prasad, M. (2013). Plant innate immunity: an updated insight into defense mechanism. *J. Biosci.* 38, 433–449. doi: 10.1007/s12038-013-9302-2
- Nagano, A. J., Fukao, Y., Fujiwara, M., Nishimura, M., and Hara-Nishimura, I. (2008). Antagonistic jacalin-related lectins regulate the size of ER body-type β -glucosidase complexes in *Arabidopsis thaliana*. *Plant Cell Physiol.* 49, 969–980. doi: 10.1093/pcp/pcn075
- Navarro-Gochicoa, M. T., Camut, S., Timmers, A. C. J., Niebel, A., Herye, C., Boutet, E., et al. (2003). Characterization of four lectin-like receptor kinases expressed in roots of *Medicago truncatula* structure, location, regulation of expression, and potential role in the symbiosis with *Sinorhizobium meliloti*. *Plant Physiol.* 133, 1893–1910. doi: 10.1104/pp.103.027680
- Newman, M.-A., Sundelin, T., Nielsen, J. T., and Erbs, G. (2013). MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Front. Plant Sci.* 4:139. doi: 10.3389/fpls.2013.00139
- Nühse, T. S. (2012). Cell wall integrity signaling and innate immunity in plants. *Front. Plant Sci.* 3:280. doi: 10.3389/fpls.2012.00280
- Paparella, C., Savatin, D. V., Marti, L., De Lorenzo, G., and Ferrari, S. (2014). The *Arabidopsis* LYSIN MOTIF-CONTAINING RECEPTOR-LIKE KINASE 3 regulates the cross talk between immunity and abscisic acid responses. *Plant Physiol.* 165, 262–276. doi: 10.1104/pp.113.233759
- Park, C. J., and Ronald, P. C. (2012). Cleavage and nuclear localization of the rice XA21 immune receptor. *Nat. Commun.* 3:920. doi: 10.1038/ncomms1932
- Petutschnig, E. K., Jones, A. M., Serazetdinova, L., Lipka, U., and Lipka, V. (2010). The lysin motif receptor-like kinase (LysM-RLK) CERK1 is a major chitin-binding protein in *Arabidopsis thaliana* and subject to chitin-induced phosphorylation. *J. Biol. Chem.* 285, 28902–28911. doi: 10.1074/jbc.M110.116657
- Peumans, W. J., and Van Damme, E. J. M. (1995). Lectins as plant defense proteins. *Plant Physiol.* 109, 347–352. doi: 10.1104/pp.109.2.347
- Puthoff, D. P., Sardesai, N., Subramanyam, S., Nemacheck, J. A., and Williams, C. E. (2005). Hfr-2, a wheat cytolytic toxin-like gene, is up-regulated by virulent Hessian fly larval feeding. *Mol. Plant Pathol.* 6, 411–423. doi: 10.1111/j.1364-3703.2005.00289.x
- Qin, Q. M., Zhang, Q., Zhao, W. S., Wang, Y. Y., and Peng, Y. L. (2003). Identification of a lectin gene induced in rice in response to *Magnaporthe grisea* infection. *Acta Bot. Sin.* 45, 76–81.
- Qiu, Y., Xi, J., Du, L., Roje, S., and Poovaiah, B. W. (2012). A dual regulatory role of *Arabidopsis* calreticulin-2 in plant innate immunity. *Plant J.* 69, 489–500. doi: 10.1111/j.1365-313X.2011.04807.x
- Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., et al. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425, 585–592. doi: 10.1038/nature02039
- Regente, M., Taveira, G. B., Pinedo, M., Elizalde, M. M., Ticchi, A. J., Diz, M. S. S., et al. (2014). A sunflower lectin with antifungal properties and putative medical mycology applications. *Curr. Microbiol.* 69, 88–95. doi: 10.1007/s00284-014-0558-z
- Roberts, D. D., and Goldstein, I. J. (1983). Adenine binding sites of the lectin from lima beans (*Phaseolus lunatus*). *J. Biol. Chem.* 258, 13820–13824.
- Ron, M., and Avni, A. (2004). The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. *Plant Cell* 16, 1604–1615. doi: 10.1105/tpc.022475
- Rotblat, B., Enshell-Seijffers, D., Gershoni, J. M., Schuster, S., and Avni, A. (2002). Identification of an essential component of the elicitation active site of the EIX protein elicitor. *Plant J.* 32, 1049–1055. doi: 10.1046/j.1365-313X.2002.01490.x
- Ryan, C. A. (2000). The systemin signaling pathway: differential activation of plant defensive genes. *Biochim. Biophys. Acta* 1477, 112–121. doi: 10.1016/S0167-4838(99)00269-1
- Scheer, J. M., and Ryan, C. A. Jr. (2002). The systemin receptor SR160 from *Lycopersicon peruvianum* is a member of the LRR receptor kinase family. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9585–9590. doi: 10.1073/pnas.132266499
- Schouuppe, D., Ghesquière, B., Menschaert, G., De Vos, W. H., Bourque, S., Trooskens, G., et al. (2011). Interaction of the tobacco lectin with histone proteins. *Plant Physiol.* 155, 1091–1102. doi: 10.1104/pp.110.170134
- Schulze, B., Mentzel, T., Jehle, A. K., Mueller, K., Beeler, S., Boller, T., et al. (2010). Rapid heteromerization and phosphorylation of ligand-activated plant transmembrane receptors and their associated kinase BAK1. *J. Biol. Chem.* 285, 9444–9451. doi: 10.1074/jbc.M109.096842
- Shahidi-Noghabi, S., Van Damme, E. J. M., De Vos, W. H., and Smagghe, G. (2011). Internalization of *Sambucus nigra* agglutinins I and II in insect midgut cells. *Arch. Insect Biochem. Physiol.* 76, 211–222. doi: 10.1002/arch.20405
- Shahidi-Noghabi, S., Van Damme, E. J. M., and Smagghe, G. (2008). Carbohydrate-binding activity of the type-2 ribosome-inactivating protein SNA-I from elderberry (*Sambucus nigra*) is a determining factor for its insecticidal activity. *Phytochemistry* 69, 2972–2978. doi: 10.1016/j.phytochem.2008.09.012
- Shahidi-Noghabi, S., Van Damme, E. J. M., and Smagghe, G. (2009). Expression of *Sambucus nigra* agglutinin (SNA-I') from elderberry bark in transgenic tobacco plants results in enhanced resistance to different insect species. *Transgenic Res.* 18, 249–259. doi: 10.1007/s11248-008-9215-2
- Shang, C., and Van Damme, E. J. M. (2014). Comparative analysis of carbohydrate binding properties of *Sambucus nigra* lectins and ribosome-inactivating proteins. *Glycoconj. J.* 31, 345–354. doi: 10.1007/s10719-014-9527-952
- Sherman-Broyles, S., Boggs, N., Farkas, A., Liu, P., Vrebalov, J., Nasrallah, M. E., et al. (2007). S locus genes and the evolution of self-fertility in *Arabidopsis thaliana*. *Plant Cell* 19, 94–106. doi: 10.1105/tpc.106.048199
- Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., et al. (2010). Two LysM receptor molecules, CEBIP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J.* 64, 204–214. doi: 10.1111/j.1365-313X.2010.04324.x

- Shinya, T., Motoyama, N., Ikeda, A., Wada, M., Kamiya, K., Hayafune, M., et al. (2012). Functional characterization of CEBiP and CERK1 homologs in *Arabidopsis* and rice reveals the presence of different chitin receptor systems in plants. *Plant Cell Physiol.* 53, 1696–1706. doi: 10.1093/pcp/pcs113
- Singh, P., Kuo, Y. C., Mishra, S., Tsai, C. H., Chien, C. C., Chen, C. W., et al. (2012). The lectin-receptor kinase-VI.2 is required for priming and positively regulates *Arabidopsis* pattern-triggered immunity. *Plant Cell* 24, 1256–1270. doi: 10.1105/tpc.112.095778
- Singh, P., and Zimmerli, L. (2013). Lectin receptor kinases in plant innate immunity. *Front. Plant Sci.* 4:124. doi: 10.3389/fpls.2013.00124
- Smith, J. M., Salamango, D. J., Leslie, M. E., Collins, C. A., and Heese, A. (2014). Sensitivity to Flg22 is modulated by ligand-induced degradation and de novo synthesis of the endogenous flagellin-receptor FLAGELLIN-SENSING2. *Plant Physiol.* 164, 440–454. doi: 10.1104/pp.113.229179
- Song, M., Xu, W., Xiang, Y., Jia, H., Zhang, L., and Ma, Z. (2014). Association of jacalin-related lectins with wheat responses to stresses revealed by transcriptional profiling. *Plant Mol. Biol.* 84, 95–110. doi: 10.1007/s11103-013-0121-5
- Stefanowicz, K., Lannoo, N., Proost, P., and Van Damme, E. J. M. (2012). *Arabidopsis* F-box protein containing a Nictaba-related lectin domain interacts with N-acetyllactosamine structures. *FEBS Open Bio* 2, 151–158. doi: 10.1016/j.fob.2012.06.002
- Stirpe, F., and Battelli, M. G. (2006). Ribosome-inactivating proteins: progress and problems. *Cell. Mol. Life Sci.* 63, 1850–1866. doi: 10.1007/s00018-006-6078-7
- Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., et al. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417, 959–962. doi: 10.1038/nature00841
- Subramanyam, S., Smith, D. F., Clemens, J. C., Webb, M. A., Sardesai, N., and Williams, C. E. (2008). Functional characterization of HFR1, a high-mannose N-glycan-specific wheat lectin induced by Hessian fly larvae. *Plant Physiol.* 147, 1412–1426. doi: 10.1104/pp.108.116145
- Sun, X., Cao, Y., Yang, Z., Xu, C., Li, X., Wang, S., et al. (2004). Xa26, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. *Plant J.* 37, 517–527. doi: 10.1046/j.1365-313X.2003.01976.x
- Sun, X.-L., Yu, Q.-Y., Tang, L.-L., Ji, W., Bai, X., Cai, H., et al. (2013). GsSRK, a G-type lectin S-receptor-like serine/threonine protein kinase, is a positive regulator of plant tolerance to salt stress. *J. Plant Physiol.* 170, 505–515. doi: 10.1016/j.jplph.2012.11.017
- Takahashi, N., Kuroda, H., Kuromori, T., Hirayama, T., Seki, M., Shinozaki, K., et al. (2004). Expression and interaction analysis of *Arabidopsis* Skp1-related genes. *Plant Cell Physiol.* 45, 83–91. doi: 10.1093/pcp/pch009
- Tanaka, K., Nguyen, C. T., Liang, Y., Cao, Y., and Stacey, G. (2013). Role of LysM receptors in chitin-triggered plant innate immunity. *Plant Signal. Behav.* 8:e22598. doi: 10.4161/psb.22598
- Thelin, L., Mutwil, M., Sommarin, M., and Persson, S. (2011). Diverging functions among calreticulin isoforms in higher plants. *Plant Signal. Behav.* 6, 905–910. doi: 10.4161/psb.6.15339
- Thomma, B. P., Nurnberger, T., and Joosten, M. H. (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *Plant Cell* 23, 4–15. doi: 10.1105/tpc.110.082602
- Tintor, N., and Saijo, Y. (2014). ER-mediated control for abundance, quality, and signaling of transmembrane immune receptors in plants. *Front. Plant Sci.* 5:65. doi: 10.3389/fpls.2014.00065
- Vaid, N., Macovei, A., and Tuteja, N. (2013). Knights in action: lectin receptor-like kinases in plant development and stress responses. *Mol. Plant* 6, 1405–1418. doi: 10.1093/mp/sst033
- Vaid, N., Pandey, P. K., and Tuteja, N. (2012). Genome-wide analysis of lectin receptor-like kinase family from *Arabidopsis* and rice. *Plant Mol. Biol.* 80, 365–388. doi: 10.1007/s11103-012-9952-8
- Van Damme, E. J. M., Fouquaert, E., Lannoo, N., Vandenborre, G., Schouppe, D., and Peumans, W. J. (2011). “Novel concepts about the role of lectins in the plant cell,” in *The Molecular Immunology of Complex Carbohydrates-3*, Vol. 705, ed. A. M. Wu (New York: Springer), 295–324.
- Van Damme, E. J. M., Hao, Q., Chen, Y., Barre, A., Vandenbussche, F., Desmyter, S., et al. (2001). Ribosome-inactivating proteins: a family of plant proteins that do more than inactivate ribosomes. *Crit. Rev. Plant Sci.* 20, 395–465. doi: 10.1080/07352689.2001.10131826
- Van Damme, E. J. M., Lannoo, N., and Peumans, W. J. (2008). Plant lectins. *Adv. Bot. Res.* 48, 107–209. doi: 10.1016/S0065-2296(08)00403-5
- Vandenborre, G., Groten, K., Smaghe, G., Lannoo, N., Baldwin, I. T., and Van Damme, E. J. M. (2010). *Nicotiana tabacum* agglutinin is active against Lepidopteran pest insects. *J. Exp. Bot.* 61, 1003–1014. doi: 10.1093/jxb/erp365
- Vandenborre, G., Smaghe, G., and Van Damme, E. J. M. (2011). Plant lectins as defense proteins against phytophagous insects. *Phytochemistry* 72, 1538–1550. doi: 10.1016/j.phytochem.2011.02.024
- Vandenbussche, F., Desmyter, S., Ciani, M., Proost, P., Peumans, W. J., and Van Damme, E. J. M. (2004a). Analysis of the in planta antiviral activity of elderberry ribosome-inactivating proteins. *Eur. J. Biochem.* 271, 1508–1515. doi: 10.1111/j.1432-1033.2004.04059.x
- Vandenbussche, F., Peumans, W. J., Desmyter, S., Proost, P., Ciani, M., and Van Damme, E. J. M. (2004b). The type-1 and type-2 ribosome-inactivating proteins from Iris confer transgenic tobacco plants local but not systemic protection against viruses. *Planta* 220, 211–221. doi: 10.1007/s00425-004-1334-2
- Van Hove, J., Fouquaert, E., Smith, D. F., Proost, P., and Van Damme, E. J. M. (2011). Lectin activity of the nucleocytoplasmic EUL protein from *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 414, 101–105. doi: 10.1016/j.bbrc.2011.09.031
- Wan, J., Tanaka, K., Zhang, X. C., Son, G. H., Brechenmacher, L., Nguyen, T. H., et al. (2012). LYK4, a lysine motif receptor-like kinase, is important for chitin signaling and plant innate immunity in *Arabidopsis*. *Plant Physiol.* 160, 396–406. doi: 10.1104/pp.112.201699
- Wei, G. Q., Liu, R. S., Wang, Q., and Liu, W. Y. (2004). Toxicity of two type II ribosome-inactivating proteins (cinnamomin and ricin) to domestic silkworm larvae. *Arch. Insect Biochem. Physiol.* 57, 160–165. doi: 10.1002/arch.20024
- Williams, C. E., Collier, C. C., Nemacheck, J. A., Liang, C., and Cambron, S. E. (2002). A lectin-like wheat gene responds systemically to attempted feeding by avirulent first-instar Hessian fly larvae. *J. Chem. Ecol.* 28, 1411–1428. doi: 10.1023/A:1016200619766
- Williams, D. B. (2006). Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *J. Cell Sci.* 119, 615–623. doi: 10.1242/jcs.02856
- Willmann, R., Lajunen, H. M., Erbs, G., Newman, M. A., Kolb, D., Tsuda, K., et al. (2011). *Arabidopsis* lysine-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19824–19829. doi: 10.1073/pnas.1112862108
- Wirthmueller, L., Maqbool, A., and Banfield, M. J. (2013). On the front line: structural insights into plant-pathogen interactions. *Nat. Rev. Microbiol.* 11, 761–776. doi: 10.1038/nrmicro3118
- Wu, J., Luo, X., Guo, H., Xiao, J., and Tian, Y. (2006). Transgenic cotton, expressing *Amaranthus caudatus* agglutinin, confers enhanced resistance to aphids. *Plant Breed.* 125, 390–394. doi: 10.1111/j.1439-0523.2006.01247.x
- Xiang, Y., Song, M., Wei, Z., Tong, J., Zhang, L., Xiao, L., et al. (2011). A jacalin-related lectin-like gene in wheat is a component of the plant defence system. *J. Exp. Bot.* 62, 5471–5483. doi: 10.1093/jxb/err226
- Xin, Y., Xiangrong, Z., Mingjiu, Z., Wenchoao, G., Yingchuan, T., Qizhong, X., et al. (2011). Transgenic potato overexpressing the *Amaranthus caudatus* agglutinin gene to confer aphid resistance. *Crop Sci.* 51, 2119–2124. doi: 10.2135/cropsci2010.11.0650
- Yamaji, Y., Maejima, K., Ozeki, J., Komatsu, K., Shiraishi, T., Okano, Y., et al. (2012). Lectin-mediated resistance impairs plant virus infection at the cellular level. *Plant Cell* 24, 778–793. doi: 10.1105/tpc.111.093658
- Yong, W. D., Xu, Y. Y., Xu, W. Z., Wang, X., Li, N., Wu, J. S., et al. (2003). Vernalization-induced flowering in wheat is mediated by a lectin-like gene VER2. *Planta* 217, 261–270.
- Zhang, C., Shi, H., Chen, L., Wang, X., Lu, B., Zhang, S., et al. (2011). Harpin-induced expression and transgenic overexpression of the phloem protein gene AtPP2-A1 in *Arabidopsis* repress phloem feeding of the green peach aphid *Myzus persicae*. *BMC Plant Biol.* 11:11. doi: 10.1186/1471-2229-11-11
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J. D., Boller, T., et al. (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell* 125, 749–760. doi: 10.1016/j.cell.2006.03.037

Zhang, W., Peumans, W. J., Barre, A., Astoul, C. H., Rovira, P., Rougé, P., et al. (2000). Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. *Planta* 210, 970–978. doi: 10.1007/s004250050705

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 June 2014; accepted: 25 July 2014; published online: 13 August 2014.

Citation: Lannoo N and Van Damme EJM (2014) Lectin domains at the frontiers of plant defense. *Front. Plant Sci.* 5:397. doi: 10.3389/fpls.2014.00397

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Lannoo and Van Damme. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Recombinant production of plant lectins in microbial systems for biomedical application – the frutalin case study

Carla Oliveira, José A. Teixeira and Lucília Domingues*

Centre of Biological Engineering, University of Minho, Braga, Portugal

Edited by:

Els Jm Van Damme, Ghent University, Belgium

Reviewed by:

Wim Van Den Ende, Katholieke Universiteit Leuven, Belgium
Alexander D. Frey, Aalto University, Finland

***Correspondence:**

Lucília Domingues, Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
e-mail: luciliad@deb.uminho.pt

Frutalin is a homotetrameric partly glycosylated α -D-galactose-binding lectin of biomedical interest from *Artocarpus incisa* (breadfruit) seeds, belonging to the jacalin-related lectins family. As other plant lectins, frutalin is a heterogeneous mixture of several isoforms possibly with distinct biological activities. The main problem of using such lectins as biomedical tools is that “batch-to-batch” variation in isoforms content may lead to inconstant results. The production of lectins by recombinant means has the advantage of obtaining high amounts of proteins with defined amino-acid sequences and more precise properties. In this mini review, we provide the strategies followed to produce two different forms of frutalin in two different microbial systems: *Escherichia coli* and *Pichia pastoris*. The processing and functional properties of the recombinant frutalin obtained from these hosts are compared to those of frutalin extracted from breadfruit. Emphasis is given particularly to recombinant frutalin produced in *P. pastoris*, which showed a remarkable capacity as biomarker of human prostate cancer and as apoptosis-inducer of cancer cells. Recombinant frutalin production opens perspectives for its development as a new tool in human medicine.

Keywords: recombinant frutalin, lectin isoforms, *Pichia pastoris* expression system, glycosylation, *Escherichia coli* expression system, biomedical application, tumor biomarker, apoptosis-inducer

OUTLINE

Plant lectins have attracted much attention for biomedical applications, especially due to their remarkable anti-tumor properties, resulting from their ability to induce programmed cell death and/or autophagocytosis in cancer cells (Liu et al., 2010; Fu et al., 2011). Plant lectins are also relevant for biomedical diagnosis (Mislovicova et al., 2009).

Frutalin is a plant lectin with reported immunomodulatory (Brando-Lima et al., 2005, 2006), anti-tumor (Oliveira et al., 2011), and tumor biomarker (Oliveira et al., 2009b) properties, among other capacities (de Vasconcellos Abdon et al., 2012), and is a good example of how recombinant production of plant lectins can be challenging but also advantageous for obtaining bioactive derivatives for biomedical application. Frutalin is found in extracts of *Artocarpus incisa* (breadfruit) seeds (Pineau et al., 1990), from which it can be purified by affinity chromatography on cross-linked *Adenanthera pavonina* galactomannan (Moreira et al., 1998). The name “frutalin” (hereinafter referred to as FTL) is a composite of part of the Portuguese common name of the lectin source plant (“fruta” of “fruta-pão”) followed by the suffix “-lin”. Although having sugar-binding preference toward D-galactose, FTL presents a rather broad sugar-binding activity, interacting also with other sugars, as D-mannose and D-glucose. FTL belongs to the jacalin-related lectins family (JRLs, found in the Moraceae plant family), specifically to the sub-family of the galactose-specific lectins (gJRLs), as it presents high structural homology, sugar specificity and sequential identity with jacalin (the galactose-specific lectin of *Artocarpus integrifolia* seeds – jack-fruit, the first member of this family to be identified; Pineau et al.,

1990; Moreira et al., 1998; Oliveira et al., 2009a). FTL is characterized by a strong and identical agglutinating activity with human erythrocytes of the ABO system and rabbit erythrocytes, which has no requirements for divalent metal cations (Moreira et al., 1998). Interestingly, the hemagglutination activity (HA) of FTL is three times higher than that of jacalin (Nobre et al., 2010). FTL has a sophisticated processing. The conversion of the primary translation product of gJRL-mRNA into the protein includes a complex series of co- and post-translational modifications including the removal of the signal peptide (vacuolar targeting), a (partial) glycosylation, removal of the N-terminal propeptide, the excision of a linker tetrapeptide, to separate two polypeptide chains (α and β), proper folding and oligomer assembly. Molecular cloning of the FTL cDNA (excluding signal and propeptide), revealed that, as jacalin, it may be encoded by a family of genes, each of them containing 471 bp, corresponding to a protein of 157 amino-acids, with a calculated molecular weight of 17.1 kDa (Oliveira et al., 2009a). Twenty amino-acids correspond to the β -chain, 4 amino-acids to the linker “T-S-S-N” and 133 amino-acids correspond to the α -chain (from N- to C-terminal). Several gJRLs conserved regions of amino-acids were found in FTL sequences, including the linker. The linker, and its processing, is specific for the sub-group of the gJRLs, being absent in the other sub-group (mannose-specific JRLs; Houles Astoul et al., 2002). FTL is a heterogeneous mixture of several slightly different amino-acid sequences sharing 93–97% of identity, with or without consensus sequences for N-glycosylation (Asn-X-Thr/Ser) in the α -chain (one of these potential N-glycosylation sites was also reported for jacalin; Oliveira et al., 2009a). In fact, FTL is a partly glycosylated

protein, with 2.1% of carbohydrates (Moreira et al., 1998). Different FTL isoforms or iso-lectins (i.e., different mature sequences) may have distinct biological activities, as reported for other plant lectins (Raemaekers et al., 1999; Ohba et al., 2003). Under denaturing conditions (SDS-PAGE), FTL presents two bands: the upper band (15.5 kDa) corresponds to the highly glycosylated isoforms of α chain, whereas the lower band (12 kDa) represents the slightly or non-glycosylated isoforms of the same chain (Oliveira et al., 2008). The β chain is not visible due to its low molecular weight (2.1 kDa; Oliveira et al., 2008). In its native form, FTL is a tetrameric molecule, consisting of four monomers bound by non-covalent linkages, each containing one β and one α chain, forming four sugar-binding sites, with a predominantly β sheet conformation (Moreira et al., 1998; Campana et al., 2002) and an apparent molecular mass of 48–49 kDa (Moreira et al., 1998; Oliveira et al., 2008). FTL is a robust protein as it is stable up to 60°C and very resistant to chemical denaturation (Campana et al., 2002).

Plant lectins are commonly isolated from their natural sources, although this presents several disadvantages, as the resulting isoforms. Recombinant production, mainly in microbial hosts, is an interesting way to overcome this problem, whilst it may allow to improve availability, ensure continuous supply and facilitate purification of lectins with interesting activities or improved/tailor-made functionalities, particularly for biomedical application (for a recent review see Oliveira et al., 2013).

This review describes the case study of the different strategies applied for the production of FTL in the bacterium *Escherichia coli* (Oliveira et al., 2009a; Costa, 2013; Costa et al., 2013a) and in the yeast *Pichia pastoris* (Oliveira et al., 2008). Several variables were considered for optimization: codon usage, strains, fusion partners, induction conditions, and purification methodology. Both microorganisms are well-established platforms for the production of recombinant proteins, including several approved biopharmaceutical products (Berlec and Strukelj, 2013; Gasser et al., 2013). These are also the most employed hosts for the production of recombinant lectins, namely plant lectins for biomedical purposes, such as jacalin (Sahasrabuddhe et al., 2004, 2006), aviscumin (from *Viscum album*; Zwierzina et al., 2011), PCL (from *Polygonatum cyrtonema*; Li et al., 2011), Oryzata (from *Oryza sativa*; Al Atalah et al., 2011), and GNA_{maize} (from *Galanthus nivalis*; Fouquaert et al., 2009). *E. coli* is commonly used to produce non-glycosylated lectins, while *P. pastoris* is mainly employed to overcome problems of insoluble expression of the bacterial system and to produce glycosylated lectins. Thus, *E. coli* and *P. pastoris* were chosen to produce non-glycosylated recombinant frutalin (EcrFTL) and glycosylated recombinant frutalin (PprFTL), respectively. The bio-molecular characterization of the recombinant FTL obtained from each host in terms of processing, molecular weight, HA and sugar-binding activity is herein presented. Finally, a main focus is given to PprFTL due to its demonstrated anti-tumor and tumor biomarker activities (Oliveira et al., 2009b, 2011).

PRODUCTION OF RECOMBINANT FRUTALIN IN *E. coli*

A FTL cDNA sequence was used for production of recombinant FTL in *E. coli* by different strategies (Oliveira et al., 2009a). The first attempts to produce soluble EcrFTL in *E. coli* focused in the use of engineered *E. coli* strains that have extra copies of rare

tRNAs and in the optimization of the induction conditions, but resulted in low yields (Oliveira et al., 2009a; Costa, 2013). The soluble production of EcrFTL from strain *E. coli* BL21 Codon Plus RIPL (DE3), harboring the pET-25b(+) expression vector (Novagen), was maximized to 16 mg/l by the implementation of an experimental factorial design (Oliveira et al., 2009a; Figure 1). However, all the experimental conditions resulted in EcrFTL produced predominantly as insoluble protein. Even though, EcrFTL was purified from crude *E. coli* extracts by sequential size exclusion (SEC) and cation ion exchange chromatography (IEC) that yielded 76 μ g of protein per liter of *E. coli* culture. Purified EcrFTL migrated in SDS-PAGE gel as a homogeneous single-band protein with a molecular mass of about 17 kDa, indicating that the linker was not cleaved. Nevertheless, EcrFTL presented HA against rabbit erythrocytes, although it required more time to develop this activity than FTL. Thus, the HA of FTL is not strictly dependent on linker cleavage. In assays of HA inhibition by different sugars, EcrFTL presented specificity for galactose; however, it could not be purified by affinity chromatography on *A. pavonina* galactomanan, thus revealing lower sugar-binding affinity than FTL. The biomedical properties of this EcrFTL were not evaluated since the amounts obtained through this strategy were unsatisfactory and we were willing to improve them.

Taking into account the low production yields previously obtained, fusion protein technology was afterward considered to improve the soluble production and purification of recombinant FTL in *E. coli*. Eight fusion tags (His₆, Trx, GST, NusA, MBP, SUMO, H, and Fh8), included in pETM vectors (EMBL), were evaluated in small-scale screening assays for recombinant FTL solubility in four *E. coli* strains (Costa et al., 2013a; Figure 1). All vectors provided a His₆ tag for purification of the EcrFTL fusions by nickel affinity (immobilized metal ion affinity chromatography – IMAC). The Rosetta strain (DE3) was selected for scale-up protein processing, namely purification and solubility evaluation before and after tag cleavage. The solubility enhancer partners NusA, Trx, and Fh8 tags considerably improved the soluble production of EcrFTL (in the order: NusA~Fh8 > Trx), being the protein soluble after their removal by TEV cleavage (Tobacco Etch Virus protease). Interestingly, the cleaved and purified EcrFTL from the Fh8 and Trx fusions presented higher amounts than that cleaved from the NusA fusion protein (Costa et al., 2013a). The fusion protein strategy boosted the availability of EcrFTL by increasing its yield from μ g to mg of active protein per liter of *E. coli* culture whilst simplifying the complete production and purification protocol (Costa, 2013). IMAC revealed to be a simpler, easier and quicker procedure than SEC or IEC for EcrFTL purification, and it also decreased EcrFTL losses during purification. Moreover, EcrFTL kept its HA when fused to this partner (our unpublished results). However, the His₆FTL fusion (produced from vector pETM-11) was found to be a very unstable protein, precipitating at physiological pH, and thus being incompatible with cell culture conditions for evaluation of its anti-tumor activity (our unpublished results). Among the fusion partners tested, the Fh8 tag was pointed as a good option for the production of soluble EcrFTL in *E. coli* because of its advantageous low molecular weight and combined solubility enhancer and purification handle activities (Costa, 2013). The Fh8

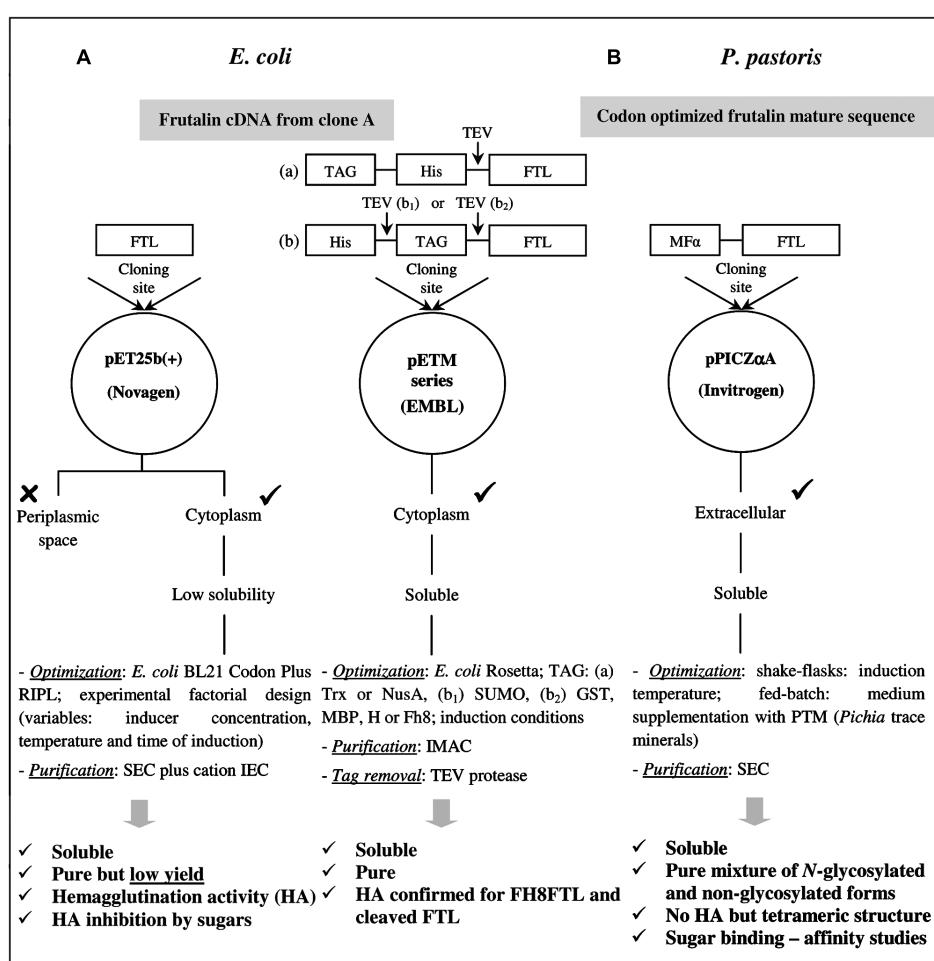


FIGURE 1 | Main strategies for the production of soluble recombinant FTL in *Escherichia coli* (A) and *Pichia pastoris* (B). Different frutalin codifying genes were cloned in *E. coli* and *P. pastoris*, which deduced amino-acid sequences share 93% of identity (Table 2 in Oliveira et al., 2009a).

tag (*Fasciola hepatica* 8-kDa antigen) has been recently ranked among the best solubility enhancer partners for recombinant protein production in *E. coli* (Costa et al., 2013a, 2014). Fh8 was also shown as a suitable fusion partner for purification of recombinant proteins in *E. coli* by HIC (hydrophobic interaction chromatography), with efficiencies comparable to those of IMAC (Costa et al., 2013b). Besides improving EcrFTL solubility, the Fh8 tag increased EcrFTL stability, compared to the His₆ tag, and it did not interfere with the HA and structure (in respect to β-sheet conformation) of EcrFTL, before and after its cleavage (Costa, 2013). The anti-tumor activity of these two versions of EcrFTL, the Fh8FTL and Fh8-cleaved FTL, is under evaluation.

PRODUCTION OF RECOMBINANT FRUTALIN IN *P. pastoris* PRODUCTION AND BIO-MOLECULAR CHARACTERIZATION

Frutalin has highly glycosylated isoforms and the presence of the glycans may be important for its functional properties. Thus, we planned to produce FTL in a microorganism capable of performing glycosylation, namely using the strain *P. pastoris*

KM71H (Oliveira et al., 2008; Figure 1). FTL gene (encoding a mature sequence; Oliveira et al., 2009a) was optimized based on the codon usage of *P. pastoris* and integrated into the yeast genome *in frame* at the C-terminal of the *Saccharomyces* α-factor preprosequence (MFα), to direct PprFTL into the secretory pathway, and under the control of the strong methanol inducible AOX1 promoter. PprFTL was produced in shake-flasks and purified from supernatants by SEC yielding 18–20 mg per liter of culture.

Important differences between the molecular and biological properties of PprFTL and FTL were found (Oliveira et al., 2008). The reason for that was the processing of FTL in *P. pastoris*, which was different from that occurring in breadfruit. As also observed in *E. coli* (Oliveira et al., 2009a), the FTL linker was not cleaved in *P. pastoris* (confirmed by N-terminal sequencing), thus suggesting that this processing can be specific for higher eukaryotes. Furthermore, the MFα secretion leader was incompletely removed, resulting in PprFTL with one Glu-Ala repeat at its N-terminal, decreasing its predicted PI (isoelectric point) from 8 to 5. These repeats are commonly observed in heterologous proteins secreted by *P. pastoris* using this signal sequence, and also reported for other

recombinant plant lectins (Raemaekers et al., 1999; Lannoo et al., 2007).

As expected, PprFTL was *N*-glycosylated by *P. pastoris*, since the corresponding protein sequence has one potential site for *N*-glycosylation (α -Asn74). Part of the secreted PprFTL undergone this post-translational modification, which led to an extension in its molecular weight of about 2.8 kDa (Oliveira et al., 2011). PprFTL, contrarily to FTL, did not agglutinate rabbit erythrocytes, despite also having a tetrameric structure (Oliveira et al., 2008). Thus, it was hypothesized that glycosylation pattern of *P. pastoris* inhibited this activity since non-glycosylated EcrFTL presented HA. In a previous work, *Pichia* glycosylation was also suggested to inhibit the HA of a fungal lectin (Iijima et al., 2003). Nevertheless, it should be noted that different FTL coding sequences were cloned in *P. pastoris* and *E. coli*, and hence the HA of the resulting proteins may differ. The deduced amino-acid sequences of the frutalin codifying genes cloned in *P. pastoris* and *E. coli* have 93% of sequence identity (Figure 1). In what concerns carbohydrate-binding activity, PprFTL presented a sugar preference similar to FTL, but with less affinity. The affinity constant for the binding of PprFTL to the monosaccharide methyl- α -galactose was determined and found to be 113-fold lower than that of FTL (Oliveira et al., 2008). The only other gjRL so far produced in microorganisms was jacalin. Jacalin was produced in *E. coli* also as an unprocessed protein with its sugar-binding activities reduced in the same order of magnitude (Sahasrabuddhe et al., 2004). The correct excision of the linker and consequent generation of a free glycine at the *N*-terminus of the α chain may determine gjRLs sugar-binding properties (Houles Astoul et al., 2002; Sahasrabuddhe et al., 2004; Oliveira et al., 2008).

The large-scale production of PprFTL was conducted in a 1.6 L stirred tank bioreactor operating in fed-batch mode at 28°C during 4 days (Wanderley et al., 2013). Supplementation of the culture medium (BMMH – buffered minimal methanol medium) with *Pichia* trace minerals (PTM) resulted in 2.5-fold higher PprFTL production (13.4 mg/l) than that achieved without supplementation (5.23 mg/l). Furthermore, bioreactor resulted in fourfold higher PprFTL production, comparing to shaker-flasks batch assays (3.3 mg/l), using the same culture medium (BMMH

plus PTM) and induction conditions (Wanderley et al., 2013). Nevertheless, the yield of PprFTL was higher from shake-flasks induced at 15°C (18–20 mg/l), which means that lower temperatures favor PprFTL production (Oliveira et al., 2008). However, 20 of such flasks (each containing 50 ml of BMMH medium) are needed to obtain the same amount of PprFTL as in one fed-batch experiment. Thus, bioreactor fermentation is more advantageous for the production of PprFTL.

BIOMEDICAL PROPERTIES

The relevance of JRLs, specifically jacalin, for cancer diagnostics and therapeutics is present in many recent works (e.g., Obaid et al., 2012; Lee et al., 2013; Marangoni et al., 2013; Zupancic et al., 2014). PprFTL was evaluated in terms of its tumor biomarker and anti-tumor properties, comparatively to FTL (Oliveira et al., 2009b, 2011). The cancer biomarker study was performed by immunohistochemistry with human prostate tissues (Oliveira et al., 2009b). Other plant lectins were used in the past in similar studies but with limited success (works cited in Oliveira et al., 2009b). The binding pattern of PprFTL and FTL to the prostate tissues was distinct, presumably due to their differences in carbohydrate-binding affinity (Oliveira et al., 2008). FTL bound to any type of prostate cells but more strongly to the neoplastic (malignant cells) than to the hyperplastic ones (non-malignant cells). On the other hand, PprFTL was much more specific, as it just recognized malignant cells (Figure 2). A significant positive statistical correlation between the binding intensity of PprFTL and the histological diagnosis of the tissues was obtained (not observed for FTL), although PprFTL did not recognize all the malignant cases studied (30% had negative binding), and when positive, the binding was heterogeneous. However, only a small number of prostate cases were analyzed and the histochemical methodology has still room for improvement. This study indicates that PprFTL has higher potential as cancer biomarker than FTL.

In *in vitro* assays, PprFTL showed a strong cytotoxic effect on HeLa cervical cancer cells proliferation, by inducing cell death by apoptosis (Oliveira et al., 2011). This effect was irreversible as well as time and dose dependent ($IC_{50} \sim 100 \mu\text{g/mL}$). Identical results were obtained for FTL in the same study (Oliveira et al., 2011).

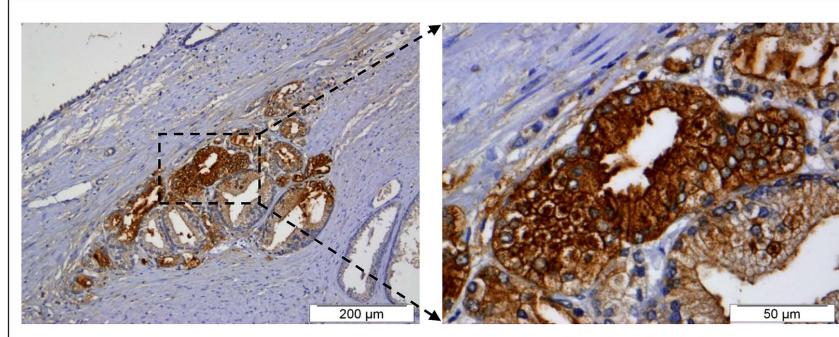


FIGURE 2 | Immunostaining pictures of a prostate cancer tissue using recombinant FTL produced in *P. pastoris* as tumor diagnostic marker.
PprFTL was able to specifically recognize carcinoma cells in middle of a

benign lesion. The staining is localized in the cells cytoplasm of the carcinoma glands (brown color). (Original publisher: BioMed Central; Adapted from Oliveira et al., 2009b).

Thus, it seems that frutalin action in tumor cells is not exclusively dependent on its carbohydrate-binding properties. Other factors, such as protein–protein interactions, may contribute to the cellular responses, as suggested for the effect of recombinant jacalin in tumor cells (Sahasrabuddhe et al., 2006). Recombinant jacalin showed a magnitude of anti-proliferative responses similar to native jacalin on human cancer cell lines, despite its inferior sugar-binding affinity (Sahasrabuddhe et al., 2006). Both, PprFTL and FTL showed nuclear migration activity on HeLa cells (Oliveira et al., 2011), a property only reported for fungal lectins (Yu et al., 1999; Francis et al., 2003; Liang et al., 2009). Plant lectins have been described to attach to cancer cells membrane [e.g., jacalin in A431 human epidermoid carcinoma cells (Sahasrabuddhe et al., 2006)] or to be internalized and located in different cellular compartments [e.g., wheat germ agglutinin, WGA, in DU-145 human prostate cancer cells (Gabor et al., 2001)]. To our knowledge, our work is the first reporting nuclear migration activity on cancer cells for a lectin from plant origin (Oliveira et al., 2011). Studies aiming to elucidate the apoptotic mechanism triggered by PprFTL on cancer cells are now being conducted.

CONCLUSION AND PROSPECTS

Escherichia coli and *Pichia pastoris* were found as suitable hosts for producing high amounts of recombinant FTL upon production and purification optimization. Optimization in *E. coli* significantly improved EcrFTL production, leading to high yields, but decreased protein stability. Furthermore, the processing of recombinant FTL in both microorganisms was different from that occurring in breadfruit, resulting in versions of FTL with inferior HA and carbohydrate-binding capacity. Nevertheless, PprFTL presented an anti-tumor activity identical to FTL and enhanced tumor biomarker capacity. The production strategies herein presented will extend the research on the biomedical properties of recombinant FTL.

The importance of amino-acids substitutions and post-translational modifications in gJRLs (e.g., linker cleavage, glycosylation) remains to be elucidated, and thus, future research might follow this direction. Within this scope, production and the availability of heterologous recombinant lectins is a valuable tool that can contribute to the fundamental understanding of the biological activity of the lectins. The production of the same lectin coding sequence in prokaryotic and eukaryotic hosts, the production of different lectin isoforms, and engineered/mutated versions, will provide insight into lectins functionality and shed light into its native physiological role. Furthermore, recombinant lectins with refined properties can be obtained. Finally, the fusion of enhanced lectins with functional moieties, by using recombinant DNA technology, for the development of functionalized drug delivery systems for site specific anti-tumor therapy, is anticipated.

AUTHOR CONTRIBUTIONS

Carla Oliveira drafted the review and carried out most of the experimental work of recombinant FTL. José A. Teixeira participated in the development of the concept. Lucília Domingues conceived the study and helped to draft the review. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

Carla Oliveira acknowledges support from Fundação para a Ciência e a Tecnologia (FCT), Portugal (by the fellowship SFRH/BDP/63831/2009). The authors thank the FCT GlycoCBMs Project REF. FCT PTDC/AGR-FOR/3090/2012 – FCOMP-01-0124-FEDER-027948, the FCT Strategic Project PEst-OE/EQB/LA0023/2013, and the Project “BioInd – Biotechnology and Bioengineering for improved Industrial and Agro-Food processes,” REF. NORTE-07-0124-FEDER-000028 Co-funded by the Programa Operacional Regional do Norte (ON.2 – O Novo Norte), QREN, FEDER.

REFERENCES

- Alataloh, B., Fouquaert, E., Vanderschaeghe, D., Proost, P., Balzarini, J., Smith, D. F., et al. (2011). Expression analysis of the nucleocytoplasmic lectin ‘Orysata’ from rice in *Pichia pastoris*. *FEBS J.* 278, 2064–2079. doi: 10.1111/j.1742-4658.2011.08122.x
- Berlec, A., and Strukelj, B. (2013). Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. *J. Ind. Microbiol. Biotechnol.* 40, 257–274. doi: 10.1007/s10295-013-1235-0
- Brando-Lima, A. C., Saldanha-Gama, R. F., Henriques, M., Monteiro-Moreira, A. C., Moreira, R. A., and Barja-Fidalgo, C. (2005). Frutalin, a galactose-binding lectin, induces chemotaxis and rearrangement of actin cytoskeleton in human neutrophils: involvement of tyrosine kinase and phosphoinositide 3-kinase. *Toxicol. Appl. Pharmacol.* 208, 145–154. doi: 10.1016/j.taap.2005.02.012
- Brando-Lima, A. C., Saldanha-Gama, R. F., Pereira, C. R., Villela, C. G., Sampaio, A. L., Monteiro-Moreira, A. C., et al. (2006). Involvement of phosphatidylinositol-3 kinase-Akt and nuclear factor kappa-B pathways in the effect of frutalin on human lymphocyte. *Int. Immunopharmacol.* 6, 465–472. doi: 10.1016/j.intimp.2005.09.008
- Campana, P. T., Moraes, D. I., Monteiro-Moreira, A. C., and Beltramini, L. M. (2002). Unfolding and refolding studies of frutalin, a tetrameric D-galactose binding lectin. *Eur. J. Biochem.* 269, 753–758. doi: 10.1046/j.0014-2956.2002.02742.x
- Costa, S. (2013). *Development of A Novel Fusion System for Recombinant Protein Production and Purification In Escherichia coli*. Ph.D. thesis, University of Minho, Braga.
- Costa, S., Almeida, A., Castro, A., and Domingues, L. (2014). Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Front. Microbiol.* 5:63. doi: 10.3389/fmicb.2014.00063
- Costa, S. J., Almeida, A., Castro, A., Domingues, L., and Besir, H. (2013a). The novel Fh8 and H fusion partners for soluble protein expression in *Escherichia coli*: a comparison with the traditional gene fusion technology. *Appl. Microbiol. Biotechnol.* 97, 6779–6791. doi: 10.1007/s00253-012-4559-1
- Costa, S. J., Coelho, E., Franco, L., Almeida, A., Castro, A., and Domingues, L. (2013b). The Fh8 tag: a fusion partner for simple and cost-effective protein purification in *Escherichia coli*. *Protein Expr. Purif.* 92, 163–170. doi: 10.1016/j.pep.2013.09.013
- de Vasconcellos Abdón, A. P., Coelho De Souza, G., Noronha Coelho De Souza, L., Prado Vasconcelos, R., Araujo Castro, C., Moreira Guedes, M., et al. (2012). Gastroprotective potential of frutalin, a D-galactose binding lectin, against ethanol-induced gastric lesions. *Fitoterapia* 83, 604–608. doi: 10.1016/j.fitote.2012.01.005
- Fouquaert, E., Smith, D. F., Peumans, W. J., Proost, P., Balzarini, J., Savvides, S. N., et al. (2009). Related lectins from snowdrop and maize differ in their carbohydrate-binding specificity. *Biochem. Biophys. Res. Commun.* 380, 260–265. doi: 10.1016/j.bbrc.2009.01.048
- Francis, F., Marty-Detraves, C., Poincloux, R., Baricault, L., Fournier, D., and Paquereau, L. (2003). Fungal lectin, XCL, is internalized via clathrin-dependent endocytosis and facilitates uptake of other molecules. *Eur. J. Cell Biol.* 82, 515–522. doi: 10.1078/0171-9335-8
- Fu, L. L., Zhou, C. C., Yao, S., Yu, J. Y., Liu, B., and Bao, J. K. (2011). Plant lectins: targeting programmed cell death pathways as antitumor agents. *Int. J. Biochem. Cell Biol.* 43, 1442–1449. doi: 10.1016/j.biocel.2011.07.004
- Gabor, F., Klausegger, U., and Wirth, M. (2001). The interaction between wheat germ agglutinin and other plant lectins with prostate cancer cells Du-145. *Int. J. Pharm.* 221, 35–47. doi: 10.1016/S0378-5173(01)00650-0

- Gasser, B., Prielhofer, R., Marx, H., Maurer, M., Nocon, J., Steiger, M., et al. (2013). *Pichia pastoris*: protein production host and model organism for biomedical research. *Future Microbiol.* 8, 191–208. doi: 10.2217/fmb.12.133
- Houles Astoul, C., Peumans, W. J., Van Damme, E. J., Barre, A., Bourne, Y., and Rouge, P. (2002). The size, shape and specificity of the sugar-binding site of the jacalin-related lectins is profoundly affected by the proteolytic cleavage of the subunits. *Biochem. J.* 367(Pt 3), 817–824. doi: 10.1042/BJ20020856
- Iijima, N., Amano, K., Ando, A., and Nagata, Y. (2003). Production of fruiting-body lectins of *Pleurotus cornucopiae* in methylotrophic yeast *Pichia pastoris*. *J. Biosci. Bioeng.* 95, 416–418. doi: 10.1016/S1389-1723(03)80079-8
- Lannoo, N., Vervecken, W., Proost, P., Rouge, P., and Van Damme, E. J. M. (2007). Expression of the nucleocytoplasmic tobacco lectin in the yeast *Pichia pastoris*. *Protein Expr. Purif.* 53, 275–282. doi: 10.1016/j.pep.2007.01.007
- Lee, C. S., Muthusamy, A., Abdul-Rahman, P. S., Bhavanandan, V. P., and Hashim, O. H. (2013). An improved lectin-based method for the detection of mucin-type O-glycans in biological samples. *Analyst* 138, 3522–3529. doi: 10.1039/c3an36258b
- Li, C. Y., Luo, P., Liu, J. J., Wang, E. Q., Li, W. W., Ding, Z. H., et al. (2011). Recombinant expression of *Polygonatum cyrtonema* lectin with anti-viral, apoptosis-inducing activities and preliminary crystallization. *Process Biochem.* 46, 533–542. doi: 10.1016/j.procb.2010.10.005
- Liang, Y., Feng, L., Tong, X., Wang, K., Li, D. F., Lin, J. C., et al. (2009). Importance of nuclear localization for the apoptosis-induced activity of a fungal galectin AAL (Agrocybe aegerita lectin). *Biochem. Biophys. Res. Commun.* 386, 437–442. doi: 10.1016/j.bbrc.2009.06.054
- Liu, B., Bian, H. J., and Bao, J. K. (2010). Plant lectins: potential antineoplastic drugs from bench to clinic. *Cancer Lett.* 287, 1–12. doi: 10.1016/j.canlet.2009.05.013
- Marangoni, V. S., Paino, I. M., and Zucolotto, V. (2013). Synthesis and characterization of jacalin-gold nanoparticles conjugates as specific markers for cancer cells. *Colloids Surf. B Biointerfaces* 112, 380–386. doi: 10.1016/j.colsurfb.2013.07.070
- Mislovicova, D., Gemeiner, P., Kozarova, A., and Kozar, T. (2009). Lectinomics I. Relevance of exogenous plant lectins in biomedical diagnostics. *Biologia* 64, 1–19. doi: 10.2478/s11756-009-0029-3
- Moreira, R. A., Castelo-Branco, C. C., Monteiro, A. C., Tavares, R. O., and Beltramini, L. M. (1998). Isolation and partial characterization of a lectin from *Artocarpus incisa* L. seeds. *Phytochemistry* 47, 1183–1188. doi: 10.1016/S0031-9427(97)00753-X
- Nobre, T. M., Pavinatto, F. J., Cominetti, M. R., Selistre De-Araujo, H. S., Zaniquelli, M. E., and Beltramini, L. M. (2010). The specificity of frutalin lectin using biomembrane models. *Biochim. Biophys. Acta* 1798, 1547–1555. doi: 10.1016/j.bbamem.2010.03.021
- Obaid, G., Chambrier, L., Cook, M. J., and Russell, D. A. (2012). Targeting the oncofetal Thomsen-Friedenreich disaccharide using jacalin-PEG phthalocyanine gold nanoparticles for photodynamic cancer therapy. *Angew. Chem. Int. Ed. Engl.* 51, 6158–6162. doi: 10.1002/anie.201201468
- Ohba, H., Bakalova, R., and Muraki, M. (2003). Cytoagglutination and cytotoxicity of Wheat Germ Agglutinin isolectins against normal lymphocytes and cultured leukemic cell lines – relationship between structure and biological activity. *Biochim. Biophys. Acta* 1619, 144–150. doi: 10.1016/S0304-4165(02)00479-8
- Oliveira, C., Costa, S., Teixeira, J. A., and Domingues, L. (2009a). cDNA cloning and functional expression of the α -D-galactose-binding lectin frutalin in *Escherichia coli*. *Mol. Biotechnol.* 43, 212–220. doi: 10.1007/s12033-009-9191-7
- Oliveira, C., Felix, W., Moreira, R. A., Teixeira, J. A., and Domingues, L. (2008). Expression of frutalin, an α -D-galactose-binding jacalin-related lectin, in the yeast *Pichia pastoris*. *Protein Expr. Purif.* 60, 188–193. doi: 10.1016/j.pep.2008.04.008
- Oliveira, C., Nicolau, A., Teixeira, J. A., and Domingues, L. (2011). Cytotoxic effects of native and recombinant frutalin, a plant galactose-binding lectin, on HeLa cervical cancer cells. *J. Biomed. Biotechnol.* 2011, 568932. doi: 10.1155/2011/568932
- Oliveira, C., Teixeira, J. A., and Domingues, L. (2013). Recombinant lectins: an array of tailor-made glycan-interaction biosynthetic tools. *Crit. Rev. Biotechnol.* 33, 66–80. doi: 10.3109/07388551.2012.670614
- Oliveira, C., Teixeira, J. A., Schmitt, F., and Domingues, L. (2009b). A comparative study of recombinant and native frutalin binding to human prostate tissues. *BMC Biotechnol.* 9:78. doi: 10.1186/1472-6750-9-78
- Pineau, N., Pousset, J. L., Preud'homme, J. L., and Aucouturier, P. (1990). Structural and functional similarities of breadfruit seed lectin and jacalin. *Mol. Immunol.* 27, 237–240. doi: 10.1016/0161-5890(90)90135-M
- Raemaekers, R. J. M., De Muro, L., Gatehouse, J. A., and Fordham-Skelton, A. P. (1999). Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris* – Correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. *Eur. J. Biochem.* 265, 394–403. doi: 10.1046/j.1432-1327.1999.00749.x
- Sahasrabuddhe, A. A., Ahmed, N., and Krishnasastri, M. V. (2006). Stress-induced phosphorylation of caveolin-1 and p38, and down-regulation of EGFr and ERK by the dietary lectin jacalin in two human carcinoma cell lines. *Cell Stress Chaperones* 11, 135–147. doi: 10.1379/CSC-160R.1
- Sahasrabuddhe, A. A., Gaikwad, S. M., Krishnasastri, M. V., and Khan, M. I. (2004). Studies on recombinant single chain jacalin lectin reveal reduced affinity for saccharides despite normal folding like native jacalin. *Protein Sci.* 13, 3264–3273. doi: 10.1110/ps.04968804
- Wanderley, M. S. O., Oliveira, C., Bruneska, D., Domingues, L., Lima, J. L., Teixeira, J. A., et al. (2013). Influence of trace elements supplementation on the production of recombinant frutalin by *Pichia pastoris* KM71H in fed-batch process. *Chem. Pap.* 67, 682–687. doi: 10.2478/s11696-013-0363-3
- Yu, L. G., Fernig, D. G., White, M. R. H., Spiller, D. G., Appleton, P., Evans, R. C., et al. (1999). Edible mushroom (*Agaricus bisporus*) lectin, which reversibly inhibits epithelial cell proliferation, blocks nuclear localization sequence-dependent nuclear protein import. *J. Biol. Chem.* 274, 4890–4899. doi: 10.1074/jbc.274.8.4890
- Zupancic, D., Kreft, M. E., and Romih, R. (2014). Selective binding of lectins to normal and neoplastic urothelium in rat and mouse bladder carcinogenesis models. *Protoplasma* 251, 49–59. doi: 10.1007/s00709-013-0524-9
- Zwierzina, H., Bergmann, L., Fiebig, H., Aamdal, S., Schoffski, P., Witthohn, K., et al. (2011). The preclinical and clinical activity of aviscumine: a potential anticancer drug. *Eur. J. Cancer* 47, 1450–1457. doi: 10.1016/j.ejca.2011.02.022

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 May 2014; accepted: 22 July 2014; published online: 08 August 2014.

Citation: Oliveira C, Teixeira JA and Domingues L (2014) Recombinant production of plant lectins in microbial systems for biomedical application – the frutalin case study. Front. Plant Sci. 5:390. doi: 10.3389/fpls.2014.00390

This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Oliveira, Teixeira and Domingues. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals

Elodie Mathieu-Rivet^{1†}, Marie-Christine Kiefer-Meyer^{1†}, Gaëtan Vanier¹, Clément Ovide¹, Carole Burel¹, Patrice Lerouge¹ and Muriel Bardor^{1,2*}

¹ Laboratoire Glyco-MEV, Faculté des Sciences et Techniques, UPRES EA 4358, Normandie Université, IRIB, VASI, Mont-Saint-Aignan, France

² Institut Universitaire de France, Paris, France

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Richard Strasser, University of Natural Resources and Life Sciences, Austria
Eva L. Decker, University of Freiburg, Germany

***Correspondence:**

Muriel Bardor, Laboratoire Glyco-MEV (Glycobiologie et Matrice Extracellulaire Végétale), UPRES EA 4358, UFR des Sciences et Techniques, Normandie Université, IRIB, VASI, Bâtiment Gadeau de Kerville, 1 rue Tesnière, 76821 Mont Saint Aignan Cedex, France
e-mail: muriel.bardor@univ-rouen.fr

[†]The two first authors equally contributed to the work.

Microalgae are currently used for the production of food compounds. Recently, few microalgae species have been investigated as potential biofactories for the production of biopharmaceuticals. Indeed in this context, microalgae are cheap, classified as Generally Recognized As Safe (GRAS) organisms and can be grown easily. However, problems remain to be solved before any industrial production of microalgae-made biopharmaceuticals. Among them, post-translational modifications of the proteins need to be considered. Especially, N-glycosylation acquired by the secreted recombinant proteins is of major concern since most of the biopharmaceuticals are N-glycosylated and it is well recognized that glycosylation represent one of their critical quality attribute. Therefore, the evaluation of microalgae as alternative cell factory for biopharmaceutical productions thus requires to investigate their N-glycosylation capability in order to determine to what extend it differs from their human counterpart and to determine appropriate strategies for remodeling the microalgae glycosylation into human-compatible oligosaccharides. Here, we review the secreted recombinant proteins which have been successfully produced in microalgae. We also report on recent bioinformatics and biochemical data concerning the structure of glycans N-linked to proteins from various microalgae phyla and comment the consequences on the glycan engineering strategies that may be necessary to render those microalgae-made biopharmaceuticals compatible with human therapy.

Keywords: microalgae, biopharmaceuticals, glycosylation pathway, glycan, *Chlamydomonas reinhardtii*, *Phaeodactylum tricornutum*, endoplasmic reticulum, Golgi apparatus

INTRODUCTION

Nowadays, biopharmaceuticals on the market consists of 200 products yielding overall global revenue greater than US\$100 billion (Walsh, 2010). Over the past 5 years, 140 biopharmaceuticals were approved in the European Union (EU) and the United States (US) markets (Walsh, 2010). These biopharmaceuticals are produced in various systems ranging from bacteria to mammalian cell cultures (Wong, 2005; Demain and Vaishnav, 2009; Huang et al., 2012). Among those, the Chinese Hamster Ovary (CHO) cells are currently the predominant industrial cell lines used for producing those drugs (Hossler et al., 2009), covering about 50% of the market (Demain and Vaishnav, 2009). However, the constant increasing needs for large amount of such therapeutic proteins, their high production cost in conventional expression systems and complicating factors related to potential virus contamination have driven scientists to explore new alternative production systems. In this context, various plant expression systems have emerged, including whole plants (*Nicotiana*, Alfalfa, Maize for examples) and *in vitro* culture systems such as plant cell suspensions (*Nicotiana tabacum*, *Lemna minor*, *Physcomitrella patens*) and hairy roots

(Stoger et al., 2005; Drake et al., 2009; Colgan et al., 2010; Decker and Reski, 2012; Parsons et al., 2012; Xu et al., 2012; Schillberg et al., 2013; Twyman et al., 2013; Buyel and Fischer, 2014). Several plant-made biopharmaceuticals have been successfully produced (De Muynck et al., 2010). This includes the Cerezyme's biosimilar (glucocerebrosidase) which has been produced in carrot cells (Shaaltiel et al., 2007) and was approved on May 2012 by the US Food and Drug Administration (Maxmen, 2012). Despite those successes, there is an increasing interest to use microalgae for biopharmaceutical production. Microalgae are unicellular photosynthetic organisms which encompass between 40,000 and probably several billion species (Cadoret and Bernard, 2008; Mata et al., 2010). As plants, microalgae are classified in Generally Recognized As Safe (GRAS) organisms. Moreover, they are cheap and easy to grow, making them potentially attractive cell factories for the large-scale production of recombinant proteins.

To date, microalgae have been mainly used for the production of food compounds or high-value added compounds like carotenoids (Spolaore et al., 2006; Sasso et al., 2012). In addition, as photosynthetic organisms, microalgae are very

efficient in converting sunlight into chemical energy, making them attractive for the production of carbohydrates, lipids, and hydrogen. Therefore, algal biomass represents a great potential for generating new sources of bioenergy such as biofuels (Beer et al., 2009; Lam and Lee, 2012; Merchant et al., 2012) and biomaterials (Hempel et al., 2011a). Several microalgae species have also been evaluated for their potential to express recombinant proteins. Among them, the model Chlorophyceae *Chlamydomonas reinhardtii* is currently the most investigated one for such a biotechnological application, due to the availability of genomic data and the existence of a powerful molecular toolkit including vectors allowing nuclear or chloroplastic transformation (Merchant et al., 2007; Harris, 2009). Additionally, the diatom *Phaeodactylum tricornutum* is also considered as an emerging system for such an application (Hempel et al., 2011b; Hempel and Maier, 2012) as it can be genetically modified and grown quite easily. Other species are also good candidates for large-scale production of recombinant proteins, especially due to their ease to obtain algal biomass and their growth rate. For example, a marine green microalga from the class of the Chlorophyceae, *Dunaliella salina* (Geng et al., 2003), but also species from the genus *Chlorella* [i.e., *Chlorella vulgaris*, recently renamed *Coccomyxa* sp. C-169 *subellipoidea* (Blanc et al., 2010), *Chlorella ellipoidea*], which belongs to the Trebouxiophyceae, are promising bio-factories for large-scale production of high-value added proteins (Hawkins and Nakamura, 1999; Chen et al., 2001; Kim et al., 2002; Bai et al., 2013). However, molecular biology tools remain poorly developed for most of those microalgae even if genomic data became available (<http://genome.jgi.doe.gov/>; <http://www.phytozome.net/>; <http://www.ncbi.nlm.nih.gov/genome>).

Whatever the considered species, many problems have still to be solved before any industrial production and commercialization of microalgae-made biopharmaceuticals. Among them, increasing the yield and secretion of recombinant protein represents a crucial issue to make these organisms competitive with traditionally used expression systems such as the CHO cell lines. Furthermore, post-translational modifications need to be considered. Especially, *N*-glycosylation acquired by the secreted recombinant proteins is crucial for biopharmaceuticals since more than one third of the approved ones are glycosylated (Gomord et al., 2010) and it represents a critical quality attribute for them (Lingg et al., 2012). Indeed, the presence and structures of the *N*-glycans are required for their biological activity, stability and half-life (Lingg et al., 2012). The evaluation of microalgae as alternative cell factory for biopharmaceutical production thus requires investigating their *N*-glycosylation capability in order to determine to what extend their *N*-glycans differ from their human counterpart. In this paper, we report on bio-informatic and biochemical data concerning the structures of glycan *N*-linked to endogenous proteins from various microalgae phyla. Based on these recent findings, strategies for the engineering of the glycosylation pathways in these new expression systems are proposed to obtain microalgae-made biopharmaceuticals that would carry oligosaccharides compatibles with human therapies.

MICROALGAE AS ALTERNATIVE SYSTEMS FOR PRODUCTION OF RECOMBINANT PROTEINS

AVAILABLE TOOLS FOR NUCLEAR TRANSFORMATION IN MICROALGAE *Chlamydomonas reinhardtii*

Currently, most of the studies reporting the expression of recombinant proteins in microalgae have been performed in the green alga *Chlamydomonas reinhardtii*. The majority of such productions has been targeted to the chloroplast. This concerned biopharmaceuticals such as the single-chain antibody directed against the glycoprotein D of the herpes simplex virus (Mayfield et al., 2003), but also the heavy and light chains of the antibody 83K7C, derived from a human IgG1 (Tran et al., 2009), human erythropoietin, domains 10 and 14 of human fibronectin, interferon β 1, proinsulin, vascular endothelial growth factor (VEGF), high mobility group protein B1(HMGB1) (Rasala et al., 2010) and two different immunotoxin proteins (Tran et al., 2012). Recently, an attenuated form of the E7 oncoprotein of the human papillomavirus (HPV) has also been produced in the same organelle (Demurtas et al., 2013). Indeed, this strategy is the most promising one to get high protein yields since *C. reinhardtii* possesses a large single chloroplast, representing about 40% of the total cell volume, and the proteins expressed in this organelle have been shown to represent 2–20% of total soluble proteins (Rasala and Mayfield, 2011). Moreover, three proteins have been shown to be expressed at levels which are sufficient for commercial production (Rasala et al., 2010). In contrast, secreted proteins produced from nuclear transformation in *C. reinhardtii* generally failed to accumulate to an equivalent level as the one observed in chloroplasts (Fuhrmann et al., 1999; Schroda et al., 2000; Specht et al., 2010). The low proteolysis in chloroplast could explain the high yields of recombinant proteins reached in this organelle (Surzycki et al., 2009; Potvin and Zhang, 2010). However, as far as glycosylated proteins such as biopharmaceuticals are concerned, chloroplast lacks the enzymatic machinery required for *N*-glycosylation. Therefore, for biopharmaceutical production, both nuclear expression and protein transport through the secretory pathway of microalgae are required.

Despite recent progress, nuclear expression remains challenging in *C. reinhardtii* (Specht et al., 2010). Indeed, low levels of expression usually observed could result from transgene silencing (Cerutti et al., 1997; Shaver et al., 2010; Rasala et al., 2012). Its GC-rich genome (Merchant et al., 2007) has been also thought to be a hindrance to the expression of foreign genes since it introduces a codon usage bias. Several strategies have been developed during the last decade to circumvent these problems. Among them, the necessity to optimize the coding sequence of the gene of interest appears to be essential for improving the nuclear expression of the foreign protein (Fuhrmann et al., 1999). With regards to promoter sequences, a few assays have been performed to express reporter genes under the control of the plant constitutive CaMV35S promoter (Ruecker et al., 2008; Díaz-Santos et al., 2013) or of other viral sequences (Ruecker et al., 2008). Recently, a high light-inducible promoter from *Dunaliella* has also been investigated in *C. reinhardtii*. This promoter has been shown to drive efficiently the expression of the luciferase reporter gene but it has not been used for expression of any biopharmaceutical yet

(Park et al., 2013). Currently, the major part of nuclear transgenes are expressed through a hybrid promoter, resulting from the fusion of the photosystem I complex (PSAD) (Fischer and Rochaix, 2001) or the Ribulose Bisphosphate Carboxylase Small Subunit (RBCS2) (Kindle, 1998) promoters with the HSP70A promoter which allows increasing the expression of the transgene (Schroda et al., 2000). In addition, several regulatory sequences have been included in the transgene coding sequence to enhance its expression. Thus, the insertion of the first intron of RBCS2 (Lumbrares et al., 1998), but also of the second and third introns (Eichler-Stahlberg et al., 2009) are required to increase the efficacy of the HSP70A/RBCS2 promoter. Compared to the CaMV35S promoter, these chimeric promoters remain actually the most efficient to drive the constitutive expression of nuclear transgenes in *C. reinhardtii* (Ruecker et al., 2008; Eichler-Stahlberg et al., 2009; Neupert et al., 2009; Rasala et al., 2012; Kumar et al., 2013). However, with a yield of secreted EPO estimated to 0.1 mg per liter of culture medium (Eichler-Stahlberg et al., 2009), the expression level of EPO under the control of the HSP70A/RBCS2A promoter remains too low for large scale production.

Rasala et al. (2012) developed a new vector in which the gene of interest is fused to the *BLE* gene of selection (which confers the resistance to bleomycin) *via* the nucleotide sequence encoding the foot-and-mouth-disease-virus 2A self-cleavage peptide (FMDV 2A) under the control of the HSP70A/RBCS2 promoter. The resulting translated product is processed into two proteins, with the 2A peptide fused to the C-terminal end of the first protein as already described for other expressions using a 2A self-cleavage peptide (Ho et al., 2013). This tandem expression allows the selection of transformants exhibiting a higher transgene expression. Finally, Neupert et al. (2009) used UV mutagenesis to generate new *C. reinhardtii* strains presenting the advantage to increase the expression of nuclear transgenes. Indeed, using one of these strains (UVM4) transformed with the specific vector pcCAgLUC allowing the expression of the luciferase protein fused to the predicted signal peptide of the extracellular carbonic anhydrase 1 (CAH1), Lauersen and coworkers demonstrated that the amount of secreted recombinant proteins could be significantly improved, reaching up to 10 mg per liter of culture (Lauersen et al., 2013a). However, despite these significant achievements, efforts are still necessary to make *C. reinhardtii* competitive with CHO cell lines, for which yield between 5 and 10 g/L of recombinant protein is currently obtained (Demain and Vaishnav, 2009).

Other microalgae

Recently, other microalgae species have also been shown to be of interest as new expression systems. For example, two recent studies reported the expression of a monoclonal human IgG antibody against the Hepatitis B and its respective antigen in the diatom *P. tricornutum*. This antibody was either secreted in the culture medium or expressed in fusion to a DDEL (instead of KDEL which is used as a standard ER retention signal in other eukaryotes) sequence allowing its retention within the endoplasmic reticulum (ER) of *Phaeodactylum tricornutum* (Hempel et al., 2011b; Hempel and Maier, 2012). In contrast to *C. reinhardtii*, the codon usage in *P. tricornutum* is much closer to that of

human (Heitzer et al., 2007), which could give advantage to this diatom for the production of biopharmaceuticals. In *P. tricornutum*, nuclear expression of transgenes is usually mediated through the promoter of the *FCPA* gene which encodes the fucoxanthin chlorophyll a/c binding protein (Apt et al., 1996; Zaslavskaya et al., 2000) or through the nitrate reductase promoter which is induced when ammonium is replaced by nitrates as nitrogen source in the culture medium (Poulsen and Kröger, 2005; Gonzalez et al., 2011; Hempel et al., 2011b; Hempel and Maier, 2012; Stork et al., 2012). This specific nitrate reductase inducible promoter was used to express the fully-assembled human IgG antibody against Hepatitis B leading to 1.5–2.5 mg of recombinant antibody per liter of culture medium depending of the different clones (Hempel and Maier, 2012) and 21 mg per gram of algal dry weight for the ER-retained form of the antibody (Hempel et al., 2011b).

Expression of commercially relevant proteins in other species of microalgae remains less reported, since tools for nuclear transformation are still missing. As for *C. reinhardtii*, the CaMV35S promoter has been shown to be efficient in several other green microalgae, including *C. ellipsoidea* (Jarvis and Brown, 1991; Chen et al., 2001; Kim et al., 2002), *C. vulgaris* (Hawkins and Nakamura, 1999), *Haematococcus pluvialis* (Kathiresan et al., 2009) and *D. salina* (Geng et al., 2003; Feng et al., 2009; Chai et al., 2013) but in most of the cases, the expression levels achieved remained low. Some studies reported the use of other foreign promoters to drive the nuclear expression of heterologous proteins in *Chlorella* sp., such as the *C. reinhardtii* RBCS2 promoter which has been used successfully in *Chlorella ellipsoidea* for transient expression of recombinant protein or resistance expression (Hawkins and Nakamura, 1999; Kim et al., 2002). In contrast, the ubiquitin maize promoter seems to be promising in *Chlorella* sp., since it allows the stable expression of a rabbit gene encoding for an α -defensin (Bai et al., 2013). In some cases, the omega element of Tobacco Mosaic Virus (TMV) which is part of the 5'UTR has been used to enhance translation efficiency (Chen et al., 2001), leading to a yield of about 11 mg/L of recombinant protein (Bai et al., 2013). This ubiquitin- Ω promoter was also successfully used in *D. salina* to stably express a gene encoding the hepatitis B surface antigen (Geng et al., 2003).

EXPRESSION OF SECRETED THERAPEUTIC RECOMBINANT PROTEINS IN MICROALGAE

As mentioned previously, most of the recombinant proteins that have been produced in *C. reinhardtii* so far were expressed in the chloroplast because of the high level of protein accumulation reached in this organelle. Thus, among the 20 recombinant proteins of industrial interest expressed in *C. reinhardtii* (for a recent review, see Rasala and Mayfield, 2014), only 3 of them have been expressed successfully through the nuclear genome: a xylanase (Rasala et al., 2012), an ice binding-protein (Lauersen et al., 2013b) and the secreted EPO (Eichler-Stahlberg et al., 2009). As the xylanase and ice binding-protein are not therapeutic proteins, we will not discuss them further in this review. In addition, by using the signal peptide of the *C. reinhardtii* gene *ARS2* encoding a periplasmic arylsulfatase, the EPO has been produced in the culture medium of *C. reinhardtii* (Eichler-Stahlberg et al., 2009; **Table 1**). Furthermore, as previously indicated, Hempel and

Table 1 | Post-translational modifications of biopharmaceuticals expressed in the secretory system of microalgae.

Expressed protein	Promoter	N-glycosylation site ^a	Glycosylation ^b	References
IN C. REINHARDTII				
Human EPO	Hsp70A/RbcS2	Yes	Yes	Eichler-Stahlberg et al., 2009
IN P. TRICORNUTUM				
Monoclonal human IgG against the Hepatitis B surface antigen	Nitrate reductase	Yes	Yes	Hempel et al., 2011b; Hempel and Maier, 2012
Hepatitis B surface antigen	Nitrate reductase	Yes	No	Hempel et al., 2011b
IN D. SALINA				
Hepatitis B surface antigen	Maize Ubiquitin +Ω TMV enhancer	Yes	No	Geng et al., 2003

^aPredicted N-glycosylation sites by bio-informatic analysis of the protein sequence.

^bExperimental evidence of the presence of N-glycans attached to the N-glycosylation site.

Maier (2012) demonstrated the capability of *P. tricornutum* to synthesize and secrete a full length functional human IgG antibody against the Hepatitis B virus surface antigen. This study clearly demonstrates that diatoms are able to produce and correctly assemble complex proteins without affecting their biological activity (**Table 1**).

Little information is available on the post-translational modifications acquired by these microalgae-made therapeutic proteins in the secretory pathway. For example, the recombinant EPO expressed in *C. reinhardtii* exhibited a molecular mass of about 33 kDa suggesting the presence of post-translational modifications such as the addition of glycans on the microalgae-made EPO. This is consistent with the fact that the EPO is known to possess 3 N-glycosylation and one O-glycosylation sites (**Table 1**) (Lingg et al., 2012). Moreover, affinoblotting with concanavalin A, a lectin specific for oligomannoside structures (Fitchette et al., 2007), of the ER-resident recombinant IgG expressed in *P. tricornutum* also suggested that this microalgae-made antibody is glycosylated (**Table 1**) (Hempel et al., 2011b). However, no structural detailed analyses are reported regarding the glycans attached to those microalgae-made biopharmaceuticals. Since the N-glycosylation of biopharmaceuticals is critical for their half-life, stability and biological activity (for a recent review, see Lingg et al., 2012), it is therefore essential to characterize the N-glycosylation process of secreted protein in microalgae.

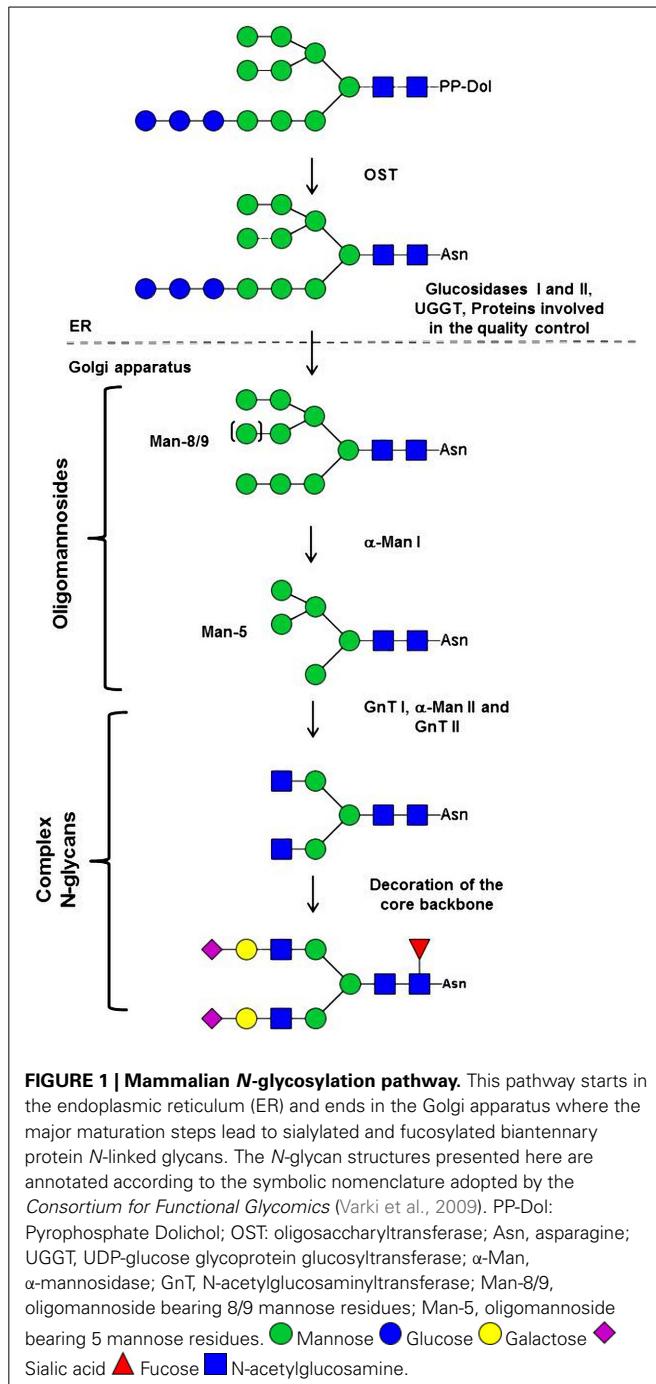
N-GLYCOSYLATION PATHWAYS IN MICROALGAE

GENERAL ASPECTS OF N-GLYCOSYLATION IN EUKARYOTES

N-glycosylation is a major co- and post-translational modification of proteins in eukaryotes occurring in both the ER and the Golgi apparatus (**Figure 1**). In this process, a lipid-linked oligosaccharide is first assembled by the stepwise addition of monosaccharides on a dolichol pyrophosphate on the cytosolic face and then in the lumen of the ER (Burda and Aebi, 1999). This precursor is then transferred by the oligosaccharyltransferase (OST) complex onto the asparagine residues of the consensus Asn-X-Ser/Thr sequences of a protein (Burda and Aebi, 1999). In 3.5% of the cases, other sequences like Asn-X-Cys, Asn-X-Val have been proven to be glycosylated on endogenous or recombinant protein produced both in mammals or plant cells (Gil et al., 2009; Zielinska et al., 2010; Matsui et al., 2011). The precursor is deglycosylated by the α-glucosidases I

and II and then reglycosylated by an UDP-glucose: glycoprotein glucosyltransferase (UGGT) to ensure the proper folding of the nascent protein through its interaction with ER-resident chaperones, such as calnexin and calreticulin. These ER events are conserved in eukaryotes because they are crucial for efficient protein folding and oligomerization (Helenius and Aebi, 2001). In contrast, evolutionary adaptation of N-glycan processing in the Golgi apparatus has given rise to a variety of organism-specific complex structures (Varki, 2011). First, α-mannosidases degrade the oligosaccharide precursor into oligomannosides ranging from Man₉GlcNAc₂ to Man₅GlcNAc₂ (Man-9 to Man-5). N-acetylglucosaminyltransferase I (GnT I) then transfers a first N-acetylglucosaminyl (GlcNAc) residue on Man-5 and initiates the synthesis of a large variety of structurally different complex-type N-glycans. In this GnT I-dependent N-glycan maturation, the processing continues by the removal of two mannose residues and the transfer of a second terminal GlcNAc residue, thus resulting in the synthesis of a core GlcNAc₂Man₃GlcNAc₂ which are common to mammals and all land plants studied so far (Lerouge et al., 1998; Wilson et al., 2001; Gomord et al., 2010; Varki, 2011) (**Figure 1**). This core is then decorated by the action of a specific repertoire of glycosyltransferases that differ from one organism to another. As a consequence, mature proteins leaving the secretory pathway carry organism-specific complex N-glycans allowing the protein to acquire a set of glycan-mediated biological functions (Varki, 1993; Gagneux and Varki, 1999). For instance, in mammals, most secreted proteins in blood circulation carry biantennary N-glycans decorated with α(1,6)-fucose residues and terminal sialic acids that impact either the protein activity or protein half-life (**Figure 1**; Lingg et al., 2012). In contrast, plant N-glycans are mainly of biantennary complex type N-glycan carrying a core-β(1,2)-xylose; a core α(1,3)-fucose and eventually terminal Lewis a antennae (Lerouge et al., 1998; Wilson et al., 2001).

By comparison with data available in pluricellular eukaryotes, information concerning protein N-glycosylation in microalgae remains very limited. A few studies using lectin blot analysis or enzymatic sequencing suggested that proteins secreted by green microalgae carry mainly oligomannosides or complex N-glycans having a core xylose residue (Balshüsemann and Jaenicke, 1990; Grunow et al., 1993; Gödel et al., 2000). More recently, a cell wall glycoprotein from the red microalgae *Porphyridium* sp. was



found to carry Man-8 and Man-9 oligomannosides containing 6-O-methyl mannoses and substituted by one or two xylose residues (Levy-Ontman et al., 2011).

STRUCTURAL INVESTIGATION OF GLYCAN N-LINKED TO MICROALGAE PROTEINS OF *CHLAMYDOMONAS REINHARDTII* AND *PHAEODACTYLUM TRICORNUTUM*

Deeper insights into the structure of glycans N-linked to proteins secreted by *C. reinhardtii* (Mathieu-Rivet et al., 2013) and *P. tricornutum* (Baïet et al., 2011) have been recently reported and N-glycosylation pathways in these two microalgae dedicated

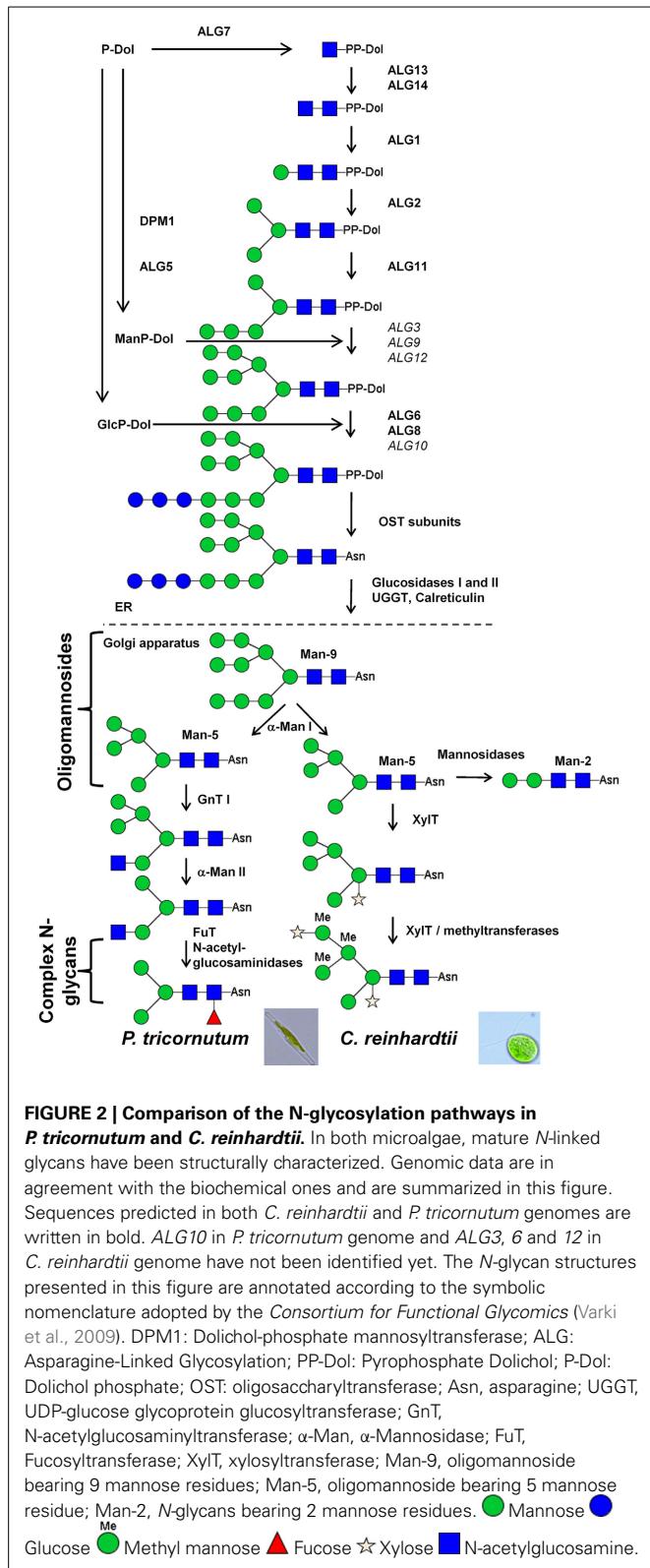
to biotechnology applications have been proposed. These studies combined data resulting from the search for genes encoding putative enzymes involved in the N-glycosylation pathway in the genomic databases and data resulting from biochemical analysis of glycan N-linked to secreted proteins.

Bio-informatic analysis

In *C. reinhardtii* and *P. tricornutum* genomes, most of the genes encoding enzymes involved in the biosynthesis of the dolichol pyrophosphate-linked oligosaccharide on the cytosolic face and in the lumen of the ER, named Asparagine-Linked Glycosylation (ALG) (Weerapana and Imperiali, 2006), are predicted (Figure 2). Although some ALG were not clearly identified in the genomes, large oligomannosides up to Man-9 were found in both *C. reinhardtii* (Mathieu-Rivet et al., 2013) and *P. tricornutum* (Baïet et al., 2011) suggesting that the synthesis of the oligosaccharide precursor is similar to the one described in other eukaryotes. In addition to ALG, genes encoding subunits of the oligosaccharyltransferase were also identified in both genomes (Figure 2). Glucosidases I and II, as well as ER-resident UGGT and chaperones such as calreticulin, are also predicted. These proteins are key elements of the quality control of proteins occurring in the ER and are crucial for acquisition of the proper folding of the nascent glycoprotein (Figure 2).

After the transfer of the glycoprotein into the Golgi apparatus, maturation of N-linked glycans starts with the trimming of Man-9/8 by α -mannosidases I (α -Man I) that generate oligomannosides ranging from Man-8 to Man-5. Genes encoding α -Man I are predicted in *C. reinhardtii* and *P. tricornutum* genomes (Figure 2) suggesting that Golgi mediated trimming of oligomannosides also occurs in both microalgae. In higher eukaryotes including animals, insects and land plants, N-glycans are then processed by GnT I that transfers a first N-acetylglucosaminyl (GlcNAc) residue on the α (1,3)-mannose arm of Man-5 (Figure 1). Thus, GnT I is a key enzyme since it is the first glycosyltransferase occurring in the Golgi apparatus in the GnT I-dependent pathway giving rise to complex N-glycans that are required for normal morphogenesis in pluricellular organisms (Ioffe and Stanley, 1994; Metzler et al., 1994). A sequence encoding for a GnT I is predicted in *P. tricornutum* (Baïet et al., 2011) but not in *C. reinhardtii* (Mathieu-Rivet et al., 2013) suggesting that N-linked glycans from these two microalgae could be processed in the Golgi apparatus according to two different pathways, referred to as GnT I-dependent and GnT I-independent pathways (Zhu et al., 2004; Crispin et al., 2006; Grass et al., 2011).

In both GnT I-dependent and independent pathways, the next steps mainly consist in the transfer of monosaccharides on the N-glycan core to synthesize highly diverse complex N-glycans. Among residues that are added to the core, fucosyl residues either α (1,3)- or α (1,6)-linked to the proximal GlcNAc are commonly observed. In both microalgae, sequences encoding for putative α (1,3)-fucosyltransferases were identified in the genomes (Baïet et al., 2011; Mathieu-Rivet et al., 2013). Search for other genes encoding for putative N-glycan processing enzymes revealed sequences exhibiting homologies with α -Man II and GnT II in *P. tricornutum* but functional characterizations remain necessary to assess their enzymatic activities.



Biochemical analysis

A detailed glycomic analysis of *C. reinhardtii* proteins indicated that both secreted and membrane-bound proteins carry oligomannosides ranging from Man-2 to Man-5 and also complex

N-glycans containing 6-O-methyl mannoses and substituted by one or two xylose residues (Figure 2) (Mathieu-Rivet et al., 2013). Similar complex structures have been previously found in a cell wall glycoprotein isolated from the red microalga *Porphyridium sp.* (Levy-Ontman et al., 2011). However, the location of the xylose residue on the core N-glycan differs in these two microalgae, since it is described to be linked to the chitobiose unit in *Porphyridium* rather than to C2 of the β -mannose as described in *C. reinhardtii* and as previously reported in land plants (Figure 2) (Lerouge et al., 1998). In agreement with bio-informatics data suggesting the absence of GnT I, these complex N-glycans are likely to result from the transfer of xylose residues onto Man-5 oligomannosides in a GnT I-independent manner. Then, methylation of mannose residues is thought to occur after xylosylation of complex N-glycans in *C. reinhardtii* as proposed in Figure 2. Alternatively, Man-5 can be trimmed by mannosidases down to Man-2 (Figure 2). Although O-methyl mannose residues was previously reported in some eukaryotes (Staudacher, 2012), complex N-glycans from *C. reinhardtii* highly differ from glycans described in animals and land plants. We postulate that xylosylation and methylation of oligomannosides in these microalgae may protect glycoproteins against deglycosylating enzymes such as endoglycosidases or peptide N-glycosidases.

In contrast to complex N-glycans from *C. reinhardtii*, glycans N-linked to proteins secreted by the diatom *P. tricornutum* can be processed through a GnT I-dependent pathway into partially fucosylated Man-3 (Figure 2) (Baiet et al., 2011). The GnT I gene predicted in the *P. tricornutum* genome was demonstrated to encode an active enzyme able to restore the maturation of N-linked glycans into complex-type N-glycans in the CHO *Lec1* mutant, defective in its endogenous GnT I (Baiet et al., 2011). The authors have proposed that this fucosylated N-linked glycan results from the addition of a terminal GlcNAc residue by GnT I on Man-5 followed by removal of two Man residues by an α -Man II and the transfer of a fucose residue by the predicted α (1,3)-FuT (Baiet et al., 2011). As illustrated in the proposed pathway depicted in Figure 2, the terminal GlcNAc introduced in the Golgi apparatus by the *P. tricornutum* GnT I is then likely removed in the secretory pathway by glucosaminidases as previously described in land plants and insect (Vitale and Chrispeels, 1984; Altmann et al., 1995). Efficient complementation of the CHO *Lec1* mutant is the main argument supporting that this diatom processes its N-glycans through a GnT I-dependent pathway. This result also indicates that the microalga transferase is properly targeted to the Golgi apparatus when expressed in mammalian cells.

In these two studies performed on *P. tricornutum* and *C. reinhardtii* N-glycan processing and as described for land plants (Séveno et al., 2004; Zeleny et al., 2006), no sialic acid residues were identified (Baiet et al., 2011; Mathieu-Rivet et al., 2013). Sialic acids are terminal residues of O- and N-glycans that are specifically involved in many biological functions in mammals, such as the half-life of blood proteins and cell-cell adhesion processes (Figure 1) (Varki, 1993; Gagneux and Varki, 1999).

GLYCAN-REMODELING STRATEGIES IN MICROALGAE: PERSPECTIVES

Differences in specificity of Golgi transferases and glycosidases between eukaryotes give rise to glycosylation profiles that differ between mammals and other eukaryotic host cells used as cell factories. As a consequence, glycans *N*-linked to recombinant proteins produced in plants, yeast or even in animal cells differ from the original therapeutic proteins. This may result in either a decrease or absence of biological activity. Furthermore, unsuitable *N*-glycan structures introduced by the expression system can induce immune responses in humans and generate adverse reactions (Van Beers and Bardor, 2012), as reported for $\alpha(1,3)$ -Gal epitope or Neu5Gc on therapeutic drugs (Chung et al., 2008; Padler-Karavani et al., 2008). Similarly, plant synthesize *N*-glycans carrying immunogenic core-xylose and core- $\alpha(1,3)$ -fucose that may induce immune responses in human treated with plant-made biopharmaceuticals (Bardor et al., 2003). In consequence, whatever the expression system that is considered, strategies have been carried out for the *in vivo* remodeling of the protein *N*-glycans in order to obtain structures that meet pharmaceutical requirements. For instance, knock-out of endogenous genes involved in the transfer of core immunogenic epitopes have been carried out to engineer oligosaccharides synthesized by plant cell into human-compatible structures (Cox et al., 2006; Schähs et al., 2007; Strasser et al., 2008). Furthermore, knock-in methodologies based on the expression in the host cells of mammalian enzyme have been developed to *in vivo* introduce missing glycan sequences. In plants, these efforts resulted in the production of plant-derived therapeutic proteins carrying *N*-glycans similar to those found on human counterpart (Palacpac et al., 1999; Bakker et al., 2001, 2006; Misaki et al., 2006; Paccalet et al., 2007; Rouwendal et al., 2007; Castilho et al., 2008, 2010, 2013).

As reported in section Microalgae as Alternative Systems for Production of Recombinant Proteins, *C. reinhardtii* and *P. tricornutum* have been evaluated for their capacity to express therapeutic proteins. The recent characterization of glycans *N*-linked to their secreted proteins allows the design of strategies to engineer their *N*-glycan pathways for the production of human-compatible therapeutic proteins.

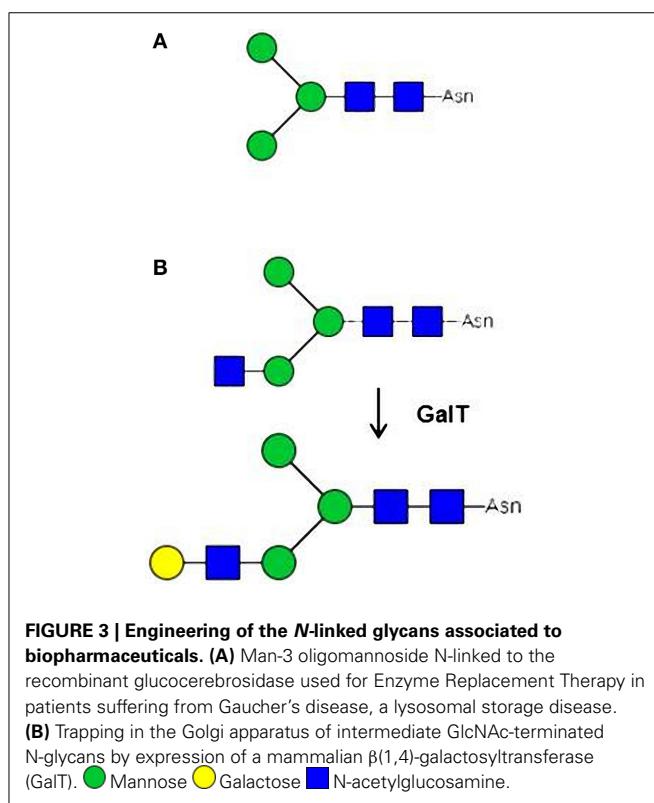
MICROALGAE FOR THE PRODUCTION OF LYSOSOMAL PROTEINS

Although complex *N*-glycans have been identified on proteins secreted by *C. reinhardtii* and *P. tricornutum*, most abundant oligosaccharides *N*-linked to proteins are oligomannosides ranging from Man-2 to Man-5 in *C. reinhardtii* (Mathieu-Rivet et al., 2013) and from Man-5 to Man-9 in *P. tricornutum* (Baïet et al., 2011) (Figure 2). These *N*-linked glycans are appropriate for the production of lysosomal therapeutics. For example, glucocerebrosidase is a glycoprotein drug administered intravenously into patients suffering from Gaucher's disease, a lysosomal storage disease. The effective targeting and internalization of this therapeutic drug into macrophages depend on terminal mannose residues of its *N*-glycans which are recognized by macrophage cell surface mannose receptors (Van Patten et al., 2007). For treatment, exogenous glucocerebrosidase is administered intravenously into patients. The preparation of the current

glucocerebrosidase expressed in CHO cells (Cerezyme®) requires *in vitro* post-purification and exoglycosidase digestions to expose the trimannose core (Man₃GlcNAc₂, Man-3) using a combination of at least three enzymes namely sialidase, galactosidase and *N*-acetylglucosaminidase (Figure 3A). These *in vitro* glyco-engineering steps increase considerably the production costs (Weinreb, 2008). Therefore, alternative expression systems that are capable of producing mannose terminated *N*-glycans have been developed as safe and cost effective production methods. These include cultured carrot cells (Shaaltiel et al., 2007) and *A. thaliana cgl* transgenic plants (He et al., 2012). However, the analysis of the *N*-glycan profiles of these glucocerebrosidase bio-similars revealed Man-5 with variable amount of Man-6 to Man-9 structures (for example low amount of Man 6 to Man-8 on glucocerebrosidase expressed in the *Arabidopsis cgl* mutant). The presence of such structures required careful biosafety studies to evaluate the possible binding to serum mannose binding lectin (MBL) and immunogenicity (Van Patten et al., 2007; Grabowski et al., 2014). The results described for *C. reinhardtii* (Mathieu-Rivet et al., 2013) suggest that this green microalgae could become in the future an appropriate platform for the production of a glucocerebrosidase bio-better which would naturally carry Man-2 to Man-5 glycan structures without any *in vitro* trimming of mannose residues (Figure 2).

MICROALGAE FOR THE PRODUCTION OF PROTEINS CARRYING HUMAN-COMPATIBLE COMPLEX-TYPE *N*-GLYCANS

Based on data reported in section N-glycosylation Pathways in Microalgae and Figure 2, the production in microalgae of



therapeutic proteins carrying human-like complex glycans will require intensive remodeling of their glycan pathways. Although no information is available on their immunogenicity, glycan epitopes identified on microalgae complex *N*-glycans will likely induce adverse effects in humans. Removal of these putative immunogenic glyco-epitopes, such as xylose and fucose residues, as well as methyl groups on mannose residues, will first require to identify gene encoding the corresponding glycosyltransferases and methyltransferases and then to carry out appropriate knock-out strategies such as RNA interference; artificial microRNA which have already been developed in microalgae (Zhao et al., 2009; Cerutti et al., 2011). If putative FuT genes have already been identified because of their high homologies with plant α (1,3)-FuT, the identification of sequences coding for xylosyltransferases (XylT) will be more tricky. Only the β (1,2)-XylT family involved in complex *N*-glycan biosynthesis in land plants have been biochemically characterized so far (Strasser et al., 2000). To date, no specific domain for XylT activity has been clearly identified and so, search by homology for XylT in microalgae genomes failed to clearly identify putative candidates.

As previously reported in land plants, “missing” residues on glycan *N*-linked to therapeutic proteins can be introduced by co-expression with appropriate glycosyltransferases. Results regarding the complementation of CHO *Lec1* mutant by *P. tricornutum* GnT I (Baïet et al., 2011) showed that the targeting of Golgi glycosyltransferases is conserved between microalgae and higher eukaryotes. As a consequence, we can expect that the expression in microalgae of heterologous glycosyltransferases would allow the *in vivo* remodeling of their *N*-glycan pathways. This heterologous expression of glycosyltransferases should be particularly suitable for the remodeling of *N*-glycans in *P. tricornutum* because this diatom possesses a functional GnT I making its *N*-glycosylation pathway closer to the mammalian one (Figures 1, 2). The terminal GlcNAc introduced by GnT I constitute the starting point for the building of antennae in mammalian complex *N*-glycans (Figure 1). We postulated that in *P. tricornutum* this terminal GlcNAc is lost during the trafficking of the glycoprotein by action of N-acetylglucosaminidases (Figure 2). However, the expression of a mammalian β (1,4)-galactosyltransferase in *P. tricornutum* would allow the transfer in the Golgi apparatus of a terminal Gal on the terminal GlcNAc, thus giving rise to a Gal β (1,4)GlcNAc extension that will protect this N-acetylglucosamine residue introduced by GnT I from degradation by N-acetylglucosaminidases downstream in the secretory pathway (Figure 3B). In addition, this extension corresponds to glycosidic motif found on some therapeutic proteins such as IgG. This strategy based on the trapping of intermediate GlcNAc-terminated *N*-glycans has already been successfully carried out in land plants (Bakker et al., 2001; Huether et al., 2005; Vézina et al., 2009) and insect cells (Hollister et al., 1998).

SEARCH FOR MICROALGAE HAVING APPROPRIATE *N*-GLYCOSYLATION PATHWAYS

As *in vivo* engineering of the *N*-glycosylation pathways in *C. reinhardtii* and *P. tricornutum* may represent a tricky work, exploration for a microalga that exhibits a more appropriate protein glycosylation has to be considered. A first overview of protein

N-glycosylation in various microalgae species can be drawn on the basis of public genomic databases. In addition to *C. reinhardtii* and *P. tricornutum*, more than 15 genomes from other microalgae are now available (e.g., *Thalassiosira pseudonana*; *Cyanidioschyzon merolae*; *Ostreococcus tauri* and *Ostreococcus lucimarinus*; *Micromonas pusilla*; *Chlorella vulgaris*; *Coccomyxa subelipsoidea*; *Nannochloropsis gaditana*; *Monoraphidium neglectum*; Table 2) (Ambrust et al., 2004; Misumi et al., 2005; Palenik et al., 2007; Worden et al., 2009; Blanc et al., 2010, 2012; Radakovits et al., 2012; Vieler et al., 2012; Bogen et al., 2013). Searches for ALG genes and other sequences encoding proteins involved in the ER protein quality control indicated that enzymes of the ER machinery are predicted in all genomes suggesting that ER steps of the *N*-glycan pathway are conserved over the microalgae phyla as discussed recently in Levy-Ontman et al. (2014). With regard to Golgi events, Table 2 summarizes and reports the genes predicted to encode enzymes involved in the maturation of the *N*-glycans. As observed for *C. reinhardtii* and *P. tricornutum*, α -mannosidases (CAZy GH 47) and α (1,3)-fucosyltransferases (CAZy GT10) are predicted in most microalgae genomes. In contrast, genes encoding GnT I are not predicted in all genomes (Table 2; Baïet et al., 2011). For instance, as for *C. reinhardtii*, no GnT I was found in *Volvox* and *Ostreococcus* species whereas this key transferase is predicted in most other microalgae, including haptophytes and cryptophytes (Table 2; Baïet et al., 2011). As demonstrated in *C. reinhardtii* and *P. tricornutum*, this indicates that the Golgi maturation of protein *N*-linked glycans into complex oligosaccharides could either occur through a GnT I-dependent pathway or a GnT I-independent pathway in microalgae. In the context of microalgae-made pharmaceutical production, this feature must be considered since the engineering of microalgae *N*-glycans into human compatible oligosaccharides would be facilitated if a GnT I-dependent machinery already exists in the strains selected as an expression system.

CONCLUSIONS

In conclusion, it can be considered that remodeling *N*-glycans as depicted in Figure 2 into human-compatible oligosaccharides would be a difficult challenge to attend. However, recent production in plants of therapeutic proteins carrying sialylated biantennary *N*-glycans demonstrates that such a glycan engineering can be achieved in a plant expression system (Castilho et al., 2013). It should also be noticed that only a few microalgae species have been investigated and so, data reported in this review should be considered as a starting point regarding protein glycosylation mechanisms occurring in these unicellular organisms and will serve as foundation for further glycobiology works. Efforts have now to be carried out to get more information concerning protein *N*-glycan structures and their pathways in different phyla to identify microalgae species that are more appropriate for glycan remodeling strategies into human-like structures. This will also necessitate the characterization of the specificity of Golgi glycosyltransferases, such as α -mannosidases and α (1,3)-fucosyltransferases predicted in the microalgae genomes, to evaluate the efficiency in microalgae of heterologous expression of glycosyltransferases as well as to identify genes encoding transferases (methyltransferases and xylosyltransferases) involved in

Table 2 | Genes predicted in microalgae genomes encoding proteins involved in N-glycan Golgi maturation such as α -Man I, mannosyl oligosaccharide alpha-1,2-mannosidase; GnT I, alpha-1,3-mannosyl-glycoprotein beta-1,2-N-acetylglucosaminyltransferase; α (1,3)-FucT, core α (1,3)-fucosyltransferase.

Microalgae	α -Man I (K01230)	GnT I (K00726)	α (1,3)-FucT (K00753)
Green algae	<i>Chlamydomonas reinhardtii</i>	XP_001700094.1	No
	<i>Volvox carteri f. nagariensis</i>	XP_002957696.1	No
	<i>Ostreococcus lucimarinus</i>	XP_001421581.1	No
	<i>Ostreococcus tauri</i>	XP_003083553.1	No
	<i>Micromonas sp RCC299</i>	XP_002505356.1 XP_002499658.1	XP_002507699.1 XP_002506141.1
	<i>Micromonas pusilla</i>	XP_003058165.1 XP_003060259.1	XP_003056751.1
	<i>Coccomyxa subellipsoidea</i> C-169	XP_005648743.1	XP_005642764.1
	<i>Chlorella variabilis</i> NC64A	XP_005852238.1	XP_005843911.1
Red algae	<i>Cyanidioschyzon merolae</i>	XP_005535714.1	No
	<i>Galdieria sulphuraria</i>	NO	XP_005709370.1 and XP_005709369.1
Stramenopile	<i>Phaeodactylum tricornutum</i>	XP_002176357.1	XP_002180610.1 XP_002180609.1 XP_002180883.1
	<i>Thalassiosira pseudonana</i>	XP_002289677.1 XP_002291430.1 XP_002289678.1	XP_002286885.1
	<i>Fragilaropsis cylindrus</i>	JGI Protein Id: 168118 JGI Protein Id: 261302	JGI Protein ID: 189180
	<i>Aureococcus anophagefferens</i>	EGB10338.1 EGB09525.1 EGB06148.1 EGB09718.1	EGB09791.1 EGB12497.1 EGB10167.1
	<i>Nannochloropsis gaditana</i> CCMP526	XP_005853502.1 XP_005853501.1	No
	<i>Emiliania huxley</i>	XP_005786216.1 XP_005777157.1 XP_005762577.1 XP_005758261.1 EOD18227.1 XP_005771362.1 XP_005771219.1	XP_005769752.1 XP_005762959.1 XP_005788156.1
	<i>Guillardia theta</i>	XP_005835818.1 XP_005827979.1	XP_005818934.1
			XP_005818444.1 XP_005834963.1 XP_005839924.1 XP_005827855.1 XP_005835841.1 XP_005829578.1
Sequences were retrieved using the KEGG orthology or by search with the protein's name (as defined in KEGG) in JGI, Phytozome or Genome at NCBI databanks. In addition, search for candidates was carried out by BLASTP or TBLASTN using human alpha-1,2-mannosidase protein sequences, rabbit GnT I (P27115) and <i>Arabidopsis thaliana</i> α (1,3)-FucT (Q9LJK1 and Q9FX97) as query sequences.			

the synthesis of putative immunogenic epitopes with the final goal to selectively inactivate their expression.

AUTHOR CONTRIBUTION

Elodie Mathieu-Rivet and Marie-Christine Kiefer-Meyer equally contributed to this work by drafting the manuscript and performing the genomic analyses of the microalgae genomes which are summarized in the Table 2 of the manuscript. Gaëtan Vanier, Clément Ovide and Carole Burel participated to the writing and took care of the figure drawing. Patrice Lerouge and Muriel Bardor came out with the idea of this review, participate to the writing and coordinated the efforts from the team prior to take care of the submission of the manuscript.

ACKNOWLEDGMENTS

We apologize to colleagues whose work is not cited due to space limitations. Part of the work described in this review was supported financially by the University of Rouen, France as well as the French National Agency for research, ANR, within the PLANT-KBBE 2008 program in the frame of the ALGALGLYCO project. Finally, authors especially thank Dr Zhang Peiqing, Bioprocessing Technology Institute, Singapore and Prof. Driouch Azeddine, Laboratoire Glyco-MEV, University of Rouen, France for critical reading of this manuscript.

REFERENCES

- Altmann, F., Schwihla, H., Staudacher, E., Glössl, J., and März, L. (1995). Insect cells contain an unusual, membrane-bound N-acetylglucosaminidase probably involved in the processing of protein N-glycans. *J. Biol. Chem.* 270, 17344–17349. doi: 10.1074/jbc.270.29.17344
- Ambrust, E. V., Berges, J., Bowler, C., Green, B., Martinez, D., Putnam, N., et al. (2004). *The Genome of the Diatom Thalassiosira Pseudonana: Ecology, Evolution, and Metabolism*. Lawrence Berkeley National Laboratory. Available online at: <http://escholarship.org/uc/item/03j415ph> [Accessed June 16, 2014].
- Apt, K. E., Grossman, A. R., and Kroth-Pancic, P. G. (1996). Stable nuclear transformation of the diatom Phaeodactylum tricornutum. *Mol. Gen. Genet.* 252, 572–579.
- Baiët, B., Burel, C., Saint-Jean, B., Louvet, R., Menu-Bouaouiche, L., Kiefer-Meyer, M.-C., et al. (2011). N-glycans of Phaeodactylum tricornutum diatom and functional characterization of its N-acetylglucosaminyltransferase I enzyme. *J. Biol. Chem.* 286, 6152–6164. doi: 10.1074/jbc.M110.175711
- Bai, L.-L., Yin, W.-B., Chen, Y.-H., Niu, L.-L., Sun, Y.-R., Zhao, S.-M., et al. (2013). A new strategy to produce a defensin: stable production of mutated NP-1 in nitrate reductase-deficient chlorella ellipoidea. *PLoS ONE* 8:e54966. doi: 10.1371/journal.pone.0054966
- Bakker, H., Bardor, M., Molthoff, J. W., Gomord, V., Elbers, I., Stevens, L. H., et al. (2001). Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2899–2904. doi: 10.1073/pnas.0314998
- Bakker, H., Rouwendal, G. J., Karnoup, A. S., Florack, D. E., Stoopen, G. M., Helsper, J. P., et al. (2006). An antibody produced in tobacco expressing a hybrid β-1, 4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7577–7582. doi: 10.1073/pnas.0600879103
- Balshüsemann, D., and Jaenicke, L. (1990). The oligosaccharides of the glycoprotein pheromone of *Volvox carteri f. nagariensis* Iyengar (Chlorophyceae). *Eur. J. Biochem.* 192, 231–237. doi: 10.1111/j.1432-1033.1990.tb19220.x
- Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., et al. (2003). Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. doi: 10.1093/glycob/cwg024
- Beer, L. L., Boyd, E. S., Peters, J. W., and Posewitz, M. C. (2009). Engineering algae for biohydrogen and biofuel production. *Curr. Opin. Biotechnol.* 20, 264–271. doi: 10.1016/j.copbio.2009.06.002
- Blanc, G., Agarkova, I., Grimwood, J., Kuo, A., Brueggeman, A., Dunigan, D. D., et al. (2012). The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol.* 13:R39. doi: 10.1186/gb-2012-13-5-r39
- Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., et al. (2010). The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell Online* 22, 2943–2955. doi: 10.1105/tpc.110.076406
- Bogen, C., Al-Dilaimi, A., Albersmeier, A., Wichmann, J., Grundmann, M., Rupp, O., et al. (2013). Reconstruction of the lipid metabolism for the microalga *Monoraphidium neglectum* from its genome sequence reveals characteristics suitable for biofuel production. *BMC Genomics* 14:926. doi: 10.1186/1471-2164-14-926
- Burda, P., and Aebi, M. (1999). The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta* 1426, 239–257. doi: 10.1016/S0304-4165(98)00127-5
- Buyel, J. F., and Fischer, R. (2014). Downstream processing of biopharmaceutical proteins produced in plants: the pros and cons of flocculants. *Bioengineered* 5, 138–142. doi: 10.4161/bioe.28061
- Cadoret, J.-P., and Bernard, O. (2008). La production de biocarburant lipidique avec des microalgues: promesses et défis. *J. Soc. Biol.* 202, 201–211. doi: 10.1051/jbio:2008022
- Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., et al. (2013). Generation of biologically active multi-sialylated recombinant human EPOFc in plants. *PLoS ONE* 8:e54836. doi: 10.1371/journal.pone.0054836
- Castilho, A., Pabst, M., Leonard, R., Veit, C., Altmann, F., Mach, L., et al. (2008). Construction of a functional CMP-sialic acid biosynthesis pathway in *Arabidopsis*. *Plant Physiol.* 147, 331–339. doi: 10.1104/pp.108.117572
- Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gattinger, P., et al. (2010). In planta protein sialylation through overexpression of the respective mammalian pathway. *J. Biol. Chem.* 285, 15923–15930. doi: 10.1074/jbc.M109.088401
- Cerutti, H., Johnson, A. M., Gillham, N. W., and Boynton, J. E. (1997). Epigenetic silencing of a foreign gene in nuclear transformants of *Chlamydomonas*. *Plant Cell* 9, 925–945. doi: 10.1105/tpc.9.6.925
- Cerutti, H., Ma, X., Msanne, J., and Repas, T. (2011). RNA-mediated silencing in algae: biological roles and tools for analysis of gene function. *Eukaryot Cell* 10, 1164–1172. doi: 10.1128/EC.05106-11
- Chai, X.-J., Chen, H.-X., Xu, W.-Q., and Xu, Y.-W. (2013). Expression of soybean Kunitz trypsin inhibitor gene SKTI in *Dunaliella salina*. *J. Appl. Phycol.* 25, 139–144. doi: 10.1007/s10811-012-9847-8
- Chen, Y., Wang, Y., Sun, Y., Zhang, L., and Li, W. (2001). Highly efficient expression of rabbit neutrophil peptide-1 gene in Chlorella ellipoidea cells. *Curr. Genet.* 39, 365–370. doi: 10.1007/s002940100205
- Chung, C. H., Mirakhur, B., Chan, E., Le, Q.-T., Berlin, J., Morse, M., et al. (2008). Cetuximab-induced anaphylaxis and IgE specific for galactose-α-1, 3-galactose. *New Engl. J. Med.* 358, 1109–1117. doi: 10.1056/NEJMoa074943
- Colgan, R., Atkinson, C. J., Paul, M., Hassan, S., Drake, P. M. W., Sexton, A. L., et al. (2010). Optimisation of contained *Nicotiana tabacum* cultivation for the production of recombinant protein pharmaceuticals. *Transgenic Res.* 19, 241–256. doi: 10.1007/s11248-009-9303-y
- Cox, K. M., Sterling, J. D., Regan, J. T., Gasdaska, J. R., Frantz, K. K., Peele, C. G., et al. (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* 24, 1591–1597. doi: 10.1038/nbt1260
- Crispin, M., Harvey, D. J., Chang, V. T., Yu, C., Aricescu, A. R., Jones, E. Y., et al. (2006). Inhibition of hybrid- and complex-type glycosylation reveals the presence of the GlcNAc transferase I-independent fucosylation pathway. *Glycobiology* 16, 748–756. doi: 10.1093/glycob/cwj119
- Decker, E. L., and Reski, R. (2012). Glycoprotein production in moss bioreactors. *Plant Cell Rep.* 31, 453–460. doi: 10.1007/s00299-011-1152-5
- Demain, A. L., and Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* 27, 297–306. doi: 10.1016/j.biotechadv.2009.01.008
- Demurtas, O. C., Massa, S., Ferrante, P., Venuti, A., Franconi, R., and Giuliano, G. (2013). A chlamydomonas-derived human papillomavirus 16 E7 vaccine induces specific tumor protection. *PLoS ONE* 8:e61473. doi: 10.1371/journal.pone.0061473

- De Muynck, B., Navarre, C., and Boutry, M. (2010). Production of antibodies in plants: status after twenty years. *Plant Biotechnol. J.* 8, 529–563. doi: 10.1111/j.1467-7652.2009.00494.x
- Díaz-Santos, E., Vega, M., Vila, M., Vigara, J., and León, R. (2013). *Efficiency of Different Heterologous Promoters in the Unicellular Microalga Chlamydomonas Reinhardtii*. Available online at: <http://onlinelibrary.wiley.com/doi/10.1002/btp.1690/full> [Accessed December 1, 2013].
- Drake, P. M. W., Barbi, T., Sexton, A., McGowan, E., Stadlmann, J., Navarre, C., et al. (2009). Development of rhizosecretion as a production system for recombinant proteins from hydroponic cultivated tobacco. *FASEB J.* 23, 3581–3589. doi: 10.1096/fj.09-131771
- Eichler-Stahlberg, A., Weisheit, W., Ruecker, O., and Heitzer, M. (2009). Strategies to facilitate transgene expression in *Chlamydomonas reinhardtii*. *Planta* 229, 873–883. doi: 10.1007/s00425-008-0879-x
- Feng, S., Xue, L., Liu, H., and Lu, P. (2009). Improvement of efficiency of genetic transformation for *Dunaliella salina* by glass beads method. *Mol. Biol. Rep.* 36, 1433–1439. doi: 10.1007/s11033-008-9333-1
- Fischer, N., and Rochaix, J.-D. (2001). The flanking regions of PsaD drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii*. *Mol. Genet. Genomics* 265, 888–894. doi: 10.1007/s004380100485
- Fitchette, A.-C., Dinh, O. T., Faye, L., and Bardot, M. (2007). “Plant Proteomics and Glycosylation,” in *Plant Proteomics Methods in Molecular Biology*, eds H. Thielmann, M. Zivy, C. Damerval, and V. Méchin (Humana Press), 317–342. Available online at: http://link.springer.com/protocol/10.1385/1-59745-227-0_9%3A317 [Accessed June 16, 2014].
- Fuhrmann, M., Oertel, W., and Hegemann, P. (1999). A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. *Plant J.* 19, 353–361. doi: 10.1046/j.1365-313X.1999.00526.x
- Gagneux, P., and Varki, A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9, 747–755. doi: 10.1093/glycob/9.8.747
- Geng, D., Wang, Y., Wang, P., Li, W., and Sun, Y. (2003). Stable expression of hepatitis B surface antigen gene in *Dunaliella salina* (Chlorophyta). *J. Appl. Phycol.* 15, 451–456. doi: 10.1023/B:JAPH.0000004298.89183.e5
- Gil, G.-C., Velander, W. H., and Van Cott, K. E. (2009). N-glycosylation microheterogeneity and site occupancy of an Asn-X-Cys sequon in plasma-derived and recombinant Protein C. *Proteomics* 9, 2555–2567. doi: 10.1002/pmic.200800775
- Gödel, S., Becker, B., and Melkonian, M. (2000). Flagellar membrane proteins of *Tetraselmis striata* butcher (Chlorophyta). *Protist* 151, 147–159. doi: 10.1078/1434-4610-00015
- Gomord, V., Fitchette, A.-C., Menu-Bouauiche, L., Saint-Jore-Dupas, C., Plisson, C., Michaud, D., et al. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.* 8, 564–587. doi: 10.1111/j.1467-7652.2009.00497.x
- Gonzalez, N. H., Felsner, G., Schramm, F. D., Klingl, A., Maier, U.-G., and Bolte, K. (2011). A single peroxisomal targeting signal mediates matrix protein import in diatoms. *PLoS ONE* 6:e25316. doi: 10.1371/journal.pone.0025316
- Grabowski, G. A., Golembio, M., and Shaaltil, Y. (2014). Taliglucerase alfa: an enzyme replacement therapy using plant cell expression technology. *Mol. Genet. Metab.* 112, 1–8. doi: 10.1016/j.ymgme.2014.02.011
- Grass, J., Pabst, M., Kolarich, D., Poltl, G., Leonard, R., Brecker, L., et al. (2011). Discovery and structural characterization of fucosylated oligomannosidic N-glycans in mushrooms. *J. Biol. Chem.* 286, 5977–5984. doi: 10.1074/jbc.M110.191304
- Grunow, A., Becker, B., and Melkonian, M. (1993). Isolation and characterization of the Golgi apparatus of a flagellate scaly green alga. *Eur. J. Cell Biol.* 61, 10–10.
- Harris, E. H. (2009). *The Chlamydomonas Sourcebook: Introduction to Chlamydomonas and Its Laboratory Use*. Oxford: Academic Press.
- Hawkins, R. L., and Nakamura, M. (1999). Expression of human growth hormone by the eukaryotic alga, *Chlorella*. *Curr. Microbiol.* 38, 335–341. doi: 10.1007/PL00006813
- He, X., Galpin, J. D., Tropak, M. B., Mahuran, D., Haselhorst, T., von Itzstein, M., et al. (2012). Production of active human glucocerebrosidase in seeds of *Arabidopsis thaliana* complex-glycan-deficient (cgl) plants. *Glycobiology* 22, 492–503. doi: 10.1093/glycob/cwr157
- Heitzer, M., Eckert, A., Fuhrmann, M., and Griesbeck, C. (2007). “Influence of codon bias on the expression of foreign genes in microalgae,” in *Transgenic Microalgae as Green Cell Factories* (Springer), 46–53. Available online at: http://link.springer.com/chapter/10.1007/978-0-387-75532-8_5 [Accessed December 1, 2013].
- Helenius, A., and Aeby, M. (2001). Intracellular functions of N-linked glycans. *Science* 291, 2364–2369. doi: 10.1126/science.291.5512.2364
- Hempel, F., Bozarth, A. S., Lindenkamp, N., Klingl, A., Zauner, S., Linne, U., et al. (2011a). Microalgae as bioreactors for bioplastic production. *Microb. Cell Fact.* 10:81. doi: 10.1186/1475-2859-10-81
- Hempel, F., Lau, J., Klingl, A., and Maier, U. G. (2011b). Algae as protein factories: expression of a human antibody and the respective antigen in the diatom *Phaeodactylum tricornutum*. *PLoS ONE* 6:e28424. doi: 10.1371/journal.pone.0028424
- Hempel, F., and Maier, U. G. (2012). An engineered diatom acting like a plasma cell secreting human IgG antibodies with high efficiency. *Microb. Cell Fact.* 11:126. doi: 10.1186/1475-2859-11-126
- Ho, S. C. L., Bardor, M., Li, B., Lee, J. J., Song, Z., Tong, Y. W., et al. (2013). Comparison of internal ribosome entry site (IRES) and Furin-2A (F2A) for monoclonal antibody expression level and quality in CHO cells. *PLoS ONE* 8:e63247. doi: 10.1371/journal.pone.0063247
- Hollister, J. R., Shaper, J. H., and Jarvis, D. L. (1998). Stable expression of mammalian beta 1,4-galactosyltransferase extends the N-glycosylation pathway in insect cells. *Glycobiology* 8, 473–480. doi: 10.1093/glycob/8.5.473
- Hossler, P., Khattak, S. F., and Li, Z. J. (2009). Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19, 936–949. doi: 10.1093/glycob/cwp079
- Huang, C.-J., Lin, H., and Yang, X. (2012). Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J. Ind. Microbiol. Biotechnol.* 39, 383–399. doi: 10.1007/s10295-011-1082-9
- Huetter, C. M., Lienhart, O., Baur, A., Stemmer, C., Gorr, G., Reski, R., et al. (2005). Glyco-Engineering of Moss Lacking Plant-Specific Sugar Residues. *Plant Biol.* 7, 292–299. doi: 10.1055/s-2005-837653
- Ioffe, E., and Stanley, P. (1994). Mice lacking N-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proc. Natl. Acad. Sci. U.S.A.* 91, 728–732. doi: 10.1073/pnas.91.2.728
- Jarvis, E. E., and Brown, L. M. (1991). Transient expression of firefly luciferase in protoplasts of the green alga *Chlorella ellipsoidea*. *Curr. Genet.* 19, 317–321. doi: 10.1007/BF00355062
- Kathiresan, S., Chandrashekhar, A., Ravishankar, G. A., and Sarada, R. (2009). Agrobacterium-mediated transformation in the green alga *Haematococcus pluvialis* (Chlorophyceae, Volvocales) 1. *J. Phycol.* 45, 642–649. doi: 10.1111/j.1529-8817.2009.00688.x
- Kim, D.-H., Kim, Y. T., Cho, J. J., Bae, J.-H., Hur, S.-B., Hwang, I., et al. (2002). Stable integration and functional expression of flounder growth hormone gene in transformed microalga, *Chlorella ellipsoidea*. *Mar. Biotechnol.* 4, 63–73. doi: 10.1007/s1012601-0070-x
- Kindle, K. L. (1998). “High-frequency nuclear transformation of *Chlamydomonas reinhardtii*,” in *Methods in Enzymology Photosynthesis: Molecular Biology of Energy Capture*, ed L. McIntosh (Academic Press), 27–38. Available online at: <http://www.sciencedirect.com/science/article/pii/S0076687998970057> [Accessed January 20, 2014].
- Kumar, A., Falcao, V. R., and Sayre, R. T. (2013). Evaluating nuclear transgene expression systems in *Chlamydomonas reinhardtii*. *Algal Res.* 2, 321–332. doi: 10.1016/j.algal.2013.09.002
- Lam, M. K., and Lee, K. T. (2012). Microalgae biofuels: a critical review of issues, problems and the way forward. *Biotechnol. Adv.* 30, 673–690. doi: 10.1016/j.biotechadv.2011.11.008
- Lauersen, K. J., Berger, H., Mussgnug, J. H., and Kruse, O. (2013a). Efficient recombinant protein production and secretion from nuclear transgenes in *Chlamydomonas reinhardtii*. *J. Biotechnol.* 167, 101–110. doi: 10.1016/j.jbiotec.2012.10.010
- Lauersen, K. J., Vanderveer, T. L., Berger, H., Kaluza, I., Mussgnug, J. H., Walker, V. K., et al. (2013b). Ice recrystallization inhibition mediated by a nuclear-expressed and -secreted recombinant ice-binding protein in the microalga *Chlamydomonas reinhardtii*. *Appl. Microbiol. Biotechnol.* 97, 9763–9772. doi: 10.1007/s00253-013-5226-x
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.-C., Gomord, V., and Faye, L. (1998). “N-glycoprotein biosynthesis in plants: recent developments and future trends,” in *Protein Trafficking in*

- Plant Cells* (Springer), 31–48. Available online at: http://link.springer.com/chapter/10.1007/978-94-011-5298-3_2 [Accessed December 1, 2013].
- Levy-Ontman, O., Arad, S. M., Harvey, D. J., Parsons, T. B., Fairbanks, A., and Tekoah, Y. (2011). Unique N-glycan moieties of the 66-kDa cell wall glycoprotein from the red microalga *Porphyridium* sp. *J. Biol. Chem.* 286, 21340–21352. doi: 10.1074/jbc.M110.175042
- Levy-Ontman, O., Fisher, M., Shotland, Y., Weinstein, Y., Tekoah, Y., and Arad, S. M. (2014). Genes involved in the endoplasmic reticulum N-glycosylation pathway of the red microalga *Porphyridium* sp.: a bioinformatic study. *Int. J. Mol. Sci.* 15, 2305–2326. doi: 10.3390/ijms15022305
- Lingg, N., Zhang, P., Song, Z., and Bardor, M. (2012). The sweet tooth of biopharmaceuticals: importance of recombinant protein glycosylation analysis. *Biotechnol. J.* 7, 1462–1472. doi: 10.1002/biot.201200078
- Lumbreras, V., Stevens, D. R., and Purton, S. (1998). Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron. *Plant J.* 14, 441–447. doi: 10.1046/j.1365-313X.1998.00145.x
- Mata, T. M., Martins, A. A., and Caetano, N. S. (2010). Microalgae for biodiesel production and other applications: a review. *Renew. Sust. Energ. Rev.* 14, 217–232. doi: 10.1016/j.rser.2009.07.020
- Mathieu-Rivet, E., Scholz, M., Arias, C., Dardelle, F., Schulze, S., Le Mauff, F., et al. (2013). Exploring the N-glycosylation pathway in *chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell. Proteomics* 12, 3160–3183. doi: 10.1074/mcp.M113.028191
- Matsui, T., Takita, E., Sato, T., Kinjo, S., Aizawa, M., Sugiura, Y., et al. (2011). N-glycosylation at noncanonical Asn-X-Cys sequences in plant cells. *Glycobiology* 21, 994–999. doi: 10.1093/glycob/cwq198
- Maxmen, A. (2012). Drug-making plant blooms. *Nature* 485, 160–160. doi: 10.1038/485160a
- Mayfield, S. P., Franklin, S. E., and Lerner, R. A. (2003). Expression and assembly of a fully active antibody in algae. *Proc. Natl. Acad. Sci.* 100, 438–442. doi: 10.1073/pnas.0237108100
- Merchant, S. S., Kropat, J., Liu, B., Shaw, J., and Warakanont, J. (2012). TAG, You're it! *Chlamydomonas* as a reference organism for understanding algal triacylglycerol accumulation. *Curr. Opin. Biotechnol.* 23, 352–363. doi: 10.1016/j.copbio.2011.12.001
- Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., et al. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318, 245–250. doi: 10.1126/science.1143609
- Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., and Marth, J. D. (1994). Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *EMBO J.* 13, 2056–2065.
- Misaki, R., Fujiyama, K., and Seki, T. (2006). Expression of human CMP-N-acetylneuraminate synthetase and CMP-sialic acid transporter in tobacco suspension-cultured cell. *Biochem. Biophys. Res. Commun.* 339, 1184–1189. doi: 10.1016/j.bbrc.2005.11.130
- Misumi, O., Matsuzaki, M., Nozaki, H., Miyagishima, S., Mori, T., Nishida, K., et al. (2005). Cyanidioschyzon merolae genome: A tool for facilitating comparable studies on organelle biogenesis in photosynthetic eukaryotes. *Plant Physiol.* 137, 567–585. doi: 10.1104/pp.104.053991
- Neupert, J., Karcher, D., and Bock, R. (2009). Generation of *Chlamydomonas* strains that efficiently express nuclear transgenes. *Plant J.* 57, 1140–1150. doi: 10.1111/j.1365-313X.2008.03746.x
- Pacalet, T., Bardor, M., Rihouey, C., Delmas, F., Chevalier, C., D'Aoust, M.-A., et al. (2007). Engineering of a sialic acid synthesis pathway in transgenic plants by expression of bacterial Neu5Ac-synthesizing enzymes. *Plant Biotechnol. J.* 5, 16–25. doi: 10.1111/j.1467-7652.2006.00211.x
- Padler-Karavani, V., Yu, H., Cao, H., Chokhawala, H., Karp, F., Varki, N., et al. (2008). Diversity in specificity, abundance, and composition of anti-Neu5Gc antibodies in normal humans: potential implications for disease. *Glycobiology* 18, 818–830. doi: 10.1093/glycob/cwn072
- Palacpac, N. Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., et al. (1999). Stable expression of human β 1, 4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4692–4697. doi: 10.1073/pnas.96.8.4692
- Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., et al. (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7705–7710. doi: 10.1073/pnas.0611046104
- Park, S., Lee, Y., Lee, J.-H., and Jin, E. (2013). Expression of the high light-inducible *Dunaliella* LIP promoter in *Chlamydomonas reinhardtii*. *Planta* 238, 1147–1156. doi: 10.1007/s00425-013-1955-4
- Parsons, J., Altmann, F., Arrenberg, C. K., Kopribova, A., Beike, A. K., Stemmer, C., et al. (2012). Moss-based production of asialo-erythropoietin devoid of Lewis A and other plant-typical carbohydrate determinants. *Plant Biotechnol. J.* 10, 851–861. doi: 10.1111/j.1467-7652.2012.00704.x
- Potvin, G., and Zhang, Z. (2010). Strategies for high-level recombinant protein expression in transgenic microalgae: a review. *Biotechnol. Adv.* 28, 910–918. doi: 10.1016/j.biotechadv.2010.08.006
- Poulsen, N., and Kröger, N. (2005). A new molecular tool for transgenic diatoms. *Febs J.* 272, 3413–3423. doi: 10.1111/j.1742-4658.2005.04760.x
- Radakovits, R., Jinkerson, R. E., Fuerstenberg, S. I., Tae, H., Settlage, R. E., Boore, J. L., et al. (2012). Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat. Commun.* 3:686. doi: 10.1038/ncomms1688
- Rasala, B. A., Lee, P. A., Shen, Z., Briggs, S. P., Mendez, M., and Mayfield, S. P. (2012). Robust expression and secretion of Xylanase1 in *chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. *PLoS ONE* 7:e43349. doi: 10.1371/journal.pone.0043349
- Rasala, B. A., and Mayfield, S. P. (2011). The microalga *Chlamydomonas reinhardtii* as a platform for the production of human protein therapeutics. *Bioengineered* 2, 50–54. doi: 10.4161/bbug.2.1.13423
- Rasala, B. A., and Mayfield, S. P. (2014). Photosynthetic biomanufacturing in green algae: production of recombinant proteins for industrial, nutritional, and medical uses. *Photosyn. Res.* doi: 10.1007/s11120-014-9994-7. [Epub ahead of print].
- Rasala, B. A., Muto, M., Lee, P. A., Jager, M., Cardoso, R. M. F., Behnke, C. A., et al. (2010). Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*: production of therapeutic proteins in algae. *Plant Biotechnol. J.* 8, 719–733. doi: 10.1111/j.1467-7652.2010.00503.x
- Rouwendal, G. J., Wuhrer, M., Florack, D. E., Koeleman, C. A., Deelder, A. M., Bakker, H., et al. (2007). Efficient introduction of a bisecting GlcNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase III. *Glycobiology* 17, 334–344. doi: 10.1093/glycob/cwl078
- Ruecker, O., Zillner, K., Groebner-Ferreira, R., and Heitzer, M. (2008). Gaussia-luciferase as a sensitive reporter gene for monitoring promoter activity in the nucleus of the green alga *Chlamydomonas reinhardtii*. *Mol. Genet. Genomics* 280, 153–162. doi: 10.1007/s00438-008-0352-3
- Sasso, S., Pohnert, G., Lohr, M., Mittag, M., and Hertweck, C. (2012). Microalgae in the postgenomic era: a blooming reservoir for new natural products. *FEMS Microbiol. Rev.* 36, 761–785. x doi: 10.1111/j.1574-6976.2011.00304.x
- Schähs, M., Strasser, R., Stadlmann, J., Kunert, R., Rademacher, T., and Steinkellner, H. (2007). Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol. J.* 5, 657–663. doi: 10.1111/j.1467-7652.2007.00273.x
- Schillberg, S., Raven, N., Fischer, R., Twyman, R., and Schiermeyer, A. (2013). Molecular farming of pharmaceutical proteins using plant suspension cell and tissue cultures. *Curr. Pharm. Des.* 19, 5531–5542. doi: 10.2174/1381612811319310008
- Schroda, M., Blöcker, D., and Beck, C. F. (2000). The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. *Plant J.* 21, 121–131. doi: 10.1046/j.1365-313x.2000.00652.x
- Séveno, M., Bardor, M., Paccalet, T., Gomord, V., Lerouge, P., and Faye, L. (2004). Glycoprotein sialylation in plants? *Nat. Biotechnol.* 22, 1351–1352. doi: 10.1038/nbt1104-1351 doi: 10.1038/nbt1104-1351
- Shaaltiel, Y., Bartfeld, D., Hashmueli, S., Baum, G., Brill-Almon, E., Galili, G., et al. (2007). Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant Biotechnol. J.* 5, 579–590. doi: 10.1111/j.1467-7652.2007.00263.x
- Shaver, S., Casas-Mollano, J. A., Cerny, R. L., and Cerutti, H. (2010). Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga *Chlamydomonas*. *Epigenetics* 5, 301–312. doi: 10.4161/epi.5.4.11608
- Specht, E., Miyake-Stoner, S., and Mayfield, S. (2010). Micro-algae come of age as a platform for recombinant protein production. *Biotechnol. Lett.* 32, 1373–1383. doi: 10.1007/s10529-010-0326-5

- Spolaore, P., Joannis-Cassan, C., Duran, E., and Isambert, A. (2006). Commercial applications of microalgae. *J. Biosci. Bioeng.* 101, 87–96. doi: 10.1263/jbb.101.87
- Staudacher, E. (2012). Methylation—an uncommon modification of glycans. *Biol. Chem.* 393, 675–685. doi: 10.1515/hsz-2012-0132
- Stoger, E., Ma, J. K.-C., Fischer, R., and Christou, P. (2005). Sowing the seeds of success: pharmaceutical proteins from plants. *Curr. Opin. Biotechnol.* 16, 167–173. doi: 10.1016/j.copbio.2005.01.005
- Stork, S., Moog, D., Przyborski, J. M., Wilhelmi, I., Zauner, S., and Maier, U. G. (2012). Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation. *Eukaryotic Cell* 11, 1472–1481. doi: 10.1128/EC.00183-12
- Strasser, R., Mucha, J., Mach, L., Altmann, F., Wilson, I. B., Glössl, J., et al. (2000). Molecular cloning and functional expression of beta1, 2-xylosyltransferase cDNA from *Arabidopsis thaliana*. *FEBS Lett.* 472, 105–108. doi: 10.1016/S0014-5793(00)01443-5
- Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., et al. (2008). Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* 6, 392–402. doi: 10.1111/j.1467-7652.2008.00330.x
- Surzycki, R., Greenham, K., Kitayama, K., Dibal, F., Wagner, R., Rochaix, J.-D., et al. (2009). Factors affecting expression of vaccines in microalgae. *Biologicals* 37, 133–138. doi: 10.1016/j.biologicals.2009.02.005
- Tran, M., Van, C., Barrera, D. J., Pettersson, P. L., Peinado, C. D., Bui, J., et al. (2012). Production of unique immunotoxin cancer therapeutics in algal chloroplasts. *Proc. Natl. Acad. Sci. U.S.A.* 110, E15–E22. doi: 10.1073/pnas.1214638110
- Tran, M., Zhou, B., Pettersson, P. L., Gonzalez, M. J., and Mayfield, S. P. (2009). Synthesis and assembly of a full-length human monoclonal antibody in algal chloroplasts. *Biotechnol. Bioeng.* 104, 663–673. doi: 10.1002/bit.22446
- Twyman, R., Schillberg, S., and Fischer, R. (2013). Optimizing the yield of recombinant pharmaceutical proteins in plants. *Curr. Pharm. Des.* 19, 5486–5494. doi: 10.2174/1381612811319310004
- Van Beers, M. M. C., and Bardor, M. (2012). Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. *Biotechnol. J.* 7, 1473–1484. doi: 10.1002/biot.201200065
- Van Patten, S. M., Hughes, H., Huff, M. R., Piepenhagen, P. A., Waire, J., Qiu, H., et al. (2007). Effect of mannose chain length on targeting of glucocerebrosidase for enzyme replacement therapy of Gaucher disease. *Glycobiology* 17, 467–478. doi: 10.1093/glycob/cwm008
- Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3, 97–130. doi: 10.1093/glycob/3.2.97
- Varki, A. (2011). Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells. *Cold Spring Harb. Perspect. Biol.* 3:a005462. doi: 10.1101/cshperspect.a005462
- Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Marth, J. D., et al. (2009). Symbol nomenclature for glycan representation. *Proteomics* 9, 5398–5399. doi: 10.1002/pmic.200900708
- Vézina, L.-P., Faye, L., Lerouge, P., D'Aoust, M.-A., Marquet-Blouin, E., Burel, C., et al. (2009). Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnol. J.* 7, 442–455. doi: 10.1111/j.1467-7652.2009.00414.x
- Vieler, A., Wu, G., Tsai, C.-H., Bullard, B., Cornish, A. J., Harvey, C., et al. (2012). Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *nannochloropsis oceanica* CCMP1779. *PLoS Genet.* 8:e1003064. doi: 10.1371/journal.pgen.1003064
- Vitale, A., and Chrispeels, M. J. (1984). Transient N-acetylglucosamine in the biosynthesis of phytohemagglutinin: attachment in the Golgi apparatus and removal in protein bodies. *J. Cell Biol.* 99, 133–140. doi: 10.1083/jcb.99.1.133
- Walsh, G. (2010). Biopharmaceutical benchmarks 2010. *Nat. Biotechnol.* 28, 917–924. doi: 10.1038/nbt0910-917
- Weerapana, E., and Imperiali, B. (2006). Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* 16, 91R–101R. doi: 10.1093/glycob/cwj099
- Weinreb, N. J. (2008). *Imiglucerase and its Use for the Treatment of Gaucher's Disease*. Available online at: <http://informahealthcare.com/doi/abs/10.1517/14656566.9.11.1987> [Accessed December 1, 2013].
- Wilson, I. B. H., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J. M., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core α 1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Wong, C.-H. (2005). Protein glycosylation: new challenges and opportunities. *J. Org. Chem.* 70, 4219–4225. doi: 10.1021/jo050278f
- Worden, A. Z., Lee, J.-H., Mock, T., Rouze, P., Simmons, M. P., Aerts, A. L., et al. (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoplankton *micromonas*. *Science* 324, 268–272. doi: 10.1126/science.1167222
- Xu, J., Dolan, M. C., Medrano, G., Cramer, C. L., and Weathers, P. J. (2012). Green factory: plants as bioproduction platforms for recombinant proteins. *Biotechnol. Adv.* 30, 1171–1184. doi: 10.1016/j.biotechadv.2011.08.020
- Zaslavskaja, L. A., Lippmeier, J. C., Kroth, P. G., Grossman, A. R., and Apt, K. E. (2000). Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J. Phycol.* 36, 379–386. doi: 10.1046/j.1529-8817.2000.99164.x
- Zeleny, R., Kolarich, D., Strasser, R., and Altmann, F. (2006). Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta* 224, 222–227. doi: 10.1007/s00425-005-0206-8
- Zhao, T., Wang, W., Bai, X., and Qi, Y. (2009). Gene silencing by artificial microRNAs in *Chlamydomonas*. *Plant J.* 58, 157–164. doi: 10.1111/j.1365-313X.2008.03758.x
- Zhu, S., Hanneman, A., Reinhold, V. N., Spence, A. M., and Schachter, H. (2004). *Caenorhabditis elegans* triple null mutant lacking UDP-N-acetyl-D-glucosamine: α -3-D-mannoside β 1,2-N-acetylglucosaminyltransferase I. Biochem. J. 382, 995–1001. doi: 10.1042/BJ20040793
- Zielinska, D. F., Gnäd, F., Wiśniewski, J. R., and Mann, M. (2010). Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141, 897–907. doi: 10.1016/j.cell.2010.04.012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 May 2014; **accepted:** 04 July 2014; **published online:** 28 July 2014.

Citation: Mathieu-Rivet E, Kiefer-Meyer M-C, Vanier G, Ovide C, Burel C, Lerouge P and Bardor M (2014) Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Front. Plant Sci.* 5:359. doi: 10.3389/fpls.2014.00359

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Mathieu-Rivet, Kiefer-Meyer, Vanier, Ovide, Burel, Lerouge and Bardor. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Glyco-engineering for biopharmaceutical production in moss bioreactors

Eva L. Decker^{1*}, Juliana Parsons¹ and Ralf Reski^{1,2,3}

¹ Department of Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany

² BIOSS Centre for Biological Signalling Studies, Freiburg, Germany

³ Freiburg Institute for Advanced Studies, Freiburg, Germany

Edited by:

Nausicaä Lannoo, Ghent University, Belgium

Reviewed by:

Markus Pauly, University of California, Berkeley, USA

Richard Strasser, University of Natural Resources and Life Sciences, Austria

***Correspondence:**

Eva L. Decker, Department of Plant Biotechnology, Faculty of Biology, University of Freiburg,

Schaenzlestraße 1, 79104 Freiburg, Germany

e-mail: eva.decker@biologie.uni-freiburg.de

The production of recombinant biopharmaceuticals (pharmaceutical proteins) is a strongly growing area in the pharmaceutical industry. While most products to date are produced in mammalian cell cultures, namely Chinese hamster ovary cells, plant-based production systems gained increasing acceptance over the last years. Different plant systems have been established which are suitable for standardization and precise control of cultivation conditions, thus meeting the criteria for pharmaceutical production. The majority of biopharmaceuticals comprise glycoproteins. Therefore, differences in protein glycosylation between humans and plants have to be taken into account and plant-specific glycosylation has to be eliminated to avoid adverse effects on quality, safety, and efficacy of the products. The basal land plant *Physcomitrella patens* (moss) has been employed for the recombinant production of high-value therapeutic target proteins (e.g., Vascular Endothelial Growth Factor, Complement Factor H, monoclonal antibodies, Erythropoietin). Being genetically excellently characterized and exceptionally amenable for precise gene targeting via homologous recombination, essential steps for the optimization of moss as a bioreactor for the production of recombinant proteins have been undertaken. Here, we discuss the glyco-engineering approaches to avoid non-human N- and O-glycosylation on target proteins produced in moss bioreactors.

Keywords: *Physcomitrella patens*, moss bioreactor, plant-made pharmaceuticals, glycosylation, posttranslational modifications

INTRODUCTION

Biopharmaceuticals are indispensable in modern medicine. In 2010 more than 200 biopharmaceuticals were available on the market and around 10–20 more are approved every year (Walsh, 2010a). As the biggest group of biopharmaceuticals consists of pharmaceutical recombinant proteins, this term is often used as a synonym for the former. The biochemical and pharmacological properties of a protein are not only determined by its amino acid sequence but also largely influenced by a palette of modifications that proteins undergo co- or posttranslationally (Mann and Jensen, 2003), usually grouped together and referred to as posttranslational modifications (PTMs). Common PTMs found in pharmaceutical proteins are glycosylation, hydroxylation, carboxylation, amidation, sulfatation, disulfide bond formation, and proteolytic processing (Walsh and Jefferis, 2006). Among these, glycosylation is the most frequent PTM, being present in at least 40% of the pharmaceutical recombinant proteins available on the market (Walsh, 2010b). The presence and quality of glycosylation plays a crucial role for the pharmacological properties of a therapeutic protein by influencing protein folding and stability, serum half-life, *in vivo* activity, pharmacokinetics, and immunogenicity (Li and d'Anjou, 2009). Approximately 50% of all eukaryotic proteins are predicted to be glycosylated and this proportion increases substantially with respect to human serum proteins, which are main targets as biopharmaceuticals (Apweiler et al., 1999).

The workhorse for the production of simple proteins is *Escherichia coli*, the best characterized expression system offering high product yields at low costs (Walsh, 2010a). However, this microorganism is not able to perform some PTMs, which are indispensable for recombinant therapeutical proteins (Kamionka, 2011). Consequently, mammalian cell lines are the preferred expression systems for the production of recombinant glycoproteins, as their protein glycosylation patterns largely resemble those of humans (Schmidt, 2004). Among the mammalian cell lines, Chinese hamster ovary (CHO) cells comprise the leading host system for current biopharmaceuticals, even though several CHO-derived products presented non-human glycosylation (Chung et al., 2008; Hossler et al., 2009; Omasa et al., 2010; Kim et al., 2012).

As higher eukaryotes, plants are able to synthesize complex multimeric proteins and perform PTMs in a similar manner as humans do. Therefore, plants and plant cell cultures are gradually gaining acceptance as production hosts for recombinant biopharmaceuticals. The first plant-made pharmaceutical (PMP) received market approval in 2012 (<http://www.protalix.com/products/eleyso-taliglucerase-alfa.asp>) and several additional PMPs are being tested in clinical trials (reviewed in Paul and Ma, 2011). The host system for Eleyso™, a recombinant glucocerebrosidase for the treatment of the lysosomal storage disease Morbus Gaucher, is a carrot-based cell line established by Protalix (Shaaltiel et al., 2007). It is cultured in

bioreactors based on disposable plastic bags. While other frequently used plant systems like alfalfa, tobacco, and *Nicotiana benthamiana* need to be grown in greenhouses, bioreactor cultivation is established for the aquatic plant *Lemna minor* and for the moss *Physcomitrella patens* (Decker and Reski, 2007; Paul and Ma, 2011; Paul et al., 2013). Within the following sections we will focus on the special features for biopharmaceutical production and achievements within glyco-engineering of the moss system.

MOSS CULTIVATION AND ENGINEERING CHARACTERISTICS

The non-seed plant *P. patens*, a moss, is a well-established model system for evolutionary and functional genomics approaches (Cove et al., 2006; Menand et al., 2007; Mosquna et al., 2009; Khraiwesh et al., 2010; Sakakibara et al., 2013). It can be grown throughout its complete life cycle under contained conditions *in vitro* in a simple mineral medium (Frank et al., 2005; Strotbek et al., 2013). The germination of the haploid spores leads to the growth of protonema (**Figure 1A**), a branched filamentous tissue which comprises two distinct cell types, chloronema and caulinema. Every cell is in direct contact with the culture medium, allowing efficient nutrient uptake and product secretion (Schillberg et al., 2013). This young tissue can be maintained in suspension cultures without any addition of phytohormones, only by mechanical disruption of the filaments. In contrast to immortalized or de-differentiated mammalian or higher-plant cell cultures, which are prone to instability or somaclonal variation (Larkin and Scowcroft, 1981; Xu et al., 2011; Bailey et al., 2012), the fully differentiated protonema tissue is genetically stable (Reutter and Reski, 1996). In the next developmental step, buds differentiating from protonema cells give rise to the adult plant, the leafy gametophore, consisting of shoot-like, leaf-like, and root-like tissues (**Figure 1B**). After fertilization of the gametes, the sporophyte, the only diploid tissue in the life cycle of mosses, grows on and is sustained by the gametophore (Reski et al., 1998). *In vitro* cultivation of all stages can be performed either on agar plates or as suspension cultures in liquid media. The availability of efficient protocols for protoplast isolation (**Figure 1C**) and transfection (Rother et al., 1994; Strotbek et al., 2013) and an excellent regeneration capacity of single transfected cells to whole plants make genetic engineering of moss a straight-forward and frequently used approach (e.g., Lorenz et al., 2003; Qudeimat et al., 2008; Ludwig-Müller et al., 2009; Mosquna et al., 2009; Khraiwesh et al., 2010; Sakakibara et al., 2013). The created moss strains can be preserved by cryo-conservation (Schulte and Reski, 2004), and thus can serve as Master Cell Banks. The International Moss Stock Center IMSC, a reference center for moss ecotypes and transgenic lines, provides a service for long-term storage (<http://www.moss-stock-center.org>).

The employment of *in vitro* axenic plant cell or tissue cultures offers an environment in which contamination with human pathogens is rather unlikely (Schillberg et al., 2013). Moreover, only in these systems culture conditions can be precisely controlled and standardized (Hohe and Reski, 2005), which is essential for the production of pharmaceuticals according to good manufacturing practice (GMP) guidelines (Fischer et al., 2012).

Various scales of highly controllable cultivation devices were developed for *Physcomitrella*, ranging from simple shaking

flasks (**Figure 1D**) and 5 L aerated flasks to diverse forms of photobioreactors, including stirred glass tank bioreactors with a volume of up to 15 L (**Figure 1E**; Hohe and Reski, 2002) and a modular tubular bioreactor with a working volume of up to 100 L (reviewed in Decker and Reski, 2008, 2012). More recently, disposable wave-bag reactors (**Figure 1F**) were employed for high-density protein production purposes under full cGMP compliance (www.greenovation.com). Several pharmaceutically interesting proteins have been synthesized in moss bioreactors, among them the growth factors vascular endothelial growth factor (VEGF; Baur et al., 2005) and erythropoietin (EPO; Weise et al., 2007) as well as the first marketed product for research use, human FGF7/keratinocyte growth factor (KGF; www.greenovation.com). In addition, proteins with a function in immune responses like IgGs (Schuster et al., 2007; Kircheis et al., 2012) and the complement-regulatory protein factor H (Büttner-Mainik et al., 2011) were produced in moss. Furthermore, two products for enzyme-replacement

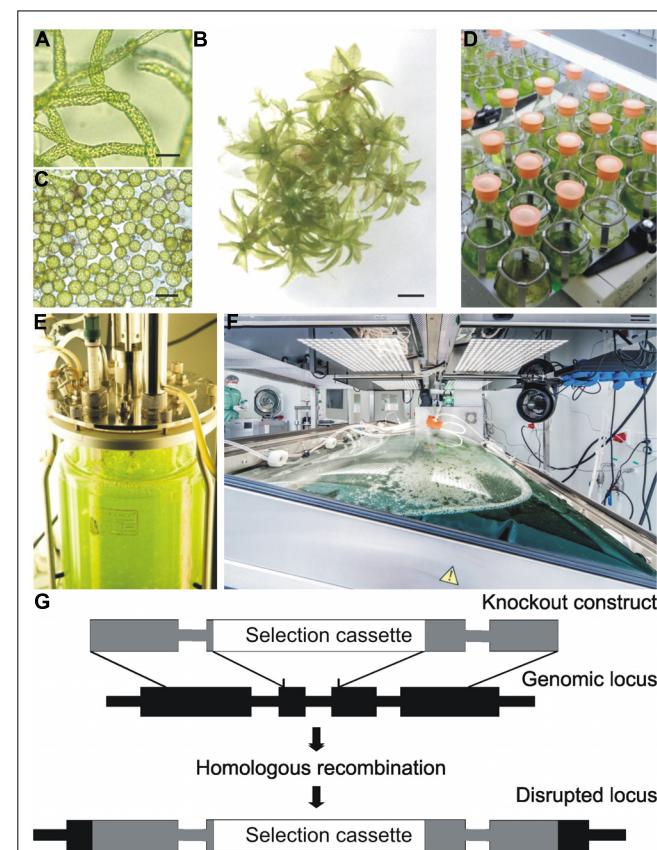


FIGURE 1 | *Physcomitrella patens* *in vitro* cultivation and schematic representation of a knockout construct for gene targeting. **(A)** young filamentous tissue, protonema, ideal for suspension cultures; **(B)** adult leafy moss plant (gametophore); **(C)** protoplasts; **(D)** small scale liquid culture in flasks; **(E)** stirred tank bioreactor; **(F)** wave bioreactor (image courtesy of greenovation Biotech GmbH); **(G)** illustration of allele replacement via homologous recombination. The regions homologous to the targeted gene, which are used for the knockout construct, are shown in gray, and the inserted selection cassette is depicted in white. Thick lines represent introns and rectangles exons. Scale bars 50 μm (**A,B**), 500 μm (**C**).

therapies, human alpha-galactosidase and glucocerebrosidase are expected to reach clinical trial phases by the end of 2014 (www.greenovation.com).

The use of *Physcomitrella* as a production host for recombinant biopharmaceuticals was facilitated by well-developed molecular toolboxes. Heterologous as well as endogenous promoters were characterized for their suitability to achieve high levels of recombinant product (Horstmann et al., 2004; Weise et al., 2006). In addition, several moss-derived signal peptides were evaluated for improved secretion of the recombinant product to the surrounding medium (Schaaf et al., 2004, 2005; Weise et al., 2006). The moss genome sequence is available since 2008 (Rensing et al., 2008), and together with nearly 400,000 expressed sequence tags (ESTs) obtained from different experimental conditions, tissues, and developmental stages (Nishiyama et al., 2003; Lang et al., 2008) it allows a reliable prediction of gene structures. The internet resource www.cosmoss.org provides access to a high-quality functional annotation including more than 32,000 protein-coding genes (Zimmer et al., 2013). This resource was very convenient for the identification of genes involved in the glycosylation of recombinant proteins synthesized in moss bioreactors.

However, the main driver for moss functional genomics approaches in general and glyco-engineering in particular was the unique accessibility of this organism for gene targeting approaches via homologous recombination. Displaying an exceptionally high rate of homologous recombination in mitotic cells (Strepp et al., 1998; Schaefer, 2001; Hohe et al., 2004; Kamisugi et al., 2006), base-specific precise genetic engineering is feasible with high efficiency. Undesirable gene functions can be completely eliminated by targeted knockout approaches. The knockout construct used for the transfection of moss protoplasts regularly consists of 700–1000 bp genomic DNA (homologous regions) flanking each side of a selection cassette, which interrupts or replaces the target gene when indicated (**Figure 1G**). Glyco-engineering of moss was successfully accomplished by various gene targeting approaches (see below).

PROTEIN GLYCOSYLATION AND MOSS GLYCO-ENGINEERING

Protein glycosylation is a complex and heterogeneous modification which can be classified in two main categories, N- and O-glycosylation. In the former, the carbohydrates are attached to the amide group of asparagine (N) in the consensus sequence N-X-S/T (where X can be any amino acid except proline, and the third amino acid can be either serine or threonine; Mononen and Karjalainen, 1984; Gavel and von Heijne, 1990). O-glycans, on the other hand, are attached to the hydroxyl group of serine (S), threonine (T), hydroxylysine or hydroxyproline (Hyp; Varki et al., 2009). In contrast to N-glycosylation, consensus sequences for O-glycosylation in mammals are not well defined or non-existing (Hansen et al., 1998; Julenius et al., 2005).

N-glycosylation in animals is a largely cell-type and species-specific feature (Raju et al., 2000; Croset et al., 2012). Moreover, potential glycosylation sites on a given protein can be either unmodified or occupied by varying glycan structures which result from the maturation of the glycan throughout the endoplasmic

reticulum (ER) and the Golgi apparatus (GA), leading to microheterogeneity of glycoproteins (Kolarich et al., 2012). Compared to other higher eukaryotes, plants display more conserved glycan patterns between different species and a less diverse palette of N-glycans (Bosch et al., 2013), facilitating the production of homogeneous glycoproteins.

As higher eukaryotes, plants are able to produce N-glycans of the complex type with the core sugar structure Man3GlcNAc2 consisting of two N-acetylglucosamine and three mannose residues that is identical to humans (Lerouge et al., 1998; Wilson, 2002). Up to two terminally attached GlcNAc residues are also common between plant and bi-antennary mammalian complex-type glycoprotein oligosaccharides (reviewed in Gomord et al., 2010). Differing from the human structure, which displays a fucose residue 1,6-linked to the proximal GlcNAc moiety, most plant N-glycans carry a β 1,2 xylose and an α 1,3 fucose linked to the glycan core. These sugar structures are common for all land plants analyzed so far, including mosses as the evolutionary oldest group (Koprivova et al., 2003; Viëtor et al., 2003). Their presence raised concerns about plant-produced biopharmaceuticals as they were shown to induce antibody formation in mammals (van Ree et al., 2000; Bardor et al., 2003; Westphal et al., 2003; Bencúrová et al., 2004; Jin et al., 2008). The consequence of an immune response against the pharmaceutical can lead to antibody-mediated reduction of product efficacy as well as to severe clinical complications (Schellekens, 2002).

Consequently, first plant glyco-engineering approaches aimed at targeting the glycosyltransferases responsible for the addition of these two residues. Ten years ago *Arabidopsis thaliana* as well as moss lines lacking β 1,2 xylosylation and α 1,3 fucosylation have been generated (Koprivova et al., 2004; Strasser et al., 2004). The predominant glycan type of the double knockout moss line for β 1,2 xylosyltransferase (XylT) and α 1,3 fucosyltransferase (FucT) was the GnGn form (GlcNAc2Man3GlcNAc2; Koprivova et al., 2004). A Δ XylT/FucT genotype is currently in use as genetic background for most of the recombinant products described from moss bioreactors.

Lacking the core fucose, the engineered moss N-glycans differ from the human ones which contain an α 1,6-linked fucose residue. However, lack of this residue proved to be advantageous for the efficacy of antibodies targeting tumor cells (Shields et al., 2002; Shinkawa et al., 2003; Cox et al., 2006; Schuster et al., 2007). The underlying phenomenon, antibody-dependent cellular cytotoxicity (ADCC), comprises receptor binding and activation of a natural killer cell by an antigen–antibody complex on the target cell surface resulting in lysis of the target cell. Binding and activation of the killer cells was up to 40 \times more efficient with a monoclonal antibody produced in glyco-engineered, fucose-lacking moss cells compared to the same antibody produced in CHO cells (Schuster et al., 2007).

In addition to the GnGn N-glycan form, many plant species display α 1,4 fucose, and β 1,3 galactose linked to the terminal GlcNAc residues on one or both of the antennae (Wilson, 2001). This trisaccharide Fuc α 1-4(Gal β 1-3)GlcNAc is known as Lewis A (Le^a) structure. It is synthesized by β 1,3 galactosyltransferases (GalT) and α 1,4 fucosyltransferases as the last steps of N-glycan

maturity in the plant GA (reviewed in Gomord et al., 2010). In contrast to the high prevalence of xylose and core fucose residues on plant *N*-glycans, Le^a structures are found in a much lower proportion (Fitchette-Lainé et al., 1997; Koprivova et al., 2003; Viétor et al., 2003; Strasser et al., 2007). However, Le^a epitopes were described on recombinant proteins produced in plants (Petrucelli et al., 2006; Weise et al., 2007; Castilho et al., 2013). The production of recombinant human EPO (rhEPO) in both moss and *N. benthamiana*, lead to proteins decorated with high amounts of Le^a structures (Weise et al., 2007; Castilho et al., 2013). Although Le^a epitopes are found in humans as part of the Lewis-positive histo-blood groups (Henry et al., 1995), they are rarely present in healthy adults, but increased in patients with certain types of cancer (Zhang et al., 1994). Furthermore, antibodies against Le^a epitopes are frequent (Wilson et al., 2001). Therefore, it is advisable to remove the respective β 1,3 galactose and α 1,4 fucose residues from plant-produced recombinant products.

A single putative α 1,4 fucosyltransferase gene was detected in the moss genome. While the targeted knockout of this gene resulted in the loss of terminal α 1,4 fucose residues, β 1,3-linked galactoses were still present on moss *N*-glycans. In contrast to the single-copy α 1,4 fucosyltransferase gene, 13 putative *galt* homologs were identified in *P. patens*. Out of these, exclusively one gene (*galt1*) was shown to be responsible for the synthesis of Le^a in moss. The disruption of *galt1* alone resulted in the absence of the complete Le^a epitope, not only of the galactose residue but also of the terminal fucose, both in the total moss *N*-glycan pool and on the moss-produced rhEPO (Parsons et al., 2012). The absence of the α 1,4 fucose in the *galt1* knockout line ($\Delta galt1$) with intact α 1,4 fucosyltransferase activity confirmed that this is the last enzyme in the plant N-glycosylation pathway and that the presence of galactose on the substrate is indispensable for the fucosyltransferase activity (Parsons et al., 2012). The lack of GalT1 activity did not affect the moss growth rate. The homogeneity of the rhEPO glycosylation achieved in the moss *galt1* knockout line was remarkable, with almost only one glycosylation form, the aimed core structure with terminal GlcNAc residues (Parsons et al., 2012; Figure 2).

In humans, the GnGn glycan is frequently further elongated with galactose added in β 1,4 linkage and this is often capped with sialic acid residues. The targeted insertion ("knockin") of the human β 1,4 *galt* into the moss genome demonstrated the general feasibility of β 1,4 galactosylation of moss *N*-glycans (Huether et al., 2005; Parsons et al., 2012). Further terminal elongation of plant *N*-glycans has been demonstrated for *N. benthamiana* which transiently produced glycoproteins with human-like sialylation (Castilho et al., 2013; Jez et al., 2013). This will be a future task for moss glyco-engineering.

Outstanding success has been achieved so far by engineering plant N-glycosylation patterns for the production of humanized glycoproteins. In contrast, the issue of adverse O-glycosylation in PMPs has not been addressed in the same detail. Concerning plant O-glycan engineering, recombinant proteins displaying human so-called mucin-type O-glycosylation were generated recently (Daskalova et al., 2010; Castilho et al., 2012; Yang et al., 2012). In contrast to rather conserved N-glycosylation patterns,

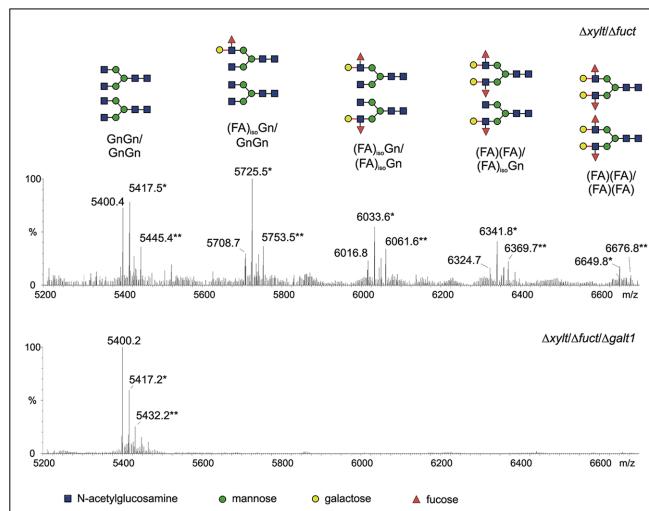


FIGURE 2 | Glycopeptides of moss-produced rhEPO from two

glyco-engineered moss lines. Comparison of the mass spectra for the rhEPO tryptic peptide harboring two glycosylation sites (EAEN_nTTGCAEH-CSLNENITVPTDK) produced in glyco-engineered moss lines: double KO Δ XylT/FucT and triple KO Δ XylT/FucT/GalT (based on Parsons et al., 2012). Salt adducts are marked with asterisks. The glycosylation patterns are schematized with sugar symbols above each peak. Le^a structures are totally absent in the triple KO line.

plant-typical O-glycosylation differs fundamentally from the typical human mucin-type O-glycosylation (reviewed by Gomord et al., 2010), and was shown to induce the formation of antibodies (Léonard et al., 2005). In plants, the main attachment site for O-glycosylation is 4-trans-Hyp (reviewed by Showalter, 2001), while no glycosylation of Hyp occurs in animals (Gorres and Raines, 2010). Hyp is generated posttranslationally by prolyl 4-hydroxylases (P4H) via hydroxylation of proline. Prolyl-hydroxylation is a very common modification both in mammals and in plants, though recognition sequences differ. In plants, the target motif for O-glycosylation after P4H-catalyzed hydroxylation, the so-called glycomodules present on Hyp-rich glycoproteins (HRGPs), are defined (Kieliszewski and Lamport, 1994) and validated (Tan et al., 2003; Shimizu et al., 2005). *In silico* analysis of the human proteome revealed that 30% of the human proteins bear a recognition sequence for plant P4Hs (Gomord et al., 2010), thus being putative candidates for non-human prolyl-hydroxylation when recombinantly produced in plant-based systems. In fact, undesired plant-typical prolyl-hydroxylation and in some cases even non-human O-glycosylation of biopharmaceuticals was reported (Karnoup, 2005; Weise et al., 2007; Pinkhasov et al., 2011). The most direct strategy to avoid non-human O-glycosylation in PMPs is the elimination of the anchor Hyp, which itself is an undesired PTM performed by plant P4H enzymes.

After systematic disruption of each of the six *p4h* genes in *Physcomitrella*, targeted deletion of *p4h1* resulted in the complete elimination of the previously reported prolyl-hydroxylation of moss-produced rhEPO (Parsons et al., 2013). As prolyl-hydroxylation and further glycosylation of plant extracellular matrix and cell wall proteins play important roles for growth,

cell differentiation, and stress adaption (Lampert et al., 2006; Velasquez et al., 2011) we expected a negative impact on the growth rate of the lines. However, the $\Delta p4h1$ moss lines were not impaired neither in growth or development nor in protein productivity (Parsons et al., 2013).

The ease of gene targeting in moss enabled glyco-engineering approaches for the elimination of any plant-typical immunogenic residues. This provides a plant-based system offering the stable production of safe protein therapeutics.

AUTHOR CONTRIBUTIONS

Eva L. Decker, Juliana Parsons, and Ralf Reski were involved in gathering and interpretation of data, writing the manuscript and revising the work. Ralf Reski is co-inventor of the moss bioreactor and co-founder of greenovation Biotech. He currently serves as advisory board member of this company. Eva L. Decker, Juliana Parsons, and Ralf Reski are co-inventors of patents and patent applications related to the topic discussed here.

ACKNOWLEDGMENTS

This work was supported by contract research “Glykobiologie/Glykomik” of the Baden-Württemberg Stiftung and by the Excellence Initiative of the German Federal and State Governments (EXC294 to Ralf Reski). We thank Anne Katrin Prowse for proof-reading of the manuscript.

REFERENCES

- Apweiler, R., Hermjakob, H., and Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8. doi: 10.1016/S0304-4165(99)00165-168
- Bailey, L. A., Hatton, D., Field, R., and Dickson, A. J. (2012). Determination of Chinese hamster ovary cell line stability and recombinant antibody expression during long-term culture. *Biotechnol. Bioeng.* 109, 2093–2103. doi: 10.1002/bit.24485
- Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., et al. (2003). Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. doi: 10.1093/glycob/cwg024
- Baur, A., Reski, R., and Gorr, G. (2005). Enhanced recovery of a secreted recombinant human growth factor using stabilizing additives and by co-expression of human serum albumin in the moss *Physcomitrella patens*. *Plant Biotechnol. J.* 3, 331–340. doi: 10.1111/j.1467-7652.2005.00127.x
- Bencúrová, M., Hemmer, W., Focke-Tejkl, M., Wilson, I. B. H., and Altmann, F. (2004). Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin. *Glycobiology* 14, 457–466. doi: 10.1093/glycob/cwh058
- Bosch, D., Castilho, A., Loos, A., Schots, A., and Steinkellner, H. (2013). N-glycosylation of plant-produced recombinant proteins. *Curr. Pharm. Des.* 19, 5503–5512. doi: 10.2174/1381612811319310006
- Büttner-Mainik, A., Parsons, J., Jérôme, H., Hartmann, A., Lamer, S., Schaaf, A., et al. (2011). Production of biologically active recombinant human factor H in *Physcomitrella*. *Plant Biotechnol. J.* 9, 373–383. doi: 10.1111/j.1467-7652.2010.00552.x
- Castilho, A., Neumann, L., Daskalova, S., Mason, H. S., Steinkellner, H., Altmann, F., et al. (2012). Engineering of sialylated mucin-type O-glycosylation in plants. *J. Biol. Chem.* 287, 36518–36526. doi: 10.1074/jbc.M112.402685
- Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., et al. (2013). Generation of biologically active multi-sialylated recombinant human EPOFc in plants. *PLoS ONE* 8:e54836. doi: 10.1371/journal.pone.0054836
- Chung, C. H., Mirakhur, B., Chan, E., Le, Q.-T., Berlin, J., Morse, M., et al. (2008). Cetuximab-induced anaphylaxis and IgE specific for galactose- α -1,3-galactose. *N. Engl. J. Med.* 358, 1109–1117. doi: 10.1056/NEJMoa074943
- Cove, D., Bezanilla, M., Harries, P., and Quatrano, R. (2006). Mosses as model systems for the study of metabolism and development. *Annu. Rev. Plant Biol.* 57, 497–520. doi: 10.1146/annurev.arplant.57.032905.105338
- Cox, K. M., Sterling, J. D., Regan, J. T., Gasdaska, J. R., Frantz, K. K., Peele, C. G., et al. (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* 24, 1591–1597. doi: 10.1038/nbt1260
- Croset, A., Delafosse, L., Gaudry, J.-P., Arod, C., Glez, L., Losberger, C., et al. (2012). Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *J. Biotechnol.* 161, 336–348. doi: 10.1016/j.biote.2012.06.038
- Daskalova, S. M., Radler, J. E., Cichacz, Z. A., Olsen, S. H., Tsaprailis, G., Mason, H., et al. (2010). Engineering of *N. benthamiana* L. plants for production of N-acetylgalactosamine-glycosylated proteins—towards development of a plant-based platform for production of protein therapeutics with mucin type O-glycosylation. *BMC Biotechnol.* 10:62. doi: 10.1186/1472-6750-10-62
- Decker, E. L., and Reski, R. (2007). Moss bioreactors producing improved biopharmaceuticals. *Curr. Opin. Biotechnol.* 18, 393–398. doi: 10.1016/j.copbio.2007.07.012
- Decker, E. L., and Reski, R. (2008). Current achievements in the production of complex biopharmaceuticals with moss bioreactors. *Bioprocess. Biosyst. Eng.* 31, 3–9. doi: 10.1007/s00449-007-0151-y
- Decker, E. L., and Reski, R. (2012). Glycoprotein production in moss bioreactors. *Plant Cell Rep.* 31, 453–460. doi: 10.1007/s00299-011-1152-1155
- Fischer, R., Schillberg, S., Hellwig, S., Twyman, R. M., and Drossard, J. (2012). GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol. Adv.* 30, 434–439. doi: 10.1016/j.biotechadv.2011.08.007
- Fitchette-Lainé, A.-C., Gomord, V., Cabanes, M., Michalski, J.-C., Saint Macary, M., Foucher, B., et al. (1997). N-glycans harboring the Lewis A epitope are expressed at the surface of plant cells. *Plant J.* 12, 1411–1417. doi: 10.1046/j.1365-313x.1997.12061411.x
- Frank, W., Decker, E. L., and Reski, R. (2005). Molecular tools to study *Physcomitrella patens*. *Plant Biol.* 7, 220–227. doi: 10.1055/s-2005-865645
- Gavel, Y., and von Heijne, G. (1990). Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng.* 3, 433–442. doi: 10.1093/protein/3.5.433
- Gomord, V., Fitchette, A.-C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., et al. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.* 8, 564–587. doi: 10.1111/j.1467-7652.2009.00497.x
- Gorres, K. L., and Raines, R. T. (2010). Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol.* 45, 106–124. doi: 10.3109/10409231003627991
- Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L., and Brunak, S. (1998). NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj. J.* 15, 115–130. doi: 10.1023/A:1006960004440
- Henry, S., Oriol, R., and Samuelsson, B. (1995). Lewis histo-blood group system and associated secretory phenotypes. *Vox Sang.* 69, 166–182. doi: 10.1111/j.1423-0410.1995.tb02591.x
- Hohe, A., Egener, T., Lucht, J. M., Holtorf, H., Reinhard, C., Schween, G., et al. (2004). An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, *Physcomitrella patens*. *Curr. Genet.* 44, 339–347. doi: 10.1007/s00294-003-0458-4
- Hohe, A., and Reski, R. (2002). Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass production of protoplasts. *Plant Sci.* 163, 69–74. doi: 10.1016/S0168-9452(02)00059-6
- Hohe, A., and Reski, R. (2005). Control of growth and differentiation of bioreactor cultures of *Physcomitrella* by environmental parameters. *Plant Cell Tissue Organ Cult.* 81, 307–311. doi: 10.1007/s11240-004-6656-z
- Horstmann, V., Huether, C. M., Jost, W., Reski, R., and Decker, E. L. (2004). Quantitative promoter analysis in *Physcomitrella patens*: a set of plant vectors activating gene expression within three orders of magnitude. *BMC Biotechnol.* 4:13. doi: 10.1186/1472-6750-4-13
- Hossler, P., Khattak, S. F., and Li, Z. J. (2009). Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19, 936–949. doi: 10.1093/glycob/cwp079
- Huether, C. M., Lienhart, O., Baur, A., Stemmer, C., Gorr, G., Reski, R., et al. (2005). Glyco-engineering of moss lacking plant-specific sugar residues. *Plant Biol.* 7, 292–299. doi: 10.1055/s-2005-837653
- Jez, J., Castilho, A., Grass, J., Vorauer-Uhl, K., Sterovsky, T., Altmann, F., et al. (2013). Expression of functionally active sialylated human erythropoietin in plants. *Biotechnol. J.* 8, 371–382. doi: 10.1002/biot.201200363

- Jin, C., Altmann, F., Strasser, R., Mach, L., Schähs, M., Kunert, R., et al. (2008). A plant-derived human monoclonal antibody induces an anti-carbohydrate immune response in rabbits. *Glycobiology* 18, 235–241. doi: 10.1093/glycob/cwm137
- Julenius, K., Mølgaard, A., Gupta, R., and Brunak, S. (2005). Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15, 153–164. doi: 10.1093/glycob/cwh151
- Kamionka, M. (2011). Engineering of therapeutic proteins production in *Escherichia coli*. *Curr. Pharm. Biotechnol.* 12, 268–274. doi: 10.2174/138920111794295693
- Kamisugi, Y., Schlink, K., Rensing, S. A., Schween, G., von Stackelberg, M., Cuming, A. C., et al. (2006). The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. *Nucleic Acids Res.* 34, 6205–6214. doi: 10.1093/nar/gkl832
- Karnoup, A. S. (2005). O-linked glycosylation in maize-expressed human IgA1. *Glycobiology* 15, 965–981. doi: 10.1093/glycob/cwi077
- Khrawesh, B., Arif, M. A., Seumel, G. I., Ossowski, S., Weigel, D., Reski, R., et al. (2010). Transcriptional control of gene expression by microRNAs. *Cell* 140, 111–122. doi: 10.1016/j.cell.2009.12.023
- Kieliszewski, M. J., and Lampert, D. T. A. (1994). Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J.* 5, 157–172. doi: 10.1046/j.1365-313X.1994.05020157.x
- Kim, J. Y., Kim, Y.-G., and Lee, G. M. (2012). CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl. Microbiol. Biotechnol.* 93, 917–930. doi: 10.1007/s00253-011-3758-3755
- Kircheis, R., Halanek, N., Koller, I., Jost, W., Schuster, M., Gorr, G., et al. (2012). Correlation of ADCC activity with cytokine release induced by the stably expressed, glyco-engineered humanized Lewis Y-specific monoclonal antibody MB314. *Mabs* 4, 532–541. doi: 10.4161/mabs.20577
- Kolarich, D., Jensen, P. H., Altmann, F., and Packer, N. H. (2012). Determination of site-specific glycan heterogeneity on glycoproteins. *Nat. Protoc.* 7, 1285–1298. doi: 10.1038/nprot.2012.062
- Koprivova, A., Altmann, F., Gorr, G., Kopriva, S., Reski, R., and Decker, E. L. (2003). N-glycosylation in the moss *Physcomitrella patens* is organized similarly to that in higher plants. *Plant Biol.* 5, 582–591. doi: 10.1055/s-2003-44721
- Koprivova, A., Stemmer, C., Altmann, F., Hoffmann, A., Kopriva, S., Gorr, G., et al. (2004). Targeted knockouts of *Physcomitrella* lacking plantspecific immunogenic N-glycans. *Plant Biotechnol. J.* 2, 517–523. doi: 10.1111/j.1467-7652.2004.00100.x
- Lampert, D. T., Kieliszewski, M. J., and Showalter, A. M. (2006). Salt stress upregulates periplasmic arabinogalactan proteins: using salt stress to analyse AGP function. *New Phytol.* 169, 479–492. doi: 10.1111/j.1469-8137.2005.01591.x
- Lang, D., Zimmer, A. D., Rensing, S. A., and Reski, R. (2008). Exploring plant biodiversity: the *Physcomitrella* genome and beyond. *Trends Plant Sci.* 13, 542–549. doi: 10.1016/j.tplants.2008.07.002
- Larkin, P. J., and Scowcroft, W. R. (1981). Somaclonal variation—a novel source of variability from cell cultures for plant improvement. *Theor. Appl. Genet.* 60, 197–214. doi: 10.1007/BF02342540
- Leonard, R., Lherould, S., Carlué, M., Fleurat, P., Maftah, A., and Costa, G. (2005). Biochemical characterization of *Silene alba* α 4-fucosyltransferase and Lewis a products. *Glycoconj. J.* 22, 71–78. doi: 10.1007/s10719-005-0404-404
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.-C., Gomord, V., and Faye, L. (1998). N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* 38, 31–48. doi: 10.1023/A:1006012005654
- Li, H., and d'Anjou, M. (2009). Pharmacological significance of glycosylation in therapeutic proteins. *Curr. Opin. Biotechnol.* 20, 678–684. doi: 10.1016/j.copbio.2009.10.009
- Lorenz, S., Tintelnot, S., Reski, R., and Decker, E. L. (2003). Cyclin D-knockout uncouples developmental progression from sugar availability. *Plant Mol. Biol.* 53, 227–236. doi: 10.1023/B:PLAN.0000009295.07016.87
- Ludwig-Müller, J., Jülke, S., Bierfreund, N. M., Decker, E. L., and Reski, R. (2009). Moss (*Physcomitrella patens*) GH3 proteins act in auxin homeostasis. *New Phytol.* 181, 323–338. doi: 10.1111/j.1469-8137.2008.02677.x
- Mann, M., and Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21, 255–261. doi: 10.1038/nbt0303-255
- Menand, B., Yi, K., Jouanic, S., Hoffmann, L., Ryan, E., Linstrand, P., et al. (2007). An ancient mechanism controls the development of cells with a root-ing function in land plants. *Science* 316, 1477–1480. doi: 10.1126/science.1142618
- Mononen, I., and Karjalainen, E. (1984). Structural comparison of protein sequences around potential N-glycosylation sites. *Biochim. Biophys. Acta* 788, 364–367. doi: 10.1016/0167-4838(84)90050-90055
- Mosquna, A., Katz, A., Decker, E. L., Rensing, S. A., Reski, R., and Ohad, N. (2009). Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. *Development* 136, 2433–2444. doi: 10.1242/dev.035048
- Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., et al. (2003). Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8007–8012. doi: 10.1073/pnas.0932694100
- Omasa, T., Onitsuka, M., and Kim, W.-D. (2010). Cell engineering and cultivation of Chinese hamster ovary (CHO) cells. *Curr. Pharm. Biotechnol.* 11, 233–240. doi: 10.2174/138920110791111960
- Parsons, J., Altmann, F., Arrenberg, C. K., Koprivova, A., Beike, A. K., Stemmer, C., et al. (2012). Moss-based production of asialo-erythropoietin devoid of Lewis A and other plant-typical carbohydrate determinants. *Plant Biotechnol. J.* 10, 851–861. doi: 10.1111/j.1467-7652.2012.00704.x
- Parsons, J., Altmann, F., Graf, M., Stadlmann, J., Reski, R., and Decker, E. L. (2013). A gene responsible for prolyl-hydroxylation of moss-produced recombinant human erythropoietin. *Sci. Rep.* 3:3019. doi: 10.1038/srep03019
- Paul, M., and Ma, J. K.-C. (2011). Plant-made pharmaceuticals: leading products and production platforms. *Biotechnol. Appl. Biochem.* 58, 58–67. doi: 10.1002/bab.6
- Paul, M. J., Teh, A. Y., Twyman, R. M., and Ma, J. K. (2013). Target product selection—where can molecular pharming make the difference? *Curr. Pharm. Des.* 19, 5478–5485. doi: 10.2174/1381612811319310003
- Petrucelli, S., Otegui, M. S., Lareu, F., Tran Dinh, O., Fitchette, A.-C., Circosta, A., et al. (2006). A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol. J.* 4, 511–527. doi: 10.1111/j.1467-7652.2006.00200.x
- Pinkhasov, J., Alvarez, M. L., Rigano, M. M., Piensook, K., Larios, D., Pabst, M., et al. (2011). Recombinant plant-expressed tumour-associated MUC1 peptide is immunogenic and capable of breaking tolerance in MUC1.Tg mice. *Plant Biotechnol. J.* 9, 991–1001. doi: 10.1111/j.1467-7652.2011.00614.x
- Qudeimat, E., Faltusz, A. M. C., Wheeler, G., Lang, D., Holtorf, H., Brownlee, C., et al. (2008). A PIIB-type Ca^{2+} -ATPase is essential for stress adaptation in *Physcomitrella patens*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19555–19560. doi: 10.1073/pnas.0800864105
- Raju, T. S., Briggs, J. B., Borge, S. M., and Jones, A. J. S. (2000). Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* 10, 477–486. doi: 10.1093/glycob/10.5.477
- Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., et al. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69. doi: 10.1126/science.1150646
- Reski, R., Reynolds, S., Wehe, M., Kleber-Janke, T., and Kruse, S. (1998). Moss (*Physcomitrella patens*) expressed sequence tags include several sequences which are novel for plants. *Bot. Acta* 111, 143–149. doi: 10.1111/j.1438-8677.1998.tb00689.x
- Reutter, K., and Reski, R. (1996). Production of a heterologous protein in bioreactor cultures of fully differentiated moss plants. *Plant Tissue Cult. Biotechnol.* 2, 142–147.
- Rother, S., Hadeler, B., Orsini, J. M., Abel, W. O., and Reski, R. (1994). Fate of a mutant macrochloroplast in somatic hybrids. *J. Plant Physiol.* 143, 72–77. doi: 10.1016/S0176-1617(11)82099-7
- Sakakibara, K., Ando, S., Yip, H. K., Tamada, Y., Hiwatashi, Y., Murata, T., et al. (2013). KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants. *Science* 339, 1067–1070. doi: 10.1126/science.1230082
- Schaaf, A., Reski, R., and Decker, E. L. (2004). A novel aspartic proteinase is targeted to the secretory pathway and to the vacuole in the moss *Physcomitrella patens*. *Eur. J. Cell Biol.* 83, 145–152. doi: 10.1078/0171-9335-00371
- Schaaf, A., Tintelnot, S., Baur, A., Reski, R., Gorr, G., and Decker, E. L. (2005). Use of endogenous signal sequences for transient production and efficient secretion

- by moss (*Physcomitrella patens*) cells. *BMC Biotechnol.* 5:30. doi: 10.1186/1472-6750-5-30
- Schaefer, D. G. (2001). Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol.* 4, 143–150. doi: 10.1016/S1369-5266(00)00150-153
- Schellekens, H. (2002). Bioequivalence and the immunogenicity of biopharmaceuticals. *Nat. Rev. Drug Discov.* 1, 457–462. doi: 10.1038/nrd818
- Schillberg, S., Raven, N., Fischer, R., Twyman, R. M., and Schiermeyer, A. (2013). Molecular farming of pharmaceutical proteins using plant suspension cell and tissue cultures. *Curr. Pharm. Des.* 19, 5531–5542. doi: 10.2174/1381612811319310008
- Schmidt, F. R. (2004). Recombinant expression systems in the pharmaceutical industry. *Appl. Microbiol. Biotechnol.* 65, 363–372. doi: 10.1007/s00253-004-1656-1659
- Schulte, J., and Reski, R. (2004). High throughput cryopreservation of 140,000 *Physcomitrella patens* mutants. *Plant Biol.* 6, 119–127. doi: 10.1055/s-2004-817796
- Schuster, M., Jost, W., Mudder, G. C., Wiederkum, S., Schwager, C., Janzel, E., et al. (2007). In vivo glyco-engineered antibody with improved lytic potential produced by an innovative non-mammalian expression system. *Biotechnol. J.* 2, 700–708. doi: 10.1002/biot.200600255
- Shaaltiel, Y., Bartfeld, D., Hashmueli, S., Baum, G., Brill-Almon, E., Galili, G., et al. (2007). Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant Biotechnol. J.* 5, 579–590. doi: 10.1111/j.1467-7652.2007.00263.x
- Shields, R. L., Lai, J., Keck, R., O'Connell, L. Y., Hong, K., Meng, Y. G., et al. (2002). Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc γ RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* 277, 26733–26740. doi: 10.1074/jbc.M202069200
- Shimizu, M., Igasaki, T., Yamada, M., Yuasa, K., Hasegawa, J., Kato, T., et al. (2005). Experimental determination of proline hydroxylation and hydroxyproline arabinogalactosylation motifs in secretory proteins. *Plant J.* 42, 877–889. doi: 10.1111/j.1365-313X.2005.02419.x
- Shinkawa, T., Nakamura, K., Yamane, N., Shoji-Hosaka, E., Kanda, Y., Sakurada, M., et al. (2003). The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem.* 278, 3466–3473. doi: 10.1074/jbc.M210665200
- Showalter, A. M. (2001). Arabinogalactan-proteins: structure, expression and function. *Cell. Mol. Life Sci.* 58, 1399–1417. doi: 10.1007/PL00000784
- Strasser, R., Altmann, F., Mach, L., Glössl, J., and Steinkellner, H. (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β 1,2-linked xylose and core α 1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4
- Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007). A unique β 1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis A structures in *Arabidopsis thaliana*. *Plant Cell Online* 19, 2278–2292. doi: 10.1105/tpc.107.052985
- Strepp, R., Scholz, S., Kruse, S., Speth, V., and Reski, R. (1998). Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4368–4373. doi: 10.1073/pnas.95.8.4368
- Strotbek, C., Krimmer, S., and Frank, W. (2013). The moss *Physcomitrella patens*: methods and tools from cultivation to targeted analysis of gene function. *Int. J. Dev. Biol.* 57, 553–564. doi: 10.1387/ijdb.130189wf
- Tan, L., Leykam, J. F., and Kieliszewski, M. J. (2003). Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins. *Plant Physiol.* 132, 1362–1369. doi: 10.1104/pp.103.021766
- van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.-P., Loutelier-Bourhis, C., Rayon, C., et al. (2000). β (1,2)-xylose and α (1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.* 275, 11451–11458. doi: 10.1074/jbc.275.15.11451
- Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., et al. (2009). *Essentials of Glycobiology*, 2nd Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Velasquez, S. M., Ricardi, M. M., Dorosz, J. G., Fernandez, P. V., Nadra, A. D., Pol-Fachin, L., et al. (2011). O-glycosylated cell wall proteins are essential in root hair growth. *Science* 332, 1401–1403. doi: 10.1126/science.1206657
- Viétor, R., Loutelier-Bourhis, C., Fitchette, A.-C., Margerie, P., Gonneau, M., Faye, L., et al. (2003). Protein N-glycosylation is similar in the moss *Physcomitrella patens* and in higher plants. *Planta* 218, 269–275. doi: 10.1007/s00425-003-1099-z
- Walsh, G. (2010a). Biopharmaceutical benchmarks 2010. *Nat. Biotechnol.* 28, 917–924. doi: 10.1038/nbt0910-917
- Walsh, G. (2010b). Post-translational modifications of protein biopharmaceuticals. *Drug Discov. Today* 15, 773–780. doi: 10.1016/j.drudis.2010.06.009
- Walsh, G., and Jefferis, R. (2006). Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* 24, 1241–1252. doi: 10.1038/nbt1252
- Weise, A., Altmann, F., Rodriguez-Franco, M., Sjoberg, E. R., Bäumer, W., Launhardt, H., et al. (2007). High-level expression of secreted complex glycosylated recombinant human erythropoietin in the *Physcomitrella* Δ -fuc-t Δ -xyl-t mutant. *Plant Biotechnol. J.* 5, 389–401. doi: 10.1111/j.1467-7652.2007.00248.x
- Weise, A., Rodriguez-Franco, M., Timm, B., Hermann, M., Link, S., Jost, W., et al. (2006). Use of *Physcomitrella patens* actin 5' regions for high transgene expression: importance of 5' introns. *Appl. Microbiol. Biotechnol.* 70, 337–345. doi: 10.1007/s00253-005-0087-8
- Westphal, S., Kolarich, D., Foetisch, K., Lauer, I., Altmann, F., Conti, A., et al. (2003). Molecular characterization and allergenic activity of Lyc e 2 (β -fructofuranosidase), a glycosylated allergen of tomato. *Eur. J. Biochem.* 270, 1327–1337. doi: 10.1046/j.1432-1033.2003.03503.x
- Wilson, I. B. H. (2001). Identification of a cDNA encoding a plant Lewis-type α 1,4-fucosyltransferase. *Glycoconj. J.* 18, 439–447. doi: 10.1023/A:101603000527
- Wilson, I. B. H. (2002). Glycosylation of proteins in plants and invertebrates. *Curr. Opin. Struct. Biol.* 12, 569–577. doi: 10.1016/S0959-440X(02)00367-366
- Wilson, I. B. H., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J. M., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis A, core α 1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., et al. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* 29, 735–741. doi: 10.1038/nbt.1932
- Yang, Z., Bennett, E. P., Jørgensen, B., Drew, D. P., Arigi, E., Mandel, U., et al. (2012). Toward stable genetic engineering of human O-glycosylation in plants. *Plant Physiol.* 160, 450–463. doi: 10.1104/pp.112.198200
- Zhang, K., Baeckström, D., and Hansson, G. C. (1994). A secreted mucin carrying Sialyl-Lewis A from colon carcinoma cells binds to E-selectin and inhibits HL-60 cell adhesion. *Int. J. Cancer* 59, 823–829. doi: 10.1002/ijc.2910590619
- Zimmer, A. D., Lang, D., Buchta, K., Rombauts, S., Nishiyama, T., Hasebe, M., et al. (2013). Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. *BMC Genomics* 14:498. doi: 10.1186/1471-2164-14-498

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 May 2014; paper pending published: 16 June 2014; accepted: 27 June 2014; published online: 09 July 2014.

Citation: Decker EL, Parsons J and Reski R (2014) Glyco-engineering for biopharmaceutical production in moss bioreactors. *Front. Plant Sci.* 5:346. doi: 10.3389/fpls.2014.00346

This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Decker, Parsons and Reski. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Trafficking of endoplasmic reticulum-retained recombinant proteins is unpredictable in *Arabidopsis thaliana*

Thomas De Meyer^{1,2} and Ann Depicker^{1,2*}

¹ Department of Plant Systems Biology, VIB, Plant-made Antibodies and Immunogens, Gent, Belgium

² Department of Plant Biotechnology and Bioinformatics, Ghent University, Gent, Belgium

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Richard Strasser, University of Natural Resources and Life Sciences, Austria

Jurgen Denecke, The University of Leeds, UK

Els J. M. Van Damme, Ghent University, Belgium

***Correspondence:**

Ann Depicker, Department Plant Systems Biology, VIB, Technologiepark 927, Gent 9052, Belgium
e-mail: anpic@psb.vib-ugent.be

A wide variety of recombinant proteins has been produced in the dicot model plant, *Arabidopsis thaliana*. Many of these proteins are targeted for secretion by means of an N-terminal endoplasmic reticulum (ER) signal peptide. In addition, they can also be designed for ER retention by adding a C-terminal H/KDEL-tag. Despite extensive knowledge of the protein trafficking pathways, the final protein destination, especially of such H/KDEL-tagged recombinant proteins, is unpredictable. In this respect, glycoproteins are ideal study objects. Microscopy experiments reveal their deposition pattern and characterization of their N-glycans aids in elucidating the trafficking. Here, we combine microscopy and N-glycosylation data generated in *Arabidopsis* leaves and seeds, and highlight the lack of a decent understanding of heterologous protein trafficking.

Keywords: molecular farming, antibody production, dense vesicle, ER-derived vesicle, protein storage vacuole, KDEL, seed-specific expression

INTRODUCTION

Recombinant proteins are often produced in eukaryotic host organisms to ensure proper folding, disulfide bridge formation and N-glycan processing. By fusing the protein of interest to an N-terminal endoplasmic reticulum (ER) signal peptide, they co-translationally enter the ER and travel along the secretory pathway, where N-glycosylation takes place on the consensus Asn-X-Ser/Thr motif. The N-glycan composition is of crucial importance for the protein structure, stability, half-life and function, and is primarily determined by the production host (Jacobs and Callewaert, 2009). In molecular farming, where plants are used as production systems, heterologous proteins targeted for secretion typically contain complex-type N-glycans with β -1,2-xylose and core α -1,3-fucose residues. These are potentially immunogenic and are hence unwanted for therapeutic protein production (Gomord et al., 2005). Therefore, significant efforts went into the development of glyco-engineered production platforms that prevent plant-specific N-glycosylation in the Golgi complex (Gomord et al., 2010; Castilho and Steinkellner, 2012). Alternatively, heterologous glycoproteins can be retained in the ER by adding a C-terminal H/KDEL-tag. Whereas both HDEL- and KDEL-tagged endogenous proteins are found in plants (Napier et al., 1992), the vast majority of heterologously expressed, ER-retained proteins carry the KDEL-tag. Typical for H/KDEL-tagging is the formation of Man₈GlcNAc₂ (Man8) N-glycans. However, as ER retention is based on retrograde trafficking from cis-Golgi to ER, the glycoproteins transiently encounter cis-Golgi processing enzymes, such as α -1,2-mannosidase, resulting in partially trimmed N-glycans, such as Man7 and Man6. In addition to the desired N-glycan profile, H/KDEL-tagging can also enhance protein accumulation levels, presumably because

the ER is a favorable compartment for protein folding and storage (Fiedler et al., 1997; Petruccielli et al., 2006). However, such an increased accumulation is not always observed (Loos et al., 2011a,b).

In contrast to this black and white distinction between secretion and H/KDEL-mediated ER retention, numerous protein localization studies reported unexpected outcomes. Drawing clear conclusions from these experiments proved hard, because of the different proteins of interest, plant species, tissues, promoters, regulatory sequences, targeting signals and achieved accumulation levels (De Muynck et al., 2010). Moreover, protein trafficking has also been shown to change throughout development (Arcalis et al., 2010; Wang et al., 2012). In this review, we provide a detailed overview of protein localization studies in leaves and seeds of the dicot model plant, *Arabidopsis thaliana*. By limiting ourselves to *Arabidopsis*, we eliminate organismal specificity and highlight tissue (i.e., leaves vs. seeds) and protein specificity in heterologous protein trafficking.

In *Arabidopsis*, most recombinant proteins have been produced in seeds, providing the advantage of long-term storage capacity, high protein content and productivity, and no interference with vegetative plant growth (Stoger et al., 2005; Kermode, 2012). In head-to-head comparisons with the same protein of interest, *Arabidopsis* seeds were more positively evaluated than those of tobacco, petunia and maize in terms of protein accumulation levels (Loos et al., 2011a; Morandini et al., 2011). However, the impact of such comparisons is limited due to the different efficiencies of regulatory sequences and codon usage across organisms. Nevertheless, one of the highest accumulation levels achieved in plants is still that of the seed-produced G4 scFv in *Arabidopsis* (i.e., 36.5% of total soluble protein (TSP) in

homozygous seeds) (De Jaeger et al., 2002). De Wilde et al. (2013) showed that VHH-Fc and scFv-Fc accumulation levels of 1% or more in *Arabidopsis* seeds trigger an unfolded protein response, because an enhanced expression of genes involved in protein folding, glycosylation, protein translocation, degradation and vesicle trafficking was observed. However, despite such an altered gene expression profile, *Arabidopsis* seeds often fail to provide a 100% N-glycan site occupancy (**Table 1B**; 3–18 and 26).

TRAFFICKING OF PROTEINS TARGETED FOR SECRETION

In leaves, heterologous proteins that carry an N-terminal ER signal peptide, are efficiently secreted to the apoplast (De Wilde et al., 1996; Peeters et al., 2001) (**Table 1A**; 1–3) and mainly carry complex-type N-glycans (Schahs et al., 2007) (**Table 1A**; 5 and 6). Of note, Loos et al. (2011b) found that an anti-hepatitis A virus scFv-Fc (HA78) contained complex-type N-glycans as expected, while an anti-HIV scFv-Fc (2G12) was completely covered with oligomannosidic N-glycans (**Table 1A**; 7 and 8). Because the authors could not detect antigen-binding activity for this 2G12 scFv-Fc, they postulated that it was not folded properly and activated the ER-associated protein degradation pathway, hence preventing further N-glycan maturation in the Golgi apparatus.

In seeds, despite successful examples of protein secretion with complex-type N-glycans, some exceptions stress the lack of a decent understanding of secreted heterologous protein trafficking. For example, HA78 and 2G12 monoclonal antibodies (mAbs) were both found in the apoplast and in electron-opaque Golgi-attached dense vesicles (DVs) in developing seeds (Loos et al., 2011a) (**Table 1B**; 8 and 11). DVs are distinct from clathrin-coated vesicles that normally mediate protein secretion, and are considered the main pathway for massive seed storage protein transport from the trans-Golgi network to the protein storage vacuole (PSV) (Robinson et al., 2005; Vitale and Hinz, 2005; Otegui et al., 2006; Wang et al., 2012) (**Figure 1**; blue stars). Their electron-opaque content reflects the aggregated state of the storage proteins. Possibly, the highly abundant storage proteins, such as globulins, exhibit a dominant sorting effect that leads to partial trapping of the heterologous proteins in DVs. A similar mechanism, imposed by endogenous seed storage proteins, has been proposed for recombinant phytase in ER-derived prolamin bodies of rice endosperm (Drakakaki et al., 2006). In agreement with the co-sorting hypothesis to the PSV via DVs in *Arabidopsis*, glucocerebrosidase that was targeted for secretion, was mainly located in the apoplast and to a minor extent in PSVs in mature seeds (He et al., 2012a) (**Table 1B**; 22). Alternatively, partial mislocalization of secreted proteins to DVs and PSVs can also be explained by the presence of cryptic vacuolar specific sequences (Jolliffe et al., 2005).

Both the HA78 and 2G12 mAbs were produced as scFv-Fc moieties, using the same targeting and regulatory sequences (Loos et al., 2011b) (**Table 1B**; 13 and 16). On the one hand, labeling in apoplast and Golgi-attached DVs was obtained for HA78 scFv-Fc (identical as for HA78 mAb) in developing *Arabidopsis* seeds. In mature seeds, PSVs were devoid of label, so the question remains where the DV-localized HA78 scFv-Fcs of the developing embryos ended up. Instead, the final destinations of HA78 scFv-Fc were the apoplast and “globular, membrane-delimited structures of

around 200 to 400 nm in diameter.” The latter were termed ER-derived vesicles (ERVs), because ribosomes were observed on their surface, but their specific formation in later developmental stages was unclear. This dual deposition pattern was in accordance with the presence of both complex-type and oligomannosidic N-glycans. On the other hand, 2G12 scFv-Fc exclusively contained Man7 and Man8 N-glycans, and was observed in ERVs and the swollen nuclear envelope. This aberrant localization is in agreement with the proposed improper folding of 2G12 scFv-Fc (see above in *Arabidopsis* leaves).

TRAFFICKING OF PROTEINS TARGETED FOR ER-RETENTION

Only one study has been performed in *Arabidopsis* leaves, in which a KDEL-tagged Fab fragment was detected intracellularly, most likely in the endomembrane system (Peeters et al., 2001) (**Table 1A**; 4).

In seeds, only Loos et al. (2011a) conclusively demonstrated successful ER retention, more in particular for a minor fraction of KDEL-tagged 2G12 mAb (**Table 1B**; 10). All other studies describe distinct deposition patterns (**Figure 1**). First, the most prevalent observation is the formation of ERVs, in a process that is not fully understood (Van Droogenbroeck et al., 2007; Loos et al., 2011b; Morandini et al., 2011; He et al., 2012b) (**Table 1B**; 3, 14, 17, 20 and 24). Their origin resembles KDEL-vesicles (KV) of *Vigna mungo* seeds, by which SH-EP, a KDEL-tagged vacuolar proteinase, is shuttled from the ER to the PSV upon germination (Toyooka et al., 2000). Moreover, the C-terminal KDEL-tag of SH-EP was shown to be essential for KV formation (Okamoto et al., 2003). Similarly, after producing GFP-KDEL in tobacco leaves, so-called protein bodies were observed in most transformants with a GFP accumulation level of at least 0.2% of TSP (Conley et al., 2009; Gutiérrez et al., 2013). Taken together, it seems that the KDEL-tag ensures a local protein build-up in the ER lumen, from which ERVs, KVs or protein bodies are formed. From results obtained in mature *Arabidopsis* seeds, two hypotheses were made. On the one hand, ERVs can represent the end-stage of heterologous protein trafficking (Van Droogenbroeck et al., 2007; Morandini et al., 2011; Loos et al., 2011b), which sometimes are observed together with equal amounts of protein deposited in the swollen nuclear envelope (Loos et al., 2011b) (**Table 1B**, 14 and 17). On the other hand, ERVs can mediate a Golgi-independent pathway to the PSV. This was demonstrated by the EndoH sensitivity of the GAD67/65 glycoprotein (Morandini et al., 2011) (**Table 1B**; 19) and the large fraction of oligomannosidic N-glycans on 2G12 mAb (Loos et al., 2011a) (**Table 1B**; 10). A similar ER-derived Golgi-independent pathway toward the PSV has been described for endogenous seed storage proteins in pumpkin seeds, where the shuttle vesicles were termed precursor-accumulating (PAC) vesicles (Hara-Nishimura et al., 1998). The authors hypothesized that the PAC pathway has evolved for efficient, massive transport of unglycosylated seed storage proteins to the PSV. Of note, vacuole biogenesis might also represent an ER-to-vacuole route, because the ER was recently proposed as the main membrane source for lytic vacuole formation (Viotti et al., 2013). Although such a mechanism has not yet been observed for PSVs, it is tempting to state that some ER-retained heterologous proteins are trapped into a PSV precursor in a similar process (**Figure 1**).

Table 1 | Overview of recombinant protein production in *Arabidopsis* leaves (A) and seeds (B), in which white boxes indicate recombinant proteins targeted for secretion, and gray boxes correspond to KDEL-tagged proteins.

production host ^a	protein of interest ^b	Localization ^c	N-glycosylation ^d	N-glycan occupancy ^e	accumulation level ^f	reference ^g
(A) PROTEIN ANALYSIS IN ARABIDOPSIS LEAVES						
1 WT	MAK33 mAb	apoplast	n.d.	n.d.	n.d.	De Wilde et al., 1996
2 WT	MAK33 Fab	apoplast	n.a.	n.a.	n.d.	De Wilde et al., 1996
3 WT	MAK33 Fab	apoplast	n.a.	n.a.	up to 6.5% of TSP	Peeters et al., 2001
4 WT	MAK33 Fab	endomembrane system	n.a.	n.a.	up to 5.9% of TSP	Peeters et al., 2001
5 WT	2G12 mAb	n.d.	GnGnXF, MGnXF, MMXF, Man7–9	100%	0.05–0.2% of TSP in young plants	Schahs et al., 2007
6 XT/FT k.o.	2G12 mAb	n.d.	GnGn, MGn, Man4–9	100%	0.05–0.2% of TSP in young plants	Schahs et al., 2007
7 WT	2G12 scFv-Fc	n.d.	Man4–9	n.d.	n.d.	Loos et al., 2011b
8 WT	HA78 scFv-Fc	n.d.	GnGnXF, MGnXF, MMXF	n.d.	n.d.	Loos et al., 2011b
(B) PROTEIN ANALYSIS IN ARABIDOPSIS SEEDS						
1 WT	human α-L-iduronidase	n.d.	N-glycans with β-1,2-xylosylation, oligomannosidic N-glycans	n.d.	up to 0.006% of TSP	Downing et al., 2006
2 cgl	human α-L-iduronidase	apoplast (developing seeds)	mainly oligomannosidic N-glycans	n.d.	up to 1.8% of TSP	Downing et al., 2006
3 WT	MBP10 scFv-Fc	periplasmic space, ER-derived spherical bodies	Man5–9	60–65%	up to 12.4% of TSP	Van Droogenbroeck et al., 2007
4 WT	HA78 scFv-Fc	n.d.	Man5–9	60–65%	up to 13.1% of TSP	Van Droogenbroeck et al., 2007
5 WT	EHF34 scFv-Fc	n.d.	Man5–8	60–65%	up to 13.9% of TSP	Van Droogenbroeck et al., 2007
6 WT	MBP10 scFv-Fc	n.d.	Man7–8	78%	n.d.	Henquet et al., 2011
7 alg3-2	MBP10 scFv-Fc	n.d.	Man5, Man7/Glc2Man5	69%	n.d.	Henquet et al., 2011
8 WT	2G12 mAb	apoplast, Golgi-attached dense vesicles (developing seeds)—apoplast	GnGnXF, MGnXF, Man7–8	<100%	up to 3.6 μg/mg DW	Loos et al., 2011a; Arcalis et al., 2013
9 XT/FT k.o.	2G12 mAb	n.d.	GnGn, MGn, Man7–8	<100%	up to 2.1 μg/mg DW	Loos et al., 2011a
10 WT	2G12 mAb	PSV, ER (developing seeds)	GnGnXF, Man6–8	<100%	up to 3.0 μg/mg DW	Loos et al., 2011a
11 WT	HA78 mAb	apoplast, Golgi-attached dense vesicles (developing seeds)	GnGnXF, Man5–8	<100%	up to 8.8 μg/mg DW	Loos et al., 2011a
12 XT/FT k.o.	HA78 mAb	n.d.	GnGn, MGn, Man5–8	<100%	up to 9.8 μg/mg DW	Loos et al., 2011a
13 WT	2G12 scFv-Fc	ER-derived vesicles, nuclear envelope (developing and mature seeds)	Man7–8	<100%	up to 0.8 μg/mg DW	Loos et al., 2011b

(Continued)

Table 1 | Continued

production host ^a	protein of interest ^b	Localization ^c	N-glycosylation ^d	N-glycan occupancy ^e	accumulation level ^f	reference ^g
14 WT	2G12 scFv-Fc	ER-derived vesicles, nuclear envelope	Man7–8	<100%	up to 0.8 µg/mg DW	Loos et al., 2011b
15 XT/FT k.o.	2G12 scFv-Fc	n.d.	Man7–8	<100%	up to 3.5 µg/mg DW	Loos et al., 2011b
16 WT	HA78 scFv-Fc	apoplast, Golgi, Golgi-attached dense vesicles (developing seeds)—apoplast, ER-derived vesicles	GnGnXF, MGnXF, Man5–9	70%	up to 8.0 µg/mg DW	Loos et al., 2011b
17 WT	HA78 scFv-Fc	PSV, ER-derived vesicles (developing seeds)—ER-derived vesicles, nuclear envelope	Man7–8	<100%	up to 3.9 µg/mg DW	Loos et al., 2011b
18 XT/FT k.o.	HA78 scFv-Fc	n.d.	GnGn, MGn, Man6–8	66%	up to 9.4 µg/mg DW	Loos et al., 2011b
19 WT	glutamic acid decarboxylase GAD67/65	PSV	oligomannosidic N-glycans	100%	1.5–5.4% of TSP (up to 4.5 µg/mg DW)	Morandini et al., 2011
20 WT	interleukin-10	apoplast, ER-like membrane compartments	n.a.	n.a.	0.1–0.7% of TSP (up to 0.82 µg/mg DW)	Morandini et al., 2011
21 WT	proinsulin	PSV	n.a.	n.a.	< 0.01% of TSP (up to 0.005 µg/mg DW)	Morandini et al., 2011
22 cgl	glucocerebrosidase	apoplast, PSV	MGnXF, MMXF, Man5F, Man4–8	n.d.	up to 0.1% of TSP	He et al., 2012a
23 WT	human α-L-iduronidase	n.d.	complex-type N-glycans, Man5–8	n.d.	up to 0.46% of TSP	He et al., 2012b
24 cgl	human α-L-iduronidase	punctate vesicles, apoplast (developing seeds)—ER or ER-derived compartments, PSV	GnGnXF, MGnXF, MMXF, Man5–9	n.d.	up to 0.61% of TSP	He et al., 2012b
25 WT	pVHH7-hGFc (VHH-Fc)	n.d.	N-glycans without α-1,3-fucosylation	100%	up to 16.3% of TSP	De Buck et al., 2013
26 WT	sV3A (sVHH-Fc)	n.d.	oligomannosidic N-glycans	<100%	up to 5 µg/mg DW	Virdi et al., 2013
27 WT	GP3, GP4 and GP5 antigens	n.d.	oligomannosidic N-glycans	n.d.	up to 2.7% of TSP	Piron et al., 2014

n.d., not determined; n.a., not applicable.

^aalg3-2, α-1,3-mannosyltransferase mutant (Henquet et al., 2008); cgl, complex glycan mutant (Von Schaewen et al., 1993); WT, wild-type; XT/FT k.o., β-1,2-xylosyltransferase and core α-1,3-fucosyltransferase knockout line (Strasser et al., 2004).

^bFab, fragment antigen-binding; mAb, monoclonal antibody; scFv-Fc, single-chain variable fragment fused to a fragment crystallisable; VHH-Fc, variable domain of the heavy chain of the heavy-chain antibody fused to a fragment crystallisable.

^cIn mature, dry seeds unless mentioned otherwise.

^dGlc₂Man5, Glc₂Man₅GlcNAc₂; GnGn, GlcNAc₂Man₃GlcNAc₂; GnGnXF, GlcNAc₂Man₃GlcNAc₂XylFuc; Man4–9, Man_{4–9}GlcNAc₂; Man5F, Man₅GlcNAc₂Fuc; MMXF, Man₃GlcNAc₂XylFuc; MGn, GlcNAc₁Man₃GlcNAc₂; MGnXF, GlcNAc₁Man₃GlcNAc₂XylFuc.

^eSometimes determined ourselves based on available figures in the corresponding references.

^fIn our experience, 1% of TSP corresponds to about 2.5 µg/mg DW. DW, dry weight; TSP, total soluble protein.

^gStudies were ordered chronologically.

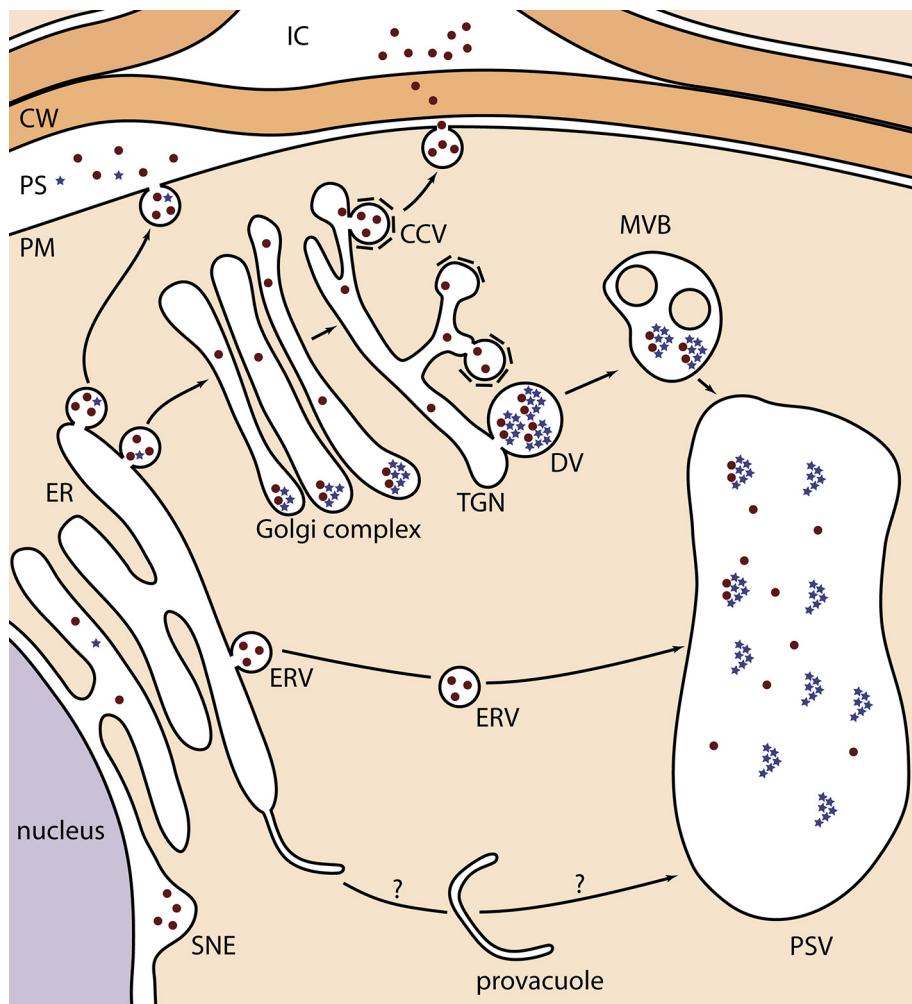


FIGURE 1 | Schematic representation of the different localizations of KDEL-tagged recombinant proteins in *Arabidopsis* seeds. Recombinant proteins are depicted as brown dots, and globulin seed storage proteins as blue stars. During their transport, globulins aggregate in the periphery of the Golgi cisternae, from where they bud into DVs toward the PSV. In one occasion, they were observed in the PS because of a disturbed protein trafficking of endogenous proteins (Van Droogenbroeck et al., 2007). The

ER-to-provacuole hypothesis, as depicted at the bottom, is based on recent observations of lytic vacuole biogenesis (Viotti et al., 2013). CCV, clathrin-coated vesicle; CW, cell wall; DV, dense vesicle; ER, endoplasmic reticulum; ERV, ER-derived vesicle; IC, intercellular space; MVB, multivesicular body; PS, periplasmic space; PSV, protein storage vacuole; PM, plasma membrane; SNE, swollen nuclear envelope; TGN, trans-Golgi network.

Second, partial leakage to the Golgi has been observed because KDEL-tagged proteins accumulated in the apoplast (Morandini et al., 2011) (**Table 1B**; 20) or carried complex-type N-glycans (Loos et al., 2011a) (**Table 1B**; 10). Such ER leakage has also been described in *Medicago* and tobacco (Triguero et al., 2005; Petrucelli et al., 2006; Abrançches et al., 2008). The authors suggested several factors that might influence successful ER retention. For example, one should consider the amount of KDEL-tags per assembled molecule, and the accessibility and integrity of the KDEL-tag. Remarkably, based on western blot analysis of subcellular fractions, He et al. (2012b) suggested a Golgi-dependent route toward the PSV for KDEL-tagged human α -L-iduronidase (**Table 1B**; 24). Unfortunately, electron microscopy localization studies on mature seeds confirming this hypothesis, were lacking.

Third, after producing MBP10 scFv-Fc in *Arabidopsis* seeds, an electron-opaque periplasmic space (PS) between the plasma membrane and cell wall was observed in which most of the MBP10 was deposited (Van Droogenbroeck et al., 2007) (**Table 1B**; 3). Moreover, MBP10 exclusively contained oligomannosidic N-glycans, pointing to a Golgi-independent route from the ER to the PS. Because ER-resident proteins, such as calreticulin and binding protein, were also present in the PS, and because of the very high MBP10 accumulation level (up to 12.4% of TSP), the authors proposed an overcharge of the ER storage capacity. Interestingly, globulin storage proteins were also deposited in the PS.

In several other studies, localization experiments were out-of-scope. However, N-glycan analyses were performed and only oligomannosidic N-glycans were detected (Henquet et al., 2011;

De Buck et al., 2013; Virdi et al., 2013; Piron et al., 2014 (**Table 1B**; 6, 7 and 25–27). Although the authors probably assumed successful ER retention, we conclude that, based on the aforementioned detailed localization studies, these KDEL-tagged glycoproteins can also reside in ERVs, the PSV (by bypassing the Golgi complex) or the PS.

CONCLUDING REMARKS

It is critical for recombinant protein production that the platform is reliable and predictable in terms of product quality. In this respect, the ER retention of KDEL-tagged recombinant proteins in *Arabidopsis*, has proven to be unpredictable, and similar observations were made in other plant production systems. Therefore, H/KDEL-tagged proteins, produced in plants, should always be analyzed in terms of final destination and N-glycan composition, unless of course, if the actual N-glycan composition is of less importance (e.g., for diagnostic proteins) and does not influence the final product performance.

Obviously, additional investigations are needed. For example, one could analyse the influence of the protein accumulation level by comparing protein localizations in low and high expressing lines. Alternatively, the promoter sequences used might also severely impact the observed protein localization. To this end, promoters with different temporal and spatial expression patterns, especially during seed development, could be worthwhile to study. Further, deletion and swapping experiments can reveal whether particular protein domains of the heterologous protein contain certain localization motifs resulting in the lack of ER retention. Finally, most of the manuscripts discussed here, did not verify KDEL-tag accessibility or integrity. For future reports on H/KDEL-tagged protein production in plants, it would thus be valuable to include, for example, western blot or immunoprecipitation analyses with commercially available anti-H/KDEL antibodies. In case recognition by these antibodies fails, linker sequences can be used to improve H/KDEL-tag accessibility. We conclude that such investigations will result in a much better predictability of the ER retention of overexpressed H/KDEL-tagged proteins in plants, and eventually contribute to the further establishment of the field of plant molecular farming.

ACKNOWLEDGMENTS

The authors thank Annick Bleys, Daniel Van Damme and Steffen Vanneste for help in preparing the manuscript.

REFERENCES

- Abrançhes, R., Arcalis, E., Marcel, S., Altmann, F., Ribeiro-Pedro, M., Rodriguez, J., et al. (2008). Functional specialization of *Medicago truncatula* leaves and seeds does not affect the subcellular localization of a recombinant protein. *Planta* 227, 649–658. doi: 10.1007/s00425-007-0647-3
- Arcalis, E., Stadlmann, J., Marcel, S., Drakakaki, G., Winter, V., Rodriguez, J., et al. (2010). The changing fate of a secretory glycoprotein in developing maize endosperm. *Plant Physiol.* 153, 693–702. doi: 10.1104/pp.109.152363
- Arcalis, E., Stadlmann, J., Rademacher, T., Marcel, S., Sack, M., Altmann, F., et al. (2013). Plant species and organ influence the structure and subcellular localization of recombinant glycoproteins. *Plant Mol. Biol.* 83, 105–117. doi: 10.1007/s11103-013-0049-9
- Castilho, A., and Steinkellner, H. (2012). Glyco-engineering in plants to produce human-like N-glycan structures. *Biotechnol. J.* 7, 1088–1098. doi: 10.1002/biot.201200032
- Conley, A. J., Joensuu, J. J., Menassa, R., and Brandle, J. E. (2009). Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* 7:48. doi: 10.1186/1741-7007-7-48
- De Buck, S., Nolf, J., De Meyer, T., Virdi, V., De Wilde, K., Van Lerberghe, E., et al. (2013). Fusion of an Fc chain to a VHH boosts the accumulation levels in *Arabidopsis* seeds. *Plant Biotechnol. J.* 11, 1006–1016. doi: 10.1111/pbi.12094
- De Jaeger, G., Scheffer, S., Jacobs, A., Zambre, M., Zobell, O., Goossens, A., et al. (2002). Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. *Nat. Biotechnol.* 20, 1265–1268. doi: 10.1038/nbt755
- De Muynck, B., Navarre, C., and Boutry, M. (2010). Production of antibodies in plants: status after twenty years. *Plant Biotechnol. J.* 8, 529–563. doi: 10.1111/j.1467-7652.2009.00494.x
- De Wilde, C., De Neve, M., De Rycke, R., Bruyns, A. M., De Jaeger, G., Van Montagu, M., et al. (1996). Intact antigen-binding MAK33 antibody and Fab fragment accumulate in intercellular spaces of *Arabidopsis thaliana*. *Plant Sci.* 114, 233–241. doi: 10.1016/0168-9452(96)04331-2
- De Wilde, K., De Buck, S., Vanneste, K., and Depicker, A. (2013). Recombinant antibody production in *Arabidopsis* seeds triggers an unfolded protein response. *Plant Physiol.* 161, 1021–1033. doi: 10.1104/pp.112.209718
- Downing, W. L., Galpin, J. D., Clemens, S., Lauzon, S. M., Samuels, A. L., Pidkowich, M. S., et al. (2006). Synthesis of enzymatically active human alpha-L-iduronidase in *Arabidopsis* cgl (complex glycan-deficient) seeds. *Plant Biotechnol. J.* 4, 169–181. doi: 10.1111/j.1467-7652.2005.00166.x
- Drakakaki, G., Marcel, S., Arcalis, E., Altmann, F., Gonzalez-Melendi, P., Fischer, R., et al. (2006). The intracellular fate of a recombinant protein is tissue dependent. *Plant Physiol.* 141, 578–586. doi: 10.1104/pp.106.076661
- Fiedler, U., Phillips, J., Artsen, O., and Conrad, U. (1997). Optimization of scFv antibody production in transgenic plants. *Immunotechnology* 3, 205–216. doi: 10.1016/S1380-2933(97)00014-6
- Gomord, V., Chamberlain, P., Jefferis, R., and Faye, L. (2005). Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol.* 23, 559–565. doi: 10.1016/j.tibtech.2005.09.003
- Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plisson, C., Michaud, D., et al. (2010). Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.* 8, 564–587. doi: 10.1111/j.1467-7652.2009.00497.x
- Gutiérrez, S. P., Saberianfar, R., Kohalmi, S. E., and Menassa, R. (2013). Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. *BMC Biotechnol.* 13:40. doi: 10.1186/1472-6750-13-40
- Hara-Nishimura, I., Shimada, T., Hatano, K., Takeuchi, Y., and Nishimura, M. (1998). Transport of storage proteins to protein storage vacuoles is mediated by large precursor-accumulating vesicles. *Plant Cell* 10, 825–836. doi: 10.1105/tpc.10.5.825
- He, X., Galpin, J. D., Tropak, M. B., Mahuran, D., Haselhorst, T., Von Itzstein, M., et al. (2012a). Production of active human glucocerebrosidase in seeds of *Arabidopsis thaliana* complex-glycan-deficient (cgl) plants. *Glycobiology* 22, 492–503. doi: 10.1093/glycob/cwr157
- He, X., Haselhorst, T., Von Itzstein, M., Kolarich, D., Packer, N. H., and Kermode, A. R. (2012b). Influence of an ER-retention signal on the N-glycosylation of recombinant human α-L-iduronidase generated in seeds of *Arabidopsis*. *Plant Mol. Biol.* 79, 157–169. doi: 10.1007/s11103-012-9902-5
- Henquet, M., Eigenhuijsen, J., Hesselink, T., Spiegel, H., Schreuder, M., Van Duijn, E., et al. (2011). Characterization of the single-chain Fv-Fc antibody MBP10 produced in *Arabidopsis alg3* mutant seeds. *Transgenic Res.* 20, 1033–1042. doi: 10.1007/s11248-010-9475-5
- Henquet, M., Lehle, L., Schreuder, M., Rouwendal, G., Molthoff, J., Helsper, J., et al. (2008). Identification of the gene encoding the alpha1,3-mannosyltransferase (ALG3) in *Arabidopsis* and characterization of downstream N-glycan processing. *Plant Cell* 20, 1652–1664. doi: 10.1105/tpc.108.060731
- Jacobs, P. P., and Callewaert, N. (2009). N-glycosylation engineering of biopharmaceutical expression systems. *Curr. Mol. Med.* 9, 774–800. doi: 10.2174/156652409789105552
- Jolliffe, N. A., Craddock, C. P., and Frigerio, L. (2005). Pathways for protein transport to seed storage vacuoles. *Biochem. Soc. Trans.* 33, 1016–1018. doi: 10.1042/BST20051016

- Kermode, A. R. (2012). "Seed expression systems for molecular farming," in *Molecular Farming in Plants: Recent Advances and Future Prospects*, eds A. Wang and S. Ma (Dordrecht; Heidelberg; London; New York, NY: Springer), 89–123.
- Loos, A., Van Droogenbroeck, B., Hillmer, S., Grass, J., Kunert, R., Cao, J., et al. (2011a). Production of monoclonal antibodies with a controlled N-glycosylation pattern in seeds of *Arabidopsis thaliana*. *Plant Biotechnol. J.* 9, 179–192. doi: 10.1111/j.1467-7652.2010.00540.x
- Loos, A., Van Droogenbroeck, B., Hillmer, S., Grass, J., Pabst, M., Castilho, A., et al. (2011b). Expression of antibody fragments with a controlled N-glycosylation pattern and induction of endoplasmic reticulum-derived vesicles in seeds of *Arabidopsis*. *Plant Physiol.* 155, 2036–2048. doi: 10.1104/pp.110.171330
- Morandini, F., Avesani, L., Bortesi, L., Van Droogenbroeck, B., De Wilde, K., Arcalis, E., et al. (2011). Non-food/feed seeds as biofactories for the high-yield production of recombinant pharmaceuticals. *Plant Biotechnol. J.* 9, 911–921. doi: 10.1111/j.1467-7652.2011.00605.x
- Napier, R. M., Fowke, L. C., Hawes, C., Lewis, M., and Pelham, H. R. B. (1992). Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. *J. Cell Sci.* 102, 261–271.
- Okamoto, T., Shimada, T., Hara-Nishimura, I., Nishimura, M., and Minamikawa, T. (2003). C-terminal KDEL sequence of a KDEL-tailed cysteine proteinase (sulfhydryl-endopeptidase) is involved in formation of KDEL vesicle and in efficient vacuolar transport of sulfhydryl-endopeptidase. *Plant Physiol.* 132, 1892–1900. doi: 10.1104/pp.103.021147
- Otegui, M. S., Herder, R., Schulze, J., Jung, R., and Staehelin, L. A. (2006). The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *Plant Cell* 18, 2567–2581. doi: 10.1105/tpc.106.040931
- Peeters, K., De Wilde, C., and Depicker, A. (2001). Highly efficient targeting and accumulation of a Fab fragment within the secretory pathway and apoplast of *Arabidopsis thaliana*. *Eur. J. Biochem.* 268, 4251–4260. doi: 10.1046/j.1432-1327.2001.02340.x
- Petrucelli, S., Otegui, M. S., Lareu, F., Tran Dinh, O., Fitchette, A.-C., Circosta, A., et al. (2006). A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol. J.* 4, 511–527. doi: 10.1111/j.1467-7652.2006.00200.x
- Piron, R., De Koker, S., De Paepe, A., Goossens, J., Grooten, J., Nauwynck, H., et al. (2014). Boosting in planta production of antigens derived from the Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) and subsequent evaluation of their immunogenicity. *PLoS ONE* 9:e91386. doi: 10.1371/journal.pone.0091386
- Robinson, D. G., Oliviusson, P., and Hinz, G. (2005). Protein sorting to the storage vacuoles of plants: a critical appraisal. *Traffic* 6, 615–625. doi: 10.1111/j.1600-0854.2005.00303.x
- Schahs, M., Strasser, R., Stadlmann, J., Kunert, R., Rademacher, T., and Steinkellner, H. (2007). Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol. J.* 5, 657–663. doi: 10.1111/j.1467-7652.2007.00273.x
- Stoger, E., Ma, J. K.-C., Fischer, R., and Christou, P. (2005). Sowing the seeds of success: pharmaceutical proteins from plants. *Curr. Opin. Biotechnol.* 16, 167–173. doi: 10.1016/j.copbio.2005.01.005
- Strasser, R., Altmann, F., Mach, L., Glossl, J., and Steinkellner, H. (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta 1,2-linked xylose and core alpha 1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4
- Toyooka, K., Okamoto, T., and Minamikawa, T. (2000). Mass transport of proform of a KDEL-tailed cysteine proteinase (SH-EP) to protein storage vacuoles by endoplasmic reticulum-derived vesicle is involved in protein mobilization in germinating seeds. *J. Cell Biol.* 148, 453–464. doi: 10.1083/jcb.148.3.453
- Triguero, A., Cabrera, G., Cremata, J. A., Yuen, C.-T., Wheeler, J., and Ramírez, N. I. (2005). Plant-derived mouse IgG monoclonal antibody fused to KDEL endoplasmic reticulum-retention signal is N-glycosylated homogeneously throughout the plant with mostly high-mannose-type N-glycans. *Plant Biotechnol. J.* 3, 449–457. doi: 10.1111/j.1467-7652.2005.00137.x
- Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F., Colanesi, S., Hillmer, S., et al. (2007). Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic *Arabidopsis* seeds. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1430–1435. doi: 10.1073/pnas.0609997104
- Viotti, C., Kruger, F., Krebs, M., Neubert, C., Fink, F., Lupanga, U., et al. (2013). The endoplasmic reticulum is the main membrane source for biogenesis of the lytic vacuole in *Arabidopsis*. *Plant Cell* 25, 3434–3449. doi: 10.1105/tpc.113.114827
- Virdi, V., Coddens, A., De Buck, S., Millet, S., Goddeeris, B. M., Cox, E., et al. (2013). Orally fed seeds producing designer IgAs protect weaned piglets against enterotoxigenic *Escherichia coli* infection. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11809–11814. doi: 10.1073/pnas.1301975110
- Vitale, A., and Hinz, G. (2005). Sorting of proteins to storage vacuoles: how many mechanisms? *Trends Plant Sci.* 10, 316–323. doi: 10.1016/j.tplants.2005.05.001
- Von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M. J. (1993). Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase-I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol.* 102, 1109–1118. doi: 10.1104/pp.102.4.1109
- Wang, J. Q., Tse, Y. C., Hinz, G., Robinson, D. G., and Jiang, L. W. (2012). Storage globulins pass through the Golgi apparatus and multivesicular bodies in the absence of dense vesicle formation during early stages of cotyledon development in mung bean. *J. Exp. Bot.* 63, 1367–1380. doi: 10.1093/jxb/err366

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 June 2014; **accepted:** 28 August 2014; **published online:** 15 September 2014.

Citation: De Meyer T and Depicker A (2014) Trafficking of endoplasmic reticulum-retained recombinant proteins is unpredictable in *Arabidopsis thaliana*. *Front. Plant Sci.* 5:473. doi: 10.3389/fpls.2014.00473

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 De Meyer and Depicker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Plant glyco-biotechnology on the way to synthetic biology

Andreas Loos and Herta Steinkellner*

Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria

Edited by:

Els J. M. Van Damme, Ghent University, Belgium

Reviewed by:

Beronda L. Montgomery, Michigan State University, USA

Björn Usadel, RWTH Aachen University, Germany

Thomas De Meyer, Vlaams Instituut voor Biotechnologie, Belgium

***Correspondence:**

Herta Steinkellner, Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria

e-mail: herta.steinkellner@boku.ac.at

Plants are increasingly being used for the production of recombinant proteins. One reason is that plants are highly amenable to glycan engineering processes and allow the production of therapeutic proteins with increased efficacies due to optimized glycosylation profiles. Removal and insertion of glycosylation reactions by knock-out/knock-down approaches and introduction of glycosylation enzymes have paved the way for the humanization of the plant glycosylation pathway. The insertion of heterologous enzymes at exactly the right stage of the existing glycosylation pathway has turned out to be of utmost importance. To enable such precise targeting chimeric enzymes have been constructed. In this short review we will exemplify the importance of correct targeting of glycosyltransferases, we will give an overview of the targeting mechanism of glycosyltransferases, describe chimeric enzymes used in plant *N*-glycosylation engineering and illustrate how plant glycoengineering builds on the tools offered by synthetic biology to construct such chimeric enzymes.

Keywords: glycoengineering, plant, glycosyltransferase, CTS, sub-Golgi targeting

INTRODUCTION

The medicinal use of proteins and blood products has a long history. Already in the 15th century ailing pope Innocent VIII was reportedly infused with blood from three healthy boys to give him back vigor and strength, however, with a fatal outcome for all four of them (Rivera et al., 2005). The first successful blood transfusion was made in 1665 between dogs (Felts, 2000) and it took over 150 years for the first successful transfusion between humans (Blundell, 1818). Proteins purified from animal or human tissues (growth hormones, insulin, clotting factors, or other blood components) have been used for medicinal purposes since the beginning of the 20th century (e.g., Eibl, 2008; Blizzard, 2012) and with the advent of recombinant protein production possibilities, many of those proteins are now produced recombinantly. The market for pharmaceutical proteins is assumed to currently amount to roughly 150–200 billion US\$, and develops strongly with growth rates of ~10% and more (Walsh, 2010; Elvin et al., 2013; Aggarwal, 2014). Special drivers of this growth are antibodies and antibody related products, but also other types of proteins are selling well, like insulin, vaccines, erythropoietin, etc. (Aggarwal, 2014).

A current major concern in producing biopharmaceuticals is a special type of post-translation modification (PTM), namely *N*-glycosylation. This PTM is found on a large proportion of pharmaceutically relevant proteins (Walsh, 2010) and can influence protein characteristics like folding and assembly, solubility and charge, serum half-life, functionality, etc. (e.g., Varki, 1993; Roth et al., 2010; Solá and Griebel, 2010). As different cell types attach different glycans, the characteristics of the *N*-glycosylated protein can be strongly affected by the expression host – a fact that should be considered carefully when choosing the production system. For example, bacteria generally do not glycosylate proteins and yeasts attach larger glycan structures than mammals. Insect cells decorate proteins with paucimannosidic *N*-glycans which are

normally not present in humans. Plants produce complex type *N*-glycans similar to humans, however, certain non-mammalian epitopes are attached and more complex human-type glycosylation cannot be produced (for reviews on typical *N*-glycosylation patterns and glycoengineering of different expression hosts see, e.g., Jacobs and Callewaert, 2009; Loos and Steinkellner, 2012). Another concern is glycan microheterogeneity, i.e., attachment of different *N*-glycans to the same *N*-glycosylation site, as homogeneously glycosylated products are required by the regulatory authorities. Thus, research has focused on modifying the glycosylation characteristics of a variety of expression systems to allow homogeneous, human-type *N*-glycosylation (Umana et al., 1999; Yamane-Ohnuki et al., 2004; Schuster et al., 2005; Cox et al., 2006; Li et al., 2006; Strasser et al., 2009; Pandhal and Wright, 2010; Meuris et al., 2014) and resulted in the production of proteins carrying modified glycans and often showing improved *in vivo* functions.

Plants have proven their capability regarding production speed, ease of scale up and to meet quality standards demanded by regulatory agencies for clinical applications (Gleba et al., 2014; Stoger et al., 2014). Also governmental agencies like the Defense Advanced Research Projects Agency [DARPA] (2012) have recognized the advantages of this technology for the quick manufacturing of vaccines, difficult to produce biopharmaceuticals, etc. This has led to massive investments in research, production facilities complying with current quality standards (Defense Advanced Research Projects Agency [DARPA], 2012; www.federalgrants.com, 2012; Stoger et al., 2014) and the first products on the market. Glucocerebrosidase, an enzyme to treat Gaucher's disease, has been approved by the FDA in 2013 as the first plant-produced, parenterally applied biopharmaceutical (Zimran et al., 2011; van Dussen et al., 2013). Additionally, several plant-made pharmaceuticals have received approval for clinical trials and other plant-produced products are

already marketed as research/diagnostic reagent, medical device, cosmetic product etc. (recently reviewed by Gleba et al., 2014; Goodman, 2014; Stoger et al., 2014). Many of these proteins are glycosylated.

In this review we will discuss the approaches taken to engineer the *N*-glycosylation pathway in *Nicotiana benthamiana* and put a strong focus on recently developed and applied semi-synthetic strategies using chimeric glycosyltransferases.

PLANT GLYCOSYLATION

In plants as in other eukaryotes, the endoplasmic reticulum (ER) and the Golgi apparatus play the central role in protein glycosylation and contain the majority of glycan modifying enzymes (reviewed by, e.g., Helenius and Aebi, 2001). While the ER and its glycan processing repertoire are largely conserved between phyla (and kingdoms), morphology and function of the Golgi differ to some extent (Loos and Steinkellner, 2012; Aebi, 2013). For example, a main function of the plant but not the mammalian Golgi is to provide large amounts of polysaccharides, a fundamental component of the cell wall (Oikawa et al., 2013). Early *N*-glycosylation steps that take place in ER and *cis*-Golgi are virtually identical in higher eukaryotes, while further processing differs (recently reviewed by, e.g., Loos and Steinkellner, 2012; Bosch et al., 2013). This is mainly due to a drastically reduced repertoire of glycosylation enzymes in plants, where a small number of Golgi-located *N*-glycan processing enzymes gives rise to typically two different glycan structures (Castilho and Steinkellner, 2012). By comparison, over 2000 different *N*-glycans have been described on mammalian proteins which arise from several 100 enzymes in the secretory pathway (Campbell and Yarema, 2005; Ohtsubo and Marth, 2006; Varki, 2006). Notwithstanding these differences, the Golgi of higher eukaryotes shares a remarkably high degree of homology, especially with respect to organization, proteome, and *N*-glycosylation capabilities.

Plant proteins typically carry two major *N*-glycans, complex GnGnXF and paucimannosidic MMXF (Strasser et al., 2004a, 2007a; for glycan nomenclature see http://www.proglycan.com/upload/nomen_2007.pdf). These two glycans contain core α 1,3-fucose and β 1,2-xylose which are plant-specific glyco-epitopes. They are not produced by mammalian cells and up to 50% of humans have been shown to carry substantial amounts of antibodies directed against these epitopes in their blood (Bardor et al., 2003). The abundantly present paucimannosidic structures (MMXF, truncated glycans lacking terminal GlcNAc residues; Dirnberger et al., 2001; Strasser et al., 2007a; Liebminger et al., 2011) are also a plant peculiarity, otherwise only found in insect cells (Altmann et al., 1999, 2001). In some cases plant proteins carry so-called Lewis A epitopes, terminally β 1,3-galactosylated and α 1,4-fucosylated structures (Fitchette-Laine et al., 1997; Strasser et al., 2007b). The abundance of this epitope differs strongly between species (Fitchette et al., 1999; Wilson et al., 2001) and organs (Strasser et al., 2007b), but seems low in *Arabidopsis* and *Nicotiana* (Fitchette et al., 1999; Strasser et al., 2007b, 2008; Matsuo and Matsumura, 2011). Noteworthy is also the absence of core α 1,6-fucosylation in plants, a glycan residue present on the vast majority of proteins produced in mammalian

cells. Removal of this residue from immunoglobulin G (IgG) glycans increases functional activities of antibodies due to a higher affinity to the antibody-dependent cell-mediated cytotoxicity (ADCC) inducing IgG receptor Fc γ RIIIa (Shinkawa et al., 2003; Jefferis, 2009).

The limited glycosylation capacity of plants has turned out to be an advantage for the generation of proteins that need homogeneous glycosylation. For example, IgG antibodies produced in plants carry usually 1–2 different glycan structures (mainly GnGnXF) while the same antibodies produced in Chinese hamster ovary (CHO) cells bear 5–7 structures (Strasser et al., 2008, 2009). For some applications, like testing of functional activities, and according to the demands from regulatory agencies, homogeneous glycosylation is required.

Plants display a remarkable tolerance toward the manipulation of their intrinsic glycan biosynthetic pathways. Elimination of complex glycans, knock-out of plant-specific xylosyl- and fucosyltransferases (XT and FT) or reduction/overproduction of the Lewis A epitope did not lead to any obvious phenotype in *Arabidopsis thaliana* under standard growth conditions (Von Schaewen et al., 1993; Strasser et al., 2004b, 2007b). Also *Lemna minor* and *N. benthamiana*, one of the major plant-based protein production platforms, tolerate a variety of glycoengineering steps without obvious phenotypes or impact on development (Cox et al., 2006; Strasser et al., 2008; Nagels et al., 2011). Only few cases of sensitive reactions to glycosylation changes have been described (Fanata et al., 2013). This general tolerance for glycoengineering was a prerequisite for humanizing the plant *N*-glycosylation pathway. A combination of knock-out/knock-down and knock-in approaches together with transient expression techniques has allowed the removal of potentially immunogenic residues, and the addition of new, human-type glycostructures. Modular, semi-synthetic constructs assembled on multi-gene vectors enable the efficient manipulation of the glycosylation pathway. These glycoengineering strategies are addressed below.

PLANT GLYCOENGINEERING

Engineering of plant glycans toward human structures requires two main types of modification: (i) plant-specific reactions have to be eliminated and (ii) reactions taking place in humans but not in plants have to be introduced. Reducing the unwanted plant-specific modifications, i.e., β 1,2-xylosylation and core α 1,3-fucosylation, has initially been achieved by targeting the recombinant protein to the ER or co-overexpressing glycosylation enzymes competing for the same substrate (e.g., Palpac et al., 1999; Bakker et al., 2006; Frey et al., 2009; Vézina et al., 2009; Karg et al., 2010; see below). However, as these approaches interfere with the execution of endogenous glycosylation processes and cause the attachment of oligomannosidic or incompletely processed and aberrant glycans they are only of limited use. RNAi approaches targeting the transcript of the unwanted glycosyltransferases or complete knock-outs by T-DNA insertion have proven more successful (Koprivova et al., 2004; Strasser et al., 2004b, 2008; Cox et al., 2006; Sourrouille et al., 2008; Shin et al., 2011; Parsons et al., 2012). Importantly, such plants produce human-type GnGn glycans, which serve as an

acceptor substrate for further mammalian modifications and were thus important milestones in the engineering of the plant *N*-glycosylation pathway toward the production of human-type structures.

Consequently, work over the past decade on the controlled expression of mammalian glycosyltransferases has established plant-based systems that synthesize a series of defined human-type glycan structures (Castilho et al., 2012; recently reviewed by Bosch et al., 2013). Recent studies demonstrate how even entire glycosylation-associated biosynthetic pathways can be introduced. Plants do not have the machinery to synthesize the sugar nucleotide precursor CMP-sialic acid (CMP-*N*-acetylneurameric acid) necessary for sialylation. The simultaneous overexpression of six mammalian genes enabled the *in planta* generation of activated sialic acid, the transfer of the activated sugar nucleotide to the Golgi, the production of terminally galactosylated glycans and the transfer of sialic acid to these terminal galactoses (see Figures 1A–D; Castilho et al., 2008, 2010).

In planta sialylation of glycans thereby highlights some of the reasons why simple overexpression of a mammalian glycosyltransferase in plants has not always proven successful in generating human-type glycans: acceptor as well as donor substrates need to be present. For example, when Wee et al. (1998) expressed the human α 2,6-sialyltransferase in *Arabidopsis*, activity of the enzyme could only be shown after applying donor and acceptor substrates in *trans*, as plants lack both. The achievements by Palacpac et al. (1999), Bakker et al. (2006) and others pointed out one more challenge, namely how delicate the glycosylation system is – coexpression of the human β 1,4-galactosyltransferase (β 1,4-GalT) had led to the production of galactosylated, but also of unusual, hybrid-type glycans. The latter was due to activity of the galactosyltransferase at a suboptimal stage of the glycosylation pathway and interference with the endogenous glycosylation reactions (described in detail below). Similar findings were reported upon overexpression of *N*-acetylglucosaminyltransferase (GnT) III (Rouwendal et al.,

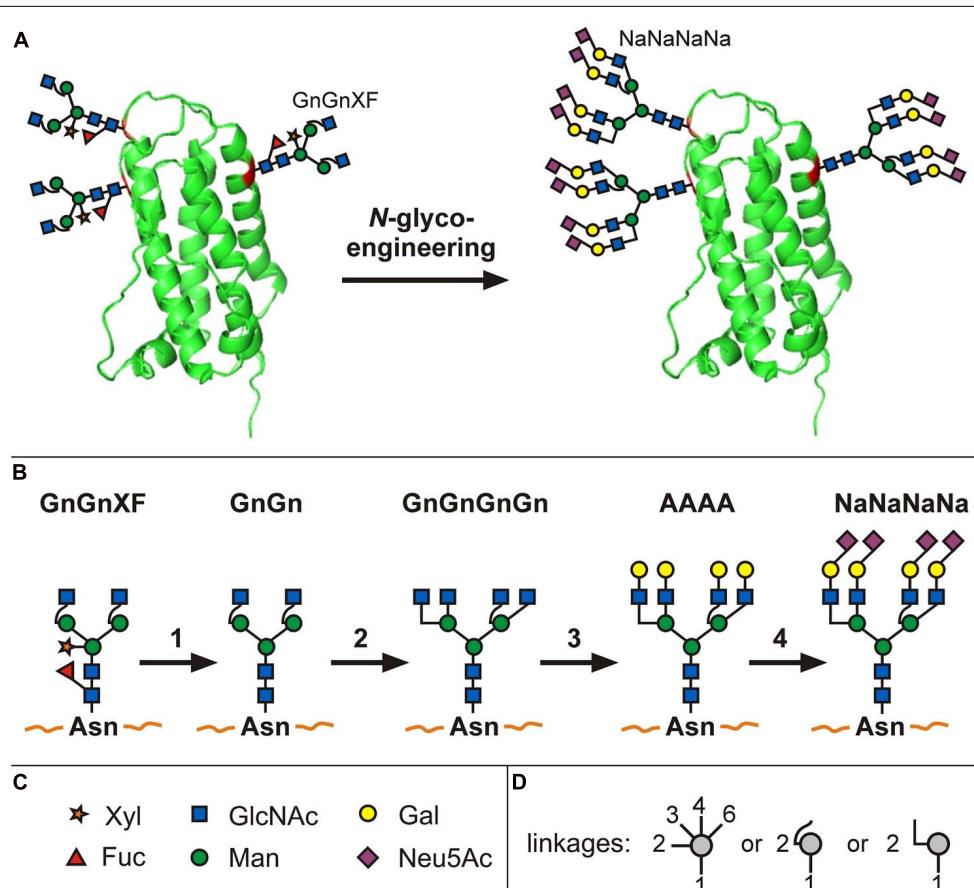


FIGURE 1 | *N*-glycoengineering in plants to produce tetra-sialylated proteins. Schematic representation of an extensively glycoengineered plant-produced glycoprotein (erythropoietin; **A**). It represents the front of plant glycan engineering and illustrates, in short, the transformation of glycans present in *Nicotiana benthamiana* wild-type plants (GnGnXF) to finally obtain a glycosylation profile present on human serum EPO (NaNaNana). For such intensive *N*-glycoengineering, a series of individual steps are necessary (**B**; 1): knock-out or knock-down of plant-specific β 1,2-XT and core α 1,3-FT, (2)

introduction of GnTIV and GnTV responsible for branching, (3) introduction of β 1,4-GalT, (4) introduction of sialyltransferase in combination with the biosynthetic pathway to produce activated sialic acid (not shown). Symbols for monosaccharides are given in (**C**), symbols depicting which monosaccharide atoms are involved in the linkage are given in (**D**). For detailed description see publications by Castilho et al. (2010, 2011, 2012, 2013). A more detailed explanation of *N*-glycan nomenclature and graphical illustrations can be found at http://www.proglycan.com/upload/nomen_2007.pdf

2007; Frey et al., 2009; Karg et al., 2010; Castilho et al., 2011; detailed description see below). These examples show that the final glycosylation profile of a co-expression approach depends on various factors, including the availability of (i) acceptor glycan and (ii) donor substrate as well as (iii) the correct subcellular targeting of the recombinant glycosyltransferase in order to avoid interference with the endogenous glycosylation machinery.

TARGETING MECHANISM OF GLYCOSYLTTRANSFERASES

The glycosylation reactions within the Golgi are carried out in a sequential, stepwise manner, and one reaction can be the prerequisite for another one – or inhibit it. Therefore, the ordered sequential arrangement of enzymatic activities, i.e., the correct subcellular localization of the involved enzymes is of utmost importance. This tight regulation has consequences for the expression and targeting of heterologous glycosylation enzymes, as they need to fit precisely into the existing pathway. Fine-tuning the subcellular localization of heterologously expressed glycosylation enzymes requires vast knowledge of the underlying targeting mechanisms.

All known Golgi-resident *N*-glycosyltransferases are type II transmembrane proteins (reviewed by, e.g., Schoberer and Strasser, 2011). Their N-terminus is exposed to the cytoplasm, followed by a transmembrane domain, a stem, and the catalytic domain (see **Figure 2**). The cytoplasmic part, transmembrane domain, and stem are referred to as CTS region and are responsible

for targeting the enzyme to the correct compartment (Essl et al., 1999), and even sub-compartment. This was shown by different Golgi-localized glycosyltransferases not present within the same sub-compartment (Saint-Jore-Dupas et al., 2006; Schoberer et al., 2009, 2010).

Research activities to elucidate the targeting mechanism(s) of type II transmembrane proteins revealed a basic conservation of processes between plants and mammals (Schoberer et al., 2010). However, the question on how the fine-tuning of targeting works has not been answered definitively, but several factors have been identified. For example, the cytoplasmic tail influences ER export of the enzyme (Schoberer et al., 2009) and can change the final destination of a protein (Jiang and Rogers, 1998). In mammalian cells it has been shown to relocate the catalytic domain from one to another Golgi subcompartment (Uliana et al., 2006). This might be due to formation of homo- or heterodimers of the enzymes or interaction with other proteins (Schoberer et al., 2013). Also the length of the transmembrane domain might influence targeting (Pagny et al., 2003), as the membrane thickness of the Golgi changes from *cis* to *trans*. Finally, the composition of the lipid bilayer surrounding the enzyme might influence targeting, too (reviewed by Schoberer and Strasser, 2011). This makes the CTS regions of glycosyltransferases key regulators for precise sub-Golgi targeting. Theoretically, CTS domains from any eukaryotic organism may serve as a suitable targeting domain. Recent genome sequencing projects provided an abundance of such sequences that may be used for targeted sub-Golgi localization (Ohtsubo and Marth, 2006; Varki, 2006). Nevertheless, as the molecular mechanisms that lie behind this fine-tuning of targeting are not fully understood, it is not entirely predictable how CTS sequences actually perform when fused to the catalytic domain of another glycosyltransferase and expressed in a foreign cell. Thus, experimental testing is required. Also the prediction of the exact size of the individual glycosyltransferase-domains (cytosolic tail, transmembrane domain, stem, catalytic domain) is difficult and even though bioinformatics prediction technology has improved in recent years, different algorithms can lead to different results. Therefore, the reliability of the identification of the separate domains should be considered carefully.

CHIMERIC GLYCOSYLTTRANSFERASES USED FOR IN PLANTA GLYCOENGINEERING

First steps to shift the plant *N*-glycosylation pattern from plant-specific *N*-glycans (i.e., GnGnXF, MMXF, etc.; see **Figure 3A**) toward the production of human-like structures were based on full-length mammalian glycosyltransferases. Expressing a human β 1,4-GalT in tobacco plants (Bakker et al., 2001) and tobacco BY2 cells (Palacpac et al., 1999) in fact resulted in galactosylated glycans, however, other oligosaccharides were produced as well. In BY2 cells, unusual and incompletely processed glycans lacking xylose and fucose (e.g., Man5A) were abundant (see **Figure 3B**), indicating interference of the heterologously expressed GalT with endogenous glycan processing enzymes like mannosidase II, GnTII and plant-specific β 1,2-xylosyltransferase and α 1,3-fucosyltransferase (XT and FT). In tobacco plants, xylosylated, and fucosylated GnGnXF remained the main oligosaccharide (see

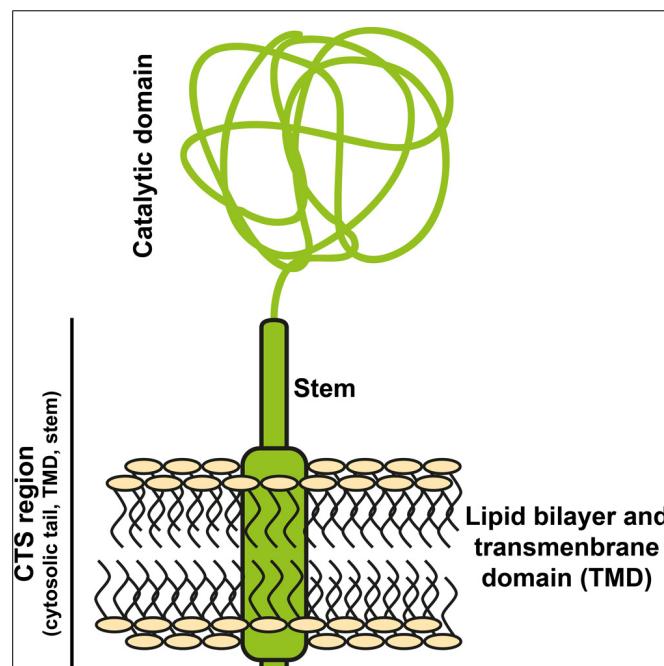


FIGURE 2 | Structure of *N*-glycosyltransferases. Golgi-located glycosyltransferases are type II transmembrane proteins. Their localization within the Golgi depends on the N-terminal CTS region, consisting of the cytosolic tail, the transmembrane domain and a stem. The C-terminal catalytic domain is directed to the Golgi lumen.

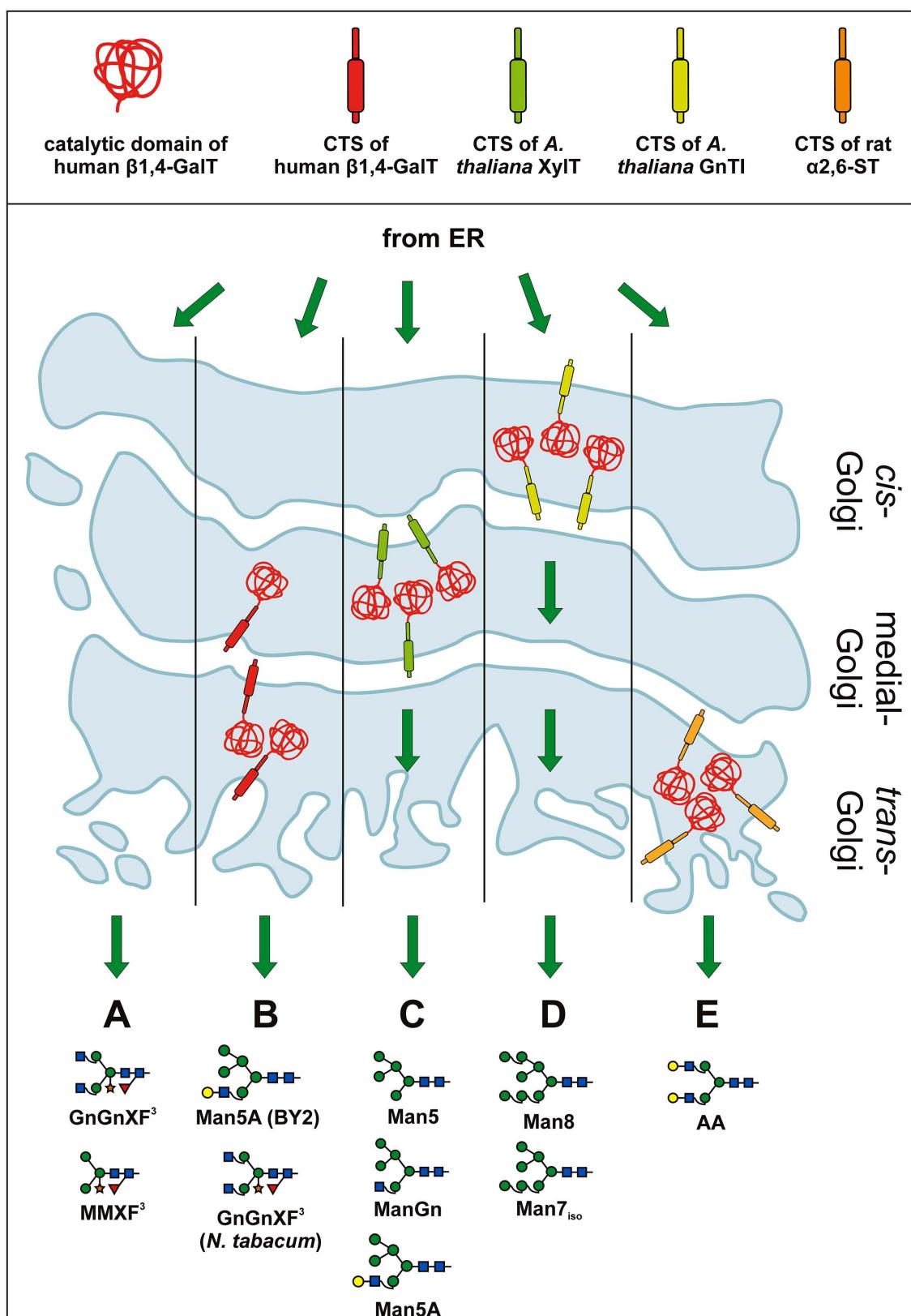


FIGURE 3 | Continued

FIGURE 3 | Continued**Expression of β 1,4-GaIT and chimeric versions thereof in plants.**

Schematic presentation of various β 1,4-GaIT constructs expressed in plants and the consequences on the glycosylation profile of total and recombinantly expressed proteins. In the top panel β 1,4-GaIT catalytic domain and various CTS regions are illustrated in different colors. The color code is used to better visualize various chimeric fusion constructs. The bottom panel shows a Golgi stack and the hypothetical localization of different β 1,4-GaIT constructs. Green arrows indicate cargo flow from ER through the Golgi. Major glycan structures produced under the given conditions are shown. **(A)** Major glycoforms detected in wild-type plants (without the expression of β 1,4-GaIT) are complex N-glycans carrying xylose and fucose (i.e., GnGnXF, MMXF, etc.; e.g., Bakker et al., 2001). **(B)** Expression of full-length human β 1,4-GaIT in BY2 tobacco cells (Palacpac et al., 1999) and tobacco plants (Bakker et al., 2001) led to different results. In BY2 cells, mainly galactosylated, hybrid-type glycans (like Man5A) as well as oligomannosidic glycans were found (Palacpac et al., 1999). In tobacco plants (Bakker et al., 2001) GnGnXF remained the major glycoform and only small amounts of galactosylated glycans were found. These results indicate that β 1,4-GaIT acted in BY2 cells at an earlier stage of the glycosylation pathway than in tobacco plants, leading to interference with endogenous glycosylation reactions in cells, but not in plants. **(C)** Major glycoforms detected upon expression of a chimeric GalT, that carries the CTS region of *A. thaliana* β 1,2-xylosyltransferase (indicated in pale green) and targets the enzyme to a medial stage of the glycan processing pathway: Man5, ManGn, Man5A. A drastically reduced amount of xylosylated and fucosylated glycans was detected (Bakker et al., 2006). The results point to an early activity of the chimeric β 1,4-GaIT, most probably in medial Golgi stacks. **(D)** Targeting the GalT to an even earlier compartment by fusing it to the CTS of the *cis*-Golgi acting GnTI (indicated in yellow; Vézina et al., 2009) induced the production of nearly exclusively oligomannosidic structures. Only minute amounts of galactosylated, hybrid Man5A were present. **(E)** Upon expression in a XT/FT knock-down plant line of a chimeric GalT carrying the late-Golgi CTS of rat α 2,6-sialyltransferase (indicated in orange) proteins carrying mainly galactosylated glycans (e.g., AA) were generated (Strasser et al., 2009). These results indicate that the ST-GaIT fusion is indeed located in a late Golgi stack where final N-glycan processing takes place.

Figure 3B) and only minor amounts of galactosylated oligosaccharides like AAXF or GnAXF were found. This points to activity of the GalT at a later stage in tobacco plants, after completion of the endogenous glycosylation reactions. At this later stage, interference with the endogenous glycosylation reactions did not take place and therefore the amount of β 1,2-xylose and α 1,3-fucose – residues unwanted on proteins needed for human applications – remained basically unchanged. In order to transfer the down-regulation effect observed in BY2 cells to tobacco plants and reduce/eliminate the two plant-specific glycan residues (β 1,2-xylose and α 1,3-fucose), a chimeric version of the human GalT was constructed that contained the CTS region of the *A. thaliana* β 1,2-xylosyltransferase (Bakker et al., 2006). The intention was to generate a chimeric enzyme that acts in the medial Golgi, simultaneously or prior to the endogenously present β 1,2-xylosyltransferase and core α 1,3-fucosyltransferase (XT and FT). As β 1,4-galactosylated proteins are no longer substrates for XT and FT (Staudacher et al., 1995; Kajiura et al., 2012), the expression in tobacco plants led to the intended, drastic decrease in plant-specific glycans. However, the early activity of the β 1,4-GaIT also led to the inhibition of other enzymes – like mannosidase II and GnTII – and thus to the generation of substantial amounts of unusual, incompletely processed glycans (oligomannosidic glycans, Man5A, etc.; see **Figure 3C**). Altogether, the glycan profile was similar to the profile of BY2 cells expressing the full-length

human β 1,4-GaIT (Palacpac et al., 1999). This indicates that the chimeric construct in tobacco plants and the full-length, human GalT in BY2 cells show activity at a comparable stage of the glycosylation pathway.

Targeting the human β 1,4-GaIT to an even earlier compartment (ER/*cis*-Golgi) by fusion with the CTS region of the *A. thaliana* GnTI further increased the amount of oligomannosidic glycans (Vézina et al., 2009; see **Figure 3D**). Galactosylated oligosaccharides were hardly found, indicating increased interference with the glycosylation machinery and the secretory pathway.

With the advent of XT/FT knock-down or knock-out lines (Koprieva et al., 2004; Cox et al., 2006; Schähs et al., 2007; Strasser et al., 2008; Shin et al., 2011) a more elegant way to prevent plant-specific glycosylation had been established and the aim of co-expressing human GalT in plants shifted from “interfering with endogenous reactions” to “generating homogeneously galactosylated, human-type glycans.” This was achieved by a rationally designed construct targeting the GalT to a late Golgi compartment. Fusions of the catalytic GalT domain to the CTS region of α 2,6-sialyltransferase, an enzyme acting in the final steps of the mammalian glycosylation pathway, indeed resulted in the generation of human-type, mono- and di-galactosylated glycans in XT/FT knock-down plants (Strasser et al., 2009; see **Figure 3E**). This was an important step for the *in planta* generation of proteins carrying fully human glycans.

This semi-synthetic approach was applied to GnTs to further explore the consequences of generating hybrid constructs carrying foreign CTS regions. One of the GnTs, β 1,4-mannosyl- β 1,4-N-acetylglucosaminyltransferase (GnTIII), catalyzes the formation of so-called bisected glycans (Carver et al., 1981; Narasimhan, 1982), a modification frequently found on human proteins but not present in plants (Rouwendal et al., 2007). Importantly, bisected – as well as the previously mentioned β 1,4-galactosylated glycans – cannot be modified with plant-specific xylose or fucose residues (Rouwendal et al., 2007). In order to produce such bisected glycans and thus prevent the addition of plant-specific glyco-epitopes, fully human GnTIII and a hybrid construct (the catalytic domain was fused to the CTS region of *A. thaliana* α -mannosidase II) were expressed in tobacco plants and BY-2 tobacco cells. The chimeric constructs led to a stronger decrease in plant-specific glyco-epitopes most probably due to targeting to an early/medial Golgi subcompartment (Rouwendal et al., 2007; Frey et al., 2009; Karg et al., 2010). However, targeting of GnTIII to an early compartment not only inhibited unwanted reactions but also led to the generation of non-standard, mainly hybrid-type glycans. Transiently expressing a series of GnTIII-constructs containing different CTS regions (*A. thaliana* Golgi mannosidase II, *A. thaliana* core α 1,3-fucosyltransferase, *A. thaliana* β 1,2-xylosyltransferase, and rat α 2,6-sialyltransferase) in XT/FT knock-down *N. benthamiana* plants identified late targeting sequences as preferential for the production of naturally occurring, bisected N-glycans (Castilho et al., 2011).

In a similar approach, enzymes responsible for branched glycans (i.e., tri- and tetraantennary glycans; human α 1,3-mannosyl- β 1,4-N-acetylglucosaminyltransferase IVa (GnTIV) and human

α 1,6-mannosyl- β 1,6-N-acetylglucosaminyltransferase V (GnTV)) were tested with different CTS regions (endogenous CTS region, rat α 2,6-sialyltransferase, *A. thaliana* core α 1,3-fucosyltransferase, *A. thaliana* β 1,2-xylosyltransferase; Castilho et al., 2011; Nagels et al., 2011, 2012a,b). The fusions with the medial Golgi-targeting CTS region of, e.g., FT allowed the generation of a high degree of branched glycans on co-expressed reporter proteins, whereas the endogenous CTS and the late Golgi CTS from α 2,6-sialyltransferase did not lead to substantial amounts of branched glycans.

Collectively these results demonstrate that it is not sufficient to “simply” introduce a foreign glycosylation enzyme into a plant to obtain a desired glycan structure. Instead, the successful production of proteins with human-type N-glycosylation in plants harbors a large number of challenges and requires knowledge of glycosylation pathways, enzyme specificities and related topics, like subcellular protein transport. Semi-synthetic approaches serve as useful tools to approach these challenges.

FURTHER CHALLENGES

In recent years, a variety of expression hosts were glyco-engineered (recently reviewed by, e.g., Jacobs and Callewaert, 2009; Loos and Steinkellner, 2012) and the first products with enhanced properties in animal studies have reached the clinic (Ratner, 2014). Plants, with their similar-to-human yet more simple N-glycosylation machinery and their amenability to glyco-engineering have been on the forefront of this development. Despite substantial achievements, the advantages of this system have been used only in a few *in vivo* studies (Bendandi et al., 2010; Forthal et al., 2010; Zeitlin et al., 2011, 2013; Hiatt et al., 2014). So far only one plant-produced product has reached the market, i.e., glucocerebrosidase to treat Gaucher’s disease. This protein carries terminal mannose residues, a glycosylation form that confers enhanced efficacies (Grabowski et al., 2014). Another plant-produced glyco-optimized protein drug has recently been used to treat patients: ZMapp, an experimental mAb cocktail against Ebola virus was given to several individuals during the ongoing Ebola epidemic (critically reviewed by Goodman, 2014). These antibodies had not yet gone through clinical studies but due to the dire predictions of Ebola virus infection and lack of other treatment options its application had been approved by regulatory authorities in several countries.

With the appearance of efficient, transient expression methods, the rapid, scalable and cost-effective production of high-value recombinant proteins became possible (Gleba et al., 2014). However, to realize the full potential of plant biotechnology, advanced, stably glyco-engineered plant strains in combination with semi-synthetic approaches will be needed. Versatile, modular expression vectors like MoClo (Weber et al., 2011) and GoldenBraid (Sarrion-Perdigones et al., 2013) allow efficient shuffling of domains and will certainly speed up the generation of constructs. Quick assembly of multi-gene vectors also simplifies the remodeling of glycosylation pathways as recently demonstrated (Schneider et al., 2014). Moreover, new technologies for genome editing, like CRISPR and TALENs (Lozano-Juste and Cutler, 2014) allow efficient elimination of genes and facilitate metabolic engineering and reprogramming of biosynthetic processes. These developments in combination with computer modeling and simulation

approaches that predict protein–glycan interactions will accelerate the development of drugs with optimized and even new functions. In sum, the currently available gene expression systems and the new tools offered by synthetic biology create an ideal environment for establishing a plant-based biomanufacturing platform that can compete with or even surpass current industry standards.

ACKNOWLEDGMENTS

We thank Jürgen Retzer for assistance with figure preparation. This work was supported by grants from the Austrian Research Promotion Agency (FFG) Laura Bassi Centres of Expertise (Number 822757, to Herta Steinkellner) and from the Austrian Science Fund (FWF; Grant Number L575-B13, to Herta Steinkellner).

REFERENCES

- Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta* 1833, 2430–2437. doi: 10.1016/j.bbamcr.2013.04.001
- Aggarwal, R. S. (2014). What’s fueling the biotech engine-2012 to 2013. *Nat. Biotechnol.* 32, 32–39. doi: 10.1038/nbt.2794
- Altmann, F., Fabini, G., Ahorn, H., and Wilson, I. B. (2001). Genetic model organisms in the study of N-glycans. *Biochimie* 83, 703–712. doi: 10.1016/S0300-9084(01)01297-4
- Altmann, F., Staudacher, E., Wilson, I. B., and Marz, L. (1999). Insect cells as hosts for the expression of recombinant glycoproteins. *Glycoconj. J.* 16, 109–123. doi: 10.1023/A:1026488408951
- Bakker, H., Bardor, M., Molthoff, J. W., Gomord, V., Elbers, I., Stevens, L. H., et al. (2001). Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2899–2904. doi: 10.1073/pnas.03141998
- Bakker, H., Rouwendal, G. J., Karnoup, A. S., Florack, D. E., Stoopen, G. M., Helsper, J. P., et al. (2006). An antibody produced in tobacco expressing a hybrid beta-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7577–7582. doi: 10.1073/pnas.0600879103
- Bardor, M., Faveeuw, C., Fitchette, A. C., Gilbert, D., Galas, L., Trottein, F., et al. (2003). Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. doi: 10.1093/glycob/cwg024
- Bendandi, M., Marillonnet, S., Kandzia, R., Thieme, F., Nickstadt, A., Herz, S., et al. (2010). Rapid, high-yield production in plants of individualized idio-type vaccines for non-Hodgkin’s lymphoma. *Ann. Oncol.* 21, 2420–2427. doi: 10.1093/annonc/mdq256
- Blizzard, R. M. (2012). History of growth hormone therapy. *Indian J. Pediatr.* 79, 87–91. doi: 10.1007/s12098-011-0609-4
- Blundell, J. (1818). Experiments on the transfusion of blood by the syringe. *Med. Chir. Trans.* 9, 56–92.
- Bosch, D., Castilho, A., Loos, A., Schots, A., and Steinkellner, H. (2013). N-Glycosylation of plant-produced recombinant proteins. *Curr. Pharm. Des.* 19, 5503–5512. doi: 10.2174/1381612811319310006
- Campbell, C. T., and Yarema, K. J. (2005). Large-scale approaches for glycobiology. *Genome Biol.* 6:236. doi: 10.1186/gb-2005-6-11-236
- Carver, J. P., Grey, A. A., Winnik, F. M., Hakimi, J., Ceccarini, C., and Atkinson, P. H. (1981). Determination of the structure of four glycopeptides from hen ovalbumin using 360-MHz proton magnetic resonance spectroscopy. *Biochemistry* 20, 6600–6606. doi: 10.1021/bi00526a013
- Castilho, A., Gattinger, P., Grass, J., Jez, J., Pabst, M., Altmann, F., et al. (2011). N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology* 21, 813–823. doi: 10.1093/glycob/cwr009
- Castilho, A., Neumann, L., Daskalova, S., Mason, H. S., Steinkellner, H., Altmann, F., et al. (2012). Engineering of sialylated mucin-type O-glycosylation in plants. *J. Biol. Chem.* 287, 36518–36526. doi: 10.1074/jbc.M112.402685
- Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., et al. (2013). Generation of biologically active multi-sialylated recombinant human EPOFc in plants. *PLoS ONE* 8:e54836. doi: 10.1371/journal.pone.0054836

- Castilho, A., Pabst, M., Leonard, R., Veit, C., Altmann, F., Mach, L., et al. (2008). Construction of a functional CMP-sialic acid biosynthesis pathway in *Arabidopsis*. *Plant Physiol.* 147, 331–339. doi: 10.1104/pp.108.117572
- Castilho, A., and Steinkellner, H. (2012). Glyco-engineering in plants to produce human-like N-glycan structures. *Biotechnol. J.* 7, 1088–1098. doi: 10.1002/biot.201200032
- Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gattinger, P., et al. (2010). In planta protein sialylation through overexpression of the respective mammalian pathway. *J. Biol. Chem.* 285, 15923–15930. doi: 10.1074/jbc.M109.088401
- Cox, K. M., Sterling, J. D., Regan, J. T., Gasdaska, J. R., Frantz, K. K., Peele, C. G., et al. (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* 24, 1591–1597. doi: 10.1038/nbt1260
- Defense Advanced Research Projects Agency [DARPA]. (2012). DARPA Produces Over 10 Million Doses of H1N1 Vaccine Candidate in One Month. Available at: <http://www.darpa.mil/NewsEvents/Releases/2012/07/25.aspx> (accessed July 7, 2014).
- Dirnberger, D., Steinkellner, H., Abdennabi, L., Remy, J. J., and Van De Wiel, D. (2001). Secretion of biologically active glycoforms of bovine follicle stimulating hormone in plants. *Eur. J. Biochem.* 268, 4570–4579. doi: 10.1046/j.1432-1327.2001.02384.x
- Eibl, M. M. (2008). History of immunoglobulin replacement. *Immunol. Allergy Clin. North Am.* 28, 737–764. doi: 10.1016/j.iac.2008.06.004
- Elvin, J. G., Couston, R. G., and Van Der Walle, C. F. (2013). Therapeutic antibodies: market considerations, disease targets and bioprocessing. *Int. J. Pharm.* 440, 83–98. doi: 10.1016/j.ijpharm.2011.12.039
- Essl, D., Dirnberger, D., Gomord, V., Strasser, R., Faye, L., Glossl, J., et al. (1999). The N-terminal 77 amino acids from tobacco N-acetylglucosaminyltransferase I are sufficient to retain a reporter protein in the Golgi apparatus of *Nicotiana benthamiana* cells. *FEBS Lett.* 453, 169–173. doi: 10.1016/S0014-5793(99)00712-7
- Fanata, W. I. D., Lee, K. H., Son, B. H., Yoo, J. Y., Harmoko, R., Ko, K. S., et al. (2013). N-glycan maturation is crucial for cytokinin-mediated development and cellulose synthesis in *Oryza sativa*. *Plant J.* 73, 966–979. doi: 10.1111/tpj.12087
- Felts, J. H. (2000). Richard lower: anatomist and physiologist. *Ann. Intern. Med.* 132, 420–423. doi: 10.7326/0003-4819-132-5-200003070-00023
- Fitchette, A. C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., et al. (1999). Biosynthesis and immunolocalization of Lewis a-containing N-glycans in the plant cell. *Plant Physiol.* 121, 333–344. doi: 10.1104/pp.121.2.333
- Fitchette-Laine, A. C., Gomord, V., Cabanes, M., Michalski, J. C., Saint Macary, M., Foucher, B., et al. (1997). N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J.* 12, 1411–1417. doi: 10.1046/j.1365-313x.1997.12061411.x
- Forthal, D. N., Gach, J. S., Landucci, G., Jez, J., Strasser, R., Kunert, R., et al. (2010). Fc-glycosylation influences Fcγ receptor binding and cell-mediated anti-HIV activity of monoclonal antibody 2G12. *J. Immunol.* 185, 6876–6882. doi: 10.4049/jimmunol.1002600
- Frey, A. D., Karg, S. R., and Kallio, P. T. (2009). Expression of rat beta(1,4)-N-acetylglucosaminyltransferase III in *Nicotiana tabacum* remodels the plant-specific N-glycosylation. *Plant Biotechnol. J.* 7, 33–48. doi: 10.1111/j.1467-7652.2008.00370.x
- Gleba, Y. Y., Tuse, D., and Giritch, A. (2014). Plant viral vectors for delivery by *Agrobacterium*. *Curr. Top. Microbiol. Immunol.* 375, 155–192. doi: 10.1007/82_2013_352
- Goodman, J. L. (2014). Studying “Secret Serums” – toward safe, effective ebola treatments. *N. Engl. J. Med.* 371, 1086–1089. doi: 10.1056/NEJMmp1409817
- Grabowski, G. A., Golembio, M., and Shaaltiel, Y. (2014). Taliglucerase alfa: an enzyme replacement therapy using plant cell expression technology. *Mol. Genet. Metab.* 112, 1–8. doi: 10.1016/j.ymgme.2014.02.011
- Helenius, A., and Aeby, M. (2001). Intracellular functions of N-linked glycans. *Science* 291, 2364–2369. doi: 10.1126/science.291.5512.2364
- Hiatt, A., Bohorova, N., Bohorov, O., Goodman, C., Kim, D., Pauly, M. H., et al. (2014). Glycan variants of a respiratory syncytial virus antibody with enhanced effector function and in vivo efficacy. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5992–5997. doi: 10.1073/pnas.1402458111
- Jacobs, P. P., and Callewaert, N. (2009). N-glycosylation engineering of pharmaceutical expression systems. *Curr. Mol. Med.* 9, 774–800. doi: 10.2174/156652409789105552
- Jefferis, R. (2009). Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev. Drug Discov.* 8, 226–234. doi: 10.1038/nrd2804
- Jiang, L. W., and Rogers, J. C. (1998). Integral membrane protein sorting to vacuoles in plant cells: evidence for two pathways. *J. Cell Biol.* 143, 1183–1199. doi: 10.1083/jcb.143.5.1183
- Kajiura, H., Okamoto, T., Misaki, R., Matsuura, Y., and Fujiyama, K. (2012). *Arabidopsis* beta1,2-xylosyltransferase: substrate specificity and participation in the plant-specific N-glycosylation pathway. *J. Biosci. Bioeng.* 113, 48–54. doi: 10.1016/j.jbiosc.2011.09.011
- Karg, S. R., Frey, A. D., and Kallio, P. T. (2010). Reduction of N-linked xylose and fucose by expression of rat beta1,4-N-acetylglucosaminyltransferase III in tobacco BY-2 cells depends on Golgi enzyme localization domain and genetic elements used for expression. *J. Biotechnol.* 146, 54–65. doi: 10.1016/j.biote.2010.01.005
- Koprivova, A., Stemmer, C., Altmann, F., Hoffmann, A., Kopriva, S., Gorr, G., et al. (2004). Targeted knockouts of *Phycomitrella* lacking plant-specific immunogenic N-glycans. *Plant Biotechnol. J.* 2, 517–523. doi: 10.1111/j.1467-7652.2004.00100.x
- Li, H., Sethuraman, N., Stadheim, T. A., Zha, D., Prinz, B., Ballew, N., et al. (2006). Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. *Nat. Biotechnol.* 24, 210–215. doi: 10.1038/nbt1178
- Liebminger, E., Veit, C., Pabst, M., Batoux, M., Zipfel, C., Altmann, F., et al. (2011). Beta-N-acetylhexosaminidases HEXO1 and HEXO3 are responsible for the formation of paucimannosidic N-glycans in *Arabidopsis thaliana*. *J. Biol. Chem.* 286, 10793–10802. doi: 10.1074/jbc.M110.178020
- Loos, A., and Steinkellner, H. (2012). IgG-Fc glycoengineering in non-mammalian expression hosts. *Arch. Biochem. Biophys.* 526, 167–173. doi: 10.1016/j.abb.2012.05.011
- Lozano-Juste, J., and Cutler, S. R. (2014). Plant genome engineering in full bloom. *Trends Plant Sci.* 19, 284–287. doi: 10.1016/j.tplants.2014.02.014
- Matsuo, K., and Matsumura, T. (2011). Deletion of fucose residues in plant N-glycans by repression of the GDP-mannose 4,6-dehydratase gene using virus-induced gene silencing and RNA interference. *Plant Biotechnol. J.* 9, 264–281. doi: 10.1111/j.1467-7652.2010.00553.x
- Meuris, L., Santens, F., Elson, G., Festjens, N., Boone, M., Dos Santos, A., et al. (2014). GlycoDelete engineering of mammalian cells simplifies N-glycosylation of recombinant proteins. *Nat. Biotechnol.* 32, 485–489. doi: 10.1038/nbt.2885
- Nagels, B., Van Damme, E. J., Callewaert, N., and Weterings, K. (2012a). Introduction of tri-antennary N-glycans in *Arabidopsis thaliana* plants. *Plant Sci.* 185–186, 161–168. doi: 10.1016/j.plantsci.2011.10.002
- Nagels, B., Van Damme, E. J., Callewaert, N., Zabeau, L., Tavernier, J., Delanghe, J. R., et al. (2012b). Biologically active, magnICON®-expressed EPO-Fc from stably transformed *Nicotiana benthamiana* plants presenting tetra-antennary N-glycan structures. *J. Biotechnol.* 160, 242–250. doi: 10.1016/j.biote.2012.03.003
- Nagels, B., Van Damme, E. J., Pabst, M., Callewaert, N., and Weterings, K. (2011). Production of complex multiantennary N-glycans in *Nicotiana benthamiana* plants. *Plant Physiol.* 155, 1103–1112. doi: 10.1104/pp.110.168773
- Narasimhan, S. (1982). Control of glycoprotein synthesis. UDP-GlcNAc:glycopeptide beta 4-N-acetylglucosaminyltransferase III, an enzyme in hen oviduct which adds GlcNAc in beta 1-4 linkage to the beta-linked mannose of the trimannosyl core of N-glycosyl oligosaccharides. *J. Biol. Chem.* 257, 10235–10242.
- Ohtsubo, K., and Marth, J. D. (2006). Glycosylation in cellular mechanisms of health and disease. *Cell* 126, 855–867. doi: 10.1016/j.cell.2006.08.019
- Oikawa, A., Lund, C. H., Sakuragi, Y., and Scheller, H. V. (2013). Golgi-localized enzyme complexes for plant cell wall biosynthesis. *Trends Plant Sci.* 18, 49–58. doi: 10.1016/j.tplants.2012.07.002
- Pagny, S., Bouissonne, F., Sarkar, M., Follet-Gueye, M. L., Driouch, A., Schachter, H., et al. (2003). Structural requirements for *Arabidopsis* beta1,2-xylosyltransferase activity and targeting to the Golgi. *Plant J.* 33, 189–203. doi: 10.1046/j.0960-7412.2002.01604.x
- Palacpac, N. Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., et al. (1999). Stable expression of human beta1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4692–4697. doi: 10.1073/pnas.96.8.4692
- Pandhal, J., and Wright, P. C. (2010). N-Linked glycoengineering for human therapeutic proteins in bacteria. *Biotechnol. Lett.* 32, 1189–1198. doi: 10.1007/s10529-010-0289-6
- Parsons, J., Altmann, F., Arrenberg, C. K., Koprivova, A., Beike, A. K., Stemmer, C., et al. (2012). Moss-based production of asialo-erythropoietin devoid of Lewis

- A and other plant-typical carbohydrate determinants. *Plant Biotechnol. J.* 10, 851–861. doi: 10.1111/j.1467-7652.2012.00704.x
- Ratner, M. (2014). Genentech's glyco-engineered antibody to succeed Rituxan. *Nat. Biotechnol.* 32, 6–7. doi: 10.1038/nbt0114-6b
- Rivera, A. M., Strauss, K. W., Van Zundert, A., and Mortier, E. (2005). The history of peripheral intravenous catheters: how little plastic tubes revolutionized medicine. *Acta Anaesthesiol. Belg.* 56, 271–282.
- Roth, J., Zuber, C., Park, S., Jang, I., Lee, Y., Kysela, K. G., et al. (2010). Protein N-glycosylation, protein folding, and protein quality control. *Mol. Cells* 30, 497–506. doi: 10.1007/s10059-010-0159-z
- Rouwendal, G. J., Wuhrer, M., Florack, D. E., Koeleman, C. A., Deelder, A. M., Bakker, H., et al. (2007). Efficient introduction of a bisecting GlcNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase III. *Glycobiology* 17, 334–344. doi: 10.1093/glycob/cwl078
- Saint-Jore-Dupas, C., Nebenfuhr, A., Boulaflous, A., Follet-Gueye, M. L., Plisson, C., Hawes, C., et al. (2006). Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. *Plant Cell* 18, 3182–3200. doi: 10.1105/tpc.105.036400
- Sarrion-Perdigones, A., Vazquez-Vilar, M., Palaci, J., Castelijns, B., Forment, J., Ziarsolo, P., et al. (2013). GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol.* 162, 1618–1631. doi: 10.1104/pp.113.217661
- Schähs, M., Strasser, R., Stadlmann, J., Kunert, R., Rademacher, T., and Steinkellner, H. (2007). Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol. J.* 5, 657–663. doi: 10.1111/j.1467-7652.2007.00273.x
- Schneider, J. D., Marillonnet, S., Castilho, A., Gruber, C., Werner, S., Mach, L., et al. (2014). Oligomerization status influences subcellular deposition and glycosylation of recombinant butyrylcholinesterase in *Nicotiana benthamiana*. *Plant Biotechnol. J.* 12, 832–839. doi: 10.1111/pbi.12184
- Schoberer, J., Liebminger, E., Botchway, S. W., Strasser, R., and Hawes, C. (2013). Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in planta. *Plant Physiol.* 161, 1737–1754. doi: 10.1104/pp.112.210757
- Schoberer, J., Runions, J., Steinkellner, H., Strasser, R., Hawes, C., and Osterrieder, A. (2010). Sequential depletion and acquisition of proteins during Golgi stack disassembly and reformation. *Traffic* 11, 1429–1444. doi: 10.1111/j.1600-0854.2010.01106.x
- Schoberer, J., and Strasser, R. (2011). Sub-compartmental organization of Golgi-resident N-glycan processing enzymes in plants. *Mol. Plant* 4, 220–228. doi: 10.1093/mp/ssq082
- Schoberer, J., Vavra, U., Stadlmann, J., Hawes, C., Mach, L., Steinkellner, H., et al. (2009). Arginine/lysine residues in the cytoplasmic tail promote ER export of plant glycosylation enzymes. *Traffic* 10, 101–115. doi: 10.1111/j.1600-0854.2008.00841.x
- Schuster, M., Umana, P., Ferrara, C., Brünker, P., Gerdes, C., Waxenecker, G., et al. (2005). Improved effector functions of a therapeutic monoclonal lewis Y-specific antibody by glycoform engineering. *Cancer Res.* 65, 7934–7941.
- Shin, Y. J., Chong, Y. J., Yang, M. S., and Kwon, T. H. (2011). Production of recombinant human granulocyte macrophage-colony stimulating factor in rice cell suspension culture with a human-like N-glycan structure. *Plant Biotechnol. J.* 9, 1109–1119. doi: 10.1111/j.1467-7652.2011.00636.x
- Shinkawa, T., Nakamura, K., Yamane, N., Shoji-Hosaka, E., Kanda, Y., Sakurada, M., et al. (2003). The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem.* 278, 3466–3473. doi: 10.1074/jbc.M210665200
- Solá, R. J., and Griebenow, K. (2010). Glycosylation of therapeutic proteins: an effective strategy to optimize efficacy. *BioDrugs* 24, 9–21. doi: 10.2165/11530550-000000000-00000
- Sourrouille, C., Marquet-Blouin, E., D'Aoust, M. A., Kiefer-Meyer, M. C., Seveno, M., Pagny-Salehabadi, S., et al. (2008). Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnol. J.* 6, 702–721. doi: 10.1111/j.1467-7652.2008.00353.x
- Staudacher, E., Dalik, T., Wawra, P., Altmann, F., and Marz, L. (1995). Functional purification and characterization of a GDP-fucose: beta-N-acetylglucosamine (Fuc to Asn linked GlcNAc) alpha 1,3-fucosyltransferase from mung beans. *Glycoconj. J.* 12, 780–786. doi: 10.1007/BF00731239
- Stoger, E., Fischer, R., Moloney, M., and Ma, J. K. (2014). Plant molecular pharming for the treatment of chronic and infectious diseases. *Annu. Rev. Plant Biol.* 65, 743–768. doi: 10.1146/annurev-aplant-050213-035850
- Strasser, R., Altmann, F., Glössl, J., and Steinkellner, H. (2004a). Unaltered complex N-glycan profiles in *Nicotiana benthamiana* despite drastic reduction of beta1,2- N-acetylglucosaminyltransferase I activity. *Glycoconj. J.* 21, 275–282. doi: 10.1023/B:GLYC.0000045099.29038.04
- Strasser, R., Altmann, F., Mach, L., Glössl, J., and Steinkellner, H. (2004b). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4
- Strasser, R., Bondili, J. S., Schoberer, J., Svoboda, B., Liebminger, E., Glossl, J., et al. (2007a). Enzymatic properties and subcellular localization of *Arabidopsis* beta-N-acetylhexosaminidases. *Plant Physiol.* 145, 5–16. doi: 10.1104/pp.107.101162
- Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glossl, J., et al. (2007b). A unique beta1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* 19, 2278–2292. doi: 10.1105/tpc.107.052985
- Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gattinger, P., et al. (2009). Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile. *J. Biol. Chem.* 284, 20479–20485. doi: 10.1074/jbc.M109.014126
- Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., et al. (2008). Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* 6, 392–402. doi: 10.1111/j.1467-7652.2008.00330.x
- Uliana, A. S., Giraudo, C. G., and Maccioni, H. J. (2006). Cytoplasmic tails of SialT2 and GalNAcT impose their respective proximal and distal Golgi localization. *Traffic* 7, 604–612. doi: 10.1111/j.1600-0854.2006.00413.x
- Umana, P., Jean-Mairet, J., Moudry, R., Amstutz, H., and Bailey, J. E. (1999). Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat. Biotechnol.* 17, 176–180. doi: 10.1038/6179
- van Dussen, L., Zimran, A., Akkerman, E. M., Aerts, J. M., Petakov, M., Elstein, D., et al. (2013). Taliglucerase alfa leads to favorable bone marrow responses in patients with type I Gaucher disease. *Blood Cells Mol. Dis.* 50, 206–211. doi: 10.1016/j.bcmd.2012.11.001
- Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3, 97–130. doi: 10.1093/glycob/3.2.97
- Varki, A. (2006). Nothing in glycobiology makes sense, except in the light of evolution. *Cell* 126, 841–845. doi: 10.1016/j.cell.2006.08.022
- Vézina, L. P., Faye, L., Lerouge, P., D'Aoust, M. A., Marquet-Blouin, E., Burel, C., et al. (2009). Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnol. J.* 7, 442–455. doi: 10.1111/j.1467-7652.2009.00414.x
- Von Schaewen, A., Sturm, A., O'Neill, J., and Chrispeels, M. J. (1993). Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase-I and is unable to synthesize golgi-modified complex N-linked glycans. *Plant Physiol.* 102, 1109–1118. doi: 10.1104/pp.102.4.1109
- Walsh, G. (2010). Biopharmaceutical benchmarks 2010. *Nat. Biotechnol.* 28, 917–924. doi: 10.1038/nbt0910-917
- Weber, E., Gruetzner, R., Werner, S., Engler, C., and Marillonnet, S. (2011). Assembly of designer TAL effectors by Golden Gate cloning. *PLoS ONE* 6:e19722. doi: 10.1371/journal.pone.0019722
- Wee, E. G., Sherrier, D. J., Prime, T. A., and Dupree, P. (1998). Targeting of active sialyltransferase to the plant Golgi apparatus. *Plant Cell* 10, 1759–1768. doi: 10.1105/tpc.10.10.1759
- Wilson, I. B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core alpha1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- www.federalgrants.com. (2012). *Butyrylcholinesterase Expression in Plants Federal Grant*. Available at: <http://www.federalgrants.com/Butyrylcholinesterase-Expression-in-Plants-33798.html> [accessed July 7, 2014].
- Yamane-Ohnuki, N., Kinoshita, S., Inoue-Urakubo, M., Kusunoki, M., Iida, S., Nakano, R., et al. (2004). Establishment of FUT8 knockout Chinese hamster ovary

- cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnol. Bioeng.* 87, 614–622. doi: 10.1002/bit.20151
- Zeitlin, L., Bohorov, O., Bohorova, N., Hiatt, A., Kim, D. H., Pauly, M. H., et al. (2013). Prophylactic and therapeutic testing of *Nicotiana*-derived RSV-neutralizing human monoclonal antibodies in the cotton rat model. *MAbs* 5, 263–269. doi: 10.4161/mabs.23281
- Zeitlin, L., Pettitt, J., Scully, C., Bohorova, N., Kim, D., Pauly, M., et al. (2011). Enhanced potency of fucose-free monoclonal antibody being developed as an Ebola virus immunoprotectant. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20690–20694. doi: 10.1073/pnas.1108360108
- Zimran, A., Brill-Almon, E., Chertkoff, R., Petakov, M., Blanco-Favela, F., Munoz, E. T., et al. (2011). Pivotal trial with plant cell-expressed recombinant glucocerebrosidase, taliglucerase alfa, a novel enzyme replacement therapy for Gaucher disease. *Blood* 118, 5767–5773. doi: 10.1182/blood-2011-07-66955

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 July 2014; accepted: 16 September 2014; published online: 08 October 2014.

*Citation: Loos A and Steinkellner H (2014) Plant glyco-biotechnology on the way to synthetic biology. *Front. Plant Sci.* 5:523. doi: 10.3389/fpls.2014.00523*

*This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.*

Copyright © 2014 Loos and Steinkellner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.