
Impact of Cellulose Synthase-Like Enzymes on Heteromannan Biodiversity

Dissertation
zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der Naturwissenschaftliche Fakultät I
“Biowissenschaften”

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von

Frau Annika Grieß-Osowski

1. Gutachter: Prof. Dr. Cătălin Voiniciuc
2. Gutachter: Prof. Dr. Steffen Abel
3. Gutachter: Prof. Dr. Marcel Quint

Datum der Verteidigung: 24.11.2023

This dissertation was executed in the independent research group “Designer Glycans” of Prof. Dr. Cătălin Voiniciuc at the Leibniz Institute of Plant Biochemistry. The supervision was carried out by Prof. Dr. Cătălin Voiniciuc in the period from February 2019 to April 2022.

Contents

List of Figures	iii
List of Tables	v
Publication	vi
Declaration of Authorship	vi
Acknowledgements	vii
Zusammenfassung	1
Abstract	2
1. Introduction	3
1.1 Structure and Components of the Plant Cell Wall	3
1.2 Physicochemical Properties of Heteromannans	6
1.3 Natural Variety of Heteromannans in the Plant Cell Wall	8
1.4 Heteromannan Biosynthesis	13
1.4.1 Generation of Activated Nucleotide Sugars for Heteromannan Synthesis	13
1.4.2 Elongation of Heteromannans	16
1.4.3 Substitution of Heteromannans and Transport to the Plasma Membrane	16
1.5 The Cellulose Synthase Superfamily	18
1.6 Comparative Analysis of GGM and XyG Synthesis	21
1.7 <i>Pichia pastoris</i> as a Model Organisms for HM Elongation	23
1.8 Selection of CSLK/A Enzymes	25
1.9 Aim of my Thesis	27
2. Methods	28
2.1 Tools Used for Selection of CSLK and CSLA Sequences	28
2.2 Cloning of Genes and Fluorescent Proteins into pPICZ-based Vectors	28
2.3 Cloning of Genes into GoldenPiCs-based Vectors	32
2.4 <i>E. coli</i> and <i>P. pastoris</i> transformation	33
2.5 <i>Pichia</i> Growth, Optical Density and Fluorescence Measurements	35
2.6 Confocal Microscopy of <i>Pichia</i> Cells	36
2.7 <i>Pichia</i> Harvest and Cell Wall Extraction	37
2.8 Enzymatic Digestion of <i>Pichia</i> Wall Polymers	38
2.9 Monosaccharide analysis	39
2.10 Glycosyl Linkage Analysis	40
2.11 Polysaccharide Analysis by Gel Electrophoresis	42
2.12 In-Gel Fluorescence of Microsomal Membrane Fractions	43
2.13 Cloning of CSLA genes into plant transformation vectors	44

2.14	Plant Growth Conditions	46
2.15	Transient expression in <i>N. benthamiana</i> leaves	47
2.16	Stable expression in <i>A. thaliana</i>	48
2.17	Mucilage Staining and Extraction	49
2.18	Monosaccharide Analysis of Plant AIR material	49
3.	Results.....	51
3.1	Detailed Analysis of Fluorescently-Tagged AkCSLA3.....	51
3.2	Assessing the Subcellular Localization of AkCSLA3	55
3.3	AkCSLA3 mis-localization.....	64
3.4	Altered sugar metabolism in <i>Pichia</i> expressing AkCSLA3.....	72
3.5	Protease Influence on fluorescently tagged AkCSLA3.....	76
3.6	Detailed analysis of fluorescently tagged <i>Arabidopsis thaliana</i> CSLAs	79
3.7	Structural analysis of <i>Arabidopsis thaliana</i> CSLAs.....	86
3.8	<i>Pichia</i> strains expressing multiple AtCSLA isoforms.....	88
3.9	Characterization of additional CSLK/As across the plant kingdom	90
3.10	Detailed characterization of CSLK/As in the plant kingdom.....	95
3.11	CSLA expression <i>in planta</i>	100
3.12	Characterization of <i>csla</i> mutants in <i>Arabidopsis</i> and moss.....	102
4.	Discussion.....	105
4.1	Elucidating the Mechanistic Function of AkCSLA3	105
4.2	Specificity of several CSLA isoforms of <i>Arabidopsis</i>	109
4.3	Function of non-characterized CSLA and CSLK.....	112
4.4	Conclusion	114
5.	References	116
A.	Appendices	139
B	Abbreviations.....	158
C	<i>Curriculum vitae</i>	161

List of Figures

Figure 1: Schematic structure of heteromannans.	5
Figure 2: Schematic overview of heteromannan biosynthesis.....	13
Figure 3: Generation of activated sugar nucleotides involved in heteromannan synthesis.	15
Figure 4: Phylogenetic tree of genes from the cellulose synthase superfamily and their corresponding enzymatic function.	20
Figure 5: Schematic overview and comparison of GT families involved in β -GGM and XyG synthesis.	22
Figure 6: Schematic overview of selected plant species throughout the plant kingdom with sequence similarity to AtCSLA2.	27
Figure 7: Overview of the cloning strategy used to fuse CSLK/A proteins with a fluorescent protein into the pPICZ cloning system.	31
Figure 8: Pichia expressing AkCSLA3 with C- or N-terminal tagged sfGFP.	52
Figure 9: Pichia expressing different fluorescent proteins fused with AkCSLA3 at the C-terminal end.	54
Figure 10: Fluorescence of Pichia co-expressing AkCSLA3-sfGFP and ScGOS1-mRuby2.	57
Figure 11: Fluorescent screening of ScGOS1-mRuby2 expressing Pichia strains during growth in YPD medium over time.	58
Figure 12: Constitutive expression of ScGOS1-mRuby2 in YPD medium for 48 h.	60
Figure 13: Fluorescence of AkCSLA3-sfGFP and ScGOS1-mRuby2 co-expressing Pichia strains after 48 h of pre-growth in YPD, followed by 48 h of methanol induction.	61
Figure 14: AkCSLA3-sfGFP co-localizes with ScGOS1-mRuby2 expressed in Pichia pastoris.	63
Figure 15: Screening of Pichia expressing RGS4:pmVenus and RGS4:AkCSLA3:pmVenus.	64
Figure 16: Localization of screened Pichia colonies expressing RGS4:pmVenus and RGS4:AkCSLA3:pmVenus.	65
Figure 17: Fluorescence and localization of Pichia expressing pmVenus or RGS4:pmVenus over time.	67
Figure 18: Fluorescence and localization of Pichia expressing RGS4:AkCSLA3:pmVenus over time.	68
Figure 19: Fluorescence and localization of Pichia expressing AkCSLA3:pmVenus.	69
Figure 20: Impact of AkCSLA3-sfGFP without the N-terminal sequence and TMD1 on fluorescence, localization and HM production in Pichia.	71
Figure 21: Impact of carbon supplementation on glucomannan production of AkCSLA3-Venus.	72

Figure 22: Fluorescence and localization of <i>Pichia</i> expressing AkCSLA3-sfGFP in X-33 and SuperMan5 (his+) strains.	74
Figure 23: Monosaccharide composition of <i>Pichia</i> expressing AkCSLA3-sfGFP in X-33 and SuperMan5 (his+) strains.	76
Figure 24: Fluorescent protease protection assay of sfGFP and AkCSLA3 fused with sfGFP on its N- and C-terminus with proteinase K and trypsin..	79
Figure 25: Screening of AtCSLA-Venus isoforms expressed in <i>Pichia</i>	81
Figure 26: AtCSLA proteins fluorescently tagged to Venus, mCitrine and 2A-sfGFP.	83
Figure 27: <i>Pichia</i> expressing AkCSLA3- and HM-producing AtCSLA-Venus isoforms.....	85
Figure 28: Verification of Heteromannans produced by AtCSLA-Venus expressing <i>Pichia</i> strains.	86
Figure 29: Conserved motif domains of AtCSLA isoforms and AkCSLA3.	88
Figure 30: AtCSLA2 and AtCSLA9 co-expression in <i>Pichia</i>	90
Figure 31: Screening of different CSLK/A enzymes with coding sequences synthesized by Geneart expressed in <i>Pichia</i>	93
Figure 32: Screening of native CSLA enzymes with coding sequences extracted from plant samples expressed in <i>Pichia</i>	94
Figure 33: Detailed analysis of <i>Pichia</i> expressing different CSLK/A enzymes in HM production..	97
Figure 34: Polysaccharide analysis by carbohydrate gel electrophoresis of active CSLA strains in <i>Pichia</i>	98
Figure 35: Detailed analysis of polysaccharides in plant material.	100
Figure 36: Transient overexpression of CSLK/A in <i>Nicotiana benthamiana</i> leaves.	102
Figure 37: AkCSLA3 expression in Col-0 and <i>csla239</i> mutant <i>Arabidopsis thaliana</i>	103
Figure 38: Influence of <i>csla</i> mutants of <i>Physcomitrium patens</i> mutants on Mannose abundance in gametophores.	104
Figure S1: Verification of AtCSLA15 (At15)-Venus DNA from <i>E. coli</i>	154
Figure S2: Localization of AkCSLA3-Venus and non-functional AtCSLA-Venus isoforms in <i>Pichia</i> cells.	155
Figure S3: Relative monosaccharide composition of fluorescently tagged and untagged AtCSLA10 and AtCSLA15 expressing <i>Pichia</i>	156
Figure S4: Comparison of conserved glucan chain coordinating motif between AtCSLA isoforms and glucomannan.....	156
Figure S5: Multiple <i>Arabidopsis thaliana</i> CSLA expression in <i>Pichia pastoris</i>	157

List of Tables

Table 1: Structural composition of heteromannan polymers from several plant tissues..	10
Table 2: Overview of fluorescent proteins used for C-terminal fusion with <i>AkCSLA3</i>	53
Table 3: Overview of selected proteins for Golgi localization in <i>Pichia</i>	56
Table 4: Putative (gluco)mannan synthases from species throughout the plant kingdom homologous to <i>AtCSLA2</i>	91
Table S1: Overview of primers used for Phusion PCR from cDNA and cloning into the pPICZ-based vectors.	139
Table S2: Overview of primers used for cloning into the pPICZ cloning system from plasmid DNA.	139
Table S3: Overview of pPICZ-based plasmids used for transformation into <i>Pichia X-33</i>	141
Table S4: Overview of primers used for Phusion PCR from cDNA and cloning into the Golden PiCs vectors.	142
Table S5: Overview of primers used for amplification and domestication (Dom) of transgenes into the Golden PiCs cloning system.	142
Table S6: Overview of Golden Gate Assemblies and created plasmids used for transformation into <i>Pichia X-33</i>	145
Table S7: Overview of primers for Phusion PCRs for expression in planta.	148
Table S8: Overview of created plasmids for overexpression in planta.	149
Table S9: Overview of primers used for genotyping and sequencing.....	150
Table S10: Glycosyl linkages in <i>Pichia</i> expressing CSLA-Venus.....	152
Table S11: Glycosyl linkages in <i>Pichia</i> expressing CSLK-Venus.....	152
Table S12 Structural comparison of CSLK/CSLA proteins.....	153

Publication

Grieff-Osowski A. and Voiniciuc C. (2023): Branched mannan and xyloglucan as a dynamic duo in plant cell walls. *Cell Surface*: 100098.

A. G. wrote the original draft, reviewed and edited changes made by the corresponding author C.V. C.V. reviewed and edited the original draft and additionally made the illustration. Certain parts of this review are reproduced in chapter 1.6 in this thesis.

Declaration of Authorship

I hereby declare that I, **Annika Grieff-Osowski**, have independently and without external assistance written the present dissertation entitled "Impact of Cellulose Synthase-Like Enzymes on Heteromannan Biodiversity." Only the specified sources and aids were used for the creation of this work, and any passages or content taken verbatim or in essence from other sources have been appropriately acknowledged. Data originating from the work of others have also been duly identified.

I am aware that providing false statements will result in the examination being considered as failed. With my signature, I affirm the accuracy of these statements and acknowledge the legal implications.

Duisburg, 29.06.2023

Annika Grieff-Osowski

Acknowledgements

I am very grateful to my supervisor Prof. Dr. Cătălin Voiniciuc who gave me the opportunity to do my doctorate in the interesting field of glycan synthesis combined with synthetic biology approaches. You have always been a great role model for me with your structured way of working and your immense efficiency. Thank you for taking the time to ask questions and for always being there with good advice, constructive feedback and instructions when needed. You showed me that a good scientist can overcome a lot of hurdles through a combination of knowledge, inventiveness and creativity. I thank you for your belief in me and my abilities, which helped me to take on a wide range of responsibilities in my project as well in the lab. Your mentorship has been instrumental in shaping my scientific thinking which helped me to complete my dissertation, but will also accompany me on my further journey.

Furthermore, I am thankful for the time I spent with my former colleagues. Bo Yang and Stefanie Müller, the three of us started nearly at the same time and built up the lab together. Thank you, Bo, for helping and sharing your opinions and advice, especially in the beginning, but it never ended. Stefanie Müller and Stefanie Clauß, I want to thank you for helping in the lab with cloning, establishing the protein and microsomal isolations from *Pichia* material. Madalen Robert, you were everytime there for others, for example driving me on Friday evenings to the main station or plating out the transformation mixtures when I was getting late to drive back to Duisburg for the weekend. In these three years, all of you have been a constant and integral part – not only in scientific matters, but also in my personal life! The friendly, welcoming and supportive working and personal environment with all of you made time pass very quickly.

I would like to express my profound appreciation to the Leibniz Institute of Plant Biochemistry which was not only a working area for me, it was more an institution where I conducted my research. The resources, facilities, and academic environment provided by the Leibniz Institute of Plant Biochemistry have been integral to the successful completion of this dissertation.

Finally, I would like to thank my husband for being the whole time on my side, taking care about me (even if there were only disgusting smoothies made for the evening, when I arrived in the middle of the night) and sometimes more or less understanding why I needed to take on later train or arriving back on Saturdays. Thanks for your support!

In conclusion, I deeply appreciate the opportunity to participate in this scientific project. I am grateful to all those who have supported and believed in me, and I look forward to the exciting future ahead as I continue my scientific endeavors. Thank all of you!

Zusammenfassung

Pflanzenzellwände bieten mechanische Festigkeit und spielen entscheidende Rollen in der Pflanzenentwicklung und -differenzierung. Im Verlauf der Evolution von Wasser- zu Landpflanzen haben deren Zellwände bedeutende strukturelle und zusammensetzungsbezogene Veränderungen durchlaufen. Heteromannan (HM), betrachtet als älteste Hemicellulose, besteht hauptsächlich aus β -1,4-verknüpfter Mannose (Man) und kann Glukose (Glc) in seinem Rückgrat einbauen. HMs werden von der Cellulose Synthase-Like A (CSLA) Familie der Glycosyltransferasen synthetisiert.

Um zu untersuchen, ob die strukturelle Vielfalt der HM-Polysaccharide hauptsächlich durch die Enzymspezifität bestimmt wird, wurden alle Arabidopsis CSLA-Gene und orthologe Gene aus einer Vielzahl von Nicht-Gefäßpflanzen ausgewählt. *Pichia pastoris* Hefezellen dienen als idealer orthogonaler Wirt zur Charakterisierung von Enzymen und Kofaktoren, die an der HM-Verlängerung beteiligt sind. Dieser synthetisch-biologische Ansatz ermöglicht es, die biochemische Aktivität der bisher nicht charakterisierten algenartigen CSLK-Klade zu sondieren, die den CSLA- und CSLC-Kladen der Landpflanzen sehr ähnlich ist.

Meine Ergebnisse zeigen, dass mehrere Mitglieder der CSLA-Familie die Verlängerung von HMs katalysieren. Interessanterweise produzierten verschiedene CSLAs von *Physcomitrium patens* unterschiedliche Polysaccharid-Erträge oder -Strukturen, was darauf hinweist, dass diese Isoformen einzigartige Spezifitäten in der HM-Synthese haben. Einzelne und höhergeordnete *csla*-Mutanten wurden mittels CRISPR Cas9 isoliert und werden helfen, die Rollen mehrerer Isoformen in einer einzigen Art zu erforschen. Der haploide Lebenszyklus von *P. patens* sollte dazu beitragen, die Beschränkungen von *Arabidopsis csla*-Mutanten, einschließlich genetischer Redundanz und Embryo-Letalität, zu überwinden. Die ursprüngliche CSLK-Familie, von der angenommen wurde, dass sie an der Synthese von Mannanen bei Rot- und Grünalgen beteiligt ist, konnte erfolgreich für *Porphyra umbilicalis* gezeigt werden. Alle Grünalgen-CSLK waren nicht an der HM-Synthese in *Pichia* beteiligt und deuten auf eine andere Funktion dieser Enzyme hin. s

Die C-terminale Fusion von *AkCSLA3* aus *Amorphophallus konjac* ist auf das Golgi lokalisiert und zeigt eine hochstabile Bindung an das Fluoreszenzprotein. Daher kann *AkCSLA3*-sfGFP nun für in vitro Charakterisierungen, Transportstudien und Studien zu Protein-Interaktionen verwendet werden. Insbesondere die N-terminale Sequenz bis zum ersten Transmembrandomän von *AkCSLA3* und sein Glucanketten-Koordinationsmotiv DYHF scheinen wichtig zu sein.

Stichpunkte: Zellwand, Hemicellulosen, Heteromannane, CSLA, CSLK, Fluoreszente Proteine, *Pichia pastoris*, *Arabidopsis thaliana*, *Porphyra umbilicalis*

Abstract

Plant cell walls providing mechanical strength and playing crucial roles in plant development and differentiation. As plants evolved from aquatic to terrestrial plants, their cell walls underwent significant structural and compositional transformations. Heteromannan (HM) considered as the most ancestral hemicellulose, primarily consists of β -1,4-linked mannose (Man) and can incorporate glucose (Glc) in its backbone. HMs are synthesized by the Cellulose Synthase-Like A (CSLA) family of glycosyltransferases.

To investigate if the structural diversity of HM polysaccharides is primarily determined by enzyme specificity, all *Arabidopsis* CSLA genes and orthologous genes from a variety of non-vascular plants were selected. *Pichia pastoris* yeast cells serve as an ideal orthogonal host to characterize enzymes and co-factors involved in HM elongation. This synthetic biology approach also enables us to probe the biochemical activity of the previously uncharacterized algal CSLK clade, which closely resembles the CSLA and CSLC clades of terrestrial plants.

My results indicate that several members of the CSLA family catalyze the elongation of HMs. Interestingly, different CSLAs from *Physcomitrium patens* produced distinct polysaccharide yields or structures, indicating that these isoforms have unique specificities in HM synthesis. Single and higher-order *csla* mutants using CRISPR Cas9 were isolated and will help to explore the roles of several isoforms in one single species. The haploid life cycle of *P. patens* should help to overcome the limitations of *Arabidopsis* *csla* mutants, including genetic redundancy and embryo-lethality. The ancestral CSLK family, previously thought to be involved in the synthesis of mannans in red and green algae, was successfully shown for *Porphyra umbilicalis*. All green algal CSLK were not involved in HM synthesis in *Pichia* and indicate for another function of these enzymes. The ancestral CSLK family, believed in HM synthesis of red and green algae, could only reproduced for *Porphyra umbilicalis* in this thesis.

The C-terminal fusion on *AkCSLA3* from *Amorphophallus konjac* is localized to the Golgi and shows highly stable binding to the fluorescent protein. Therefore, *AkCSLA3*-sfGFP can now be used for *in vitro* characterization, trafficking and protein interactions studies. Especially the N-terminal sequence up to the first transmembrane domain of *AkCSLA3* and its glucan chain coordination motif DYHF seems to be important.

Keywords: cell wall, hemicelluloses, heteromannan, CSLA, CSLK, Fluorescent proteins, *Pichia pastoris*, *Arabidopsis thaliana*, *Porphyra umbilicalis*

1. Introduction

1.1 Structure and Components of the Plant Cell Wall

Plant cell walls are important structural elements surrounding each cell and, therefore, protect the protoplasm from the environment and allow the plant to grow in a strong, rigid and flexible manner [Chebli and Geitmann, 2017]. Components of the plant cell wall are predominantly carbohydrates, a small amount of proteins and in certain cell types lignin. Carbohydrates are built up of monosaccharides that are linked linearly and can be branched to form different kind of oligosaccharides (usually <20 linked monosaccharides) or polysaccharides (typically larger number of monosaccharides with a repetitive oligosaccharide motif) [Varki et al., 2022]. The nomenclature of these polymers is based on the major monosaccharides of the elongated backbone and are frequently terminated with “-an”. For example, if glucose (Glc) is the major component of the backbone, the polymers are defined as glucans, whereas they are termed as mannans, if mannose (Man) is the main monosaccharide. Starting from 13 possible monosaccharides that can be linked and branched with each other, a large number of possible glycans can be assumed and underline their structural complexity and diversity. In addition, the composition of the cell wall varies depending on the species, cell type, cell wall domain and cellular differentiation [Wolf et al., 2012] which makes the field of glycobiology even more complex.

In general, two types of cell walls have been classified. First, the primary cell wall (PCW) is ubiquitously built up in growing cells and form a thin and flexible layer. When cells stopped growing and have been differentiated, a secondary cell wall (SCW) is deposited onto the inside of the PCW and thicken the layer [Cosgrove, 2005; Zeng et al., 2014]. Four polysaccharides of the plant cell wall can classically be distinguished – cellulose, hemicellulose, pectin and lignin. In both types of cell walls cellulose, hemicellulose and pectin are present, but lignin additionally occurs only in SCWs. Up to 30% of each polysaccharide is present in PCWs. Cellulose and lignin are strongly increased during SCW formation and make up over 60–80% of the wall, whereas pectin is less present (>10%) than in PCWs [Albersheim et al., 2011; Zhong et al., 2018].

Cellulose is the major polysaccharide in plant cell walls and is composed completely of Glc molecules. One chain of cellulose can polymerize >500 up to 14,000 β -D-Glc residues connected to each other by β -1,4 glycosidic linkages. Because the β -1,4-linked glucan chains of cellulose are unbranched, they can easily form intramolecular as well as intermolecular hydrogen bonds resulting in lateral aggregation and crystallization of these chains. The structure that is formed when approximately 40 β -1,4-linked glucan chains are connected to each other is called cellulose microfibril [Somerville et al., 2004; Nakano et al., 2015]. As

mentioned before, the abundance of cellulose in SCWs is higher compared to PCWs, but also the degree of polymerization (DP) and the width of the microfibrils differ in the type of cell wall. In cotton, cellulose has a relatively low DP (~2,000–6,000) and smaller microfibril widths (2–2.5 nm) in PCWs, whereas in SCWs the DP (~13,000) and microfibril widths (5–10 nm) is increased [Nakano et al., 2015].

In contrast to cellulose, hemicelluloses and pectins are “frequently heteropolymers” [Morrison, 2001], therefore they consist of different monosaccharides or of the same monosaccharide that is connected by different linkages. Hemicelluloses can be grouped into mixed-linkage glucans, xyloglucans, xylans, mannans and glucomannans. Their common feature is the β -1,4-linked backbone of Glc, Man or xylose (Xyl) that shows the same equatorial configuration at C1 and C4 [Scheller and Ulvskov, 2010]. But dependent on the kind of hemicellulose, their backbones can be further substituted with galactose (Gal), fucose (Fuc), Xyl, arabinose (Ara), glucuronic acid (GlcA), acetyl (Ac) or 4-O-methylglucuronic acid (Me-GlcA) and varies with the plant species [Scheller and Ulvskov, 2010, Pauly et al., 2013].

Mixed-linkage glucans (MLGs) are widely distributed in the *Poaceae* family (grasses and cereals) but are also found in green algae, liverworts, bacteria, fungi and lichens [Kraemer et al., 2021]. They are built up of β -1,3;1,4-glucans whereby the amount of β -1,4-linkages predominates in most of these plant species and seem to influence the physiochemical and functional properties of the primary and secondary cell wall [Burton and Fincher, 2009].

Xyloglucan (XyG) is the major hemicellulose of the PCW of dicots and non-commelinoid (ie. non-flowering) monocots. Their backbone is synthesized by cellulose synthase-like C (CSLC) enzymes [Cocuron et al., 2007; Dwivany et al., 2009], while XyG xylosyltransferases (XXTs) attach Xyl to the elongated polymer chain by α -1,6-linkages [Cavalier et al., 2008]. Although XyG is completely absent in *cs/c* quintuple mutants [Kim et al., 2020] and *xxt* triple mutants [Zabotina et al., 2012], they exhibit a nearly normal growth phenotype comparable to wildtype plants- As a result, it has been suggested that XyG may have a structural function as spacer or adapter molecule for cellulose [Pauly et al., 2013], or that other polysaccharides or wall components can compensate for the lack of XyG [Kim et al., 2020].

Xylan polymers replace XyG in the PCW of commelinoid monocots (grasses, cereals and bromeliads). Xylans possess a linear backbone consisting of only β -1,4-linked Xyl residues. In general, the xylan backbone can be branched with Ac, GlcA, Me-GlcA or Ara. The xylan backbone of the PCW of commelinoid monocots is decorated with Ara and GlcA, whereas it is branched with arabinofuranose (Araf) and only minor amounts of GlcA in SCWs [Handford, 2006; Peña et al., 2016]. In dicots, the predominant hemicellulose of the SCW is xylan branched with MeGlcA and GlcA but it lacks Ara in their side chains [Handford, 2006; Brown et al., 2007; Bromley et al., 2013; Rennie and Scheller, 2014].

Mannans and glucomannans are grouped together as heteromannans (HMs). While the backbone of mannans primarily consist of β -1,4-linked Man (>90%), glucomannans additionally incorporate β -1,4-linked Glc in its backbone. Both polymers can be substituted with Gal residues by α -1,6-linkages and are then referred to as galactomannans or galactoglucomannans (GGMs) [Scheller and Ulvskov, 2010]. The GGM backbone in PCWs of several eudicots have a regular alternating pattern of Glc-Man disaccharides which is frequently substituted with α -1,6-linked Gal and β -1,2-Gal- α -1,6-Gal disaccharides [Seymour et al., 1990; Sims et al., 1997; Schröder et al., 2001; Prakash et al., 2012; Yu et al., 2022] and is referred as β -GGM by Yu et al. (2022) (Figure 1). HMs are widely distributed throughout the plant kingdom, ranging from aqueous to terrestrial plants. However, their structure and abundance vary depending on the species, life cycle, tissue and developmental stage. HMs fulfil structural functions as shown for linear mannans that form microfibrillar structures similar to cellulose in green algae (*Codium fragile*) [Mackie and Preston, 1968]. Furthermore, they serve as an energy storage polymer as shown for GGM in a variety of seed endosperms [Ishrud et al., 2001; Redgwell et al., 2003; Dhugga et al., 2004; Buckeridge, 2010] or glucomannan in the corm of *Amorphophallus konjac* [Gille et al., 2011].

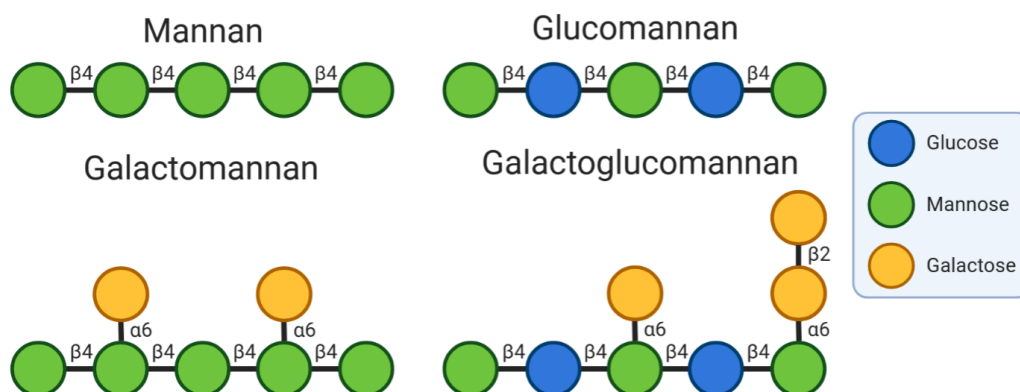


Figure 1: Schematic structure of heteromannans. The backbone of mannan consists only of β -1,4-linked mannose, whereas β -1,4-linked mannose and glucose subunits form glucomannan. Additionally, the backbone can be linked to single galactose units by α -1,6-linkages or β -Gal- α -Gal disaccharide side chains and form galactomannan and galactoglucomannan. The figure was prepared with bioRENDER (biorender.com).

Pectins are classified by a large amount (approximately 70%) of galacturonic acid (GalA) and have α -1,4 linkages of GalA at the O-1 and O-4 position in common. They comprise four different groups, namely homogalacturonan (HG), rhamnogalacturonan-I (RG-I) and the substituted galacturonans rhamnogalacturonan-II (RG-II) and xylogalacturonan (XGA). While the backbone of HG, RG-II and XGA only consist of GalA residues (HG: unbranched, RG-II: decorated with 12 different linkages, XGA: decorated with Xyl), RG-I additionally incorporates rhamnose (Rha) in its backbone by α -1,2 linkages [Ridley et al., 2001; Mohnen, 2008]. Although HG is a linear polymer, it is highly methylesterified (often >50%) and partially O-acetylated [Ishi, 1997; Ridley et al., 2001]. Due to the negative charge of unesterified GalA

residues in HG, they can cross-link to calcium (Ca^{2+}) ions, increasing their cell wall adhesion and stiffness of the gel structure in HG [Voiniciuc et al., 2013; Levesque-Tremblay et al., 2015].

The formation of lignin occurs in the SCW. Three main types of monolignols (sinapyl alcohol, coniferyl alcohol and p-coumaryl alcohol) are polymerized by peroxidases and laccases [Liu et al., 2018]. Lignin can interact with hemicelluloses resulting in an alternate order of lignin-hemicellulose and cellulose-hemicellulose networks [Zeng et al., 2014]. These covalent cross-linking interactions between lignin and hemicelluloses in SCWs affect the properties of the cell wall. Higher amounts of lignin increase the mechanical strength, rigidity and hydrophobicity of SCWs [Zhong et al., 2018; Terret and Dupree, 2019].

While cellulose is synthesized at the plasma membrane, hemicelluloses and pectins are synthesized in the Golgi apparatus and transported by vesicles to the cell wall surface. The integration of these matrix polysaccharides into cellulose microfibrils is achieved by physical interaction, enzymatic ligation and crosslinking reactions [Cosgrove, 2005]. The interaction of cellulose and matrix polysaccharides form a strong fibrous network that is referred as hemicellulose-cellulose network or pectin-cellulose network dependent on the interaction partner of cellulose [Cosgrove, 2005; Broxterman and Schols, 2018a]. Meanwhile, an interaction of pectin with xylans has also been reported in the mucilage of *Arabidopsis thaliana* and the PCW of tomato and are formed by covalently cross-links between the RG-I region of pectin and xylan [Ralet et al., 2016; Broxterman and Schols, 2018b].

1.2 Physicochemical Properties of Heteromannans

The physicochemical properties of plant cell wall polysaccharides are predominantly determined by their chemical structure. Factors such as monosaccharide composition, degree of polymerization (DP; backbone length), degree of substitution (DS; side branches) and their arrangement or distribution can significantly impact the properties of the polysaccharide.

Determining the structure of a polysaccharide can be accomplished using several techniques, such as multiangle light scattering, size exclusion chromatography, gas chromatography or NMR spectroscopy [Amos and Mohnen, 2019]. However, characterizing the polymer's structure still remains challenging due to highly insoluble or galactosylated polysaccharides forming aggregates, may leading to overestimated molecular weights [Nishinari et al., 2007; Pitkänen et al., 2011]. Physical or chemical methods, such as heating, sonication, pressure or strong solvents, can help disperse the polysaccharide in solution, but may also lead to structural damage [Nishinari et al., 2007]. Additionally, the determination of the polysaccharide size can be complicated by the diverse and irregular length distribution that exists even within a single plant species. For instance, the DP can range from below 20 to above 10,000 as shown for the green algae *Codium fragile* (with 90% of the mannan polymers

having a DP between 100 and 2,500) [Mackie and Sellen, 1969]. Therefore, averages of molecular weights and DPs are commonly calculated to compare the polymer molecular masses and lengths, but the extraction method used can also affect these values [Guo et al., 2018; Amos and Mohnen, 2019].

As already mentioned, HMs consist of the monosaccharides Glc, Man and Gal. While Glc and Man are built into the polymer backbone, Gal is commonly found substituted to mannopyranosyl residues [Scheller and Ulvskov, 2010]. The glycosidic linkages between monosaccharides can affect the polymer's flexibility and stability, due to the way the sugar residues are oriented to each other and their ability to form intra- and intermolecular hydrogen bonds. Molecular dynamic simulations coupled with nuclear magnetic resonance (NMR) spectroscopy revealed two possible types of linkages in glucomannan backbones in aqueous solution: (i) C-type linkages between 4-Glc- β -1,4-Man- β -1 and 4-Glc- β -1,4-Glc- β -1 and (ii) M-type linkages between 4-Man- β -1,4-Man- β -1 and 4-Man- β -1,4-Glc- β -1. Due to their conformational orientation, C-type linkages are stiffer and less flexible than M-type linkages [Berglund et al., 2016]. According to Berglund et al. (2016) glucomannan with a higher content of Glc in the backbone may lead to a stiffer polymer, resulting in a greater affinity to bind to cellulose. The glucomannan backbone can be further stabilized and stiffened by Gal side chain substitutions [Moreira and Filho, 2008; Berglund et al., 2019]. Determining which structural effect (Glc:Man ratio of the backbone, Gal substitutions or their distribution within the polymer) has the greatest impact on the stability is difficult to determine [Berglund et al., 2019] due to their complex interference and structural diversity in planta.

In general, the stronger the glycosidic bonds are and the more intermolecular bonds with other polymers can be formed, the more insoluble is the polysaccharide in solution. Linear polysaccharides with a high DP can easily form aggregates via hydrogen bonds between the hydroxyl groups of sugar units from the backbone and, therefore are highly insoluble in water. In contrast, the incorporation of Glc into the mannan backbone or decoration with Gal, sulphate or Ac groups makes the HM polymer more soluble to water by reducing intermolecular bonds with other polysaccharides [Dea and Morrison, 1975; Matheson, 1990; Schröder et al., 2009; Chokboribal et al., 2015]. Glucomannans are more soluble than mannans due to their trans-configured hydroxyl groups at C-2 and C-3 of Glc (OH on opposite sides) and partial acetylation, which reduce intermolecular hydrogen-bonding by steric hindrance of the polymer in solution [Koroskenyi and McCarty, 2001; Du et al., 2012; Mathur, 2012]. Assuming the same molecular weight of a polymer, it can be concluded that longer polysaccharides with a higher DP are more viscous and absorb more water than shorter polysaccharides with many side branches due to their high ability to form hydrogen bonds [Podzimek, 2011; Guo et al., 2018]. Similarly, the higher number of Ac groups in glucomannan [Gao and Nishinari, 2004;

Campestrini et al., 2013; Minjares-Fuentes et al., 2018] or Gal side chains in galactomannan [Gadkari et al., 2018] can lower the viscosity of the polymer by reducing their hydrogen bonding and hydrophobic interaction with other polymers [Nishinari et al., 2007; Mathur, 2012; Aanisah et al., 2022]. Moreover, an increasing amount of Gal side chains on the mannan backbone, increase their ability to form weak gels, as demonstrated for guar gum and fenugreek gum which exhibit higher Man:Gal ratios. In contrast, locust bean gum exhibit lower Man:Gal ratios and is unable to form gels [Gadkari et al., 2018 and 2019]. However, it is much more difficult to determine the rheological properties of substituted polysaccharides from natural varieties because their distribution within the polysaccharide can affect the polymer's properties and is not observable when only the DS is determined. If the side decorations on mannan or glucomannan are block wise distributed, longer unbranched regions of the sugar (so-called smooth regions) can form intermolecular interactions and, therefore reduce the polymers solubility and increase their viscosity – at least in this region [Kuravadi et al., 2012; Mathur, 2012]. Due to the strong interdependence of all these factors and their variability within natural resources, it become notable that identifying the specific structure that underlies a particular property of the respective polysaccharide is challenging.

Because HMs interact with other polysaccharides such as celluloses in the complex network of the plant cell wall [Yu et al., 2018] it is not enough to consider only the physicochemical properties of HMs alone. The differences in the polysaccharide composition and their interplay with each other influence their overall mechanical behaviour in the cell wall. In case of HMs, not much is known about their function in the cell wall. Despite having reduced amounts of mannan in the inflorescence stem of *Arabidopsis thaliana csla9* single mutants or *csla239* triple mutants, these mutants exhibit a normal growth phenotype similar to the wildtype plant [Goubet et al., 2009]. This suggests that other polysaccharides might compensate for this deficiency or do functional overlap.

1.3 Natural Variety of Heteromannans in the Plant Cell Wall

Man-rich polysaccharides are widely distributed throughout the plant kingdom. Although most studies on HM biosynthesis have focused on angiosperm and gymnosperm species, cell wall analysis using mannan-specific antibodies revealed high levels of β -1,4-mannan in many algae [reviewed in Popper et al., 2011; Ciancia et al., 2020], bryophytes [Ye and Zhong 2022] and lycophytes [Harholt et al., 2012]. Even cellulose seems to have been replaced with β -mannan as their major component of the cell wall in some green and red algae [Cronshaw et al., 1958; Painter, 1983; Moreira and Filho, 2008; Domozych et al., 2012]. An overview of the structural diversity of HM polymers from several plant species is summarized in Table 1.

Representative plants containing linear mannans as the dominant storage polymer in their endosperms are ivory nuts (*Phytelephas macrocarpa*, 97–98% pure Man) [Aspinall et al., 1953; Meier, 1958; Chanzy et al., 1984] and date palms (*Phoenix dactylifera*, 93% 1,4-Man) [Meier, 1958; Ishrud et al., 2001]. Their seeds are extremely hard and their mannans are highly water-insoluble underlining the structural function of mannans to resist against mechanical stresses [Singh et al., 2018]. Furthermore, linear mannans occur in the thallus of the red alga *Porphyra umbilicalis* and the green alga *Codium fragile* [Frei and Preston, 1964; Mackie and Preston, 1968]. Linear mannans from ivory nuts, dates and *Codium fragile* have been reported to occur in two forms – mannan A and mannan B. The length of mannan A is shorter (smaller DP) and appear as granular and crystalline structures whereas mannan B is longer (higher DP) and form microfibrillar structures. These granular mannans are embedded into the mannan microfibrils [Meier, 1958, Mackie and Preston, 1968; Mackie and Sellen, 1969; Chanzy et al., 1984]. In contrast, mannans from *Porphyra umbilicalis* occur only crystalline or granular in the thallus or cuticle after hot water extraction and are not structurally organised as microfibrils [Frei and Preston, 1964]. Many red and green algal mannans are partially sulphated. As shown for the green alga *Codium vermilara*, 23% of the mannose units in the mannan backbone are linked at the C-2 position to sulphate which increase the polymers solubility resulting in amorphous regions of the cell wall [Ciancia et al 2007; Usov, 2011; Fernández et al., 2014].

Table 1: Structural composition of heteromannan polymers from several plant tissues. Green surroundings illustrate pure mannan backbones, while blue surroundings illustrate glucomannan backbones. Yellow surroundings show galactose (Gal) branching of the backbone. Backbone (BB), glucose (Glc), mannose (Man), degree of substitution (DS) acetylated (Ac) or sulphated (Su), reference (Ref).

BB	Plant Species	Clade	Tissue	DP	Glc	Man	Gal	DS _{Ac/Su}	Ref	
Mannan	<i>Porphyra umbilicalis</i> (Red alga)	Rhodophyte	Thallus	12	–	1	–	–	Jones et al., 1959	
	<i>Codium fragile</i> (Green alga)	Charophyte	Thallus	100–2,500	–	1	–	Su (n.a.)	Mackie and Sellen, 1969; Fernández et al., 2012 & 2014	
	<i>Phytelephas macrocarpa</i> (Ivory nut)	Angiosperm (Monocot)	Seed	Mannan A	15–30	–	1	–	–	Meier, 1958; Chanzy et al., 1984; Heux et al., 2005
				Mannan B	80	–	1	–	–	–
	<i>Phoenix dactylifera</i> (Date palm)	Angiosperm (Monocot)	Seed	Mannan A	17–21	1	13.3	–	Ac	Meier, 1958;
				Mannan B	80	–	13.3	–	(1.0; 22%)	Ishrud et al., 2001
	<i>Coffea arabica</i>	Angiosperm (Eudicot)	Seed	early	17–24	–	7	1	Ac (11%)	Nunes and Coimbra, 2002; Redgwell et al., 2003; Nunes et al., 2005
				mature		–	40	1		
	<i>Ceratonia siliqua</i> (Locust bean)	Angiosperm (Eudicot)	Seed	1,500	–	3.5–4.3	1	–	von Freiesleben et al., 2018; Bouzouita et al., 2007	
	<i>Cyamopsis tetragonolobus</i> (Guar)	Angiosperm (Eudicot)	Seed	900	–	1.6–1.8	1	–	Shobha et al., 2005; Mudgil et al., 2011	
<i>Trigonella foenum-graecum</i> (Fenugreek)	Angiosperm (Eudicot)	Seed	4,900–6,500	–	1	1.02–1.14	–	Brummer et al., 2003		
Glucomannan	<i>Physcomitrium patens</i> (Moss)	Bryophyte	Leaf	n.a.	1	4.1	–	Ac (0.22)	Zhong et al., 2019a	
	<i>Selaginella moellendorffii</i>	Lycophyte	Stem	n.a.	1	5.3	–	Ac (0.29)	Zhong et al., 2019a	
	<i>Aloe barbadensis</i> Miller (Aloe vera)	Angiosperm (Monocot)	Leaf gel	12,000	1	15	–	Ac (0.79; 95%)	Avigad and Dey, 1997; Chow et al., 2005; Chokboribal et al., 2015	
	<i>Amorphophallus konjac</i> (Konjac)	Angiosperm (Monocot)	Tuber	>6,000	1	1.4–1.8	–	Ac (0.11; 11%)	Avigad and Dey, 1997; Mao et al., 2018; Gille et al., 2011; Zhong et al., 2018	
	<i>Dendrobium officinale</i> (Orchid)	Angiosperm (Monocot)	Stem	3–15	1	4.7–5.9	–	Ac (0.25)	Xing et al., 2014 & 2015; He et al., 2017; Zhong et al., 2019b	
	<i>Arabidopsis thaliana</i>	Angiosperm (Eudicot)	Stem	n.a.	n.a.	n.a.	n.a.	Ac (0.25)	Zhong et al., 2019a	
			Mucilage	n.a.	1	1	0.73	–	Yu et al., 2018 & 2022	
<i>Picea abies</i> (Norway spruce)	Gymnosperm	Stem	180-300	1	4	0.6	Ac (0.32; 65%)	Willför et al., 2003		

To date, glucomannan backbones have only been reported in terrestrial plants, leading to the suggestion that glucomannan polymers have evolved during land plant evolution. Glucomannans have been found in the moss *Physcomitrium patens*, the lycophyte *Selaginella moellendorffii* [Zhong et al., 2019] and in several angiosperms such as *Amorphophallus konjac* [Cescutti et al., 2002], *Dendrobium officinale* [He et al., 2017], *Aloe barbadensis* Miller (*Aloe vera*) [Chow et al., 2005] and *Arabidopsis thaliana* [Handford et al., 2003]. They are highly abundant in konjac tubers (60–80%) [Mikkonen et al., 2012], *Aloe vera* leaf gels (86.87% Man, 12.68% Glc) [Quezada et al., 2017] and *Dendrobium officinale* stems (58.3%) [Xing et al., 2014; He et al., 2017]. Man polymers are less abundant in the other species (angiosperms: 1–5%; *P.patens*: ~6%; *S. moellendorffii*: ~10% of the total cell wall) [Liepman et al., 2007; Moller et al., 2007; Sørensen et al., 2011; Harholt et al., 2012; Willför et al., 2015b].

Glucomannans can be acetylated at the position C-2, C-3 and C-6 only in Man residues in the HM backbone. While many glucomannans are predominantly O-acetylated at the C-2 and C-3 position [Gille et al., 2011; Zhong et al., 2018, 2019a], glucomannan from *Aloe vera* is highly acetylated at the C-6 position [Simões et al., 2012; Campestrini et al., 2013]. In general, the reason for the acetylation is mostly not well understood, but is believed to be important for pathogen resistance, plant development and impacts the mechanical properties of the cell wall by preventing aggregation of the polymer [Manabe et al., 2013]. Two of the most extensively studied glucomannans are konjac glucomannan and acemannan from *Aloe vera* [Shi et al., 2020]. Konjac glucomannan serves as a storage polymer in the corm, with Glc and Man randomly attached in the backbone chain. Only the mannosyl residues are slightly acetylated (11%), with equal acetylation occurring at position C-2 and C-3 [Cescutti et al., 2002; Gao and Nishinari, 2004; Gille et al., 2011]. According to ¹³C nuclear magnetic resonance (NMR) analysis, the mannosyl residues of the backbone can also be branched with α -1,6-linked Glc at the C-3 position of Man, although this occurs infrequently and varies depending on the konjac species [Katsuraya et al., 2003; Behera and Ray, 2016]. The glucomannan backbone of *Aloe vera* consists mainly of Man and incorporates only a few Glc residues but no Gal [Chow et al., 2005; Salinas et al., 2019]. It is highly acetylated (~95% at position C-2, C-3 and/or C-6), thus it is referred to as acemannan and occurs in the inner gel of the leaves [Reynolds and Dweck, 1999; Chokboribal et al., 2015]. At least two acetyl groups are assembled per Man residue [Talmadge et al., 2004; Simões et al., 2012]. The degree of acetylation of acemannan not only increases the solubility and hydrophilicity of the polymer, but also alters the topology of deacetylated acemannan from an amorphous to a packed structure. The topological change could be the reason why the wound healing of human gingival fibroblasts is impaired when acemannan was deacetylated and confirm the importance of acetylation for their biological activity [Chokboribal et al., 2015].

Galactomannans occur as storage polysaccharides in seed endosperms of various plants such as *Coffea arabica* [Wolfrom et al., 1961], *Ceratonia siliqua* (locust bean), *Cyamopsis tetragonolobus* (guar) and *Trigonella foenum-graecum* (fenugreek) [Dea and Morrison, 1975]. Man:Gal ratios vary a lot in different family members of legume species: from relatively few Gal branches (3.9–4.0:1) shown for locust bean [Bouzouita et al., 2007], to intermediate occurring Gal branches (1.6–1.8:1) in guar [Mudgil et al., 2011] and to galactomannans where every Man is branched with a Gal (1:1) as shown for fenugreek [Brummer et al., 2003]. As already mentioned, the distribution of Gal substitutions in the mannan backbone differs for all three gums. While the backbone of locust bean gum and guar gum exhibit smooth regions (locust bean: ~25 Man units; guar gum: <6 Man units), fenugreek has very few unsubstituted Man in its backbone [Mathur, 2012]. The broad variation in molecular weight, Man:Gal ratio and distribution of Gal branches within the polymer makes these galactomannans diverse for many industrial applications (see more in chapter 1.8) [Daas et al., 2002] which is probably the reason why these three species are the most studied representatives for galactomannans. The reason for variations of Gal:Man ratios between species is not well understood. Interestingly, the degree of Gal substitution in galactomannan from *Arabica caturra* coffee beans is reduced with elongating weeks after flowering, while the mannan content is increased, indicating that galactomannan synthesis is developmentally regulated [Redgwell et al., 2003].

GGMs are highly abundant in the SCWs of gymnosperms such as *Picea abies* (Norway spruce) and other conifer species, comprising about 15–25% of the dry wood [Willför et al., 2005; Song et al., 2013; Terret et al., 2019]. However, their abundance in the PCWs of gymnosperms and the PCWs and SCWs of several angiosperms is relatively low [Harris and Stone, 2008; Schröder et al., 2009; Scheller and Ulvskov, 2010]. In SCWs, GGMs exhibit a random arrangement of Glc and Man residues in the backbone, along with Gal branches [Capek et al., 2000; Hannuksela and Penhoat, 2004; Harris and Stone, 2008]. Gal sugars can attach to Man or Glc units of the backbone by α -1,6-linkages, with a preference for Man. However, Gal substitutions occur only in small quantities in SCWs. Conversely, SCW GGMs are highly acetylated at the Man residues in the backbone [Capek et al., 2000; Lundqvist et al., 2002; Willför et al., 2003] leading to their designation as acetylated GGMs (AcGGMs) [Yu et al., 2022]. In contrast to the random pattern of AcGGM, PCW GGMs in eudicot species exhibit a regular structured pattern, referred as β -GGM (see chapter 1.1) [Yu et al., 2022]. The seed mucilage GGM in *Arabidopsis thaliana* differs from other eudicot PCW GGMs due to the absence of β -1,2-Gal substitution. This is likely due to the fact that β -Gal substitutions are removed from GGMs by β -galactosidase [Dean et al., 2007; Yu et al., 2022]. Mucilage-Modified 2 (MUM2) is highly expressed in the seed mucilage and is able to remove β -1,2-Gal substitution even from β -GGMs as shown for *mum2* mutants that exhibit minor but typical

motifs for β -GGM oligosaccharides. This emphasizes that GGMs found in the mucilage of *Arabidopsis thaliana* are another example of β -GGM [Yu et al., 2022].

1.4 Heteromannan Biosynthesis

In order to synthesize HMs and facilitate their integration into the cell wall, three major steps are necessary. Firstly, activated sugar nucleotides must be produced in the cytosol and further transported to the Golgi apparatus. Secondly, glycosyltransferases (GTs), which are embedded within the luminal membrane of the Golgi, synthesize cell wall polysaccharides using activated sugar nucleotides. Finally, HMs must be transported via vesicles to the plasma membrane, where polysaccharides can become incorporated in the extracellular space of the cell wall [Voiniciuc, 2022]. A schematic overview of HM synthesis is illustrated in Figure 2.

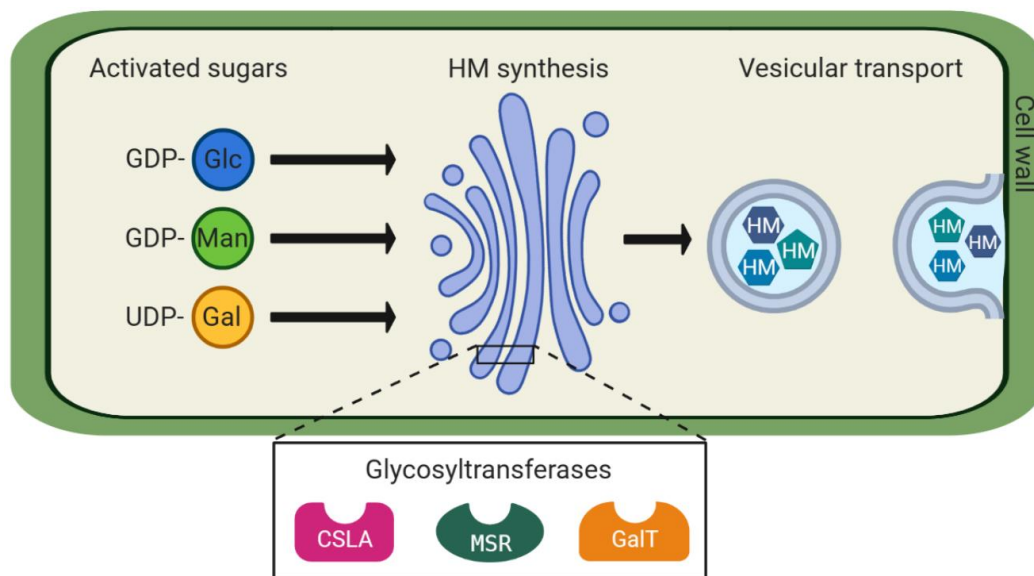


Figure 2: Schematic overview of heteromannan biosynthesis. Activated sugar nucleotides are transported into the Golgi apparatus where Cellulose Synthase-like A (CSLA), Mannan Synthase-Related (MSR) and Galactosyltransferases (GalT) are involved in HM synthesis. Afterwards, HMs are transported by vesicular transport to the plasma membrane to be incorporated in the cell wall. The figure was prepared with bioRENDER (biorender.com).

1.4.1 Generation of Activated Nucleotide Sugars for Heteromannan Synthesis

GTs catalyse the transfer of sugar moieties from activated nucleotide sugars (donors) to specific acceptor substrates. Sugar precursors are linked with either guanine diphosphate (GDP) or uridine diphosphate (UDP) to achieve their activated conformation [Lairson et al., 2008; Oikawa et al., 2013]. *In vitro* experiments with extracted enzymes possessing mannan and glucomannan synthase activity have revealed that mannan is elongated using GDP-Man, whereas glucomannan elongation needs additionally GDP-Glc and act competitively in the presence of GDP-Man [Elbein, 1969; Piro et al., 1993; Liepman et al., 2005]. Interestingly, the sugar donor UDP-Glc that is most likely used as a substrate for β -1,4-glucan backbone

elongation of cellulose [Pear et al., 1996], xyloglucan [Faik et al., 2002] and mixed-linkage glucans [Gibeaut and Carpita, 1993], cannot be utilized for glucomannan synthesis *in vitro* [Elbein, 1969; Yang et al., 2020].

Important branch point molecules for UDP- or GDP-nucleotide sugar formation are fructose-6-phosphate (Fru6P), glucose-1-phosphate (Glc1P) and mannose-1-phosphate (Man1P) that originate from products generated by photosynthesis or starch degradation (sucrose, triose phosphates and free Glc). Fru6P can be transformed into Glc1P and Man1P [Weber et al., 2000; Weber and Linka, 2011; Verbančič et al., 2018] which are both needed for glucomannan backbone formation. The enzyme UDP-Glucose Pyrophosphorylase (UGP) is able to convert Glc1P into UDP-Glc [Kleczkowski et al., 2010; Bar-Peled and O'Neill, 2011]. Although UDP-Glc is an important precursor for other UDP-linked nucleotide sugars in the cytosol [Bar-Peled and O'Neill 2011], it is still unknown how or whether UDP-Glc can be interconverted into GDP-Glc [Voiniciuc, 2022]. Furthermore, UDP-Glc can also be interconverted into UDP-Gal by UDP-Glucose 4-Epimerase (UGE) in the cytosol. In higher plants, UDP-Gal can be additionally generated from galactose-1-phosphate (Gal1P) by UDP-Sugar Pyrophosphorylase (USP). Compared to UGP, USPs are able to convert many sugar-1-phosphates into the respective UDP-sugar, but this enzyme has a higher preference for Glc1P and Gal1P [Kotake et al., 2004; Litterer et al., 2006; Altenhammer et al., 2020]. The Gal unit of UDP-Gal can then be transferred onto mannan or glucomannan backbones to form galactomannans or GGMs [Edwards et al., 1999; Yu et al., 2018].

The transportation of nucleotide sugars from the cytosol into the Golgi is crucial for HM synthesis [Davis et al., 2010] and is facilitated by many nucleotide sugar transporters (NSTs) [Temple et al., 2016; Verbančič et al., 2018]. Among them, the Golgi-Localized Nucleotide Sugar Transporter 1 (GONST1) in *Arabidopsis thaliana* was the first identified NST capable of transferring GDP-Man and GDP-Glc into the Golgi [Baldwin et al., 2001; Mortimer et al., 2013]. Based on their sequence similarity to GONST1, four additional GONST proteins (GONST 2–5) were discovered [Handford et al., 2004]. Among these proteins, only GONST2 has been demonstrated to transport GDP-Man into the Golgi *in vitro* [Handford et al., 2004; Rautengarten et al., 2016; Sechet et al., 2018; Jing et al., 2021]. However, *Arabidopsis thaliana* mutants lacking GONST1, GONST2, or both do not show reduced (gluco)mannan content in the cell wall but rather affect the mannosylation of glycosylinositol phosphorylceramides [Mortimer et al., 2013; Jing et al., 2021]. Therefore, Jing et al. (2021) proposed that CSLAs do not rely on NST for GDP-Man translocation or alternatively, the transporter lack the typical canonical GDP-binding motif (GX[L/V]NK) and therefore may be undiscovered so far.

In order to synthesize galactomannans or GGMs in the Golgi, UDP-Gal has to be transported into the Golgi lumen. The first transporter in plants was identified by gene

expression of *AtUTr1* from *Arabidopsis thaliana* in yeast cells and is capable of transferring UDP-Gal and UDP-Glc [Norambuena et al., 2002]. Since then, several other UDP-Gal transporters were identified in *Arabidopsis*, including UDP-Gal Transporter (*AtUTr2*, 3 and 7 [Norambuena et al., 2005; Reyes et al., 2010; Handford et al., 2012], *AtUDP-GalT1* [Bakker et al., 2005]) and UDP-Rha/Gal Transporter (*AtURGT1–6*) [Bakker et al., 2005; Rollwitz et al., 2006; Rautengarten 2014]. However, Reyes et al. (2006 and 2010) have shown that *AtUTr1* and *AtUTr3* transport UDP-Gal into the endoplasmic reticulum (ER) instead of the Golgi lumen. Although all the other transporters are capable of transferring UDP-Gal from the cytosol into the Golgi lumen *in vitro*, no direct association with galactomannan or GGM synthesis has been demonstrated in *Arabidopsis* mutants so far. A possible explanation might be that different UDP-Gal transporter can compensate for each other's function, or that UDP-Gal transporters can be specific to a particular polysaccharide and form complexes in the Golgi [Handford et al., 2012; Rautengarten et al., 2014; Saez-Aguayo et al., 2021]. An overview about the generation of activated nucleotide sugars from precursors derived from photosynthesis or starch degradation and their transport into the Golgi is summarized in Figure 3.

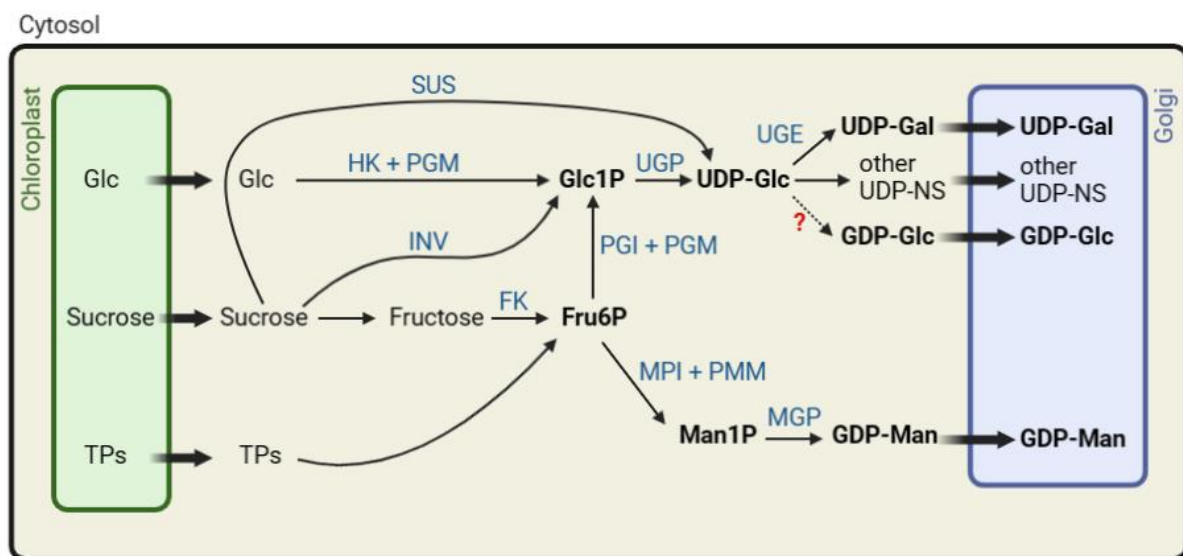


Figure 3: Generation of activated sugar nucleotides involved in heteromannan synthesis. Sugar precursors derived from photosynthesis or starch degradation are transported from the chloroplast into the cytosol. The production of activated nucleotide sugars take place in the cytosol. After nucleotide sugars were produced, they need to be translocated to the Golgi apparatus for HM biosynthesis. Fru6P: fructose-6-phosphate; GDP: guanidine diphosphate; Glc: glucose; Glc1P: glucose-1-phosphate; HK: hexokinase; INV: Invertase; Man1P: mannose-1-phosphate; MGP: mannose-1-phosphate guanylyltransferase; MPI: mannose-6-phosphate isomerase; NS: nucleotide sugars; PGI: phosphoglucose isomerase; PGM: phosphoglucomutase; PMM: phosphomannomutase; SUS: sucrose synthase; TPs: triose phosphates; UDP: uridine diphosphate; UGE: UDP-glucose glucose 4-epimerase; UGP: UDP-glucose pyrophosphorylase. Figure was created using information and illustrations of Weber and Linka (2011) and Verbančić et al. (2018). The figure was prepared with bioRender.

1.4.2 Elongation of Heteromannans

The elongation of the HM backbone is catalysed by proteins from the Cellulose Synthase-like A (CSLA) family. This family belongs to the carbohydrate active enzyme (CAZy) family GT2 [Cantarel et al., 2009; Liepman and Cavalier; 2012] comprising all members of the cellulose synthase superfamily that are involved in the synthesis of cell wall polysaccharides containing a β -linked backbone [Yin et al., 2009] (more details in chapter 1.5). All proteins of this superfamily are embedded in the Golgi apparatus [Oikawa et al., 2013] and have multiple transmembrane domains (TMDs) spanning across the Golgi [Richmond and Somerville, 2000; Davis et al., 2010]. Dependent on whether the catalytic active domain of these GTs faces the cytosolic or luminal side of the Golgi, the cytosolic or luminal nucleotide sugar pool can likely be used for the synthesis of the respective polysaccharide. In case of CSLA9 from *Arabidopsis thaliana* (*AtCSLA9*), Davis et al. (2010) showed that the catalytic active domain of this protein faces the Golgi lumen, leading to the assumption that *AtCSLA9* utilizes activated nucleotide sugars from the Golgi [Davis et al., 2010].

Another protein class involved in mannan biosynthesis was discovered by deep EST profiling of fenugreek seed endosperm and is “likely to be a glycosyltransferase”. Therefore, this protein class was named Mannan Synthesis-Related (MSR) [Wang et al., 2013]. In *Pichia pastoris*, HMs cannot be produced solely by heterologously expressed MSR1 [Voiniciuc et al., 2019]. However, co-expression of MSR1 with CSLA enzymes enhances glucomannan synthesis [Voiniciuc et al., 2019; Robert et al., 2021]. CSLA3 from *Amorphophallus konjac* (*AkCSLA3*) is a glucomannan synthase and the co-expression with MSR1 from *Arabidopsis thaliana* (*AtMSR1*) increase the glucomannan content without changing its Glc:Man ratio in yeast. Moreover, *AtMSR1* enables the incorporation of Glc into the mannan backbone that is produced by *AtCSLA2* in *Pichia pastoris*, leading to the conclusion that MSR proteins modulate CSLA activity to produce distinct types of HM [Voiniciuc et al., 2019]. The manner in which MSR proteins modulate the activity of CSLA enzymes remains unknown. Whether MSRs can glycosylate CSLA enzymes or physically interact with CSLA proteins to influence their activity [Voiniciuc et al., 2019].

1.4.3 Substitution of Heteromannans and Transport to the Plasma Membrane

Galactosyltransferases (GalT) belonging to the GT34 family decorate mannan and glucomannan backbones with α -1,6-linked Gal. The first characterized GalT involved in galactomannan synthesis was discovered in fenugreek seed endosperms (*TfGMGT*) by Edwards et al. (1999). Manno-oligosaccharide acceptors must have a chain length of five to eight sugar residues. *TfGMGT* recognizes six Man residues of the elongating mannan backbone polysaccharide and is able to transfer the Gal unit specifically to the third Man unit

[Edwards et al., 2002]. The closely related mannan α -Galactosyltransferase1/Mucilage-Related10 (MAGT1/MUCI10) decorates glucomannan in *Arabidopsis thaliana* seed coat mucilage [Voiniciuc et al., 2015b; Yu et al., 2018]. Yu et al. (2018) showed that AtMAGT1 decorates the mannose residues of an alternately Glc-Man patterned glucomannan backbone. It is also possible that other MAGTs exist that specifically transfer Gal sugars to other structural organised HM backbones, but these are still unknown [Yu et al., 2022]. In 2022, Yu et al. identified a new GalT that belongs to the GT47 family. This enzyme is capable of attaching Gal to the α -1,6-Gal side branches of glucomannan through β -1,2-linkages. This function was previously only known in xylosyltransferases. The unique feature of the enzyme prompted the researchers to name it Mannan β -Galactosyltransferase1 (MBGT1). Compared to CSLAs, GalTs have only one TMD spanning across the Golgi that is closely localized to the N-terminal end of the protein [Edwards et al., 1999; Keegstra and Raikhel, 2001; Keegstra and Cavalier, 2011]. Because their catalytic domain is localized within the Golgi lumen and UDP-Gal is likely produced in the cytosol, UDP-Gal has to be transported from the cytosol into the Golgi apparatus [Keegstra and Raikhel, 2001].

Ac groups are transferred to several matrix polysaccharides including mannan and glucomannan by members of the Trichome Birefringence-like (TBL) family. *In vitro* acetyltransferase activity assays using radioactive-labelled acetyl-CoA (donor) and mannohexaose (Man₆; acceptor) verified four *Arabidopsis thaliana* (TBL23–26, renamed as AtMOAT1–4) and one *Amorphophallus konjac* (AkMOAT1) recombinant protein(s) involved in acetylation of mannosyl residues at position O-2 and O-3 [Zhong et al., 2018]. The catalytic properties to acetylate the mannan backbone are different for these enzymes. While AtMOAT1, 2, 3 have a higher affinity towards mannohexaose, AtMOAT3, 4 and AkMOAT1 have a much higher O-acetyltransferase activity compared to AtMOAT1 and 2 [Zhong et al., 2018]. In general, the specific branching pattern of HM side chains is not well understood. Possible functions can be: (i) binding to other polysaccharides such as cellulose microfibrils (structural function) [Yu et al., 2018 and 2022], (ii) recognition site for cell wall modifying enzymes such as mannanases (developmental function and important for remodelling) [Bååth et al., 2018], or (iii) determining their biological activity or antioxidant efficiency [Chokboribal et al., 2015; Basa et al., 2020].

GTs involved in polysaccharide synthesis have been reported to be distributed in different Golgi cisternae, along with their respective products. The complexity of branching in polysaccharide chains appears to increase as they progress from the cis- to trans-Golgi cisternae [Zhang and Staehelin, 1992; Parsons et al., 2019]. To prevent aggregation of very long and linear polysaccharides during transport between cisternae, side branches that enhance the polymers solubility must be added simultaneously [Zabotina et al., 2021]. Protein-

protein interaction studies focusing on enzymes involved in XyG backbone (CSLC4) and side chain (XXT1, XXT2, XXT5, MUR3, XLT2, FUT1) synthesis have demonstrated strong interaction among these enzymes, forming homodimers or heterocomplexes [Chou et al., 2012 and 2015; Lund et al., 2015]. This leads to the hypothesis that GTs are functionally organized throughout the Golgi and are synthesized continuously in a stepwise manner during Golgi maturation [Hoffmann et al., 2021; Zabolina et al., 2021], which could also be the case for proteins involved in HM synthesis.

The transport of HMs from the trans Golgi network (TGN) to the plasma membrane is considered to occur through vesicles (conventional secretory pathway) as it was already shown for other non-cellulosic polysaccharides, such as XyG and pectin [Toyooka et al., 2009; Gendre et al., 2015; Wilkop et al., 2019]. In general, this process involves the selective sorting and packaging of polysaccharides into vesicles, aided by specific cargo receptors and sorting signals at the TGN [Kang et al., 2011]. After sorting, the vesicles bud from the TGN and target the plasma membrane along the cytoskeleton (specifically microtubules and actin filaments) [Tamura et al., 2005; Kong et al., 2015]. Once the vesicles arrived at the plasma membrane, vesicles tether and dock to the plasma membrane with the help of Snap receptors (SNAREs) and small Rab GTPase proteins. After docking, the vesicles fuse with the plasma membrane, polysaccharides are released into the extracellular space, and become incorporated into the complex network of cell wall polysaccharides [reviewed in Worden et al., 2012; Kim and Brandizzi, 2014; Gendre et al., 2015; Sinclair et al., 2018]. Nevertheless, it should be mentioned that non-conventional secretion pathways of proteins bypassing the route through the ER and Golgi have been reported [Poulsen et al., 2014; van de Meene et al., 2017; Wang et al., 2017] and could also be the case for GTs involved in HM synthesis.

1.5 The Cellulose Synthase Superfamily

The cellulose synthase superfamily is involved in the synthesis of cellulose and hemicelluloses. The genes of this superfamily are divided into 12 gene families including one *Cellulose Synthase (CesA)* and 11 *Cellulose Synthase-Like (Csl)* families which are present in all land plants [Yin et al., 2009; Little et al., 2018; Pancaldi et al., 2022]. All gene classes of the CesA superfamily encode GT2 enzymes involved in β -1,4-linked polysaccharide backbone synthesis. The first discovered gene family was *CesA* which synthesizes cellulose. This was identified in cotton through the sequence similarity of *CesA1* and *CesA2* to a bacterial *CesA* gene [Pear et al., 1996]. The best studied *CesA* gene family is that from *Arabidopsis thaliana* and contains ten different genes [Richmond, 2000]. The *AtCESA* proteins are synthesized in the ER, and multiple *AtCESA* proteins assemble as cellulose synthase complexes (CSCs) with a sixfold symmetric rosette structure in the Golgi apparatus [Kimura et al., 1999; McFarlane et al., 2014]. Afterwards, the CSCs are transported to the plasma membrane where cellulose

synthesis takes place. *AtCESA1*, 3 and 6 form a complex that is involved in PCW cellulose synthesis [Wang et al., 2006; Desprez et al., 2007; Carroll et al., 2012], while *AtCESA4*, 7 and 8 form a separate CSC specific for cellulose synthesis in SCWs [Turner and Somerville, 1997; Taylor et al., 2003; McFarlane et al., 2014].

The discovery of the *CesA* gene family set the basis to find other gene families involved in non-cellulosic polysaccharide synthesis. In 2000, Richmond and Somerville discovered 31 *CesA-like* genes in *Arabidopsis thaliana* based on sequence similarity to *CesA1* from *Arabidopsis thaliana* and cotton *CesA* and were grouped into six subfamilies (*CsIA–E* and *CsIG*). Since then, *CsI* genes have been found in several plant species and uncover other *CsI* gene families as *CsIF*, *H*, *J*, *K* and *M* (Figure 4; *CsIM* clade is not included in the phylogenetic tree) [Fincher, 2009; Yin et al., 2014; Little et al., 2018]. *CsIF*, *H* and *J* genes mediate the synthesis of MLGs [Burton et al., 2006; Doblin et al., 2009; Little et al., 2018]. Despite *CsIB* is the dicot-specific sister clade of *CsIH*, and *CsIM* is closely related to *CsIJ*, no function has yet been attributed to *CsIB* and *CsIM* genes [Little et al., 2018]. Interestingly, *CsIF* and *H* are found only in monocots, while *CsIB* and *G* appear to be confined to dicots. All other *CsI* classes occur in both monocots and dicots, suggesting that *CsIF/H* and *CsIB/G* occurred after the evolutionary divergence of these two plant groups [Daras et al., 2021].

The *CsID* clade's close relation to *CsIF* and the similarity of structural domains and motifs of CSLD and CESA proteins [Daras et al., 2021] suggest a function in MLG or cellulose synthesis. However, attributing a specific function to CSLD proteins has been challenging due to their association with the synthesis of several polysaccharides as xylan and homogalacturonan [Bernal et al., 2007], cellulose [Park et al., 2011; Wang et al., 2011; Yang et al., 2020], and mannan [Verhertbruggen et al., 2011; Yin et al., 2011; Gu et al., 2016]. While mannan synthase activity was only detected in CSLD proteins upon overexpression [Verhertbruggen et al., 2011; Yin et al., 2011] and the level of synthesized mannan was much lower than for CSLA proteins [Liepman et al., 2005; Goubet et al., 2009; Yang et al., 2020], the question arose if CSLD accumulation impacts its activity [Gu et al., 2016]. Emerging evidence increasingly points towards a role for CSLD in cellulose synthesis. First, both CSLD3 and CESA6 are able to use UDP-Glc (and not GDP-Man) as a sugar substrate *in vitro* [Yang et al., 2020]. Second, genetic complementation experiments, wherein the swapped catalytic domains of CESA6 and CSLD3 successfully rescue the phenotypical defects of *csld3*-null or *cesa6*-null mutants [Park et al., 2011; Yang et al., 2020]

As already mentioned, CSLA proteins are involved in mannan or glucomannan elongation [Dhugga et al., 2004; Liepman et al., 2005 and 2007; Suzuki et al., 2006; Gille et al., 2011], while CSLCs synthesize the β -glucan backbone of XyGs [Cocuron et al., 2007]. Both of these families are distantly related to the other *CsI* family classes [Yin et al., 2009]. A

1.6 Comparative Analysis of GGM and XyG Synthesis

Oligosaccharide profiling of *Arabidopsis thaliana* mutants by Yu et al (2022) indicated that CSLA2 and CSLA9 are required for the synthesis of two GGMs with a different structural pattern. While CSLA9 synthesizes AcGGM (random Glc and Man distribution in backbone) in PCW-rich tissues as leaves, CSLA2 is involved in β -GGM synthesis (alternating Glc and Man disaccharides in backbone) [Yu et al., 2022; Grieb-Osowski and Voiniciuc, 2023]. The overall structure of β -GGM is similar to that observed for XyG (Figure 5), indicating a functional overlap between these polysaccharides. Therefore, the specificities of several GTs involved in GGM and XyG synthesis, the polymers structure and functional roles *in planta* are summarized in this chapter. This chapter reproduce several text parts and the figure that was published by Grieb-Osowski and Voiniciuc (2023).

In *Arabidopsis thaliana* five CSLC genes (*CSLC4*, 5, 6, 8 and 12) are functionally redundant for the elongation of XyG [Kim et al., 2020], while at least four of the nine CSLA genes (*CSLA2*, 3, 7 and 9) encode (gluco)mannan synthases (Voiniciuc, 2022). Both CSLC and CSLA belong to the GT2 family and are integral membrane proteins with multiple TMDs spanning the Golgi. However, their active domain is predicted to be orientated on opposite sides (*CSLC4*: cytosol; *CSLA9*: Golgi lumen), which would influence the nucleotide sugar accessibility and the need for translocation [Davis et al., 2010]. Another difference is the usage of activated nucleotide sugars of the respective CSL class. While CSLCs use UDP-Glc for XyG elongation, CSLAs use GDP-Man and GDP-Glc for (gluco)mannan backbone synthesis [Davis et al., 2010]. Structural predictions of CSLCs identified a “VED” amino acid motif instead of a “TED” sequence which is specific for Glc elongation and translocation of cellulose synthases [Julian and Zobotina, 2022]. All active (gluco)mannan synthases share this VED motif with CSLCs and a number of other highly conserved motifs [Robert et al., 2021], suggesting that CSLAs and CSLCs could elongate polymers via similar mechanisms. An exchange of the catalytic domains of *Arabidopsis thaliana* *CSLA2* and *Tropaeolum majus* (nasturtium) *CSLC4* expressed in *Pichia pastoris* was not functional compared to the parental lines [Robert et al., 2021], which illustrates that further structural and/or *in vitro* characterization are necessary to understand the biochemical function of CSLAs and CSLCs completely.

While *CSLA2*, *CSLA9* and *MSR1* influence the structural pattern of GGM backbones, the underlying mechanistic pathway remains elusive. Through heterologous expression of *AtMSR1* (GT106 family) and *AtCSLA2* in *Pichia*, Glc units can be incorporated into the mannan backbone and glucomannan synthesis is enhanced [Voiniciuc et al., 2019, Robert et al., 2021]. However, the Glc:Man ratio has not been able to reach the high molar ratio observed in seed mucilage [Voiniciuc et al., 2015b; Yu et al., 2018]. Consequently, it is likely that additional

proteins function together *in vivo* to define the GGM pattern. A homologous gene belonging to GT106 family involved in XyG synthesis is not known, yet.

The backbones of mannans and XyGs can be both substituted with α -1,6-Gal or -Xyl and is carried out by GTs of the 34 family. MAGT/MUC10 uses UDP-Gal as an activated nucleotide sugar to transfer single Gal units onto Man residues of the HM backbone [Voiniciuc et al., 2015b; Yu et al., 2018], whereas XXTs use UDP-Xyl for the transfer of Xyl that is regularly distributed onto the XyG backbone [Culbertson et al., 2018]. Several XXTs can add additional Xyl residues on glucan oligosaccharides *in vitro* or even use other donor sugars such as UDP-Gal [Julian and Zabolina, 2022]. This suggests that GT34 activity in plant cells is partly determined by donor substrate availability, and that the mannan-related enzyme specificities should be further explored. The further substitution with β -1,2-Gal onto α -1,6-Xyl side chains in XyG or α -1,6-Gal side chains in β -GGM is synthesized by GTs of the 47 family [Pauly and Keegstra, 2016; Yu et al., 2022]. β -galactosylation of XyG is catalysed by MURUS3 (MUR3) and XyG L-Side Chain Galactosyltransferase Position2 (XLT2) [reviewed by Pauly and Keegstra, 2016]. The structure of GGM oligosaccharides was more challenging to decipher. Already in 2015, the hypothesis of β -1,2-Gal- α -1,6-Gal disaccharides decorations onto glucomannan in *Arabidopsis* mucilage was raised by Voiniciuc et al. (2015b) due to the significant reduction of terminal-Gal and 2-Gal linkages detected in the mucilage of *muci10* mutant seeds. In 2022, Yu et al. showed that MBGT1 is responsible for β -galactosylation onto GGM in *Arabidopsis thaliana*. A schematic overview of GTs involved in GGM and XyG synthesis and the structural arrangement of sugar residues in both polymers is illustrated in Figure 5.

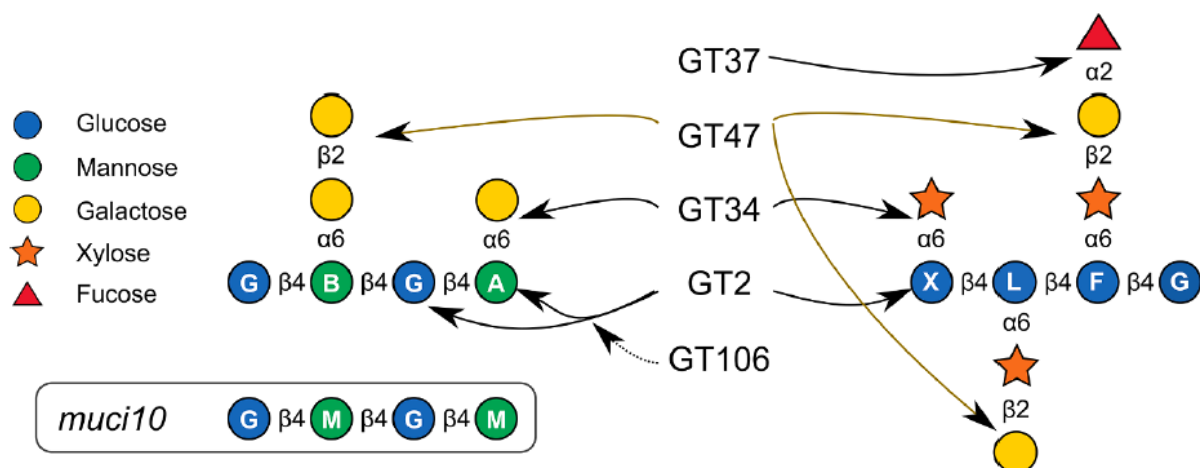


Figure 5: Schematic overview and comparison of GT families involved in β -GGM and XyG synthesis. The highly branched structure of β -GGM is similar to the structure of XyG and requires enzymes from the same GT family. The single-letter nomenclature (described in Pauly and Keegstra, 2016; Julian and Zabolina, 2022; Yu et al., 2022) denotes if the specific sugar residue in the backbone is unbranched (G for Glc; M for Man) or branched with several side chains (A for α -1,6-Gal; B for β -1,2-Gal- α -1,6-Gal; X for α -1,6-Xyl; L for β -1,2-Gal- α -1,6-Xyl; F for α -1,2-Fuc- β -1,2-Gal- α -1,6-Xyl). Figure was taken from Grieb-Osowski and Voiniciuc (2023).

The elongation of insoluble β -1,4-glucan by CSLC4 in yeast was boosted by XXT1 co-expression [Cocuron et al., 2007; Robert et al., 2021]. Similarly, it is possible that MAGT/MUCI10 and CSLA2 can form a complex to modulate its stability or activity for GGM synthesis in *Arabidopsis* seed mucilage because the glucomannan content is strongly reduced in *muci10* mutants [Voiniciuc et al., 2015b]. The *csla2* and *muci10* mutants exhibit an increased density of the mucilage and have reduced amounts of crystalline cellulose relatively to wildtype plants [Voiniciuc et al., 2015b]. Conversely, mucilage stainings of XyG-deficient *xxt1 xxt2 xxt5* triple mutants are comparable to the wildtype mucilage and show less severe cellulose defects than GGM-deficient seeds [Voiniciuc et al., 2015a; Yang et al., 2022]. Therefore, mannans (akin to XyG in tip-growing cells [Kim et al., 2020]) have unique roles in some tissues [Voiniciuc, 2022].

Despite XyG is absent in *cs/c* quintuple mutants [Kim et al., 2020] and *xxt1 xxt2* double mutants [Zabotina et al., 2012], these mutants show only minor growth defects, suggesting that other polysaccharides can compensate for their loss. However, this hypothesis has lacked substantial evidence until recently. Neither altered CSLA gene expression [Kim et al., 2020; Sowinski et al., 2022] nor altered AcGGM or β -GGM amounts [Yu et al. 2022] were reported in *xxt1 xxt2* double mutants. The elevated levels of glucomannan observed in rosette leaf cell walls, as reported by Sowinski et al. (2022) are likely a consequence of altered polysaccharide turnover rather than an increase in polymer biosynthesis rates. This can alter the polysaccharide abundance in the wall and help to maintain the structural integrity of the cell wall.

csla2 xxt1 xxt2 triple mutants, which exhibit reduced cell elongation and disrupted cellulose arrangement in the hypocotyl as well as shorter siliques and plant height in the absence of β -GGM and XyG, exceeds the phenotypes lacking only β -GGM or XyG [Yu et al., 2022; Grieb-Osowski and Voiniciuc 2023]. Moreover, double mutants involved in β -galactosylation of β -GGM (*mannan β -galactosyltransferase 1 (mbgt1)*) and XyG (*mur3-3 (mur3-3)*) in *Arabidopsis* plants show reduced rosette size, shorter inflorescence stems and a more drastic cabbage-like phenotype compared to the respective single mutants. These observations suggest that MBGT1 can partially compensate the *mur3-3* phenotype [Yu et al., 2022; Grieb-Osowski and Voiniciuc, 2023]. These observations made in higher order mutants provide strong evidence in overlapping function of GGM and XyG.

1.7 *Pichia pastoris* as a Model Organisms for HM Elongation

Many GTs involved in the synthesis of plant cell wall polysaccharides have been identified, and in some content characterized *in planta* by numerous studies using molecular biology, reverse genetics, and genomics. All these studies helped to gain better insights into

the function of proteins in cell wall formation and polysaccharide synthesis [Zabotina et al., 2021], but primarily due to the following obstacles it is quite challenging to study the protein's structure, biological role and biochemical activity directly in the respective plant system. Firstly, the cell wall is a dynamic entity whose composition is diverse in several cell types and can be modified during plant development and in response to environmental (abiotic and biotic) conditions [Houston et al., 2016; Hoffmann et al., 2021]. Therefore, it is essential to consider the timepoint of gene expression and the protein's location before the protein is extracted and purified from plant material. Furthermore, the intricate interplay between different metabolic pathways can complicate the elucidation of a precise biochemical role of the protein. Secondly, to conduct biomolecular assays and structural analyses of GTs, a large amount of enzyme is required. Obtaining these enzymes from plant material is difficult as the proteins occur in relatively low quantities and are not easy to extract and purify [Smith, 2016]. Thirdly, the analysis of GT mutants can be complicated if deletion of these genes causes lethality as shown for *AtCSLA7* or if several isoforms from the same gene family have similar functions that can compensate the mutant phenotype (functional redundancy) [Goubet et al., 2009]. Fourthly, it is still unclear whether the sugar nucleotide pool affects the structural pattern of the specific polysaccharide, and if so, how the sugar substrates specify the polysaccharide structure. Analysing the impact of sugar substrates directly in the plant system would be overly complex due to the high number of different nucleotide sugars (approximately 30) and their versatile interconversion pathways [Bar-Peled and O'Neill, 2011].

Expressing candidate genes involved in cell wall polysaccharide synthesis in a heterologous expression host can not only alleviate many of the mentioned limitations of the corresponding plant system, but also facilitate deeper insights into the enzymes specificities and their biological roles in the complex process of cell wall synthesis. To date, many host systems have been used for recombinant protein production ranging from prokaryotic to eukaryotic systems. All these expression host systems have advantages and disadvantages and their suitability has to be considered for the specific protein and its purpose [Gomes et al., 2016], but predictions for the successful establishment can only be made to a certain extent.

As the most basic requirements for recombinant protein expression, the expression host must have a clean background and possess nucleotide sugars which can be used as building blocks for HM synthesis [Pauly et al., 2019]. A clean background refers to an expression host devoid of naturally occurring HMs or HM-degrading enzymes that could potentially overlay or interfere with the HMs produced by the recombinant protein. Furthermore, it would be of great interest if transformations are easy and quick to perform. This can be achieved by modular cloning strategies as Golden Gate compatible cloning [Weber et al., 2011] with a high number of vectors, promoters, terminators and selection markers as well

as many commercially available mutant collections optimized for special applications as glycoengineered strains [Jacobs et al., 2008; Piirainen et al., 2022]. Decoupling biomass production from protein expression by selecting inducible promoters can help to avoid toxicity of the produced polysaccharides [Zepeda et al., 2018].

Bacterial systems are easy to genetically modify, exhibit a rapid growth and the most commonly used strains have the lowest cultivation costs among various other expression hosts [Gomes et al., 2016]. However, bacterial systems lack the machinery for eukaryotic post-translational modifications such as forming disulfide bonds or adding O- and N-glycosylations [Gellisen, 2005], which are crucial for the proper folding and function of membrane-bound GTs, including the CSLA proteins [Nguema-Ona et al., 2014; Welner et al., 2017].

Compared to bacterial systems, yeasts can be engineered quickly, produce recombinant proteins at a high yield, and are capable of post-translational modification of these recombinant proteins. Since yeast meets all of the criteria described above, it's generally an excellent choice for the expression GTs [reviewed by Pauly et al., 2019]. The most commonly used yeasts for recombinant protein production are *Saccharomyces cerevisiae* and *Pichia pastoris* [Darby et al., 2012]. Additionally, *Pichia pastoris* was successfully used for *AkCSLA3* and *AtCSLA2* expression and resulted in glucomannan or mannan production [Voiniciuc et al., 2019]. As this thesis aims to analyse candidate CSLK and CSLA genes responsible for HM production, *Pichia* is the perfect candidate to reach my research aim.

1.8 Selection of CSLK/A Enzymes

The function in mannan or glucomannan elongation of many CSLA members throughout the plant kingdom was identified using heterologous expression of recombinant proteins in several hosts systems [Dhugga et al., 2004; Liepman et al., 2005 and 2007; Suzuki et al., 2006; Gille et al., 2011; Voiniciuc et al., 2019; Robert et al., 2021].

Amorphophallus konjac glucomannan is one of the best studied polysaccharides because of its broad range of industrial applications and its numerous health benefits as a dietary fibre [Zhong et al., 2005]. Although it is well-known that *AkCSLA3* catalyse the synthesis of the glucomannan backbone in the developing corm [Gille et al., 2011] as well as in the heterologous host *Pichia pastoris* [Voiniciuc et al., 2009], it still has to be validated how the enzyme determines the glucomannan structure in detail.

Arabidopsis thaliana contain nine CSLA family members (*AtCSLA1*, 2, 3, 7, 9, 10, 11, 14 and 15). Until now, it is still not known why so many CSLA genes are present in *Arabidopsis thaliana*. But it is hypothesized that all of these family members are involved in glucomannan synthesis probably in different tissues or developmental stages [Liepman et al., 2005]. Up to

date, only four of these isoforms (*AtCSLA2*, 3, 7 and 9) have been identified in the production of glucomannans [Liepman et al., 2005 and 2007; Goubet et al., 2009; Voiniciuc et al., 2019]. Most *csla* single mutants from *Arabidopsis thaliana*, except *csla7* that is embryo lethal, show a normal growth phenotype no matter if the glucomannan content in the inflorescence stems is reduced as shown for *csla9* single mutants or *csla239* triple mutants [Goubet et al., 2009]. This suggests that whether glucomannan produced by other *AtCSLA* members or other polysaccharides such as XyG can complement the mutant phenotypes in the stem. Another option of course would be that the specific *AtCSLA* isoform is involved in (gluco)mannan synthesis but in another tissue or developmental stage.

While *Arabidopsis* have multiple CSLA proteins which increase the complexity to examine their biological role in planta, other plant species encode significantly fewer CSLA proteins, and some possess only a single CSLA protein. The reduced complexity presents an opportunity for a more straightforward investigation of the respective biological roles of CSLA *in planta*.

Most studies on HM biosynthesis have focused on angiosperm and gymnosperm species, although HM has been regarded as the most ancestral hemicellulose in the plant kingdom [Scheller and Ulvskov, 2010]. Investigating species beyond these groups is crucial for understanding the evolution and diversification of the putative enzymes and their role in biosynthesis of distinct types of cell wall polysaccharides. In this context, the discovery of single-copy *CSLK* genes in several green (chlorophytes) and red (rhodophytes) algal species, which are ancestral to CSLA and CSLC (evolved in charophytes) [Yin et al., 2014], provides a compelling starting point. Studying the function of these ancestral CSLK in combination with CSLAs from both early- and late-diverging terrestrial plant species can shed light on the evolutionary journey of CSLK and CSLA enzymes. This offers the opportunity to understand how these cell wall polysaccharides adapted during the transition from aquatic to terrestrial environments.

In order to find CSLK/A enzymes involved in HM biosynthesis that could give insights into their structural diversity during evolution throughout the plant kingdom, the protein sequence of *AtCSLA2* was compared to other biological sequences using the BLAST search of several databases (see section 2.1). Furthermore, plant species known to incorporate mannan or glucomannan in their cell wall from the whole plant kingdom were specifically selected to determine the function of the respective CSLK. A schematic overview of all selected organisms which possess *CSLK* or *CSLA* genes and were analysed in this thesis is illustrated in Figure 6.

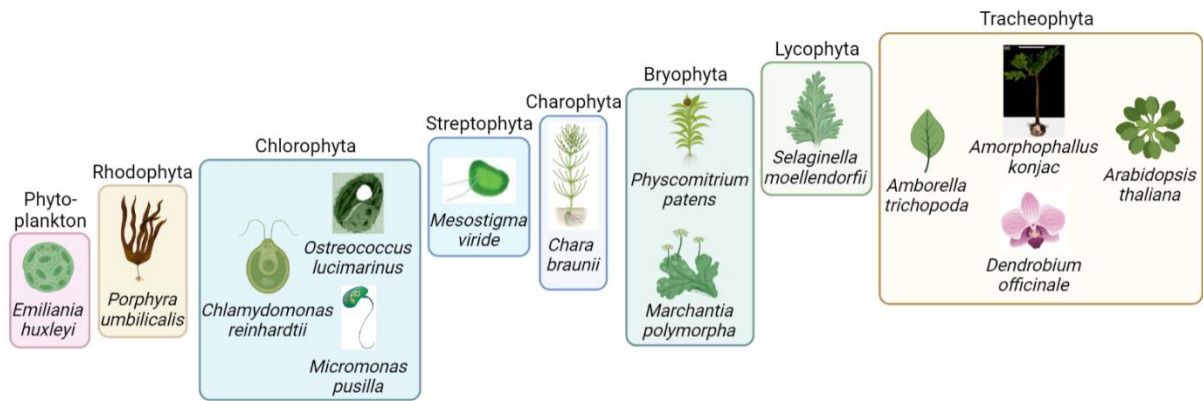


Figure 6: Schematic overview of selected plant species throughout the plant kingdom with sequence similarity to AtCSLA2. The figure was illustrated in bioRender.

1.9 Aim of my Thesis

In the past two decades several CSLA genes were discovered and characterized in HM biosynthesis. The structure of HMs used in industrial approaches such as *Amorphophallus konjac* glucomannan is studied quite well, but how the glucomannan structure is determined by AkCSLA3 still remains to be elucidated. The function of many CSLA members in HM biosynthesis is still unknown, because observations of the respective CSLA *in planta* can be subjected to a number of challenges such as lethality (shown for AtCSLA7) or redundancy. The individual investigation of the CSLA in a separate host system can help to elucidate their function separately. The identification of single-copy CSLK genes of green algae as well as selection of CSLA from many plant species in hierarchical order from aqueous to terrestrial plant colonization can help to elucidate the substrate specificity of CSLK/A during evolution. Therefore, the following questions were examined in this thesis:

- What is the function of the specific CSLK or CSLA enzyme?
- Is the glucomannan structure primarily determined by the specificity of CSLAs? Or can the sugar availability alter the HM product of CSLA?
- Can CSLA be translocated to the plasma membrane and retain their function in HM synthesis?
- Do CSLK and CSLA produce distinct types of HMs that evolved or adapt during evolution from aqueous to terrestrial plants?

2. Methods

2.1 Tools Used for Selection of CSLK and CSLA Sequences

The selection and analysis of CSLK and CSLA sequences required a systematic approach leveraging numerous databases and bioinformatic tools. The initial stage entailed selecting the sequences of all nine *At*CSLA isoforms from the Aramemnon database (<http://aramemnon.uni-koeln.de/>). This was crucial to design primers and clone the full-length coding sequence (CDS) of each *At*CSLA into the *Pichia* genome. Subsequently, conserved domains within the nine *At*CSLA isoforms were discovered using the motif discovery tool within the Multiple Em for Motif Elicitation (MEME) webtool (<https://meme-suite.org/>). The sequences were aligned with the MUSCLE function [Edgar, 2004], and sequences associated with sugar nucleotide coordination and sugar binding were analysed using the Molecular Evolutionary Genetics Analysis (MEGA; version 11) software (<https://www.megasoftware.net/>) [Tamura et al., 2021].

To identify CSLK and CSLA enzymes associated with HM biosynthesis and thereby shed light on their diversity across plant evolution, the *At*CSLA2 protein sequence was compared with other plant protein sequences. This involved utilizing the Basic Local Alignment Search Tool (BLAST) across several databases, such as the National Center for Biotechnology (NCBI) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), Phytozome (<https://phytozome-next.jgi.doe.gov/>), or PhycoCosm (<https://phycocosm.jgi.doe.gov/>).

Predictions for membrane protein topology and signal peptides in all in all CSLK and CSLA proteins considered in this study were made using TOPCONS (<https://topcons.net/>) [Tsirigos et al., 2015] prediction tool.

The *Pichia* Genome Atlas website (<http://pichiagenome-ext.boku.ac.at:8080/>) [based on Valli et al., 2020] was used to predict potential localizations of proteins within *Pichia* when different carbon sources such as Glc and methanol were used. Using this tool as a guide, proteins anticipated to be expressed in the Golgi were introduced into *Pichia*. This served as a comparative reference to confirm the Golgi localization of *Ak*CSLA3.

2.2 Cloning of Genes and Fluorescent Proteins into pPICZ-based Vectors

Genes and fluorescent proteins were cloned into vectors of the EasySelect *Pichia* expression system (Invitrogen™, Cat# K174001). Therefore, cDNA sequences of several CSLK/A proteins (without their natural stop codon) followed by the fluorophore sequence were integrated into the multiple cloning site of the pPICZB or pPICZX vector. Both vectors contain

a methanol-inducible promoter which starts the recombinant protein expression as well as a resistance marker (Zeocin) to select the transformants on plate. While the *PmeI* restriction site is present in the pAOX1 promoter of the pPICZ B vector, it has been removed in the pPICZ X vector.

In order to obtain transcripts from plant material (*Arabidopsis thaliana*: *AtCSLA1*, *AtCSLA3*, *AtCSLA14* from developing siliques, *Marchantia polymorpha*: *MpoCSLA* from leaf, *Physcomitrium patens*: *PpCSLA1–3* from protonemata), RNA was isolated using the NucleoSpin RNA Plant and Fungi Kit (Machery-Nagel; Cat# 740120.5), followed by transcription into cDNA with the SuperScript IV First Strand Synthesis System (Thermo Fisher Scientific, Cat# 10387979). Other *Arabidopsis* transcripts used in this study were previously generated by Cătălin Voiniciuc from plant cDNA.

Most of the coding sequences were amplified by Phusion PCR with the High-Fidelity Phusion DNA Polymerase (Thermo Fisher Scientific, Cat# F530L) 1 µl of DNA was mixed with 200 µM dNTPs, 0.02 U Polymerase, 10 µM forward and reverse primer and 1x Phusion HF buffer in a total volume of 20 µl. Primers for the amplification of Phusion PCR products were designed with the Multiple Primer Analyzer webtool from Thermo Fisher Scientific (<http://www.thermoscientificbio.com/webtools/multipleprimer/>) and are listed in Table S1 for cDNA and S2 for plasmid DNA templates. The PCR was performed in a thermocycler (BioRad Laboratories, Hercules, US) with one cycle for the initial denaturation (30 s, 98°C), five repetitive cycles for denaturation (10 s, 98°C), annealing (20 s, optimal temperature for primer specific combination calculated by Tm calculator (www.thermofisher.com/tmcalculator/)) and extension (20 s, 72°C), followed by 30 cycles of denaturation (30 s, 98°C) and extension (20 s, 72°C). The final extension step was performed for 5 min at 72°C and the cycler was cooled down to 12°C for storage.

Phusion products and 1 µl of a 1 kb DNA GeneRuler (Thermo Fisher Scientific, Cat# SM0311) were loaded onto 0.7% (w/v) agarose gels in 1x TAE buffer (50 mM Tris/HCl), 40 mM acetic acid, 1 mM EDTA, pH 8–8.3) mixed with ROTI®GelStain Red (Carl Roth, Cat# 0984.1) as described in the manufacturer's instructions. The amplified DNA was separated for 60 min at 100 V using a PowerPac™ Basic Power Supply (BioRad Laboratories, Hercules, US) and were visualized with a Fisherbrand™ Safeview Blue Light Transilluminator (Thermo Fisher Scientific, Waltham, US) to cut out the amplicon with the right size from gel.

The amplified PCR fragments were extracted from gel using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Cat# D4002) according to the manufacturer's instructions, instead that the DNA fragments were finally eluted in 20 µl HyClone™ water (Thermo Fisher Scientific, Cat# 10307052). 1 µl of each purified DNA was pipetted onto an Eppendorf µCuvette® G1.0 (Eppendorf, Cat# 6138000018) and the DNA concentration was determined

by measuring the absorption at 260 nm using a BioSpectrometer® basic (Eppendorf, Hamburg, DE). Additionally, other CSLK/A gene sequences with restriction sites for the introduction into pPICZ-based vectors were commercially synthesized by Geneart or Genescript (*Emiliania huxleyi*: EhCSLK, *Porphyra umbilicalis*: PumCSLK2, *Chlamydomonas reinhardtii*: CrCSLA, *Micromonas pusilla*: MicpuCSLK, *Mesostigma viride*: MvCSLK, *Chara braunii*: CbCSLK and *Selaginella moellendorffii*: SmCSLA).

All transcripts that were synthesized and made by Phusion PCR were designed that they include a specific restriction site at their N- and C-terminal end of the sequence that was used for ligation into pPICZ-based vectors. To integrate fluorescent proteins (Venus, mRuby2, mCitrine, 2A-sfGFP) into pPICZ-based vectors, the restriction sites of *Eco72I* and *XhoI* were used, while CSLK/A sequences were inserted into pPICZB + fluorescent protein vectors in front of the fluorescent tag by digestion with *Bsp119I* and *Eco72I* or only with *Eco72I*. Most of the synthesized transcripts were designed with a stop codon in the end of the C-terminal sequence of the specific CSLK/A. Only CrCSLK, MicpuCSLK, SmCSLA and MpoCSLA have not included a stop codon. Therefore, a STOP cassette was integrated into the pPICZ B vector and afterwards the four CSLK/A were integrated in front of the sequence of the STOP cassette. For this purpose, a Fusion PCR with two overlapping primers was performed to obtain the STOP cassette with restriction sites for *Eco72I* and *XhoI* on the terminal ends of the fragment that were used for ligation into the pPICZ B vector (Table S2). The four CSLK/A sequences were digested with *Bsp119I* and *Eco72I* (MicpuCSLK and SmCSLA), only *Bsp119I* (CrCSLA) or only *Eco72I* (MpoCSLA) for pPICZB + STOP cassette integration. A schematic overview of the cloning strategy to create pPICZ-based vectors integrating CSLA-Venus is illustrated as an example in Figure 2. Digestion of pPICZ-based vectors as well as DNA inserts were done with FastDigest® restriction enzymes from Thermo Fisher Scientific. While 200 ng of each DNA insert were mixed in a total volume of 20 µl with 1 µl of each restriction enzyme and 2 µl 10x FastDigest® buffer, 1 mg of vector DNA was mixed in a total volume of 30 µl with 1 µl FastAP Alkaline Phosphatase (Thermo Fisher Scientific, Cat# 10151000) for dephosphorylation, 1 µl of each restriction enzyme and 3 µl of 10x FastDigest® Green buffer (Thermo Fisher Scientific, Cat# B72). Digestion mixes were incubated at 37°C for 15 min, followed by an inactivation of the enzymes at 80°C for 5–10 min according to the manufacturer's instructions.

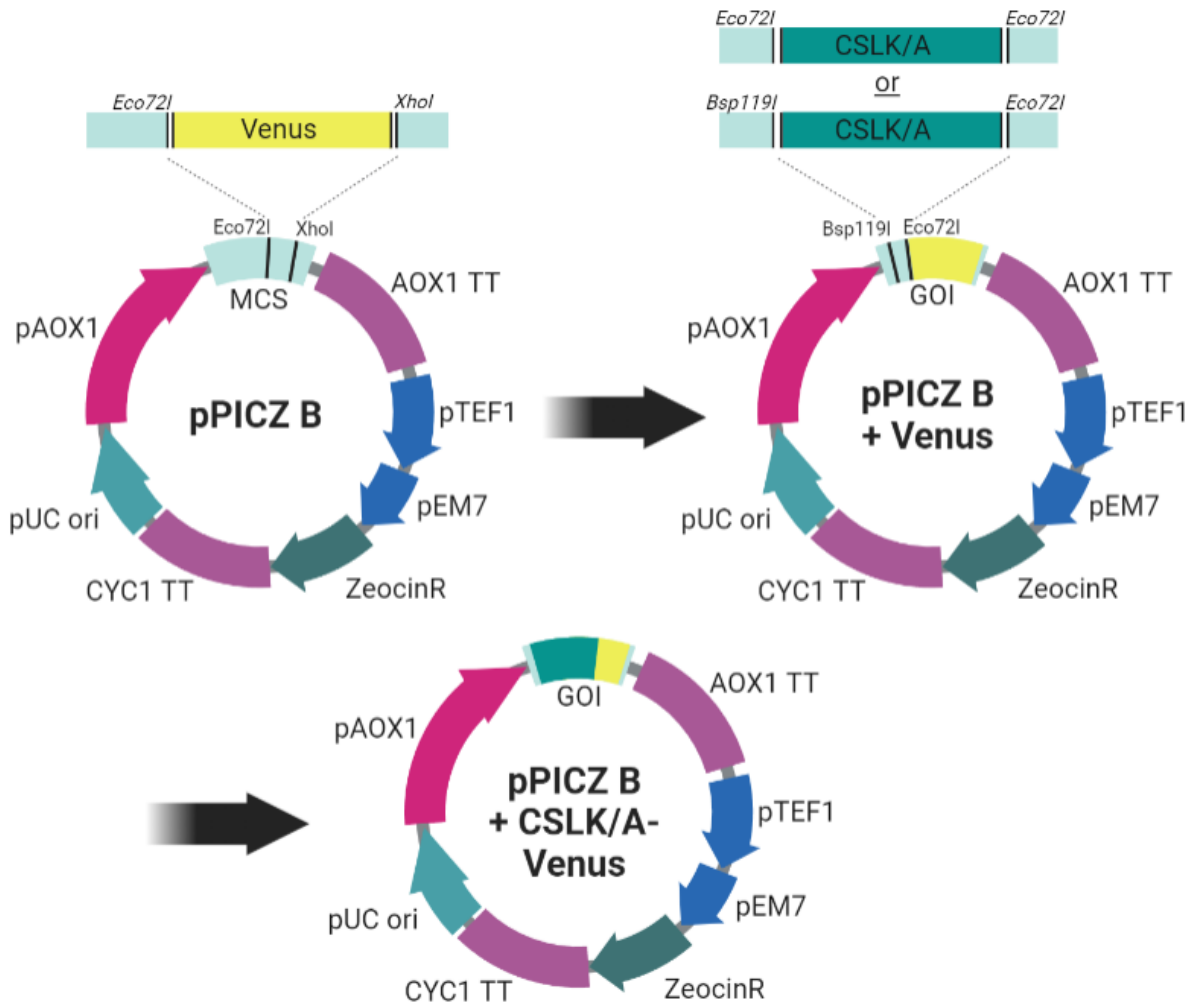


Figure 7: Overview of the cloning strategy used to fuse CSLK/A proteins with a fluorescent protein into the pPICZ cloning system. MCS: multiple cloning site, GOI: gene of interest. The figure was prepared with bioRender.

After digestion of PCR-amplified insert DNA or pPICZ-based vectors, they were purified in front of further ligation. While double digested PCR-amplified DNA inserts were purified using the DNA Clean & Concentrator[®]-5 (Zymo Research, Cat# D4014), double digested and dephosphorylated vectors were separated in a 0.7% (w/v) agarose gel and extracted from gel using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Cat# D4002) or the GeneJet Gel Extraction Kit (Thermo Fisher Scientific, Cat# K0692). The protocol for both kits was performed as described in the manufacturer's instructions, instead that DNA inserts were finally eluted in 12 µl HyClone[™] water and vectors were eluted in 20 µl HyClone[™] water. Afterwards, the DNA concentration was determined by measuring the absorption at 260 nm using a BioSpectrometer[®] basic.

The insert to vector ratio that was used for all ligations in this study was 3:1 and was performed with 1 µl T4 DNA Ligase (Thermo Fisher Scientific, Cat# EL0014) and 2 µl 10x T4 DNA Ligase Buffer in a total volume of 20 µl. Ligation mixtures were incubated at minimum for 1 h or overnight at room temperature (RT) and were afterwards transformed into *Escherichia*

coli (*E. coli*) TOP10'F competent cells (Thermo Fisher Scientific, Cat# C303003) by heat shock transformation. All pPICZ-based constructs created in this study are summarized in Table S3.

2.3 Cloning of Genes into GoldenPiCs-based Vectors

In order to use a more flexible cloning system, other transgenes were assembled into the GoldenPiCs cloning system [Prielhofer et al., 2017] which was a gift from the Gasser/Mattanovich/Sauer group (Addgene, Kit# 1000000133). The GoldenPiCs cloning system is a modular cloning system with three backbone levels (BB1 to BB3) using either *BsaI* or *BpiI* restriction sites. Consequently, the restriction cleavage sites for *BsaI* and *BpiI* were removed from their native coding sequence of the transgenes by replacing one or two nucleotides that encode for the same amino acid (AA) in this sequence. Primers used for amplification of the transgene and if necessary for their domestication of the respective coding sequence by Phusion PCR from transcribed cDNA of *Saccharomyces* are listed in Table S4 and transgenes from plasmid DNA are listed in Table S5.

All transcripts were amplified with their entire native sequence, except *Saccharomyces cerevisiae* (*Sc*) MNN2 and RGS4. Instead, ScMNN2 was cloned containing 46 AAs of its N-terminal sequence. Based on the protein prediction using the TOPCONS server (<https://topcons.cbr.su.se/>), this sequence is translated into a protein containing the first transmembrane domain (TMD) of ScMNN2. The TMD sequence is essential to determine the proteins localization and anchor the protein to the organelle membrane [Munro, 1998; Gleeson, 1998]. Therefore, ScMNN2 (AA1-46) was planned to be fused with mRuby2 that the protein can be visualized in the Golgi and determine the localization of AkCSLA3-sfGFP. RGS4 was amplified containing 33 AAs of its N-terminal sequence. This region is sufficient to translocate proteins to the plasma membrane in *Saccharomyces cerevisiae* [Srinivasa et al., 1998]. RGS4 was cloned in front of AkCSLA3-sfGFP to mis-localize the protein to the plasma membrane.

Phusion PCR settings, extraction of amplified DNA sequences from gel as well as determination of the DNA concentration was performed the same as described in section 2.2 for pPICZ-based cloning. In addition, the sequences of *PumCSLK1* and *OICSLK* were synthesized by Geneart. All Phusion products and synthesized DNA contain an adapter with the *BsaI* restriction enzyme sequence and the sequence of their fusion site to ensure their correct ligation into the BB1_23 vector DNA.

All Phusion products and synthesized DNA were assembled based on Prielhofer et al., (2017). For each assembly, 0.5 μ l T4 DNA Ligase, 0.5 μ l restriction enzyme (*BsaI*), 1 μ l 10 mM ATP, 1 μ l 10x FD buffer (all solutions from Thermo Fisher Scientific) and 25 ng of each DNA part (BB1_23 and amplified PCR products) were mixed in a total volume of 10 μ l. The mixture was incubated in a thermocycler (BioRad Laboratories, Hercules, US) as followed: at least 5

cycles for digestion (5 min, 37°C) and ligation (10 min, 22°C), followed by final digestion (5 min 37°C) and enzyme inactivation (10 min 75°C). The assembly mixtures were stored at 12°C in the thermocycler until they were used for *E. coli* transformation by heat shock transformation. For single transcription units, BB1 plasmid constructs can be directly assembled into BB3_14 plasmids, whereas multiple transcription units have to be assembled first in BB2 backbones, followed by assembling BB2 constructs into BB3 plasmids. All assemblies into the next cloning level were performed as described above, only the restriction enzyme that was used differs in the assemblies (BB1 to BB2 or BB3: *BpiI*; BB2 to BB3: *BsaI*). All created constructs cloned into GoldenPiCs-based vectors are summarized in Table S6.

2.4 *E. coli* and *P. pastoris* transformation

Each generated construct was transformed into *E. coli* TOP10F' calcium competent cells by heat shock transformation and all pipetting steps were performed under a sterile bench. The preparation of the calcium competent cells was performed as described in the protocol "How to make calcium competent cells" of the Krantz Lab (<http://mcb.berkeley.edu/labs/krantz/links.html>). For the transformation, 25 µl of *E. coli* TOP10F' cells were mixed gently by tapping with 2.5 µl of each ligation or assembly mixture. This mixture was incubated for 15 min on ice until the cells were heat shocked for 40 s at 42°C in a water bath VWB2 5 (VWR, Cat# 462-0556, Darmstadt, DE). *E. coli* cells were placed quickly on ice for 2 min before they were mixed with 250 µl low salt LB (IsLB) medium (20 g/l Lennox LB powder, Carl Roth, Cat# X964.3; adjusted with 1 M NaOH to pH 7.5) and incubated for 60 min at 250 rpm and 37°C in a MaxQ™ 6000 incubator (Thermo Fisher Scientific, Schwerte, DE). After the incubation step, 100 µl of the transformed *E. coli* strains were plated on IsLB agar plates (IsLB medium supplemented with 1.5% (w/v) bacto Agar, Carl Roth, Cat# 6494.4) containing the respective antibiotic (Kanamycin and Zeocin: 25 µg/ml; Carbenicillin, Nourseothricin and Spectinomycin: 50 µg/ml; Ampicillin: 100 µg/ml; Hygromycin: 200 µg/ml). The plates were incubated overnight at 37°C in a MaxQ™ 6000 incubator.

Colonies were verified by colony PCR with a gene- and backbone-specific primer combination. Colony PCR primers were designed with the Primer3 webtool (<https://primer3.ut.ee/>), ordered from eurofins Genomics (Ebersberg, DE) and are listed in Table S9. The colony PCR was performed in a final volume of 10 µl containing a final concentration of 1x RedTaq DNA Polymerase Master Mix (VWR, Cat# 733-2546) and 1 µM forward and reverse primer. The colony was touched with a sterile toothpick (Carl Roth, Cat# EC48.1), re-streaked onto new IsLB plates containing the respective antibiotic and tapped into the PCR mixture. PCR reactions were incubated in a thermocycler (BioRad Laboratories, Hercules, US) with one cycle for the initial denaturation (3 min, 95°C) and 25 cycles for denaturation (20 s, 95°C), annealing (30 s, 60°C) and extension (1 min/kb, 72°C). The final

extension step was performed for 5 min at 72°C and the cyclor was cooled down to 12°C for storage. 3 µl of the PCR mixture were loaded onto 0.7% (w/v) agarose gels and the gel electrophoresis was performed as described in section 2.2 except that amplified DNA was separated for 30 min.

Positively genotyped colonies were grown in 3 ml IsLB supplemented with the respective antibiotic and incubated overnight at 250 rpm and 37°C in a MaxQ™ 6000 incubator. 1 ml of the *E. coli* culture was transferred with glycerol (final concentration 15%) into cryo vials (VWR, Cat# 479-1219) and stored at -70°C for long-term storage. The rest of the culture was used to isolate plasmid DNA with the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, Cat# K0503). All steps were performed as described in the manufacturer's instructions, except that the purified DNA was eluted in 35 µl HyClone™ water. 250–500 ng of each purified plasmid DNA was mixed with 2.5 pmol of a backbone-specific primer (final volume 10 µl) to verify the inserted DNA by Sanger Sequencing (Genewiz, Leipzig, DE).

1000 ng plasmid DNA of each construct were linearized with FastDigest® restriction enzymes (pPICZ B and X: *MssI*, BB3aZ_14 and BB3rN_14: *SgsI*, BB3eH_AC: *SmaI*) as described in the manufacturer's instruction. The linearized DNA was purified using the DNA Clean & Concentrator®-5 (Zymo Research, Cat# D4014) and the DNA concentration was determined with a BioSpectrometer® basic as described in section 2.2. 150 ng of each linearized plasmid was transformed into *Pichia pastoris* X-33 competent cells (EasySelect™ Pichia Expression Kit, Invitrogen) by condensed electroporation as described in Lin-Cereghino et al. (2005). The transformation was performed in ice-cold 0.2 cm electroporation cuvettes (Carl Roth, Cat# PP39.1) in a Micropulser Electroporator (BioRad Laboratories, Hercules, US) with a charging voltage of 1.5 kV. After transformation, 1 ml YPD medium (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose) supplemented with the same volume of 1 M sorbitol was added to the shocked electrocompetent cells as recovery medium and they were incubated for 1 h at 30°C and 250 rpm in a MaxQ™ 6000 incubator. Afterwards, the transformed cells were centrifuged at 3,000 x g for 5 min. The medium was removed, the cell pellet was resuspended in 200 µl YPD medium and the whole mixture was plated onto YPDS agar (5% (w/v) agar) plates supplemented with the respective antibiotic (Nourseothrinin and Zeocin: 50 µg/ml, Hygromycin: 200 µg/ml) for the transformed plasmid. The plates were incubated for three days at 30°C in a MaxQ™ 6000 incubator. The successful integration into the *Pichia* genome was verified by colony PCR using the same settings and primer combinations as done for *E. coli* (see Table S9), except that 30 cycles of denaturation, annealing and extension were performed. Positively genotyped *Pichia* strains were grown overnight in YPD supplemented with the respective antibiotic and cryopreserved at -70°C in a final concentration of 15% glycerol.

In order to generate *Pichia* cells expressing *AtCSLA2* and *AtCSLA9*, BB3eH + pAOX1:*AtCSLA2*:RPP1Btt *Pichia* competent cells were freshly prepared and transformed with linearized pPICZB + *AtCSLA9*-Venus DNA. The preparation of competent cells and transformation leaned on Davidow et al. (1985) and some steps of this protocol were adapted to the protocols of Xuan et al. (1988) and Chen et al. (1997). 2 loops of one day earlier re-streaked *Pichia* cells (BB3eH + pAOX1:*AtCSLA2*:RPP1Btt) were resuspended in 1 ml TE buffer (10 mM Tris, 1 mM ethylenediamine tetra acetic acid disodium salt dihydrate), centrifuged for 1 min at 8,500 x g and was replaced with 600 µl lithium acetate buffer (0.1 M lithium acetate dihydrate pH 6.0 adjusted with 10% acetic acid). This mixture was incubated for 1 h at 28°C in a neoMix 7-0921 thermomixer (neoLab Migge GmbH, Heidelberg, DE) before cells were pelleted again and resuspended in 60 µl lithium acetate buffer. 40 µl of these competent cells were mixed with 2 µg single stranded carrier DNA from salmon testes (Sigma Aldrich, Cat# D7656-1ML) and 250 ng linearized pPICZB + *AtCSLA9*-Venus DNA. The transformation mix was incubated for 15 min at 28°C without shaking. 350 µl of 40% polyethylene glycol 4000 dissolved in 0.1 M lithium acetate was added, mixed thoroughly by pipetting up and down and further incubated for 1 h at 28°C. Afterwards, the transformation mixture was incubated in a water bath at 39°C for 10 min. The cells were immediately cooled by adding 600 µl 0.1 M ice-cold lithium acetate before the whole transformation mixture was plated onto a YPDS agar plate supplemented with 50 µg/ml Zeocin.

2.5 *Pichia* Growth, Optical Density and Fluorescence Measurements

At least two – four independent biological replicates of each *Pichia* transformant were grown at 30°C and 250 rpm in a MaxQ™ 6000 incubator. To enhance the productivity of heterologously expressed proteins in *Pichia*, the growth was uncoupled from the protein production by the usage of medium containing different carbon sources. Medium containing glycerol ensure biomass production of the *Pichia* culture, while addition of methanol to the medium result in protein expression of heterologously expressed proteins under the control of a methanol-inducible promoter.

In this study, *Pichia* cultures were firstly grown in buffered-glycerol-complex medium (BMGY: 1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate (pH 6.0), 1.34% (w/v) YNB, 4 x 10⁻⁵% (w/v) biotin and 1.25% glycerol (v/v)), that was switched to buffered complex medium containing 0.5% methanol instead of glycerol to induce protein expression. The *Pichia* cultures were grown in 2–4.3 ml BMGY in glass culture tubes (16 x 100 mm) for 24 h or 64 h followed by centrifugation at 2,000 x g for 2 min, resuspension mostly in an equal volume of BMMY and further growth for 24–72 h in either glass culture tubes or transferred as replicates in 24-deepwell plates (EnzyScreen CR1424a) covered with metal lids (CR1224b).

24-deepwell plates from EnzyScreen have the advantage of better shaking and higher rates of oxygen in the medium of the *Pichia* cultures that enables better growth.

In order to test if the HM structure produced by *AkCSLA3-Venus* can be altered by the sugar availability in the medium, *Pichia* were grown together with different carbon sources in the medium. Therefore, YP medium was supplemented with 1.5% (v/v) methanol (YPM) and one of the following carbon sources: 0.5% (v/v) glycerol, 0.5% (w/v) glucose or mannose. Additionally, the usage of carbon sources in the medium has the benefit to repress the gene expression of proteins under the control of the pAOX1 promoter why only biomass is produced in the beginning of the growth. With increasing time of growth, the glycerol content in the medium is depleted, resulting in heterologous expression of proteins in *Pichia* cultures [Celic and Calik, 2012]. *Pichia* cells were cultivated in 24-well plates (VWR, Cat# 734-2779) for 48 h at 250 rpm and 30°C in a MaxQ™ 6000 incubator.

Because this “co-feeding” strategy allows to minimize the handling steps and disruption of *Pichia* growth as described for growth in BMGY and BMMY medium, all *Pichia* strains with gene expression under the control of a methanol-inducible promoter were grown in YPM medium supplemented with 0.5% (v/v) glycerol (YPMG). The cultivation was performed in 2 ml YPMG in 24-deepwell plates from EnzyScreen for 48 h.

Pichia strains co-expressing genes under the control of the constitutive promoter pGAP (Golgi marker proteins: ScMNN2, ScGOS1 and ScOCH1) and the methanol-inducible promoter pAOX1 (*AkCSLA3*) were first cultivated for 24–48 h in YPD medium, followed by centrifugation at 2,000 x g for 2 min, resuspension in YP medium supplemented with 1.5% (v/v) methanol (YPM) for further 24–48 h of growth. More detailed information of growth conditions for each experiment are summarized in the legend of each figure in the result part (section 3) and the detailed recipes of prepared medium is listed in Table S10.

After protein production, each culture was diluted 1:1 with water in a total volume of 200 µl and the optical density at 600 nm (OD₆₀₀) and fluorescence (FL) were measured using a Tecan M1000 plate reader (Tecan, Grödig, AT). Furthermore, the normalized FL (FL/OD₆₀₀) or the fold change of FL/OD₆₀₀ relative to a non-fluorescent control were determined for each construct.

2.6 Confocal Microscopy of *Pichia* Cells

Confocal microscopy was performed to determine the localization and expression pattern of genes that were fused to fluorescent proteins in *Pichia* cells. For this purpose, the diluted culture that was used for the OD₆₀₀ and fluorescence measurement was mixed to an equal volume with 0.01% staining solutions. Calcofluor white (CF) and trypan blue (TB) label

β -glucans on the yeast plasma membrane. Propidium iodide (PI) binds to guanine and cytosine and, therefore is able to label DNA and RNA in *Pichia* cells but only when their membranes are leaky [Suzuki et al., 1997; Hohenblum et al., 2002]. 1 μ l of these diluted cultures was pipetted onto a diagnostic slide with 12 incorporated wells with a diameter of 5 mm (Carl Roth, Cat# L199.1) and were covered with high precision cover slips (VWR, Cat# 630-2187). All *Pichia* cells were imaged on laser scanning confocal microscopes (LSM) from Carl Zeiss (Carl Zeiss AG, Oberkochen, DE).

LSM880 images (Figure 8C, 12B, 14B, 25B, 26B) were made with a 40x/1.2 water-immersion objective, a beam splitter MBS 488/561 and AiryScan mode. The excitation and emission wavelength for Venus fluorescence was acquired at 488 nm and 523 nm, whereas TB or mRuby2 were acquired at 561 nm and 579 nm. LSM700 images (Figure 16, 17B, 18B, 19B, 36A) were done with a 40x or 63x/1.2W water-immersion objective and a beam splitter MBS 405/488/555/639. Excitation and emission wavelength were set as followed: Venus (488 nm and BP 490–555 nm), TB (639 nm and BP 640 nm). LSM900 images (Figure 20C, 22C, 24E, 24F, 27B) were acquired using a 63x/1.20 water-immersion objective with AiryScan mode and the following excitation and emission set up: CF (405 nm, 410–490 nm, SP545), Venus (488 nm, 490–560 nm), PI (639 nm, 656–700 nm, LP655). All images were processed uniformly with ImageJ and arranged in Inkscape.

2.7 *Pichia* Harvest and Cell Wall Extraction

After the OD₆₀₀ and FL was determined, the *Pichia* cultures were transferred to 2 ml SafeSeal tubes (Sarstedt, Cat# 72.695.400) and were harvested by centrifugation at 5,000 x g for 3 min. Afterwards, either alcohol-insoluble residue (AIR) or alkaline-insoluble (AKI) polymers were extracted from *Pichia* cells.

The *Pichia* AIR contains all endogenous yeast polymers from the wall which are mainly β -glucans and mannoproteins [Orlean, 1997, Magnelli et al., 2002]. Polymers of the AIR were isolated leaned on the methods described for plant cell wall polysaccharides extraction [Pettolino et al., 2012, Pauly et al., 2019]. *Pichia* cell pellets were lysed with 1 ml 70% ethanol with a steel ball using a Retsch MM400 mill (Retsch GmbH, Haan, DE) with two TissueLyser Adapters (Qiagen GmbH, Cat# 69982) mixing at 30 Hz for 1:30 min. The cell pellet was centrifuged at 20,000 x g for 2 min, the supernatant was decanted and the pellet was washed twice times with 1 ml acetone mixing for 2:30 min at 30 Hz. Afterwards, the extracted AIR pellet was transferred into pre-weighed tubes and dried in an Eppendorf concentrator 5301 (Eppendorf AG, Hamburg, DE) at 30°C. The precise weight of the AIR was documented and the AIR was resuspended in MiliQ water to a final concentration of 1–1.5 mg/ml using a steel ball and a Retsch MM400 at 30 Hz for 15 min. 300 μ g AIR slurry was transferred into 2 ml

screw-cap tubes (VWR, Cat# 211-0093), dried down with a Eppendorf concentrator 5301 and further processed for monosaccharide analysis by TFA hydrolysis (section 2.9). For detailed information of the structure of produced HM in *Pichia*, linkage analysis was performed using 1 mg of the AIR material.

AKI polymers of the *Pichia* cell wall were extracted directly after the *Pichia* cells were harvested (done for most of the growth batches in this study) or after AIR polymers were isolated (Figure 23, 28 and 33). Treatment of the cell pellet with 1 M sodium hydroxide removes most of the yeast mannoproteins [Magnelli et al., 2002] and therefore enrich HM polysaccharides in the AKI fraction. All steps for AKI preparation were performed the same as described in Voiniciuc et al. (2019), but using a neoMix 7-0921 thermomixer (neoLab Migge GmbH, Heidelberg, DE) for the incubation step with sodium hydroxide. The pellet was neutralized with 50% acetic acid (8.7 M) and was washed with water to remove the alkaline-soluble polymers. Each pellet containing the AKI, was resuspended in 600 μ l water and homogenized using glass beads (Sigma Aldrich, Cat# G8772-500G) and a Retsch MM400 mill (Retsch GmbH, Haan, DE) for 2:30 min at 30 Hz. The extracted AKI material was used directly for monosaccharide analysis or was enzymatically digested with endo-1,4- β -mannanase (from *Aspergillus niger*, Megazyme Cat# E-BMANN) before the monosaccharide composition was determined.

2.8 Enzymatic Digestion of *Pichia* Wall Polymers

50 μ l of AKI material was transferred into 1.5 ml tubes (Sarstedt, Cat# 72.690.001) and mixed with a digestion mix containing a final concentration of 0.1 U endo-1,4- β -mannanase (from *Aspergillus niger*, Megazyme Cat# E-BMANN) and 0.2 M potassium phosphate buffer (PB, pH 7.0; 0.077 M potassium dihydrogen phosphate, 0.123 M di-potassium hydrogen phosphate) in a total volume of 200 μ l per reaction. The tubes were incubated for 30 min at 40°C at 1,000 rpm shaking in a neoMix 7-0921 thermomixer. Afterwards, all tubes were centrifuged at 20,000 x g for 2 min and 100–150 μ l of each supernatant was transferred to 2 ml screw cap tubes. The supernatant contains oligosaccharides with β -1,4-Man and β -1,4-Glc bonds that were cleaved by β -mannanase and are typical for mannan and glucomannans produced by the heterologously expressed proteins in *Pichia*. The transferred supernatant was dried with an Eppendorf concentrator 5301 and the released sugars were hydrolysed with trifluoroacetic acid (TFA, 2 M (v/v)).

Additionally, AKI material of three biological replicates of *Pichia* expressing *Physcomitrium patens* transgenes (Figure 33A) were pooled, resuspended in 300 μ l 0.2 M PB (pH 7.0) and, then treated with 300 μ l zymolyase mixture containing 125 μ g Z1000 Zymolyase 20T (Lyticase from *Arthrobacter luteus*; USBiological Life Sciences) and 10 μ g sodium azide

to further enrich HMs. Zymolyase hydrolyses linear β -1,3-glucans that occur typically in the yeast cell wall and sodium azide was added as a preservative. The samples were incubated horizontally at 250 rpm and 37°C in a MaxQ™ 6000 incubator (Thermo Scientific) for 48 h. After incubation, samples were centrifuged for 5 min at 16,000 x g, washed two times with 1 ml water and one time with 300 μ l acetone before drying the enriched (gluco)mannan (EM) material overnight at RT. The EM fraction was resuspended in 1 ml water using a Retsch MM400 mill (Retsch GmbH, Haan, DE) for 5 min at 30 Hz. 50 μ l of the EM material was transferred for each sample used for monosaccharide analysis with 2 M TFA.

2.9 Monosaccharide analysis

To analyse the monosaccharide content of the Pichia AIR or AKI material, polysaccharides were hydrolysed by a one-step hydrolysis with 4% (w/v) sulfuric acid [Yeats et al., 2016] or 2 M TFA. For 4% sulfuric acid hydrolysis, 50 μ l AKI material, 800 μ l of 30 μ g/ml ribose (internal standard) and 30 μ l 72% (w/w) sulfuric acid (4% (w/v) final concentration) were transferred into 2 ml screw cap tubes and incubated for 60 min at 120°C in a Techne™ Dri-Block™ DB200/3 heater (Fisher Scientific, Schwerte, DE). Blank and sugar standards (containing Glc, Gal, Man) were prepared similarly to the samples. 0, 1, 2, 5, 10, 25, 50, 75, 100 and 125 μ l of a 1 mg/ml sugar stock were transferred to a tube and filled to a volume of 150 μ l with water. 700 μ l of 30 μ g/ml ribose and 30 μ l 72% (w/w) sulfuric acid were added on top of each standard and incubated the same as all samples. After hydrolysis, all tubes were cooled on ice, centrifuged for 15 min at 20,000 x g to remove any particles remaining in the insoluble material and 200 μ l of the supernatant was transferred into an IC vial.

2 M TFA hydrolysis was performed with 50 μ l AKI material, 100–150 μ l supernatant of the digested AKI material or 300 μ g of AIR material. In front of the hydrolysis, all samples and standards (containing glucose, galactose, mannose; concentrations: 0, 1, 2, 5, 10, 25, 50, 75, 100 or 125 μ g/ μ l) were dried using an Eppendorf concentrator 5301, followed by hydrolysis in 150 μ l 2 M TFA for 90 min at 120°C. After hydrolysis, the samples and standards were cooled on ice, briefly centrifuged and dried using an Eppendorf concentrator 5301. Samples and standards were washed with 300 μ l isopropanol and dried once again. All samples and standards were resolved in 400 μ l 30 μ g/ml ribose solution (internal standard) using a Retsch MM400 mill for 1:30 min at 30 Hz. Afterwards, samples were centrifuged for 2 min at 20,000 x g and 100 μ l of each supernatant was transferred into an IC vial.

Hydrolysed monosaccharides were determined by high-performance anion-exchange chromatography coupled with pulsed electrochemical detection (HPAEC-PAD). Therefore, 10 μ l of each sample was injected into a 940 Professional IC Vario system (Metrohm) and sugars were separated by CarboPac PA20 and Metrosep Carb 2-250/4.0 analytical and guard

columns. All steps were performed with a constant flow rate of 0.5 ml/min. The columns were equilibrated with 2 mM sodium hydroxide (NaOH) and 3.4 mM sodium acetate (NaAce) before samples and standards were injected. Using the same eluents, the neutral sugars Glc, Gal and Man were separated for 20 min. Glucosamine was also detected with this separation method, but only in negligible amounts and was therefore not quantified. After the separation, the columns were rinsed with 80 mM NaOH and 136 mM NaAce for 3 min, followed by 4 min column equilibration with 2 mM NaOH and 3.4 mM NaAce before the next sample was injected. Using the MagIC Net 3.2 software, monosaccharide peaks of all samples and standards were automatically annotated and the area under the peak was quantified relatively to the calibration curve of the standards.

2.10 Glycosyl Linkage Analysis

Linkages of oligosaccharides produced by CSLK/A expressing *Pichia* strains were determined by glycosyl linkage analysis of the isolated AIR material. Therefore, oligosaccharides were first methylated using a fine suspension of 50% NaOH in dimethyl sulfoxide (DMSO) and methyl iodide (Ciucanu et al., 1984; Anumula and Taylor 1991). Methylated oligosaccharides were hydrolysed by TFA, followed by their reduction with sodium borodeuteride and acetylation with acetic anhydride to volatilize them for GC analysis [Pettolino et al., 2012].

1 mg AIR material of a 1.5 mg/ml AIR slurry was transferred into glass tubes (VWR, Cat# 212-8215) and dried using a Techne™ Dri-Block™ DB200/3 heater and a Techne™ Tray Column and Gas Distributor (Fisher Scientific, Cat# 11593859) under pressurized air flow. A stir bar and 200 µl dried DMSO (DMSO was pre-treated with molecular sieves: Carl Roth, Cat# 8487.2) was added to the dried AIR material, sonicated for 15 min and incubated overnight at RT to ensure swelling of the material that improves the methylation reaction [Pettolino et al., 2012].

The NaOH/DMSO suspension was prepared by briefly mixing 200 µl methanol with 140 µl 50% NaOH. The suspension was thoroughly resuspended in 5 ml dried DMSO, sonicated for 5 min and centrifuged for 2 min at 2,000 x g in a swing bucket Eppendorf 5810R centrifuge (Eppendorf AG, Hamburg, DE) The supernatant was decanted and the washing steps with DMSO was repeated at least four times or until the suspension was homogeneously translucent. The final elution of the NaOH pellet was resolved in 2.8 ml dried DMSO. 200 µl NaOH/DMSO suspension was added to the AIR solubilized in 200 µl DMSO and mixed for 15 min in front of the methylation step. The methylation was performed for 2–3 h and was started by adding 100 µl methyl iodide. To avoid quenching of the methylation reaction by water addition or formation, all steps previously described were done with pre-heated glass ware and

the incubation with methyl iodide was performed under nitrogen atmosphere. After incubation, the methylation reaction was stopped by addition of 2 ml MQ water and nitrogen was gently bubbled into the solution until the milky solution have been cleared using a glass pasteur pipette (VWR, Cat# 612-1702). 2 ml dichloromethane was added onto each solution, thoroughly vortexed and centrifuged for 2 min at 2,000 x g resulting in a phase separation of the aqueous (upper) and organic (lower) phase. ~1.5–2 ml of the organic phase containing methylated oligosaccharides were transferred into new 2 ml screw cap tubes and were dried in a Techne™ Dri-Block™ DB200/3 heater under exposure to nitrogen gas.

Hydrolysis of the methylated oligosaccharides was performed with 300 µl 2 M TFA at 121°C for 90 min. The tubes were cooled on ice, centrifuged for 30 s at 20,000 x g and 200 µl of the supernatant was transferred into new glass tubes and were dried under nitrogen gassing. The dried material was washed with isopropanol and again evaporated to remove the volatile TFA from the partially methylated monosaccharides (PMM).

For the reduction of the monosaccharides, 10 mg/ml sodium borodeuteride were resuspended in 1 M ammonium hydroxide (prepared with 25% ammonia solution). 200 µl of the freshly prepared sodium borodeuteride solution was added to each tube, vortexed and incubated for 60 min at RT to convert the PMM to alditols. The samples were neutralized by adding 150 µl glacial acetic acid and 300 µl methanol:acetic acid (9:1 (v/v)) was added to wash out the boric acid by evaporation. The wash step was repeated one time more with 300 µl methanol:acetic acid (9:1 (v/v)) and three times with 300 µl methanol.

Alditols were acetylated by adding 50 µl acetic anhydride and 50 µl pyridine incubated for 20 min at 121°C. The partially methylated and acetylated monosaccharides were cooled on ice and were washed two times with 200 µl toluene. After the final evaporation, 1.2 ml ethyl acetate was added to the samples, gently swirled and 5 ml water was added to the samples to separate the organic and aqueous phase by centrifuging tubes for 2 min at 2,000 x g. 600 µl of each organic (upper) phase was transferred into 2 ml screw cap tubes, evaporated and resolved in 300 µl acetone.

2 µl of each sample containing the partially methylated alditol acetates (PMAAs) was injected into an Agilent 8890 GC system equipped with a Supelco SP-2380 column (30 m x 0.25 mm x 0.2 µm) and was detected by an Agilent 5977B Inert Plus GC/MSD (Agilent Technologies, Santa Clara, US). Each sample have run for 38.5 min with a flow of 1.502 ml/min at 1.02 bar. The initial temperature of the oven was 80°C that was held for 3 min, heated to 170°C for 3 min (rate: 30°C/min) and further heated to 240°C (rate: 4°C/min) that was held from 22.5 to 38.5 min. After each run, the column was heated to 250°C for 5 min to clean the column from injected sugars. The PMAAs were identified by their retention time of the detected sugar peaks and their mass ion profile available on the CCRC Spectral Database for PMAAs

(<https://www.ccruc.uga.edu/specdb/ms/pmaa/pframe.html>). For quantification, the PMAAs were automatically quantified with the Agilent MassHunter Quantitative Analysis software.

2.11 Polysaccharide Analysis by Gel Electrophoresis

To gain detailed structural information about the polysaccharides produced by several functional CSLK/A enzymes expressed in *Pichia*, polysaccharides were separated by gel electrophoresis leaned on Pidatala et al. (2017). Therefore, 100 μ l AKI material was digested with endo-1,4- β -mannanase (from *Aspergillus niger*, Megazyme Cat# E-BMANN) as described in section 2.8, centrifuged for 2 min at 20,000 x g and 150 μ l of the supernatant was transferred into a screw cap tube and dried with an Eppendorf concentrator 5301. In addition, the following oligosaccharide standards were digested, centrifuged, transferred into 2 ml screw cap tubes and dried as well: (i) β -1,4-linked Man sugar mix containing 1 mM of the following sugars (all parentheses show the Cat# of sugars from Megazyme International Ireland if not mentioned separately): Man (Carl Roth, Cat# 4220.1), mannobiose (O-MBI), mannotriose (O-MTR), mannotetraose (O-MTE), mannopentaose (O-MPE) and mannohexaose (O-MHE), (ii) 200 μ g ivory nut mannan (INM, P-MANIV) and (iii) 200 μ g konjac glucomannan (KGM, P-GLCML).

Samples and standards were fluorescently labelled with 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) and borane-2-methylpyridine (B2M). Therefore, 5 μ l 0.2 M ANTS, 5 μ l 0.2 M B2M and 10 μ l derivatization buffer (water:acetic acid:DMSO (17:3:20)) were added to the dried and digested samples and standards, briefly vortexed, centrifuged to collect the liquid in the bottom of the tube and incubated at 37°C for at least 16 h protected from the light. After the incubation, samples were dried in an Eppendorf concentrator 5301 for 2 h and resuspended in 100 μ l 6 M urea.

3 μ l samples and 2 μ l standards were separated by gel electrophoresis using a Mini PROTEAN Tetra Vertical Electrophoresis Cell (Bio-Rad Laboratories GmbH, Hercules, US) filled with TRIS-Glycine buffer (25 mM TRIS and 192 mM glycine) at 200 V for 2.5 h. The electrophoresis was performed on mini-gels (1.0 mm thickness with 10 pockets) using two different concentrations of polyacrylamide solution (60% (w/v) acrylamide and 1.6% (w/v) N,N'-methylene bisacrylamide, ratio 37.5:1). The upper 2 cm of the gel (stacking gel, lower concentration acrylamide) lines up all oligosaccharides of the probes, whereas the lower part of the gel (resolving gel; higher concentration of acrylamide) separates the oligosaccharides. Ammonium persulfate (APS) and N,N,N',N'-Tetramethylethylenediamin (TEMED) were added to the mixed gel solutions in the end to ensure the polymerization of the gel. First, the resolving gel was mixed containing 2.5 ml 1.5 M Tris-HCl (pH 8.5) and MQ water, 5 ml (v/v) polyacrylamide solution, 50 μ l 10% (w/v) APS and 10 μ l TEMED. The mix was pipetted in between the glass plates holded in a handcasting system (Bio-Rad Laboratories GmbH,

Hercules, US) and overlaid with isopropanol to ensure a straight surface of the gel. After 30 min of polymerization, isopropanol was removed, the cassette was completely filled with stacking gel containing 6.87 ml MQ water, 2.5 mM Tris-HCl, 630 μ l polyacrylamide solution, 100 μ l 10% APS and 10 μ l TEMED and a comb was placed on top of the stacking gel. After the gel electrophoresis was finished, the oligosaccharide bands were visualized by UV light using a UVP Gelstudio PLUS Touch (Analytik Jena GmbH, Jena, DE).

2.12 In-Gel Fluorescence of Microsomal Membrane Fractions

To gain knowledge about the topology of *AkCSLA3* in the Golgi apparatus, *AkCSLA3* was tagged on its N- and C-terminal end with sfGFP. *Pichia* cultures (sfGFP control, sfGFP-*AkCSLA3* and *AkCSLA3*-sfGFP) were first grown as pre-cultures in 3 ml YPD supplemented with the respective antibiotic and incubated at 30°C and 250 rpm overnight in a MaxQ6000 shaker (Thermo Fisher Scientific, Schwerte, DE). These pre-cultures were used to inoculate 50 ml YPD starter cultures with an OD of 0.3 for each strain. The starter cultures were incubated for further 24h at 30°C and 250 rpm. *Pichia* starter cultures were transferred to 50 ml Falcon tubes (Sarstedt, Cat# 62.547.254), centrifuged for 5 min at 3,000 x g and the supernatant was decanted. The cultures were resuspended in 50 ml YPM and were grown for further 24 h at the same conditions as described above to express heterologous proteins. After 24 h of protein expression, the growth rate and protein expression were determined by OD and fluorescence measurements using a Tecan M1000 (Tecan, Grödig, AT) as described in section 2.5. Afterwards, *Pichia* cells were harvested by centrifugation at 3,000 x g for 5 min, the medium was decanted, and the weight of the harvested *Pichia* cells was determined.

To extract microsomal membrane fractions from *Pichia* cultures, *Pichia* cells were firstly lysed with lysis buffer (50 mM Tris-HCl (pH 7.6), 150 mM sodium chloride (NaCl), 1x Pierce™ protease inhibitors (Thermo Fisher, Cat# A32955)) and secondly the cell lysate and the microsomal membrane fraction were separated by centrifugation. The protocol leaned on Drew et al. (2008) and was adapted by a master student of our group [Clauß, 2021]. Therefore, 1/3 of the *Pichia* cell pellet size was covered with 0.5–0.75 mm acid-washed glass beads and were then resuspended in a cell pellet to lysis buffer rate of 0.28:1. Cell lysis was performed in a Retsch MM400 mill with adapters for Falcon tubes using at least seven repetitive cycles of mechanical disruption by shaking at 20 Hz for 3 min and cooling the cells on ice to avoid protein degradation. After seven cycles, the lysed *Pichia* cells were centrifuged at 3,000 x g and 4°C for 20 min. The fluorescence of the pellet and supernatant was checked with a UVP Gelstudio PLUS Touch (Analytik Jena GmbH, Jena, DE). When the supernatant was not cloudy enough and the fluorescence was mostly not visible in the supernatant, the lysis was repeated with more cycles. Otherwise, the supernatant was transferred into several 2 ml Eppendorf tubes and centrifuged for 1 h at 20,000 x g and 4°C to receive the microsomal membrane fraction.

The supernatants of each strain were collected in a 15 ml Falcon tube and represent the cytosolic fraction, whereas the pellets were resolved in a total volume of 50 μ l and pooled together in a 2 ml Eppendorf tube for each strain.

The protein concentration of the cell lysate and the microsomal membrane fraction was determined by Bradford assay leaned on He (2011). Therefore, 30 μ l of different bovine serum albumin (BSA) standards (0, 5, 10, 20, 30, 40, 50, 75, 100 μ g), 1:10 dilutions of cell lysates and 1:30 dilutions of microsomal membrane fractions were transferred into 2 ml Eppendorf tubes as duplicates for standards and triplicates for samples. To all samples and standards 1.5 ml Bradford reagent (0.005% (w/v) Coomassie Brilliant Blue G 250 (Carl Roth, Cat# 9598.1), 5% methanol, 8.5% phosphoric acid) was added to each tube, vortexed and incubated for 5 min at RT. 200 μ l of each standard and sample was transferred into a 96-well plate and the absorbance at 595 nm was measured with a Tecan Spark 10M (Tecan, Grödig, AT) plate reader.

20 μ g protein of the lysate or microsome fraction were mixed in a ratio of 1:5 with 5x Laemmli protein buffer (5% (w/v) sodium dodecyl sulphate (SDS), 250 mM Tris-HCl (pH 6.8), 30% (v/v) glycerol, 5% (v/v) β -mercaptoethanol, 0.02% bromphenol blue) and were incubated for 5 min at 55°C in a thermomixer (neoLab Migge GmbH, Heidelberg, DE) without shaking. 2 μ l PageRuler™ (Thermo Scientific, Cat# 11812124) and 20 μ g protein of each sample were loaded into the pockets of a SDS mini-gel (1.0 mm thickness with 10 pockets) containing 2 cm of stacking gel (4% (v/v) ROTIPHORESE® Gel 40 (37.5:1, Carl Roth, Cat# T802.1), 125 mM Tris-HCl (pH 6.8), 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.001% (v/v) TEMED) on the top of each gel to line up the proteins in the gel and the resolving gel (12% (v/v) ROTIPHORESE® Gel 40, 375 mM Tris-HCl (pH 8.8), 0.1% SDS, 0.05% APS, 0.0005% (v/v) TEMED) that separate the proteins in the gel. The electrophoresis was performed in a Mini PROTEAN tetra Vertical Electrophoresis Cell (Bio-Rad Laboratories GmbH, Hercules, US) filled with 1x ROTIPHORESE® SDS-PAGE buffer (Carl Roth, Cat# 3060.2) at 100–150 V for 2 h. After the gel electrophoresis was finished, the proteins tagged to a fluorescent protein (sfGFP) were visualized using a GFP filter and blue LED light in a UVP Gelstudio PLUS Touch (Analytik Jena GmbH, Jena, DE).

2.13 Cloning of CSLA genes into plant transformation vectors

In order to gain knowledge about the biological role of CSLA in planta, several CSLA genes were cloned into vectors that were used for plant transformation. Vectors of the modular cloning (MoClo) system [Weber et al., 2011] were kindly provided by Dr. Sylvestre Marillonnet (IPB, Halle Saale). These vectors were used to transform Venus, *AtCSLA2*, *AtCSLA9*-Venus, *PumCSLK1*-Venus and *PumCSLK1* under the control of the p35S promoter transiently into

Nicotiana benthamiana leaves. Firstly, a Phusion PCR was performed on plasmid DNA containing the respective coding sequence with primers listed in Table S7. The primers have included the sequence of the *Bpil* restriction site (nnnnGAAGAC) and the fusion sites corresponding to AATG and GCTT sequences that are necessary to introduce the respective CSLK/A into pAGM4031 vector (level 0) in frame. CSLK/As from verified level 0 plasmids were assembled into pICH75055 vector (level 1) together with pICH51277 (Addgene, Cat# 50268) containing the cauliflower mosaic virus 35S promoter (p35S) for overexpression and pICH41432 containing the octopine synthase terminator (tOCS) by using *Bsal* restriction sites.

For stable transformation of *AkCSLA3* into *Arabidopsis thaliana*, the ligation-independent cloning (LIC) technique demonstrated by De Rybel et al. (2011) was used. *AkCSLA3* was cloned into pCV01 vector (35S:LIC-sYFP:tNOS) that was previously described in Voiniciuc et al. (2015). Primers used for the Phusion PCR of *AkCSLA3* are listed in Table S7.

Phusion PCR settings, extraction of amplified DNA sequences from gel and determination of the DNA concentration was performed the same as described in section 2.2 for pPICZ-based cloning. The assembly reactions for MoClo constructs were performed the same as described for GoldenPiCs vectors in section 2.3. For LIC cloning, 4 µg pCV01 vector was linearized with *Kspal* (Thermo Fisher Scientific, Cat# ER1031) as described in the manufacturer's instructions. Afterwards, 400 ng of the linearized pCV01 vector and 200 ng of amplified *AkCSLA3* containing LIC adapters were mixed with 0.0625 U T4 DNA Polymerase (Thermo Fisher Scientific, Cat# EP0061), 1x T4 buffer, 5 mM dithiothreitol (DTT), 0.1 µg BSA and 10 mM dCTP for pCV01 vector or dGTP for *AkCSLA3* insert (Thermo Fisher Scientific, Cat# R0181). Both reaction mixes were mixed thoroughly, centrifuged at 5,000 x g for 30 s and incubated at 11°C for 20 min, followed by inactivation of the polymerase at 75°C for 10 min. An insert to vector ratio of 3:1 was mixed to ligate the linearized and T4-treated pCV01 vector with the T4-treated *AkCSLA3* insert and was incubated for 1 h at RT. All plasmids were transformed into *E. coli* TOP10F' competent cells as described in section 2.4, with the exception that 500 µl of the transformation mix was plated onto IsLB plates supplemented with 50 µg/ml Kanamycin for LIC strains. 100 µl transformation mixture of level 0 constructs were plated on IsLB plates containing 50 µg/ml Spectinomycin, while level 1 constructs were plated onto IsLB plates supplemented with 50 µg/ml Carbenicillin. Verification by colony PCR, isolation of plasmid DNA of grown cultures and verification of the plasmid by Sanger Sequencing was performed the same as described in section 2.4 with the primers listed in Table S9. Verified pICH75055 + p35S:CSLA(-Venus):tOCS (MoClo strains) and pCV01 + p35S:*AkCSLA3*:YFP-NOS1 (LIC strain) were transformed into GV3101 *Agrobacterium tumefaciens* competent cells.

The preparation of the calcium competent cells leaned on Berestovoy et al. (2018) with some adaptations. GV3101 were re-streaked freshly onto IsLB plate containing 50 µg/ml Rifampicin and 25 µg/ml Gentamycin two day before GV3101 strain was pre-grown in 5 ml IsLB at 30°C and 250 rpm in a MaxQ™ 6000 incubator overnight. 500 µl of this pre-culture was inoculated into 50 ml IsLB medium and was grown at the same conditions for another overnight incubation. On the next day, the culture was cooled on ice for 30 min, 45 ml culture were transferred into a 50 ml Falcon tube and centrifuged for 10 min at 3,800 x g at 4°C in a swing bucket Eppendorf 5810R centrifuge. The pellet was washed with 5 ml ice-cold 20 mM calcium chloride, centrifuged for once again for 5 min and, then the pellet was resuspended in 1 ml ice-cold 20 mM calcium chloride solution. The calcium competent cells were aliquoted in 50 µl reactions and were directly used for transformation or stored at -70°C.

The transformation protocol was performed with the freeze-thaw method based on Dityatkin et al. (1972). Therefore, 200 ng of plasmid DNA and 100 ng helper plasmid pSOUP (Hellens et al., 2000) were mixed with competent cells and were hold on ice for 15 min. Afterwards, tubes were frozen in liquid nitrogen for 5 min, followed by a heat shock at 37°C for 5 min in a water bath. Then, cells were incubated for further 5 min on ice before the cells were mixed with 1 ml IsLB and incubated at 30°C and 250 rpm for 4 h. After the incubation, cells were centrifuged for 5 min at 2,000 x g, cells were resuspended in 200 µl IsLB and plated onto IsLB plates supplemented with 80 µg/ml Carbenicillin and 50 µg/ml Rifampicin for MoClo constructs and .50 µg/ml Kanamycin and Rifampicin for the pCV01 + p35S:AKCSLA3:YFP-tNOS strain. All created constructs for plant transformation in this study are summarized in Table S8. Plates were incubated for 2–3 days in a MaxQ™ 6000 incubator and Agrobacteria colonies were genotyped by colony PCR as described in section 2.4 with primers summarized in Table S9.

2.14 Plant Growth Conditions

In case of *N. benthamiana* (Tobacco), plants were directly grown in 10 cm pots filled with CL Pikier soil (Patzner, Cat# 10-00200) in a greenhouse at 18–22°C with constant light for 16 h (from 6 am to 10 pm). Tobacco plants were grown for 5 to 6 weeks until they were used for transient transformation with agrobacteria (see section 2.15).

For *A. thaliana*, seeds were poured onto Murashige and Skoog (MS) plates (0.216% (w/v) MS, Sigma Aldrich, Cat# M5519; 0.7% (w/v) plant agar, Carl Roth, Cat# 4807.2) and were stored at 4°C in the dark for 4 days to ensure their simultaneous germination. After this stratification, seeds on MS plates were placed into the phytochamber for germination (~5–7 days) and then, transferred into 8 cm round pots filled with CL Pikier soil mixed with vermiculite (ratio 4:1). All plants have grown at 100–120 µmol m⁻² s⁻¹ light, 22°C and 60%

relative humidity. After the plants have started to flower, they were staked and matured until the whole plant was completely dried (~10–12 weeks). The plants were harvested, seeds were separated from the chaff and were stored in individual paper bags for each plant.

The liverwort *Marchantia polymorpha* (Tak1: male and Tak2: female) were shared from the Quint Lab (Martin-Luther University Halle-Wittenberg, DE) and were grown on 90 x 90 cm plates containing 0.2% (w/v) Gamborg B5 media (Duchefa, Cat# G0210.0050) supplemented with 1% (w/v) sucrose (adjusted with potassium hydroxide to pH 5.5) and 1% plant agar. For propagation, *M. polymorpha* gemmae or pieces of the thalli were transferred onto new plates as described in Ishizaki et al., (2015).

Physcomitrium patens ecotype Gransden (wildtype) and generated *cs1a* mutants were a gift from the Bezanilla Lab (Dartmouth College, US). They were grown on PpNH₄ plates (1.03 mM MgSO₄, 1.86 mM KH₂PO₄, 3.3 mM CaNO₃, 2.7 mM di-ammonium tartrate, 9.93 μM H₃Bo₃, 220 nM CuSO₄, 1.966 μM MnCl₂, 231 nM CoCl₂, 191 nM ZnSO₄, 169 nM KI, 103 nM Na₂MoO₄) supplemented with 0.7% (w/v) phyto agar (Duchefa, Cat# P1003.1000). Moss tissue was propagated in two different ways. To propagate moss gametophytes, one gametophyte of each strain was transferred to four spots onto PpNH₄ plates one time per month. Protonematal tissue was propagated every week by shredding ~ 2cm of protonemata in 1 ml sterile MQ water with a Fisherbrand™ homogenizer (Fisher Scientific, Cat# 15565809) for 10 s at 4375 rpm. The homogenized protonematal tissue was transferred onto PpNH₄ plates covered with a cellophane disk (Hans Schütt, Cat# 23109000). The cellophane enables the easy removal of protonematal tissue from the plate. *M. polymorpha* and *P. patens* were grown in a cabinet in the lab equipped with a LED lamp for plant growth containing 169 LED lights with illumination in the blue and red spectra (Amazon, ASIN: B08SWCSWYF) at 100 μmol/m² s⁻¹ and 22°C.

2.15 Transient expression in *N. benthamiana* leaves

Agrobacterium strains carrying the genes of interest, or the silencing inhibitor (P19) that was a gift from the Bürstenbinder group (IPB, Halle Saale) were grown in 3–4 ml IsLB supplemented with the respective antibiotic (50 μg/ml Rifampicin and 80 μg/ml Carbenicillin (for MoClo vectors) or 50 μg/ml Kanamycin (for P19 strain) for 24 h at 30°C and 250 rpm. After growth, Agrobacterium strains were harvested by centrifugation at 3,800 x g for 5 min in a swing bucket Eppendorf 5810R centrifuge before the transient expression into *N. benthamiana* leaves was proceeded as described in Grefen et al. (2010). Agrobacterium strains were two times washed with 2 ml freshly prepared infiltration medium (10 mM MES-KOH buffer (pH 5.3–5.5), 10 mM magnesium chloride hexahydrate, 150 μg/ml acetosyringone) and finally resuspended in 1 ml infiltration medium. For each strain, the OD₆₀₀ was determined using a BioSpectrometer® basic and the gene of interest was mixed together with the P19 suppressor

to an OD₆₀₀ of 0.7 and further incubated for 1 h at RT and 250 rpm in a neoMix 7-0921 thermomixer. These agrobacteria mixtures were infiltrated randomly into the bottom of three different leaves of 5–6 weeks old *N. benthamiana* plants and was repeated for four different plants. Four days after infiltration, the leaves were harvested, and the yellow fluorescence of the whole leaf was observed firstly with a UVP Gelstudio PLUS Touch using a GFP emission filter (AnalytikJena, Cat# 849-00405-0) and blue LED light. Afterwards, the yellow fluorescence was visualized with a confocal microscope (LSM700) for each spot to see the localization of the fluorescence in more detail as described in section 2.6.

2.16 Stable expression in *A. thaliana*

GV3101 agrobacterial strain containing pCV01 + p35S:AkCSLA3:YFP-tNOS were stably transformed into *A. thaliana* plants by the floral spray method based on the protocol “*Arabidopsis* transformation by floral spraying *Agrobacterium*” [Ye, 2009]. Therefore, one colony of this agrobacterial strain was inoculated into 3 ml IsLB supplemented with 50 µg/ml Rifampicin and Kanamycin in glass culture tubes and was incubated 2 days over night at 30°C and 250 rpm. Afterwards, 1 ml of the pre-culture was transferred into 150 ml IsLB medium and further grown for 20–24 h at the same conditions. The *Agrobacterium* was harvested at 3,000 x g for 10 min and the cell pellet was resuspended in 100 ml infiltration medium (5% (w/v) sucrose, 0.05% (v/v) silwet-77). Nineteen Col-0 or *csla239* flowering *A. thaliana* plants were dipped into the infiltration solution and were placed for 24 h in the dark. After 24 h, the plants were placed back to light and grown further under continuous light conditions as described in section 2.14 until the plant was matured and completely dried. At this stage the whole plant material was harvested in paper bags, seeds were cleaned from the chaff and stored in separated bags.

Around 200 seeds of each transformed plant were poured onto 8 cm pots and germinated under constant light conditions for 2 weeks. To select positively transformed plants, germinated seedlings were sprayed with 40 mg/l glufosinate ammonium solution (BASTA® active ingredient, Sigma Aldrich, Cat# 45520-100MG). Surviving plants were transferred to single pots and screened for yellow fluorescence (488 nm excitation, 505–550 nm emission) on a LSM700 confocal microscope (see section 2.6) and were verified by Touch-and-Go PCR as described in Berendzen et al. (2005). Therefore, a leaf was punched with a 10 µl pipette tip and was mixed in 6 µl HyClone water that was transferred into a PCR tube. To this mixture, 10 µl RedTaq 2x DNA Polymerase Master Mix, 1 µl 10 mM gene-specific forward and reverse genotyping primers (see Table S9) were added and the PCR was performed as described in section 2.2 with the following adaptation: Initial denaturation for 2 min at 95°C and 30–35 cycles of denaturation, annealing and extension were implemented. 5 µl of the PCR mixture was loaded onto 0.7% (w/v) agarose gels and the gel electrophoresis was performed as described

in section 2.2. Positively genotyped plants were grown to maturity to harvest their seeds and later stain their mucilage with ruthenium red.

2.17 Mucilage Staining and Extraction

Seed mucilage of stable transformed pCV01 + p35S:AkCSLA3:YFP-tNOS in Col-0 and *csla239* mutant *A. thaliana* were stained with ruthenium red (RR, Sigma Aldrich, Cat# R2751-1G) as described in Voiniciuc and Günl (2016). Therefore, around 30 seeds of each plant were transferred into wells of a 24-well plate (VWR, Cat# 734-2779) pre-filled with 500 µl MQ water to imbibe the seeds for 5 min while shaking slightly. Afterwards, the water was replaced with 0.01% RR and was incubated for 5 min followed by washing the seeds with 300 µl of water. The mucilage of each seed was imaged using a stereo zoom microscope OZL-46 (Carl Roth, Cat# HHE8.1).

To extract the total mucilage of *A. thaliana* seeds, 4–6 mg seeds were transferred together with 1 ml of 30 µg/ml ribose into 2 ml Safe Lock tubes and were shaken for 30 min at 30 Hz using a Retsch MM400 (after 15 min of shaking, the block was rotated around its vertical axis on 180° and shaken for further 15 min). Seeds were settled down and 800 µl of each supernatant was transferred into 2 ml screw cap tubes. To determine the sugar content of the mucilage, a 9-sugar mix containing Fuc, Rha, Ara, Gal, Glc, Xyl, Man, GalA and GlcA was pipetted as several dilutions (0, 1, 2, 5, 10, 25, 50, 75, 100, 125 µg) into 2 ml Screw Cap tubes and 1 ml of 30 µg/ml ribose was added on top. Samples and standards were dried using an Eppendorf concentrator 5301 at 30°C. The hydrolysis was performed the same as described in section 2.9. Dried, hydrolysed samples and standards were finally eluted in 400 µl MQ water, mixed at 30 Hz for 1:30 min in a Retsch MM400 and 100 µl were transferred into an IC vial to analyse their monosaccharide composition using the Metrohm system. Most of the settings used in section 2.9 were the same to detect sugars from plant material, but for better separation a 60 min protocol was used. Equilibration of the column and separation of the neutral sugars was done using 2 mM NaOH and 3.4 mM NaAce for 22 min. Afterwards uronic acids were separated with 100 mM NaOH and 170 mM NaAce for 21 min, followed by rinsing the columns with 2 mM NaOH for 15 min. The quantification of the monosaccharides was performed the same as described in section 2.9.

2.18 Monosaccharide Analysis of Plant AIR material

Plant tissue was harvested in a 2 ml snap cap tube and frozen first in liquid nitrogen, followed by freeze drying the samples for 24 h in a LDplus freeze dryer (Martin Christ Gefriertrocknungsanlagen GmbH, DE). Afterwards, plant material was ground with a steel ball (3 mm, VWR, Cat# 412-2918) using a Retsch MM400 for 2 min at 30 Hz. The AIR material was extracted based on Pettolino et al. (2012) and Pauly et al. (2016). Therefore, plant material

was successively washed with 1 ml of 70% (v/v) ethanol, 1 ml of chloroform:methanol (1:1 v/v) and 1 ml of acetone using the Retsch MM400 for 1:30 min at 30 Hz. In between every washing step, the samples were centrifuged at 20,000 x g for 5 min. Finally, the AIR was re-solved in 300 µl acetone and was dried overnight in a fume hood.

In order to destarch plant material, around 2 mg of AIR material were treated with 1 ml starch digestion mixture containing 0.001% (v/v) α-amylase (Sigma Aldrich, Cat# A3403-500KU), 0.01% (v/v) pullulanase (Sigma Aldrich, Cat# E2412-50ML) and 0.5% (v/v) 0.2 M phosphate buffer). The tubes were incubated at 30°C for 30 min in a neoMix 7-0921 thermomixer. Afterwards, samples were cooled on ice, and were incubated horizontally for 20 h at 225 rpm and 37°C in a MaxQ™ 6000 incubator. After destarching, the undigested material was pelleted by centrifugation at 20,000 x g for 2 min, the starch that remain in the supernatant was removed and the pellet was washed two times with 1.5 ml MQ water and one time with 1.5 ml acetone. The acetone wash was discarded after the last centrifugation step and the AIR material was evaporated at 60°C without blowing air to the samples.

The precise AIR weight was recorded to prepare a homogenous 1–1.5 mg/ml slurry using a steel ball and a Retsch MM400 for up to 15 min at 30 Hz. 300 µg of this AIR slurry were used for hydrolysis and the protocol was performed as described in section 2.9. After hydrolysis, samples were cooled and centrifuged briefly at 20,000 x g. 150 µl supernatant were transferred into a new 2 ml screw cap tube and were evaporated at 45°C. As monosaccharide standards a 9-sugar mix was used (see section 2.17) and dried down like the samples. All samples and standards were finally eluted in 300 µl 30 µg/ml ribose solution. 150 µl samples and standards were transferred to an IC vial and loaded into the Metrohm system with a 60 min plant method (see section 2.17) to analyse the plant polysaccharides of the AIR.

3. Results

3.1 Detailed Analysis of Fluorescently-Tagged *AkCSLA3*

Labelling proteins with a fluorescent tag is a powerful tool for many applications in the molecular biological science. Attaching a fluorescent protein (FP) to a protein of interest gives insights about its expression, trafficking, localization, interaction with other proteins and degradation in living organisms [Chudakow et al., 2009]. A broad variety of FPs that cover the whole visible spectrum are available and, there are many things to consider to make this approach as useful as possible.

It plays an important role how to fuse the FP to the protein of interest. The fluorescent tag can be fused on the C- or N-terminal end of the protein of interest. Dependent on which termini of the protein the fluorescent tag is fused to it could affect the role of the protein to determine the right localization or to interact with other proteins. *Pichia* expressing *AkCSLA3* fused to sfGFP at its N-terminal end (sfGFP-*AkCSLA3*) have been yielded in less HM production compared to its untagged counterpart [Voiniciuc et al., 2019]. In this study, sfGFP was cloned additionally to the C-terminal end of *AkCSLA3* (*AkCSLA3*-sfGFP) to test whether the loss of activity caused by the N-terminal fusion can be avoided. Therefore, the fluorescence and monosaccharide content after several timepoints of protein expression was observed (Figure 8A). After 6 h of protein expression, the fluorescence of sfGFP and *AkCSLA3*-sfGFP was already detectable, while sfGFP-*AkCSLA3* shows significant changes in fluorescence only after 24 h (Figure 8A). While the fluorescence of sfGFP was detected in similar levels from 6 h up to 24 h after protein expression, the highest fluorescence was detectable after 9 h for *AkCSLA3*-sfGFP and 24 h for sfGFP-*AkCSLA3*. Because the protein expression is under the control of a methanol-inducible promoter and the methanol concentration in the media is depleted over time the fluorescence levels were decreased after 48 h and 72 h of protein expression (Figure 8A). Nevertheless, the HM content seem to be very stable in *Pichia*, meaning that HM accumulation has reached a saturation point after 24 h growth in YPM without further addition of methanol. The mannose content where the highest fluorescence was observed is similar to that of upcoming timepoints. In all timepoints, *AkCSLA3*-sfGFP showed higher fluorescence and mannose contents compared to sfGFP-*AkCSLA3* (Figure 8B).

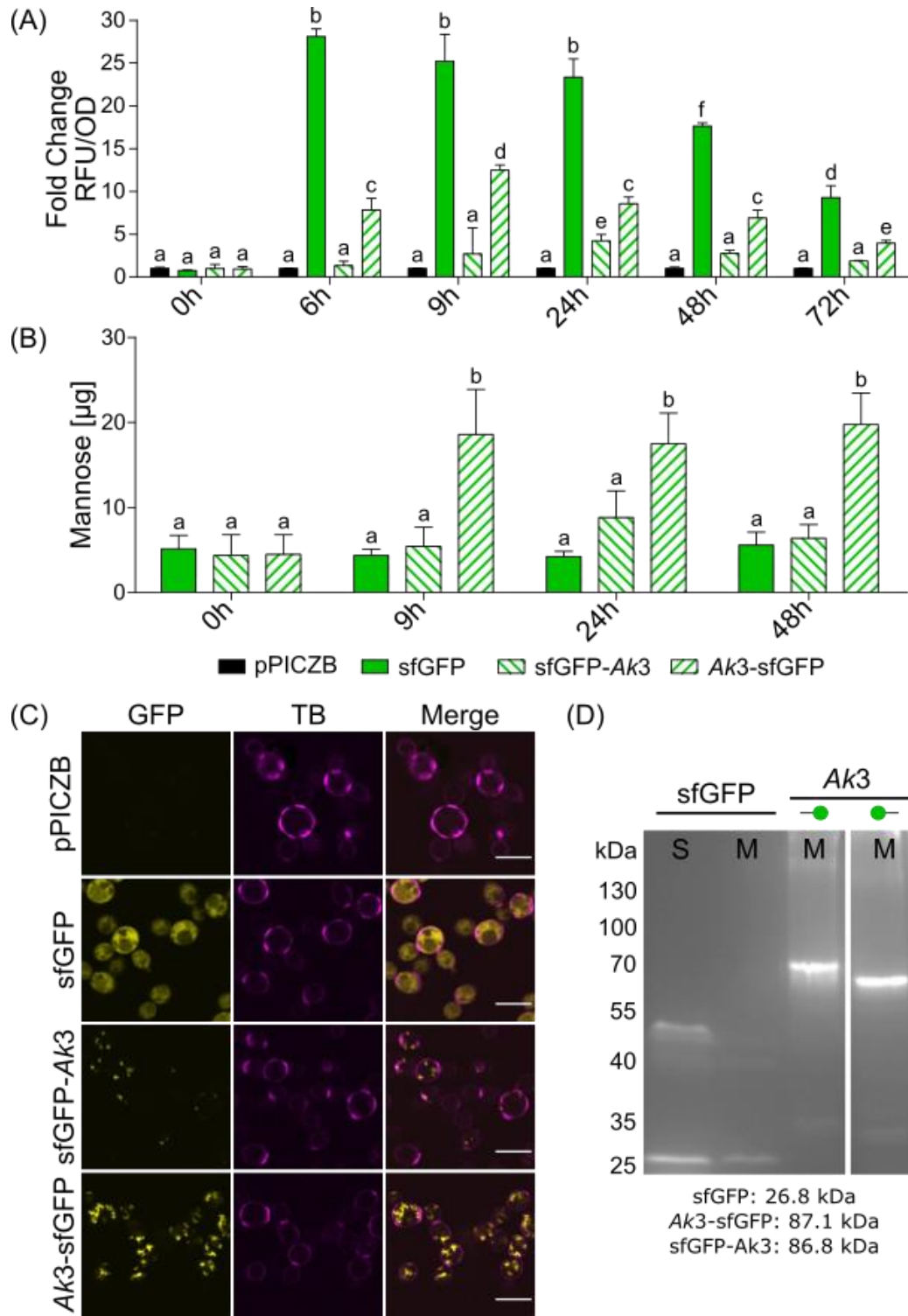


Figure 8: *Pichia* expressing AkCSLA3 with C- or N-terminal tagged sfGFP. Yeast strains were first cultivated in BMGY for 24 h to produce biomass, followed by 24 h growth in BMMY for protein expression. (A): Fold change of normalized relative fluorescence units (RFU/OD) after 0, 6, 9, 24, 48 and 72 h relative to AkCSLA3. (B) Mannose content of AKI wall polymers were hydrolysed with sulfuric acid after 0, 9, 24 and 48 h of protein expression. Data show mean + SD from three biological replicates and letters denote significant changes (one-way ANOVA with Tukey test $p < 0.05$). (C) Localization of AkCSLA3 24 h after protein expression in *Pichia*. The plasma membrane was labelled with trypan blue (TB). The intensity of green fluorescent proteins (GFP) was adjusted for all channels after the acquisition relatively to sfGFP. Scale bar: 5 μm . (D) In-gel fluorescence of the soluble (S) and microsome (M) fraction of expressed proteins in *Pichia* following SDS-PAGE under non-denaturing conditions. Green dots symbolizing sfGFP fusions on the C- or N-terminus of AkCSLA3. Ak3: *Amorphophallus konjac* CSLA3.

To ensure that the higher fluorescence and mannose contents of *AkCSLA3*-sfGFP were not observed due to cleavage or dimerization/aggregation of the FP, the localization (Figure 1C) and the protein size (Figure 8D) were determined. sfGFP was visible in the cytoplasm, while both tagged *AkCSLA3*s are localized in puncta (Figure 1C). Correlating with the levels of fluorescence and mannose content, *AkCSLA3*-sfGFP show more and brighter fluorescent puncta than sfGFP-*AkCSLA3* (Figure 1C). Soluble and microsomal fractions of sfGFP and microsomes of both fluorescently tagged *AkCSLA3* were extracted from *Pichia* cultures. sfGFP show the brightest band around 26.8 kDa in the soluble fraction representing sfGFP in the cytosol and only low abundance in the microsome fraction that represents sfGFP in the membranes. An additional band around 50 kDa was detected in sfGFP, pointing out dimer formation of sfGFP on its own. The brightest band is detectable around 70 kDa for *AkCSLA3*-sfGFP and around 65 kDa for sfGFP-*AkCSLA3* while no band at 26.8 kDa was visible, meaning that sfGFP cleavage occurs only in minor extent.

In general, *AkCSLA3*-sfGFP is the better option than sfGFP-*AkCSLA3* to analyse the protein function in detail, but it still has to be validated whether *AkCSLA3*-sfGFP has the same activity as untagged *AkCSLA3*. Additionally, two other FPs (Venus and mRuby2) were tested because the protein expression can influence the folding, maturation, brightness and stability of the FP [Johnson and Straight, 2013]. An overview of the properties from selected FPs are listed in Table 2.

Table 2: Overview of fluorescent proteins used for C-terminal fusion with *AkCSLA3*. Values are from fpbase.org (date: 2022.03.21).

	sfGFP	Venus	mRuby2
Extinction λ (nm)	485	515	559
Emission λ (nm)	510	528	600
Brightness	54.15	52.55	42.94
Photostability ($t_{1/2}$ s)	208.26	15	123
Maturation (min)	13.6	17.6	150
Oligomerization	Weak dimer	Weak dimer	Monomer
Fpbase ID	B4SOW	YUJWJ	8MJ78

Most commonly used FPs are green and yellow FPs due to their high brightness and fast maturation (as shown for sfGFP and Venus; s. Table 2). sfGFP binds extremely stable independent of the solubility and misfolding status of the fusion partner [Pedelacq et al., 2006]. Venus could be useful because its pH sensitivity is reduced compared to other yellow FPs (pKa 6.0), even though the fluorescence stability of Venus is very low ($t_{1/2}$ 15) [Nagai et al., 2002]. The formation of weakly dimers of both FPs is not optimal because dimerization can lead to mislocalization of the protein of interest, but sfGFP fusions have shown to be more soluble and, therefore, could be used for cell trafficking experiments [Pedelacq et al., 2006]. On the contrary, mRuby2 was selected because it is one of the brightest red FPs and is a monomeric FP that reduces the possibility for mislocalization and aggregation. However, most

FPs were analysed *in vitro* with purified proteins (Heppert et al., 2016), why the fluorescence of the selected FPs fused to *AkCSLA3* on the C-terminus can differ a lot from the literature and has to be evaluated.

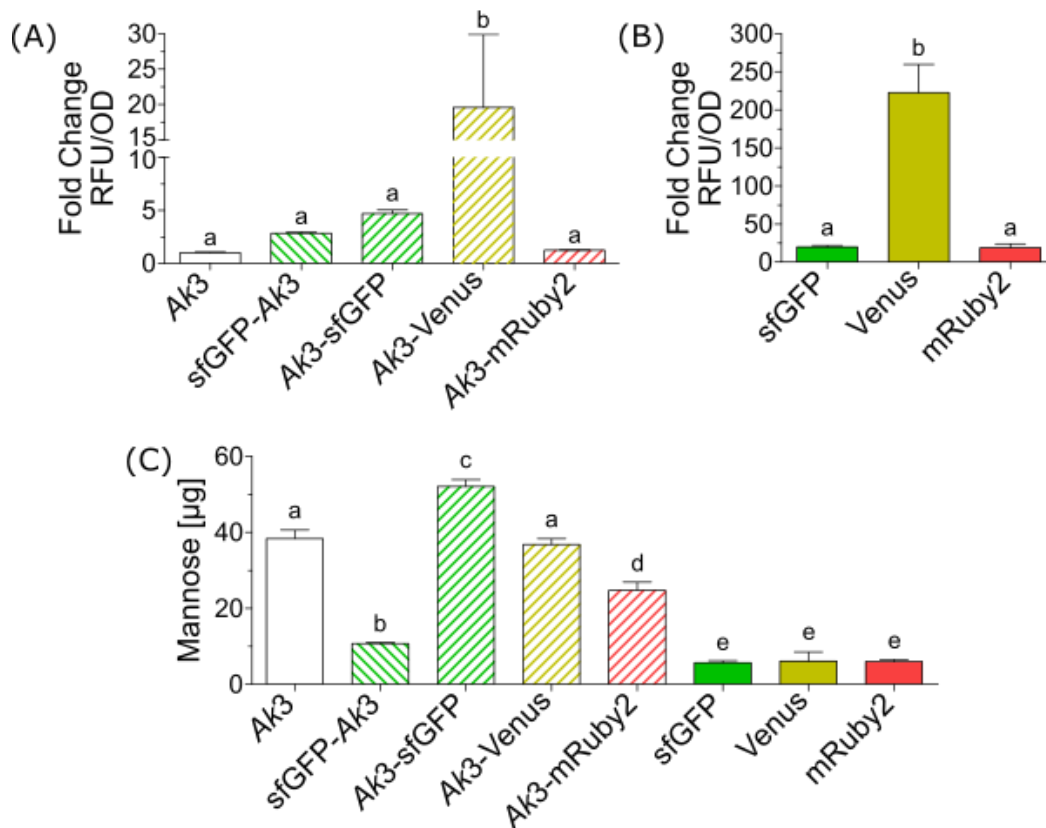


Figure 9: *Pichia* expressing different fluorescent proteins fused with *AkCSLA3* at the C-terminal end. Yeast strains were cultivated for 24 h in BMGY medium for biomass production followed by 24 h growth in BMMY for protein expression. (A) Fold change of normalized relative fluorescence units (RFU/OD) of fluorescently tagged *AkCSLA3* (N- or C-terminal end) relative to *AkCSLA3*. (B) RFU/OD of *Pichia* expressing the fluorescent protein as a negative control relative to *AkCSLA3*. (C) Mannose content of AKI wall polymers. Polymers were hydrolysed with sulfuric acid. Data show the mean + SD of three biological replicates and letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). *Ak*: *Amorphophallus konjac* *CSLA3*.

All FPs show a bright fluorescent signal when they are expressed in *Pichia* (Figure 9B). The change in fluorescence of Venus is extremely high (approximately 223 times higher compared to non-fluorescent *AkCSLA3*), while it is only 20 times higher for sfGFP and mRuby2 (Figure 9B). A similar trend is detectable for *AkCSLA3*-FPs fusions (Figure 9A). *Pichia* expressing *AkCSLA3*-Venus show the brightest fluorescence (~20x fold change) compared to *AkCSLA3*, followed by *AkCSLA3*-sfGFP (~5x fold change), and *AkCSLA3*-mRuby2 that show only marginal fluorescence (~1.25x fold change). All *Pichia* strains expressing FPs fused on the C-terminus of *AkCSLA3* produce more Mannose than sfGFP-*AkCSLA3* (Figure 9C). In contrast to the fluorescence data, the mannose content of *AkCSLA3*-sfGFP is the highest (~52 µg) and even higher than *AkCSLA3* on its own (~38.4 µg). *AkCSLA3*-Venus has the same activity as *AkCSLA3* and produce similar amounts of mannose (~36.8 µg), while *AkCSLA3*-

mRuby2 (~24.7 µg) show reduced activity in HM production (Figure 9C). These results show that sfGFP as well as Venus could be used as C-terminal fusion on AkCSLA3. Due to the fact that sfGFP fusions are more soluble than other fusions with FPs, AkCSLA3-sfGFP was used to determine its localization.

3.2 Assessing the Subcellular Localization of AkCSLA3

To investigate the subcellular localization of AkCSLA3, fluorescently tagged signal peptides or proteins from *Saccharomyces cerevisiae* (Sc; hereafter *Saccharomyces*) that are known to be localized in the Golgi apparatus were co-expressed together with AkCSLA3-sfGFP. In order to determine not only the localization but also the trafficking at earlier stages of AkCSLA3 expression, the Golgi proteins were expressed under the control of a strong constitutive promoter (pGAP). Furthermore, Golgi proteins were fused to mRuby2 at their C-terminus to differentiate the signals from the green fluorescence of sfGFP.

ScMNN2 (Accession number NC_001134.8) from *Saccharomyces* was selected as a candidate gene for heterologous expression in *Pichia* because its localization is visible as clear puncta that belong to the early Golgi [Rayner and Munro, 1998] in *Saccharomyces*. Previous studies have shown that all N-mannosyltransferases from *Saccharomyces*, including ScMNN2, are classified as a type II membrane proteins having a single transmembrane domain (TMD) at its N-terminal end [Rayner and Munro, 1998; Dean, 1999]. The TMD sequence of a protein is essential to determine the proteins localization and can serve as an anchor to retain the protein in the Golgi apparatus or used as a signal when the protein was wrongly translocated to be sent to its right destination [Munro, 1998; Gleeson, 1998]. After 24 h of growth in BMGY and further growth in BMMY for 24 h the fluorescence of ScMNN2(AA1-46)-mRuby2 was localized in a double ring pattern (data not shown) that is normally shown for proteins localized in the ER [Soderholm et al., 2004]. Therefore, ScMNN2(AA1-46)-mRuby2 could not be used to determine the localization of AkCSLA3.

Although the protein localization of *Saccharomyces* is well studied and documented in the *Saccharomyces* genome database (<https://www.yeastgenome.org/>), it does not necessarily mean that heterologously expressed proteins from *Saccharomyces* are localized in the same organelles in *Pichia* because both yeasts are phylogenetically far apart from each other. But to increase the possibility that further selected proteins from *Saccharomyces* are Golgi-localized and also to minimize possible error reasons leading to mis-localization, the entire coding sequence of the selected proteins was cloned into *Pichia*. Additionally, the *Pichia* genome atlas website (created based on the publication from Valli et al. (2020)) was used to compare whether the localization of the selected *Saccharomyces* protein is the same as the homologous protein from *Pichia*. Using these tools, ScOCH1 and ScGOS1 from

Saccharomyces were selected for cloning and expression in *Pichia*. An overview of the selected proteins is listed in Table 3.

Table 3: Overview of selected proteins for Golgi localization in *Pichia*. Localization of *Saccharomyces* proteins was evaluated using the *Saccharomyces* genome database, whereas the localization of homologous proteins from *Pichia* was evaluated by the *Pichia* genome atlas website. The website normalized the Log2FC for the signal intensity that was detected for each protein in a specific organelle, meaning the closer the value is to 1, more protein was detected in the organelle.

		ScOCH1	ScGOS1
Accession number		NC_001139.9	NC_001140.6
Size (amino acids)		480	223
Function		α -1,6-mannosyltransferase elongate backbone consisting of α -1,6-linked mannose residues	v-SNARE protein involved in Golgi vesicle transport and fusion
Localization	<i>Saccharomyces cerevisiae</i>	Cis-Golgi cisterna	Cis- to medial Golgi cisterna
	<i>Pichia</i> (Glc/MeOH metabolism)	trans-Golgi (0.65) ER (0.4)	cis-Golgi (0.22/0.26) trans-Golgi (0.66); ER (0.08)
Source		Nakayama et al. (1992 and 1997); Harris and Waters (1996)	McNew et al. (1998)

Rossanese et al. (1999) have shown Golgi-localization of ScOCH1 expressed in *Pichia* using immunoelectron microscopy. Moreover, co-localization of ScOCH1 with ScSEC7 has been visualized by antibody staining in *Pichia* cells and ScSEC7 is known to be localized in the late Golgi in *Saccharomyces*. Unfortunately, even after 43 h growth in BMGY ScOCH1-mRuby2 have shown no red fluorescence when it was expressed in *Pichia* (data not shown) and was therefore excluded.

ScGOS1 is a v-SNARE protein and is involved in Golgi vesicle transport and fusion in *Saccharomyces*. It is localized in the cis- to medial-Golgi cisterna [McNew et al., 1998]. In *Pichia*, ScGOS1 is mostly localized in the cis-Golgi when glucose is used as a carbon source but is also localized in minor abundance in the ER. In contrast, the localization of ScGOS1 is mostly shifted to the early Golgi with minor localization in the cis-Golgi when methanol is utilized by the *Pichia* cultures (Table 3). Yeast strains were grown for 48 h in BMGY followed by 24 h growth in BMMY. The fold change of normalized green and red fluorescence over time was measured after the medium was changed to BMMY and is presented in Figure 10. The green fluorescence of AkCSLA3-sfGFP and AkCSLA3-sfGFP co-expressed together with ScGOS1-mRuby2 was significantly increased after 3 h with the highest fluorescence after 9 h of methanol induction compared to AkCSLA3 and all measurements at timepoint 0 h (Figure 10A). Although ScGOS1-mRuby2 and mRuby2 were cloned under the control of the constitutive promoter pGAP and, therefore, should be expressed during the prior growth in

BMGY for 24 h (data not shown), the red fluorescence was not detectable after the media was replaced to BMMY (0 h, Figure 10B). The red fluorescence increased to significant levels only after 6 h (for ScGOS1-mRuby2) and 24 h (for mRuby2) of growth in BMMY (Figure 10B), indicating the protein expression under the control of pGAP promoter in BMGY was not optimal.

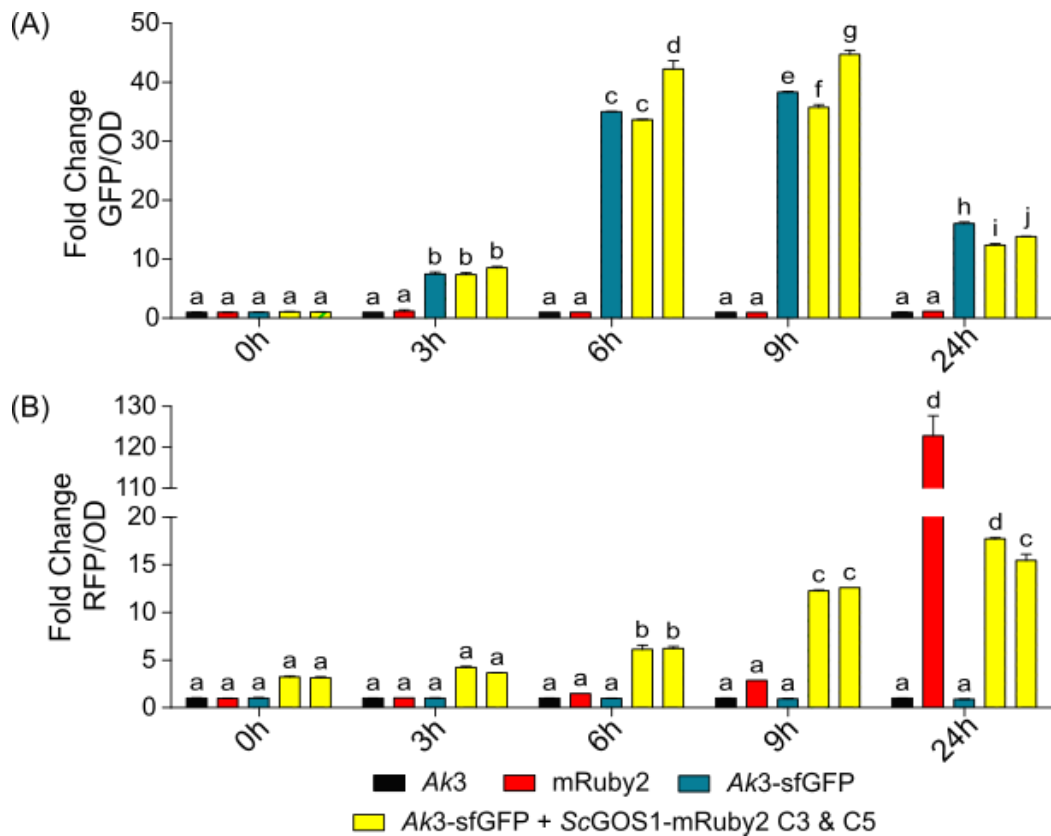


Figure 10: Fluorescence of *Pichia* co-expressing *AkCSLA3*-sfGFP and *ScGOS1*-mRuby2. *Pichia* cultures were grown for 24 h as pre-cultures in 3.5 ml BMGY in glass tubes. After 24 h the cultures were centrifuged, resolved in BMMY and 600 μ l were transferred into 24-well plates to harvest one well for each timepoint (0, 3, 6, 9 and 24 h after protein expression). (A) Fold change of normalized green fluorescence units (GFP/OD) and (B) red fluorescence units (RFP/OD) relative to *AkCSLA3*. Data show the mean + SD of two biological replicates and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). *Ak3*: *Amorphophallus konjac* CSLA3; *Sc*: *Saccharomyces cerevisiae*.

The protein expression under the control of the pGAP promoter is highly variable dependent on the carbon source that is used. Waterham et al. (1997) have shown a 30% decrease of β -lactamase expression when glycerol was used as a carbon source instead of glucose. Therefore, the green and red fluorescence development for all Golgi-marker strains and controls was observed during growth in YPD over 43 h to find the optimal expression period (Figure 11). During the time measurement the green fluorescence for all strains (including the negative control *AkCSLA3*) was increased from approximately 2500 at timepoint 0 h to 2750 or 3000 after 9 h up to 18 h of growth in YPD. In later timepoints, the green fluorescence was only slightly increased and reached its saturation point after ~33 h for each strain (Figure 11A). But no green fluorescence was visible with a confocal microscope for all

strains (data not shown). Red fluorescence was only increased over time for *Pichia* strains expressing mRuby2 or ScGOS1-mRuby2, whereas AkCSLA3, sfGFP and AkCSLA3-sfGFP strains have shown the same background level for all time points (Figure 11B). An increase in the red fluorescence was already detected after 3 h of growth in YPD for *Pichia* expressing ScGOS1-mRuby2 whereas the red fluorescent signal of mRuby2 and AkCSLA3-sfGFP co-expressed with ScGOS1-mRuby2 started to increase after 21 h of growth in YPD (Figure 11B). Furthermore, the increase in fluorescence for ScGOS1-mRuby2 and mRuby2 cells was higher than for AkCSLA3-sfGFP that are co-expressed together with ScGOS1-mRuby2. After 43 h growth in YPD, all strains have not reached their maximal fluorescence (Figure 11B), therefore 48 h growth in YPD was selected for further growth Batches for reasons of simplification.

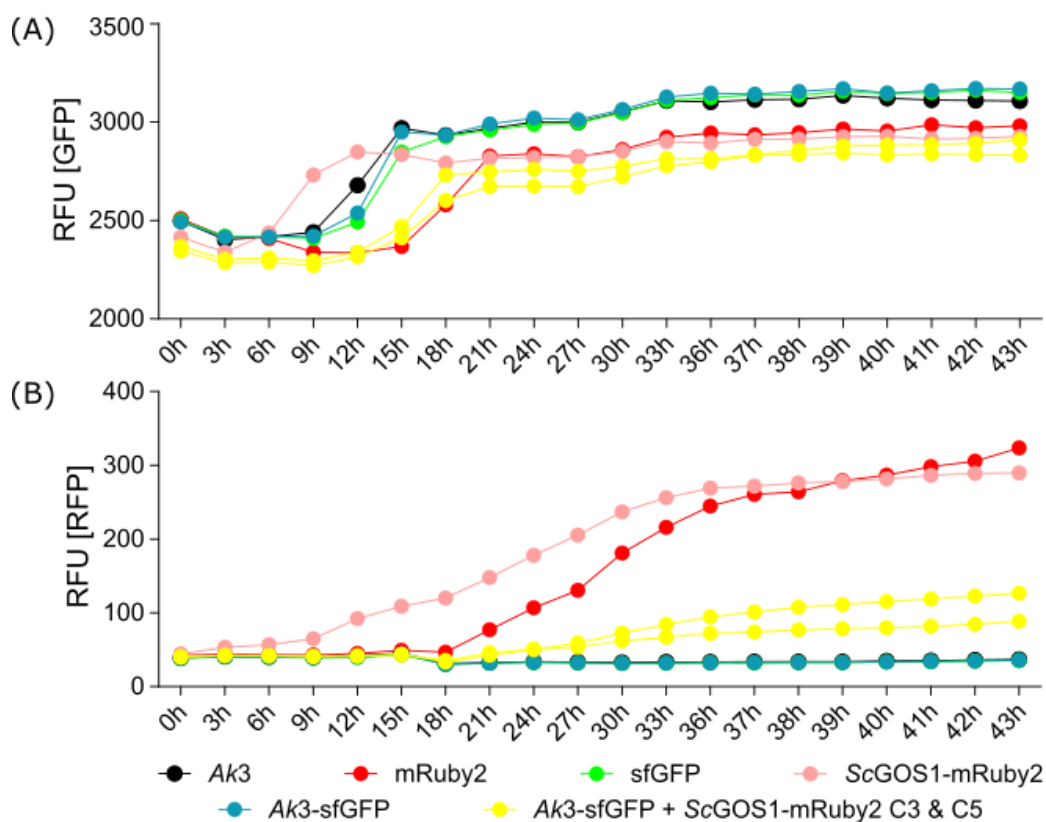


Figure 11: Fluorescent screening of ScGOS1-mRuby2 expressing *Pichia* strains during growth in YPD medium over time. *Pichia* cultures were grown in 300 μ l YPD in a 48-well plate. Relative fluorescence units (RFU) of (A) green (GFP) and (B) red (RFP) fluorescence was observed for 43 h. Data show the mean of two biological replicates, except for ScGOS1-mRuby2 only one replicate is shown. Ak3: *Amorphophallus konjac* CSLA3; Sc: *Saccharomyces cerevisiae*.

Modifying the growth conditions to YPD for 48 h led to high fold changes in normalized red fluorescence for *Pichia* expressing mRuby2 and ScGOS1-mRuby2, without significant changes in the green fluorescence for sfGFP and AkCSLA3-sfGFP expressing *Pichia* cultures (Figure 12A). mRuby2 strains have shown the highest fold change in normalized red fluorescent signal (~113), followed by AkCSLA3-sfGFP and ScGOS1-mRuby2 co-expressing strains (~54) and ScGOS1-mRuby2 strains (~36) (Figure 12A). Although there is no significant increase in the green fluorescence of sfGFP expressing strains after 48 h of growth in YPD (Figure 12A), fluorescence could be detected at low levels with a confocal microscope after 48 h of growth in YPD (Figure 12B). Localization of mRuby2 was visible in the cytosol, while it was localized in puncta for ScGOS1-mRuby2 (Figure 12B).

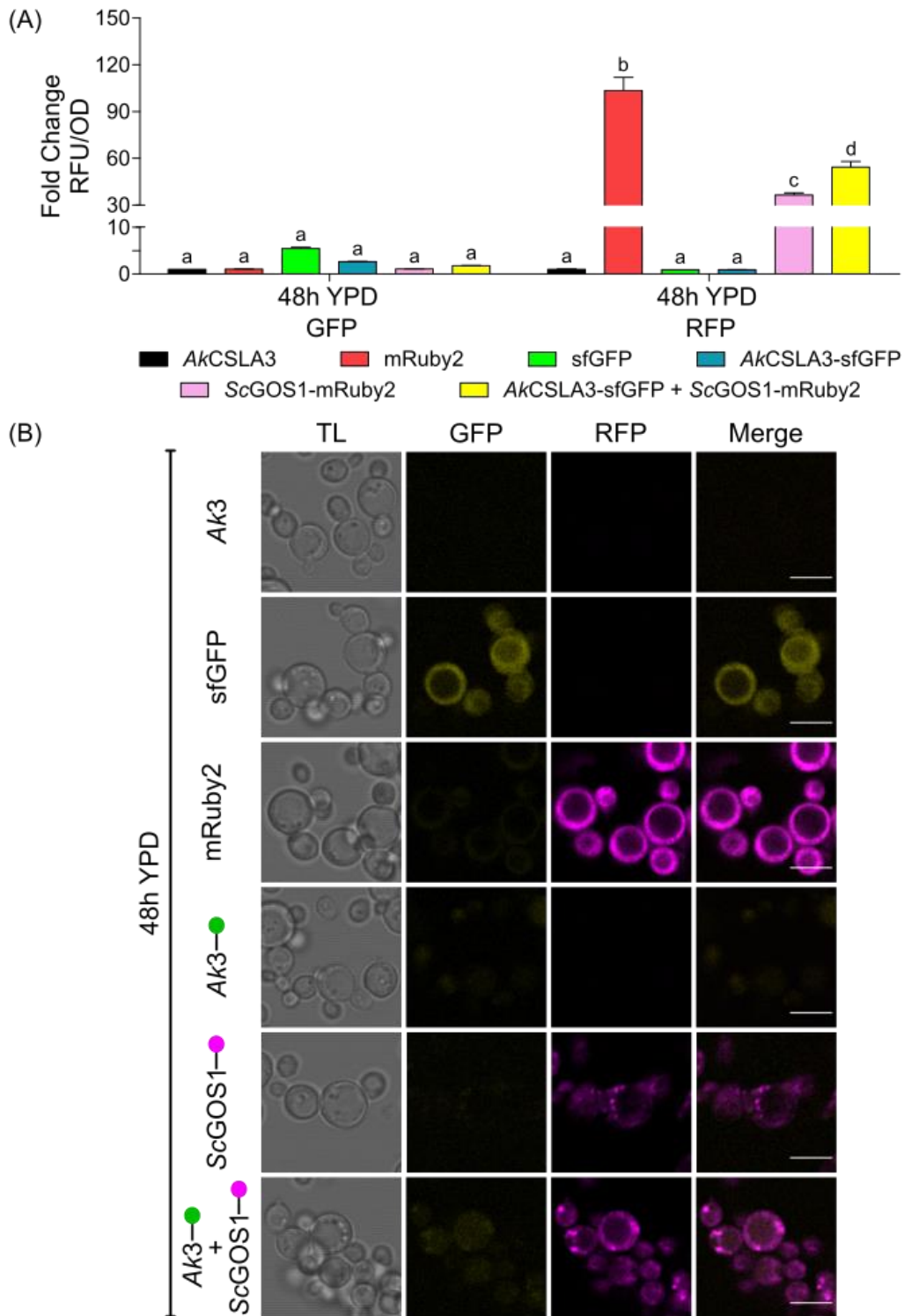


Figure 12: Constitutive expression of ScGOS1-mRuby2 in YPD medium for 48 h. *Pichia* strains were cultivated in 2 ml YPD in a 24-EnzyScreen plate for 48 h. (A) Fold Change of the normalized relative fluorescence units (RFU/OD) for green fluorescent proteins (GFP) and red fluorescent proteins (RFP) relative to AkCSLA3. Data show the mean + SD of three biological replicates and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (B) Localization of GFPs and RFPs expressed in the different *Pichia* strains. TL: transmitted light; Ak3: *Amorphophallus konjac* CSLA3; Sc: *Saccharomyces cerevisiae*; green dots symbolizing sfGFP fusions and pink dots mRuby2 fusions on the C-terminus of the respective protein. Scale bar: 5 μ m.

The change to YPM, have led to the expression of sfGFP and *AkCSLA3*-sfGFP (including the co-expressing *ScGOS1*-mRuby2 strain) in significantly amounts as visible on the fold change in normalized green fluorescence and was only minor altered between 24 h and 48 h of protein expression (Figure 13A). In contrast, the fold change of the normalized red fluorescence of mRuby2, *ScGOS1*-mRuby2 (including the strain co-expressing *AkCSLA3*-sfGFP) was ~3 times less after 24 h of growth in YPM (Figure 13B) compared to 48 h growth in YPD (Figure 13A). This drop was mostly recovered after 48 h growth in YPM (Figure 13A and 13B).

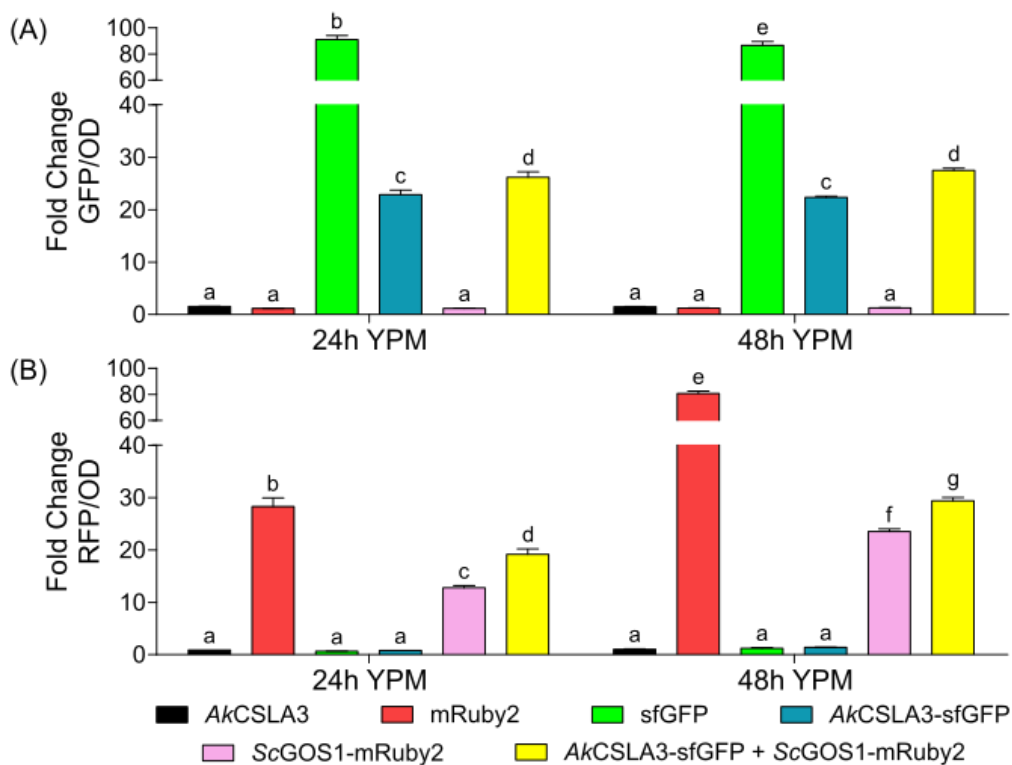


Figure 13: Fluorescence of *AkCSLA3*-sfGFP and *ScGOS1*-mRuby2 co-expressing *Pichia* strains after 48 h of pre-growth in YPD, followed by 48 h of methanol induction. *Pichia* cell were pre-grown in 2 ml YPD for 48 h, followed by 24 h and 48 h growth in YPM in a 24-EnzyScreen plate. (A) Fold change of normalized green fluorescent proteins (GFP/OD) and (B) red fluorescent proteins (RFP/OD) relative to *AkCSLA3*. Data show the mean + SD of three biological replicates and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). *Ak*: *Amorphophallus konjac*; *Sc*: *Saccharomyces cerevisiae*.

While mRuby2 was localized in the cytosol (Figure 14A), *ScGOS1*-mRuby2 expressing *Pichia* strains (including the co-expressing *AkCSLA3*-sfGFP strain) have shown puncta localized fluorescence that was more visible after 48 h of growth in YPM compared to 24 h methanol induction (Figure 14B). Even though the fluorescence of *AkCSLA3*-sfGFP expressing strains was localized in puncta after 24 h and 48 h growth in YPM, the detected puncta were smaller at 48 h compared to 24 h (Figure 14B). This indicates a degradation of the protein with time probably due to the depletion of the methanol concentration in the medium. Nevertheless, the red fluorescence of *ScGOS1*-mRuby2 expressing *Pichia* cells have

shown a lot of background fluorescence after 24 h of YPM induction and was more distinguishable after 48 h of growth in YPM media (Figure 14B). Most importantly, the green fluorescent puncta of *AkCSLA3-sfGFP* were co-localized with the red fluorescent puncta of *ScGOS1-mRuby2* in the co-expressing strain after 48 h of YPM induction. This indicates that *AkCSLA3* is expressed in the Golgi in *Pichia* strains (Figure 14B). Furthermore, it should be pointed out that the pattern of *ScGOS1-mRuby2* was not altered by the co-expression of *AkCSLA3-sfGFP* (Figure 14B).

To sum up, 48 h growth in YPD followed by 48 h growth in YPM is useful to express the constitutive-inducible *ScGOS1-mRuby2* and the methanol-inducible *AkCSLA3-sfGFP* side by side. This co-expressing strain can be used to elucidate how CSLAs are trafficked from the ER to the Golgi and what kind of post-translational modifications are necessary to translocate CSLAs to the Golgi.

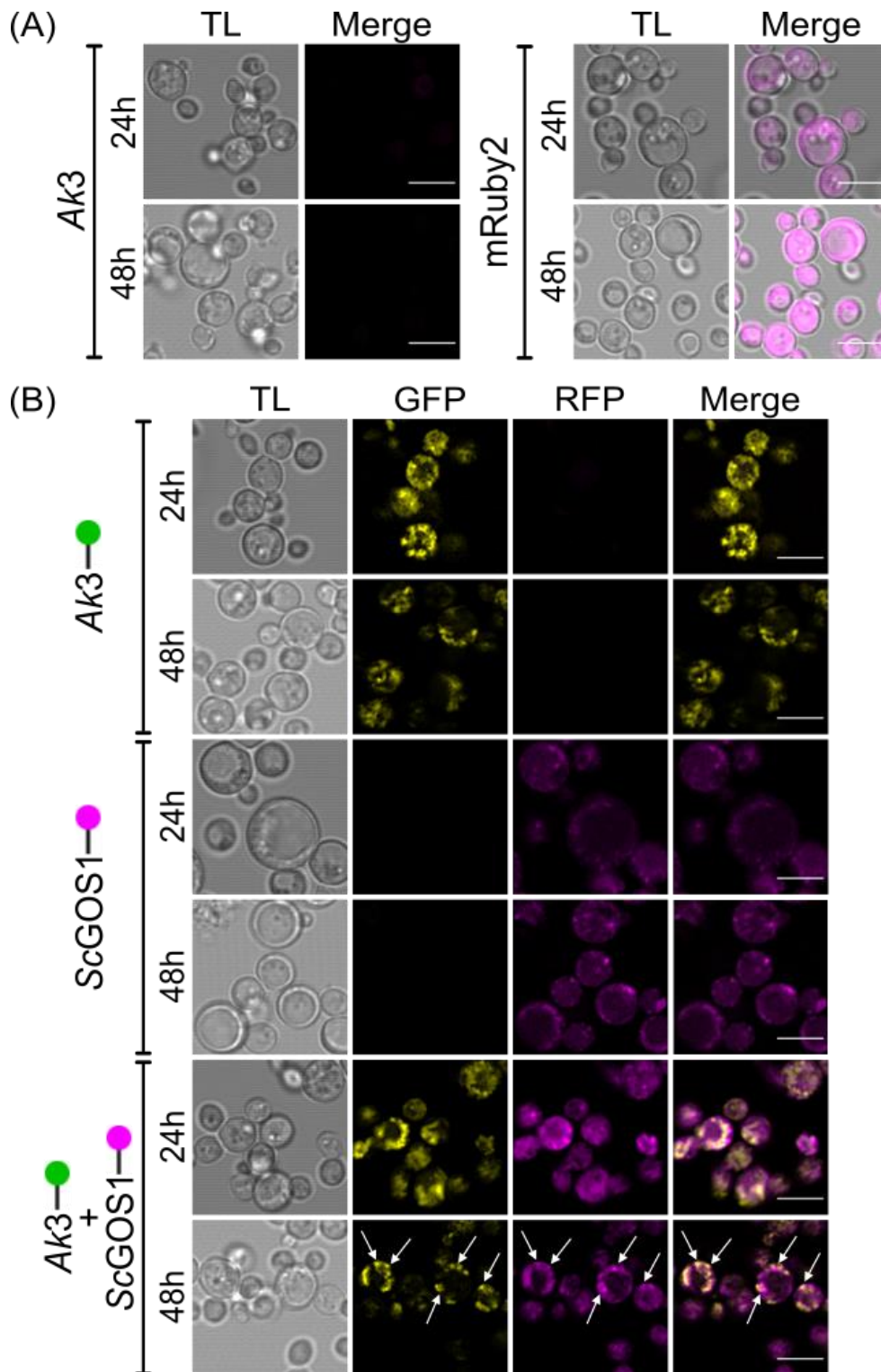


Figure 14: AkCSLA3-sfGFP co-localizes with ScGOS1-mRuby2 expressed in *Pichia pastoris*. *Pichia* cell were pre-grown in 2 ml YPD for 48 h, followed by 24 h and 48 h growth in YPM in a 24-EnzyScreen plate. Localization of (A) fluorescent controls (*AkCSLA3* and *mRuby2*) and (B) *AkCSLA3*-sfGFP and *ScGOS1*-mRuby2 expressing strains using a confocal microscope (LSM880). The intensity of green (GFP) and red fluorescent proteins (RFP) was adjusted for all channels after the acquisition relatively to sfGFP for the green signal and mRuby2 for the red signal. White arrows show spots of co-localization of *AkCSLA3*-sfGFP and *ScGOS1*-mRuby2. Gain 24 h: 700 (GFP), 800 (RFP); Gain 48 h: 700 (GFP), 700 (RFP). Green dots symbolizing sfGFP fusions and red dots mRuby2 fusions on the C-terminus of the specific protein. *Ak3*: *Amorphophallus konjac* CSLA3; *Sc*: *Saccharomyces cerevisiae*; TL: transmitted light; Scale bar: 5 μ m.

3.3 AkCSLA3 mis-localization

Cellulose is synthesized at the plasma membrane by cellulose synthase complexes (CSCs) that are assembled in the Golgi apparatus and transported to the plasma membrane [Kumar and Turner, 2015]. To investigate whether AkCSLA3 can also be incorporated into the plasma membrane and accomplish HM backbone elongation, the aim was to mis-localize AkCSLA3 to the plasma membrane of *Pichia*.

Regulator of G protein signalling 4 (RGS4) is a mammalian protein that enhance guanine triphosphatase (GTPase) activity by binding to the α -subunit of G proteins [Hooks et al., 2008]. Rat RGS4 expressed in *Saccharomyces* have shown to be localized at the plasma membrane, targeted by the N-terminal 33 AA of RGS4. Even GFP was localized to the plasma membrane when the first 33 AA of N-terminal RGS4 was fused on the N-terminus of GFP in *Saccharomyces* [Srinivasa et al. 1998]. In this study, the N-terminal 33 AA of Rat RGS4 were codon optimized for *Pichia* and cloned as N-terminal fusions to pmVenus (RGS4:pmVenus) and AkCSLA3-pmVenus (RGS4: AkCSLA3:pmVenus)

Several *Pichia* colonies expressing RGS4:pmVenus and RGS4: AkCSLA3:pmVenus were screened by fluorescence after growth in 1 ml YPMG for 72 h (Figure 15). All colonies have expressed the proteins because all of them have shown a fluorescent signal. However, their expression was variable for most of the colonies. The fluorescence of RGS4:pmVenus strains was in general higher (except colony 1) than RGS4: AkCSLA3:pmVenus expressing strains (Figure 15).

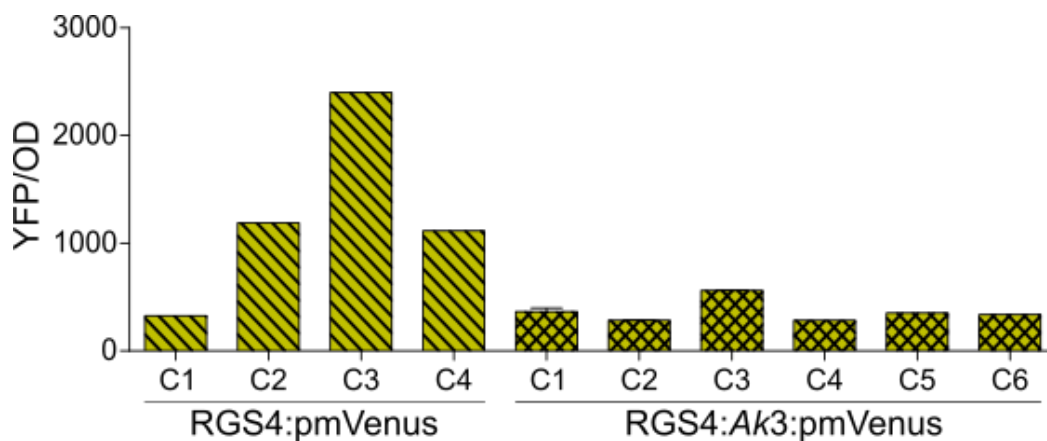


Figure 15: Screening of *Pichia* expressing RGS4:pmVenus and RGS4: AkCSLA3:pmVenus. *Pichia* cultures were grown in 1 ml YPMG for 72 h in a 24-EnzyScreen plate. Normalized yellow fluorescence (YFP/OD) is shown for one biological replicate (RGS4:pmVenus) or two biological replicates (RGS4: AkCSLA3:pmVenus) for each colony. Ak3: *Amorphophallus konjac* CSLA3; C: colony.

Additionally, the localization of the screened colonies was examined using a confocal microscope (LSM700). Representative images are shown in Figure 16. The fluorescence of RGS4:pmVenus was detected in the cytosol, whereas RGS4: AkCSLA3:pmVenus was

localized in bigger aggregates and not as expected translocated to the plasma membrane for all screened *Pichia* colonies (Figure 16).

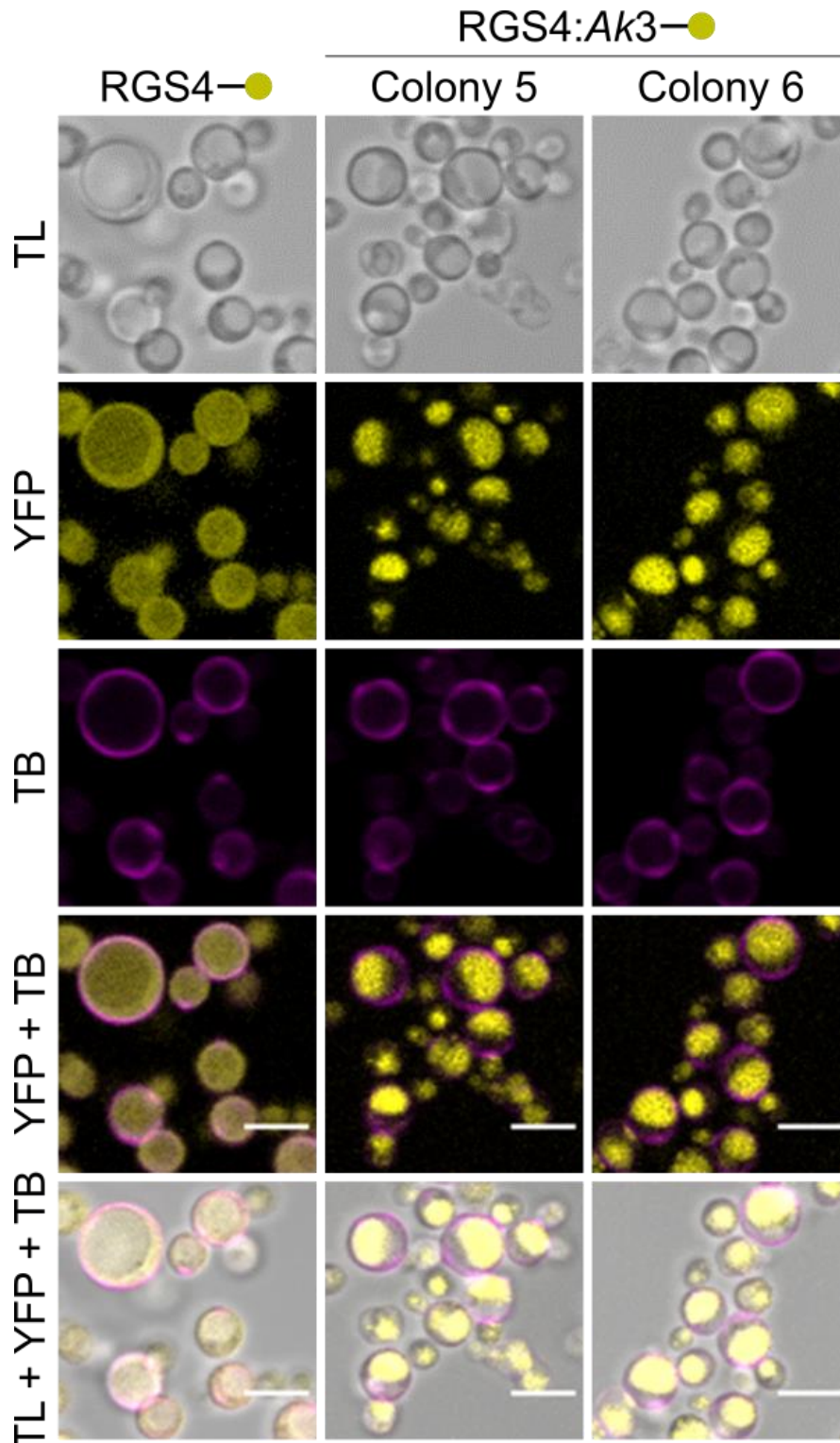


Figure 16: Localization of screened *Pichia* colonies expressing RGS4:pmVenus and RGS4: AkCSLA3:pmVenus. *Pichia* cultures were cultivated for 72 h in 1 ml YPMG in a 24-EnzyScreen plate. The plasma membrane of *Pichia* cells was labelled with trypan blue (TB). After acquisition, the yellow fluorescence (YFP) of RGS4:pmVenus was decreased to see its specific localization. Therefore, RGS4: AkCSLA3:pmVenus fluorescence seem brighter than RGS4:pmVenus. Yellow dots symbolizing pmVenus fusions on the C-terminus of the specific protein. *Ak3*: *Amorphophallus konjac* CSLA3; TL: transmitted light; Scale bar 5 μ m.

One potential explanation for the lack of visibility of proteins at the plasma membrane may be attributed to their gradual overexpression, resulting in their accumulation throughout the cytosol (RGS4:pmVenus) or in larger aggregates (RGS4:AkCSLA3:pmVenus). For this reason, *Pichia* strains were re-grown and the fluorescence and localization was examined at earlier timepoints (0, 3, 6, 9 and 25 h) of protein expression (Figure 17 and 18). Additionally, pmVenus was included as a control in this experiment to determine whether RGS4:pmVenus expression results in a different localization pattern compared to pmVenus itself. Based on the fluorescence measurements, the heterologous proteins were already expressed after 6 h (pmVenus and RGS4:pmVenus colony 3) or 9 h (RGS4:pmVenus colony 2) of methanol induction (Figure 17A). However, even at this earlier time points the localization was visible in the cytosol for both strains (Figure 17B). Therefore, *Pichia* strains expressing RGS4 are not able to translocate pmVenus to the plasma membrane and behave differently as reported for *Saccharomyces*.

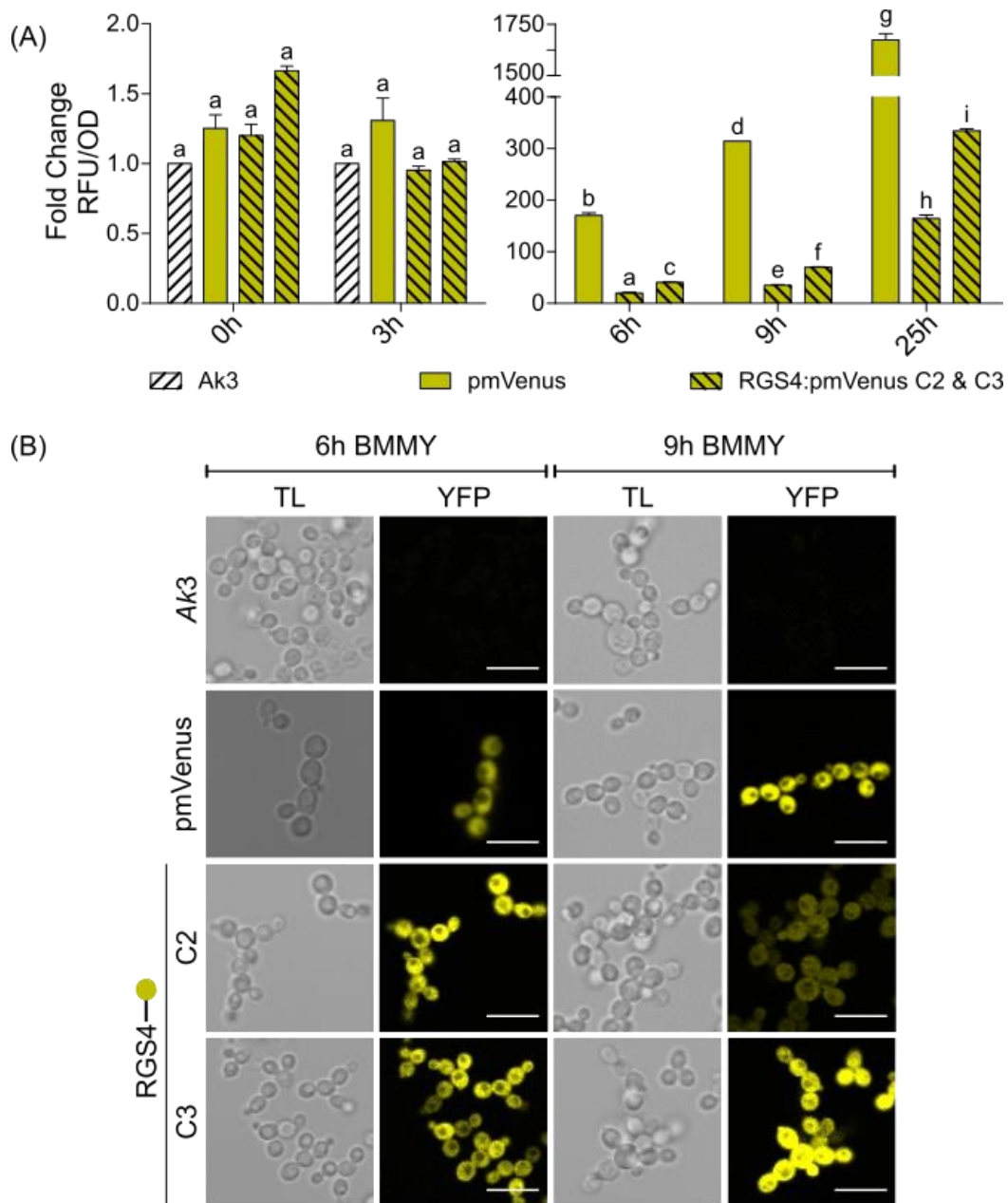


Figure 17: Fluorescence and localization of *Pichia* expressing pmVenus or RGS4:pmVenus over time. *Pichia* cultures were cultivated in 3 ml BMGY for 24 h for biomass production, resolved the cell pellet in 10 ml BMMY and transferred 1.5 ml culture to several wells of a 24-EnzyScreen plate to harvest them after several timepoints (0, 3, 6, 9, 25 h). (A) For each timepoint, the fold change of normalized yellow fluorescence (YFP/OD) was calculated relative to *AkCSLA3*. Data show the mean + SD of two biological replicates for each strain and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (B) Intracellular localization of the yellow fluorescent proteins (YFP). Fluorescence intensity of pmVenus was decreased after the acquisition to visualize the localization. Yellow dots symbolizing pmVenus fusions on the C-terminus of the specific protein. Ak3: *Amorphophallus konjac* CSLA3; C: colony; TL: transmitted light; Scale bar: 10 μ m.

Based on the fluorescence data, RGS4:*AkCSLA3*:pmVenus was also expressed in detectable amounts after 6 h of protein expression in methanol (Figure 18A). However, RGS4:*AkCSLA3*:pmVenus was already accumulated in bigger aggregates and have not been translocated to the plasma membrane (Figure 18B). This means that the deposition of

RGS4:AkCSLA3:pmVenus in bigger aggregates was not caused by overexpression and accumulation of the protein over time.

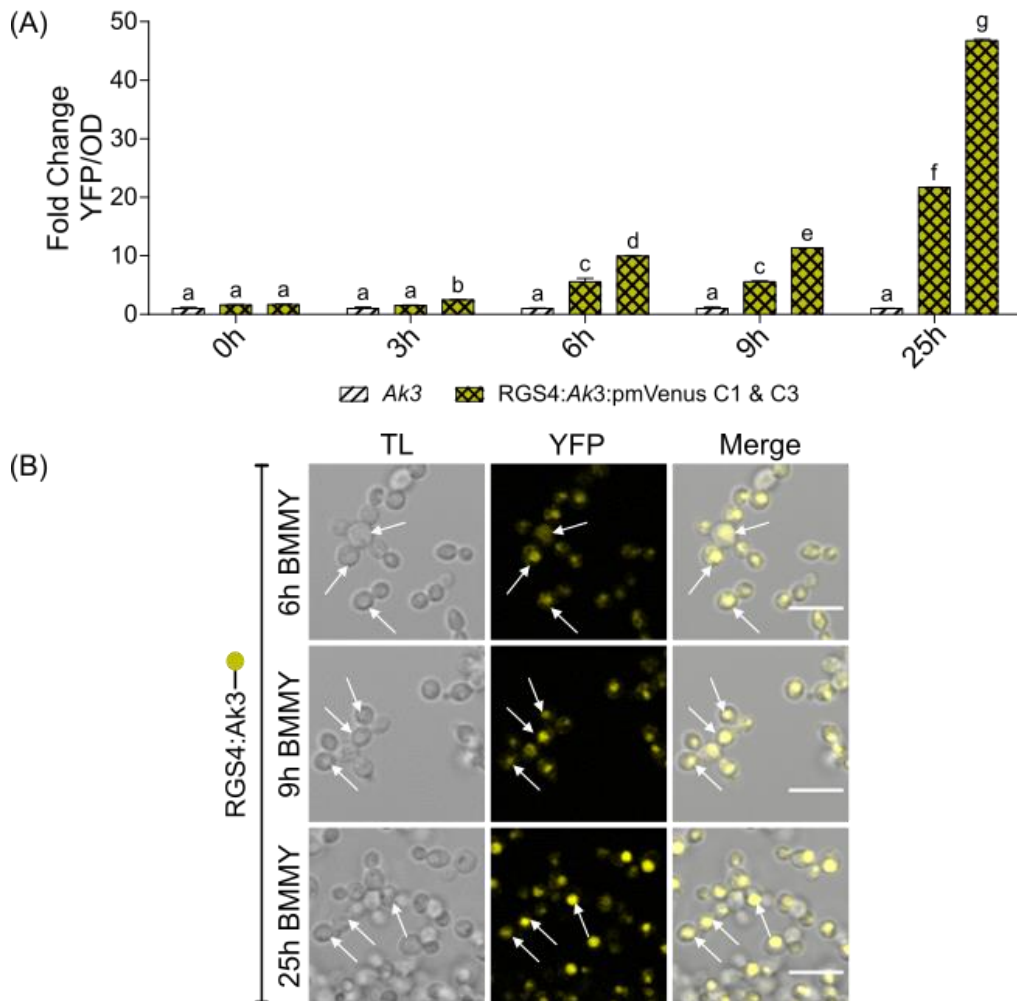


Figure 18: Fluorescence and localization of *Pichia* expressing RGS4:AkCSLA3:pmVenus over time. *Pichia* cultures were cultivated in 3 ml BMGY for 24 h for biomass production, followed by growth in 1.5 ml BMMY for several timepoint (0, 3, 6, 9 and 25 h). (A) For each timepoint, the fold change of normalized yellow fluorescence (YFP/OD) was determined. Data show the mean + SD of two biological replicates and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (B) Intracellular localization of the yellow fluorescent proteins (YFP). Yellow dots symbolizing pmVenus fusions on the C-terminus of RGS4:AkCSLA3. Arrows show the localization of RGS4:Ak3:pmVenus in bigger aggregates. Ak3: *Amorphophallus konjac* CSLA3; C: colony; TL: transmitted light; Scale bar: 10 μ m.

Because the accumulation of RGS4:AkCSLA3:pmVenus in aggregates could not be explained by overexpression of the protein, it could be mis-localized by the pmVenus tag. To test this hypothesis, *Pichia* strains expressing AkCSLA3-pmVenus were cloned and the fluorescence and localization of several colonies was determined 48 h after growth in 1 ml YPMG (Figure 19). Based on the fluorescence, pmVenus and AkCSLA3-pmVenus was expressed after 48 h growth in YPMG (Figure 19A). All colonies of AkCSLA3-pmVenus have shown a similar fluorescence, except colony 6 where a brighter fluorescent signal was detected (Figure 19A). The fluorescent signal of mostly all colonies (except colony 6) expressing

AkCSLA3-pmVenus was localized in bigger aggregates (Figure 19B), concluding that pmVenus is not the right fluorescent tag for fusion on *AkCSLA3*.

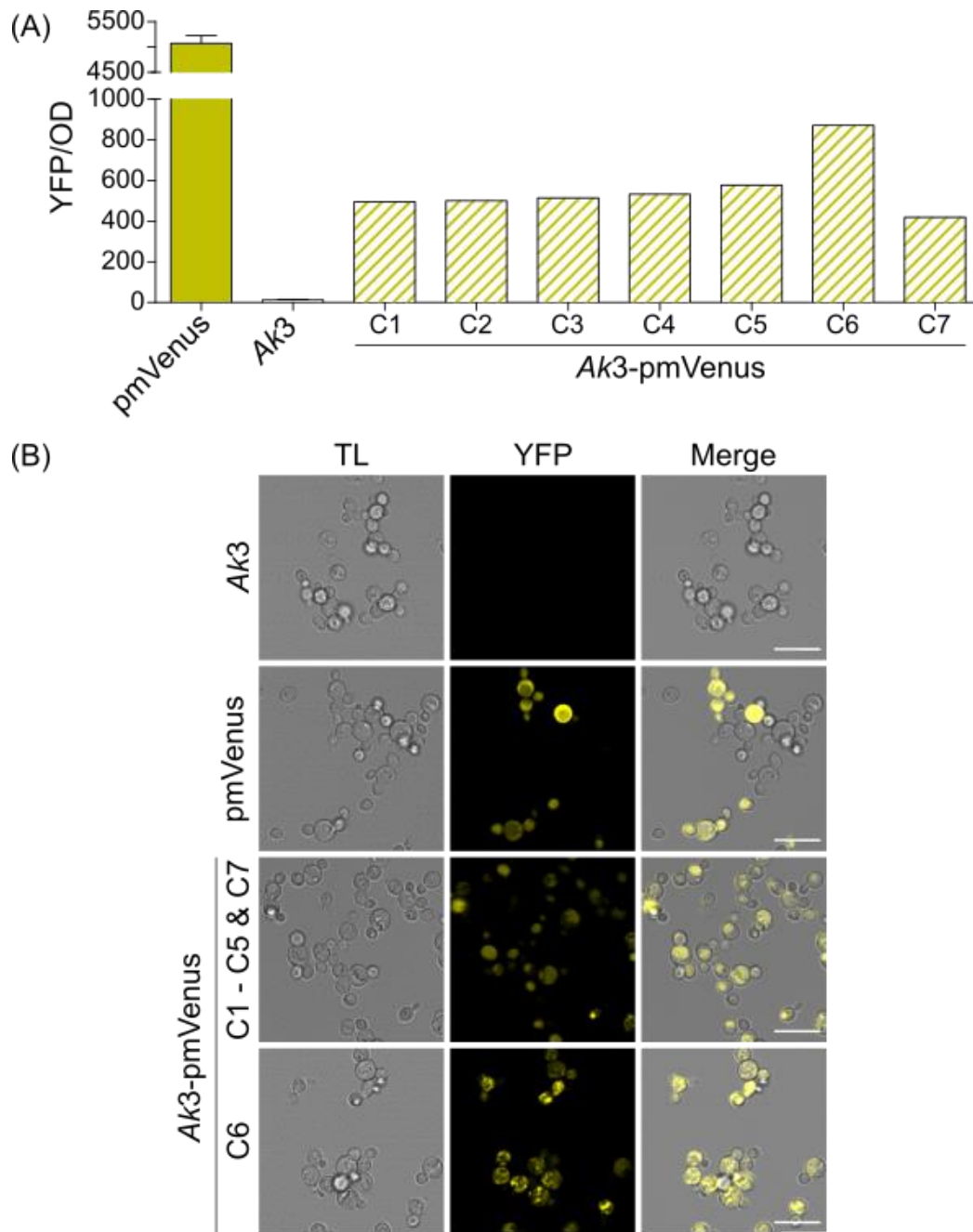


Figure 19: Fluorescence and localization of *Pichia* expressing *AkCSLA3*:pmVenus. *Pichia* cultures were cultivated in 1 ml YPMG for 48 h in a 24-EnzyScreen plate. (A) The normalized yellow fluorescence (YFP/OD) was determined. Data show the mean + SD of two biological replicates (pmVenus and *AkCSLA3*) and one biological replicate of *AkCSLA3*-pmVenus. (B) Intracellular localization of the yellow fluorescent proteins (YFP). C: colony; *Ak3*: *Amorphophallus konjac* CSLA3; C: colony; TL: transmitted light; Scale bar: 10 μm.

To sum up, RGS4 was not able to translocate neither the fluorescent protein pmVenus nor *AkCSLA3*:pmVenus to the plasma membrane of *Pichia* expressing strains as it was shown in *Saccharomyces*. The reasons for the accumulation of *AkCSLA3*:pmVenus in aggregates and their localization is unknown. For further experiments sfGFP should be preferred as a

fluorescent tag for fusion of *AkCSLA3* because it is localized in Golgi compartments (Figure 14B) and is enzymatically active in producing glucomannan in *Pichia* cells (Figure 8B).

Proteins often contain sequences that are responsible for their final localization. In general, three regions of GTs could be involved in the retention of the protein in the Golgi: (i) the N-terminal cytosolic domain, (ii) the first transmembrane-spanning domain (TMD1) and/or (iii) the luminal stem region [Pagny et al., 2003; Tu and Banfield, 2010; Cosson et al., 2013]. Therefore, another option to mis-localize *AkCSLA3* was to remove the N-terminal sequence including the first TMD. Using TOPCONS, the first TMD is predicted between AA21 and AA40. *AkCSLA3*-sfGFP was truncated by removing AA1M to AA44L and was transformed into *Pichia* (further named as *AkCSLA3*-sfGFP w/o TMD1). *Pichia* were cultivated in 2 ml YPMG for 48 h, followed by measuring the fluorescence, determining the localization and the HM content that was produced by *AkCSLA3*-sfGFP and its truncated construct (Figure 20).

Based on the fluorescence, sfGFP and *AkCSLA3*-sfGFP with and w/o TMD1 were expressed in *Pichia* after 48 h growth in YPMG (Figure 20A). *Pichia* expressing sfGFP have shown the highest fold change in fluorescence followed by *AkCSLA3*-sfGFP. Compared to *AkCSLA3*-sfGFP, all colonies of *AkCSLA3*-sfGFP (w/o TMD1) have shown tremendously lower fluorescence (Figure 20A), but the signal was still higher than the non-fluorescent negative control *AtCSLA2* (data not shown). The same trend as observed for the fluorescence could be reflected in the microscope images (Figure 20C). While the fluorescence of *AkCSLA3*-sfGFP was distributed as many small fluorescent dots with a high background fluorescence overall in the entire *Pichia* cell, *AkCSLA3*-sfGFP (w/o TMD1) was localized as one large (colony 2 and 5) or few small punctae nearly located to the plasma membrane of *Pichia* (Figure 20C). Extraction of the cell wall material of the engineered strains have shown that the loss of the TMD1 of *AkCSLA3*-sfGFP resulted in the loss of HM activity (Figure 20B). Swap experiments have shown that replacing the N-terminal sequence and first TMD of *AkCSLA3* with that from *AtCSLA2* (mannan synthase) also led to reduced glucomannan synthase activity compared to its parental *AkCSLA3* strain (Robert et al., 2021), implying that *AkCSLA3*-sfGFP cannot easily tolerate changes in this region and that it is necessary for HM function.

In conclusion, mis-localization of *AkCSLA3* to the plasma membrane expressing RGS4 was unsuccessful in *Pichia*. The loss of HM activity of *Pichia* expressing *AkCSLA3*-sfGFP (w/o TMD1) could be explained by two reasons. The first option is that the sequence up to TMD1 is essential for HM activity of *AkCSLA3*, whereas the second option is that *AkCSLA3*-sfGFP (w/o TMD1) is misfolded. In both cases, the fluorescent pattern of *AkCSLA3*-sfGFP (w/o TMD1) could still be localized in the Golgi. The lower number of fluorescent puncta can be associated with a lower protein expression, while the larger fluorescent dots can indicate misfolding.

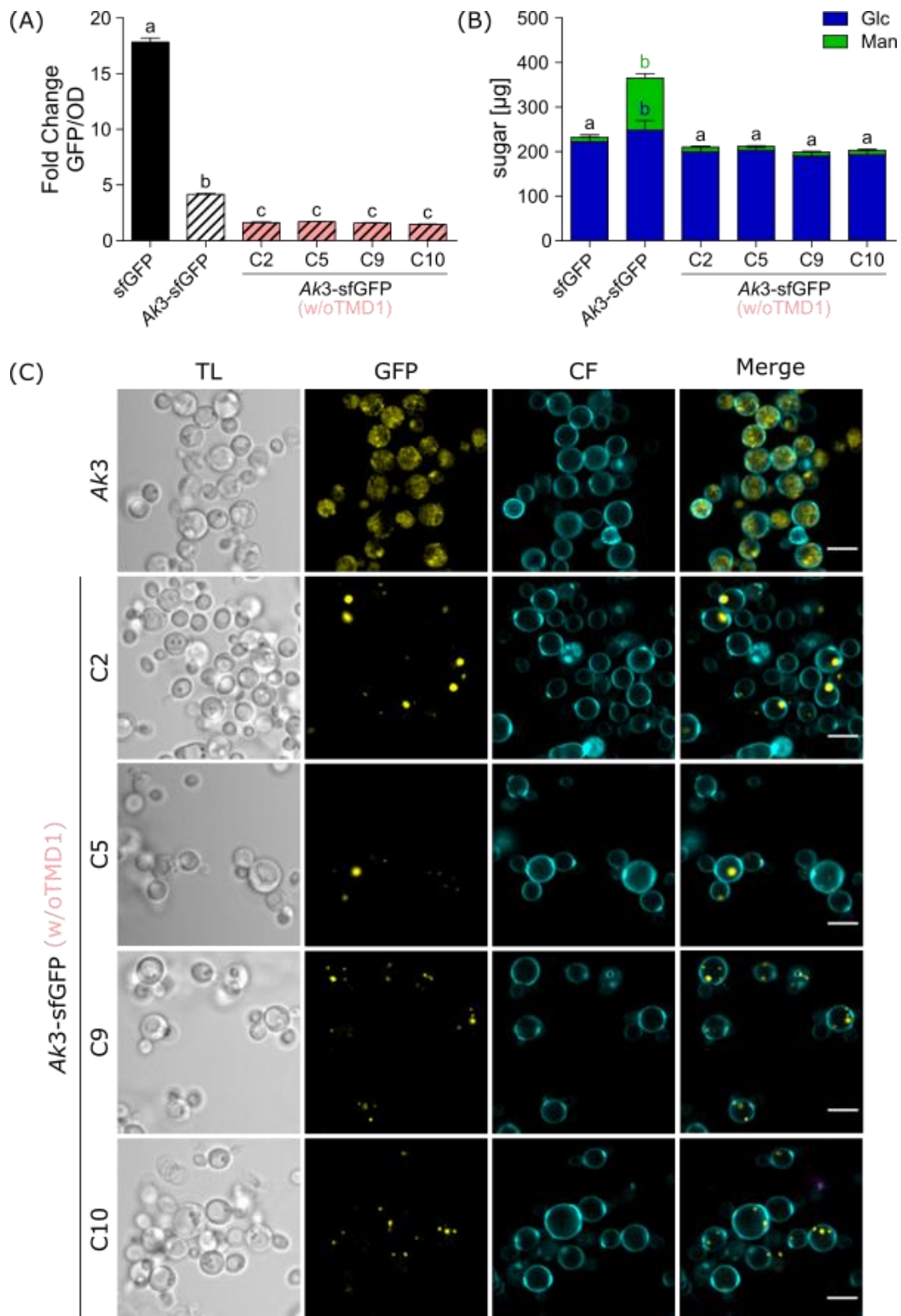


Figure 20: Impact of AkCSLA3-sfGFP without the N-terminal sequence and TMD1 on fluorescence, localization and HM production in *Pichia*. *Pichia* cultures were cultivated in 2 ml YPMG for 48 h (A) Fold change of the normalized green fluorescence (GFP/OD) and (B) mannose content of AKI wall polymers. Polymers were hydrolysed with TFA. Data show the mean + SD of three biological replicates and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (C) Intracellular localization of the green fluorescent proteins. Plasma membrane of *Pichia* was stained by calcofluor white (CF). Ak3: *Amorphophallus konjac* CSLA3; w/oTMD1: without transmembrane domain 1; C: colony; TL: transmitted light; Scale bar: 5 µm.

3.4 Altered sugar metabolism in *Pichia* expressing AkCSLA3

The expression of recombinant proteins in *Pichia* depends on many factors, including the carbon source and metabolism [Heyland et al., 2011; Prielhofer et al., 2015]. In this study, two distinct approaches were undertaken to investigate the possibility of modifying the quantity or Glc:Man ratio of glucomannans produced by AkCSLA3.

In the first approach, *Pichia* expressing AkCSLA3-Venus were grown in YPM medium supplemented with different carbon sources (glycerol, Glc and Man) for 48 h (Figure 21). Growth in YPM led to the lowest Man production and Glc was not incorporated in a significant amount. The supplementation with carbon sources not only resulted in an increase of the Glc content in AkCSLA3-Venus expressing *Pichia* strains, but also in the negative control Venus (Figure 21A), meaning *Pichia*-specific glucans were formed at the same time and led to a higher background in Glc. In YPM supplemented with glycerol, AkCSLA3-Venus strains have produced the highest concentration of Glc and Man. The addition of Glc and Man also led to significant increase in glucomannan production, but the concentration of glucose or mannose was not altered for both of them (Figure 21A).

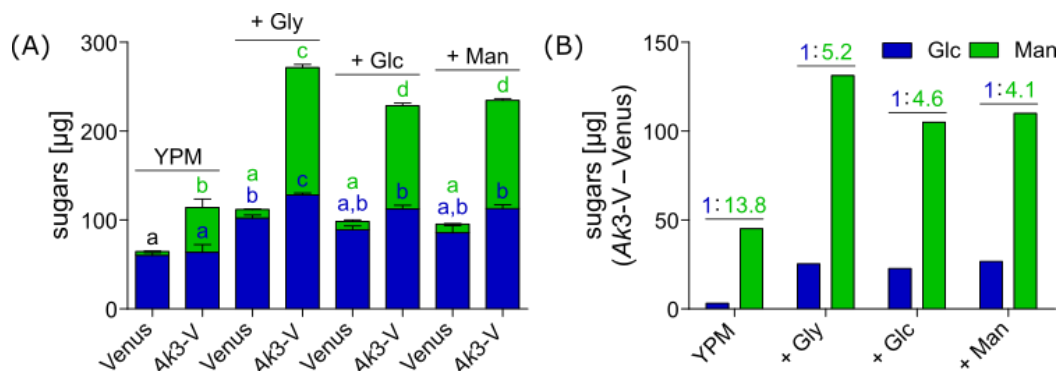


Figure 21: Impact of carbon supplementation on glucomannan production of AkCSLA3-Venus. Yeast strains were cultivated in 2 ml YPM (1.5% MeOH) supplemented with different carbon sources: glycerol (Gly), glucose (Glc) or mannose (Man) for 48 h in a 24-EnzyScreen plates. (A) Monosaccharide content of AKI wall polymers. Polymers were hydrolysed with TFA. Data show the mean + SD of three biological replicates, and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (B) Monosaccharides produced by AkCSLA3-Venus were calculated by subtracting the sugars produced by Venus controls and their glucose to mannose ratio. Ak3-V: *Amorphophallus konjac* CSLA3-Venus.

To assess potential alterations in the composition of glucomannan due to various supplementation, the Glc:Man ratio was calculated (Figure 21B). Although AkCSLA3-Venus grown in YPM produced the lowest amount of glucomannans, the highest amount of Man was incorporated in its backbone. In comparison, all AkCSLA3-Venus strains that were supplemented with carbon sources incorporated a higher proportion of Glc into their produced sugar chain, resulting in a lower Glc:Man ratio (Figure 21B). But to determine the Glc:Man ratio more accurate, it would be better to detect Glc and Man after digestion with β -mannanase to get rid of the high background of yeast α -mannoproteins or detect 4-Glc and 4-Man content of

the AKI by linkage analysis that are easily distinguishable from yeast α -mannoproteins and β -1,3- glucans. In general, the supplementation of Glc and Man did not lead to more Glc or Man incorporation in the produced polymer. Less glucomannan content of *AkCSLA3*-Venus expressing *Pichia* that were grown in YPM could probably be explained by rapid *AkCSLA3* production to a toxic abundance that limit further growth and protein expression [Cerregghino and Cregg, 2000]. Regardless of the non-altered sugar incorporation, the growth of YPM supplemented with glycerol should be used for further experiments due to their higher content in glucomannan and fewer steps of intervention in ongoing processes that should minimize sources of error.

In the second approach, the glycoengineered *Pichia* strain SuperMan5 (his+) of GlycoSwitch[®] was selected for *AkCSLA3*-sfGFP expression. The SuperMan5 (his+) strain exhibit a mutation in an endogenous glycosyltransferase gene (*OCH1*). *OCH1*, which typically encodes for an α -1,6-mannosyltransferase, facilitates the transfer of mannose residues onto branches of N-linked glycans ($\text{Man}_8\text{GlcNAc}_2$) using α -1,6-linkages in the Golgi [Nakayama et al., 1992; Nagasu et al., 1992]. This process subsequently instigates hypermannosylation [Jacobs et al., 2009; Krainer et al., 2013]. The disruption and stepwise introduction of glycosidases and glycosyltransferases in the SuperMan5 *Pichia* strains results in a shorter fragment of *OCH1* that reduce its ability to mannosylate yeast $\text{Man}_8\text{GlcNAc}_2$ and enables them to trim $\text{Man}_8\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$ [Jacobs et al., 2009]. While $\text{Man}_8\text{GlcNAc}_2$ can be mannosylated up to a dozen times, $\text{Man}_5\text{GlcNAc}_2$ cannot be modified by yeast glycosyltransferases [Jacobs et al., 2009]. Therefore, the GDP-Man availability in SuperMan5 strains should be higher than in wildtype *Pichia* strains (X-33) and could potentially incorporated into HMs synthesized by *AkCSLA3*.

Pichia expressing *AkCSLA3*-sfGFP were transformed in SuperMan5 strains and were grown together with controls in the X-33 strains. After 48 h of growth in YPMG, sfGFP and *AkCSLA3*-sfGFP were expressed based on the fold change in fluorescence. The fluorescence was not significantly different for *AkCSLA3*-sfGFP in both backgrounds (Figure 22A). Also the localization of *AkCSLA3*-sfGFP in X-33 and SuperMan5 looked similar in both backgrounds (Figure 22B). *Pichia* strains were stained with calcofluor white (CF) and propidium iodide (PI). CF label β -glucans on the yeast plasma membrane, whereas PI binds to guanine and cytosine and, therefore, is able to label DNA and RNA in *Pichia* cells but only when their membranes are leaky [Suzuki et al., 1997; Hohenblum et al., 2002]. Noticeable was the uptake of PI not only in *AkCSLA3*-sfGFP expressing *Pichia* cells, but also for the SuperMan5 strain on its own, indicating a loss of viability as shown for *AkCSLA3* expressing *Pichia* strains [Robert et al., 2021].

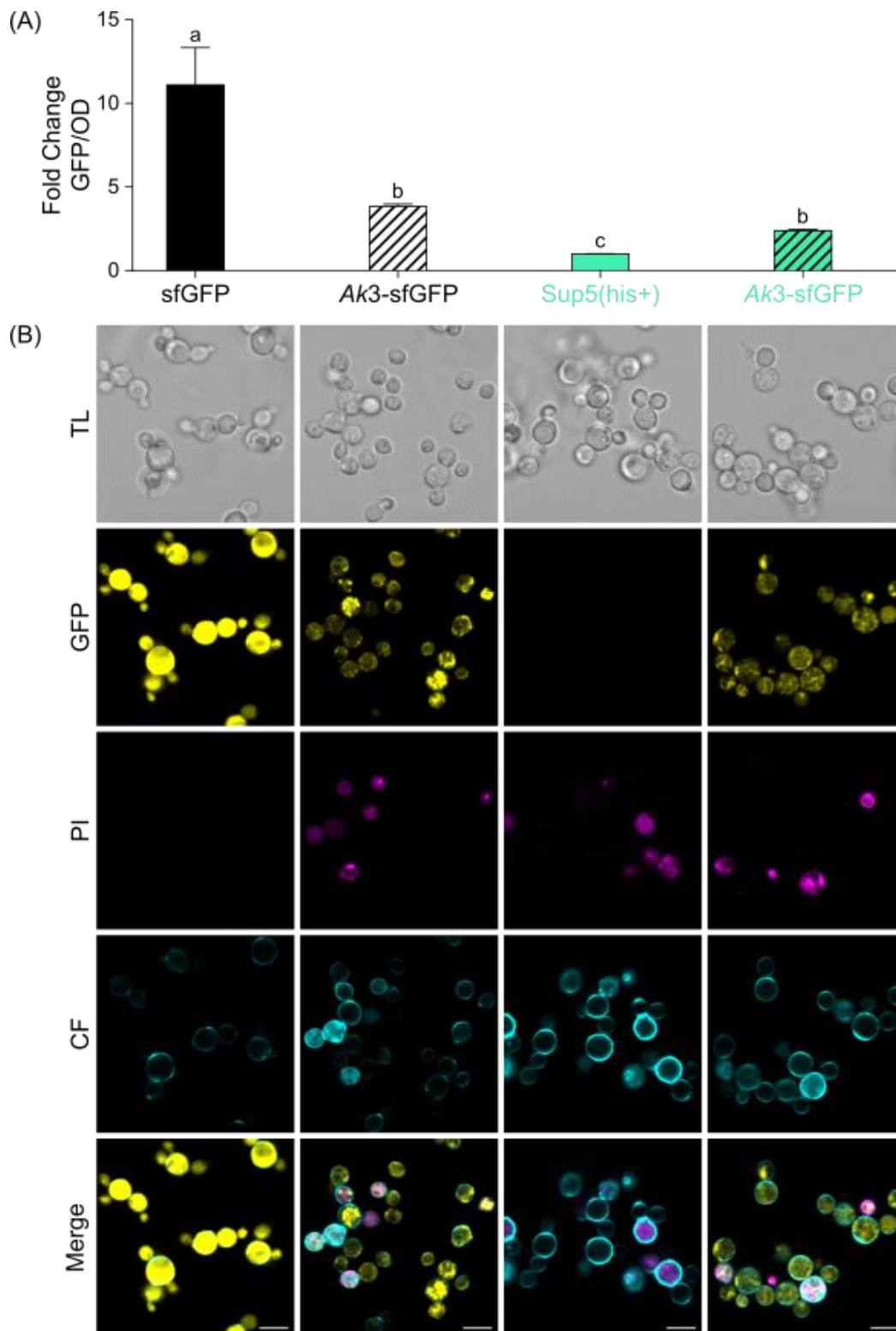


Figure 22: Fluorescence and localization of *Pichia* expressing AkCSLA3-sfGFP in X-33 and SuperMan5 (his+) strains. Yeast strains were cultivated for 48 h in 2 ml YPM (1.5% MeOH) and 0.5% glycerol. (A) Fold change of normalized green fluorescence (GFP/OD) relative to SuperMan5 control. Data show the mean + SD of six biological replicates, and different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (B) Localization and staining with calcofluor white (CF) and propidium iodide (PI) of *Pichia* strains using a confocal microscope (LSM900). Ak3: *Amorphophallus konjac* CSLA3; Sup5(his+): SuperMan5 (his+) *Pichia* strain; TL: transmitted light; GFP: green fluorescent protein; Scale bar: 5 μ m.

First, the alcohol-insoluble residue (AIR) that contains all native yeast α -mannoproteins and β -1,3-glucans was analysed to determine if the SuperMan5 strain incorporated less Man in its cell wall. Indeed, the SuperMan5 expressing *Pichia* control have shown less Man in the alcohol-insoluble residue (AIR) compared to the sfGFP control (Figure 23A). *AkCSLA3*-sfGFP expressing *Pichia* strains were able to increase the Glc as well the Man content in the AIR in both backgrounds (Figure 23A). The Glc:Man ratio of *AkCSLA3*-sfGFP expressed in the SuperMan5 strain was smaller compared to *AkCSLA3*-sfGFP expressed in the X-33 strain (Figure 23B), indicating that more Man was incorporated in the cell wall in the SuperMan5 background.

In order to verify whether the higher abundance of Man in SuperMan5 strains was incorporated into the produced glucomannan chains when *AkCSLA3*-sfGFP was expressed, the alkaline-insoluble residue (AKI) was extracted and digested with endo- β -mannanase. Endo- β -mannanase cleaves specifically β -1,4-Man and β -1,4-Glc bonds of the produced glucomannan backbone. The Glc and Man content was increased in the *Pichia* AKI of both *AkCSLA3*-sfGFP expressing *Pichia* backgrounds compared to their negative controls, whereas less glucomannan was produced in the SuperMan5 strain expressing *AkCSLA3*-sfGFP (Figure 23C). Although glucomannan was produced in both strains, the Glc:Man ratio was the same (Figure 23D), showing that the higher content of Man in the AIR was not due to increased incorporation of Man into the glucomannan chain.

In summary, neither the substrate availability of different carbon sources nor the possible increase in GDP-Man availability was able to change the glucomannan composition in *Pichia* strains expressing *AkCSLA3*-sfGFP.

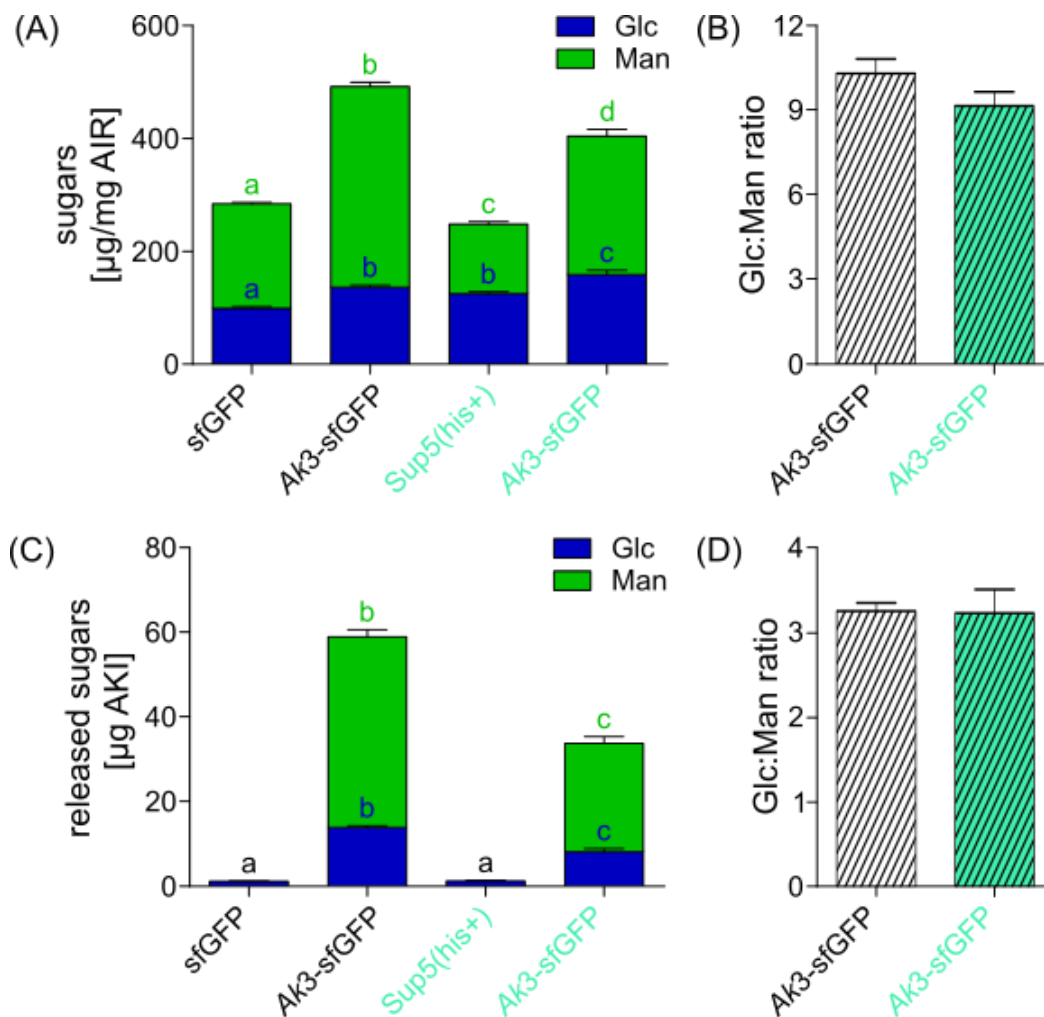


Figure 23: Monosaccharide composition of *Pichia* expressing *AkCSLA3*-sfGFP in X-33 and SuperMan5 (his+) strains. Yeast strains were cultivated for 48 h in 2 ml YPM (1.5% MeOH) and 0.5% glycerol. (A) Monosaccharide composition of the *Pichia* alcohol-insoluble residue (AIR) and (B) the glucose:mannose (Glc:Man) ratio of *AkCSLA3*-sfGFP expressing strains in both backgrounds. (C) Released monosaccharides of the alkaline-insoluble residue (AKI) and (D) the Glc:Man ratio of *AkCSLA3*-sfGFP expressing strains in both backgrounds. AIR and AKI polymers were hydrolysed with TFA. Data show the mean + SD of six (A and B) or four (C and D) biological replicates, and different letters denote significant changes (one-way ANOVA with Tukey test, $p < 0.05$). *Ak3*: *Amorphophallus konjac* CSLA3; *Sup5(his+)*: SuperMan5 his+ *Pichia* strain.

3.5 Protease Influence on fluorescently tagged *AkCSLA3*

Glycosyltransferases are integral membrane proteins that span the Golgi multiple times. If the number of membrane-spanning domains is even, the N- and C- terminal end of the protein face the same site, while an odd number result in opposite sites of the termini. *AkCSLA3* is predicted to have 5 TMDs using TOPCONS, meaning that the C- and N- terminus facing the opposite sites. It is important to know how *AkCSLA3* is integrated in the membrane, because the localization of the catalytic active domain determine which sugar pool is used for the glucomannan biosynthesis. If the catalytic domain faces the Golgi lumen, it will bind nucleotide sugars inside of the Golgi, while nucleotide sugars from the cytosol will be incorporated when the active site faces the cytosol [Davis et al., 2010]. To determine the

topology of *AkCSLA3*, a fluorescence protease protection (FPP) assay was tried to be established for living *Pichia* cultures and lysate or microsome fractions. For this reason, the membrane of *Pichia* has first to be permeabilized with a detergent and second treated with a protease that can cleave or degrade the fluorescent protein if it is facing the cytosol.

Based on the FPP assay proceeded from White et al. (2015), digitonin was tested in several concentrations (15 μ M, 30 μ M, 60 μ M, 1.5 mM) to permeabilize the plasma membrane of living *Pichia* expressing sfGFP. The incubation with digitonin should reduce the fluorescence of *Pichia* expressing sfGFP over time because soluble sfGFP can exit the cell through the pores that arise during the permeabilization. Unfortunately, *Pichia* membranes seem to be resistant to digitonin because even after 60 min of incubation at RT no loss in fluorescence was visible with the microscope for all used concentrations (data not shown).

Consequently, lysate and microsome fractions of *Pichia* expressing sfGFP and *AkCSLA3* with N- and C-terminal fused sfGFP were extracted and treated with and without detergent and proteinase K (Figure 24). DDM was used as a detergent, because a previous master student in our lab established the solubilization of CSLA membrane proteins from microsome fractions with DDM [Stefanie Clauß, master thesis, 2021]. After 60 min incubation with (i) DDM, (ii) proteinase K and (iii) both chemicals at 30°C and 1000 rpm in a thermocycler, the fluorescence in tubes was not reduced for all *Pichia* expressing strains (Figure 24A). In order to examine if sfGFP that faces the cytosol was degraded after the treatment with proteinase K, 20 μ g protein of each fraction was loaded onto a SDS gel. While the treatment of *Pichia* with the buffer control (N) should have no effect on the size of each fluorescent protein (sfGFP: 26.8 kDa; sfGFP-*AkCSLA3*: 86.8 kDa; *AkCSLA3*-sfGFP: 87.1 kDa), degradation of sfGFP should be clearly visible as a fluorescent band with a smaller size. sfGFP and sfGFP-*AkCSLA3* treated with the buffer have shown a fluorescent band at the expected size, whereas the fluorescent band of *AkCSLA3*-sfGFP was visible at a smaller size, indicating degradation of the protein during the incubation on its own (Figure 24B). Therefore, 60 min of incubation was probably too long for this assay and should be reduced. Moreover, all treatments with proteinase K have obviously led to a detection of two bands visible approximately at 35 kDa and 40 kDa. While the fluorescence of the 40 kDa band was brighter for sfGFP lysates and sfGFP-*AkCSLA3* microsomes, the fluorescent signal of the 35 kDa band was increased in *AkCSLA3*-sfGFP microsomes when they were treated with proteinase K (Figure 24B).

The FPP assay was repeated for sfGFP lysates with 30 min of incubation. Additionally, proteinase K was loaded onto the SDS gel to determine if proteinase K have a fluorescent signal on its own that should be visible at 29.8 kDa. Another possibility is that proteinase K is able to bind somehow to monomeric sfGFP. This band would be detected at a higher size and could explain the brighter intensity of the 40 kDa band due to overlapping with the band of

sfGFP dimers. To test this hypothesis, sfGFP lysate was mixed with proteinase K, but not incubated with a thermocycler. Proteinase K alone was not fluorescent, but when sfGFP and proteinase K were mixed in one tube two bands around 35 kDa and 40 kDa with an intensified fluorescent signal at 40 kDa was detected for all treatments with proteinase K independent of the incubation time in the thermocycler (Figure 24C). The 5 min denaturation step where the proteins were incubated at 50°C to break down their three-dimensional structure for SDS gel separation, may be sufficient to enhance the activity of proteinase K even in non-incubated lysate fractions. Therefore, proteinase K cannot be used to determine which termini of AkCSLA3 is facing the cytosol.

Because proteinase K cannot be used in *Pichia* to visualize degradation of cytosolic sfGFP, trypsin was tested based on the FPP assay published by Lorenz et al. (2005). The lysate fraction of soluble sfGFP was incubated with trypsin in several concentrations (2 mM, 10 mM and 50 mM) at 30°C and 1000 rpm for 30 min. The in-gel fluorescence of the separated proteins is illustrated in Figure 24D. The treatment with trypsin has led to a brighter fluorescent signal of the two bands visible at 40 kDa (for 2 mM trypsin) and 35 kDa (for 10 mM and 50 mM trypsin). This could also be a result of dimerization of sfGFP or binding of trypsin to sfGFP during the treatment with proteases.

It was not possible to detect the degradation of sfGFP in the gel, neither using proteinase K nor trypsin. However, the degradation of the cytosol facing sfGFP from AkCSLA3 microsomes should be visible with a confocal microscope and result in a loss of the fluorescence. Therefore, both AkCSLA3-tagged variants were treated with (i) buffer or (ii) 1% DDM and 50 mM trypsin for 30 min at 30°C and 1000 rpm (Figure 24E and 24F). But sfGFP-AkCSLA3 (Figure 24E) and AkCSLA3-sfGFP (Figure 24F) were still fluorescent after the treatment with DDM and trypsin. Furthermore, the fluorescence pattern was similar to the buffer treated microsomes.

To summarize, it was not possible to determine the topology of AkCSLA3 by using the FPP assay, but sfGFP seems to stabilize AkCSLA3. Even after freezing and thawing the microsomes several times the fluorescence was still visible in the microsomes. *In vitro* assays are difficult to perform with glycosyltransferases, because normally they are highly unstable [Dhugga et al., 2004; Brown et al., 2012]. Therefore, these strains can be used for *in vitro* characterization of AkCSLA3, which appears to be protease resistant.

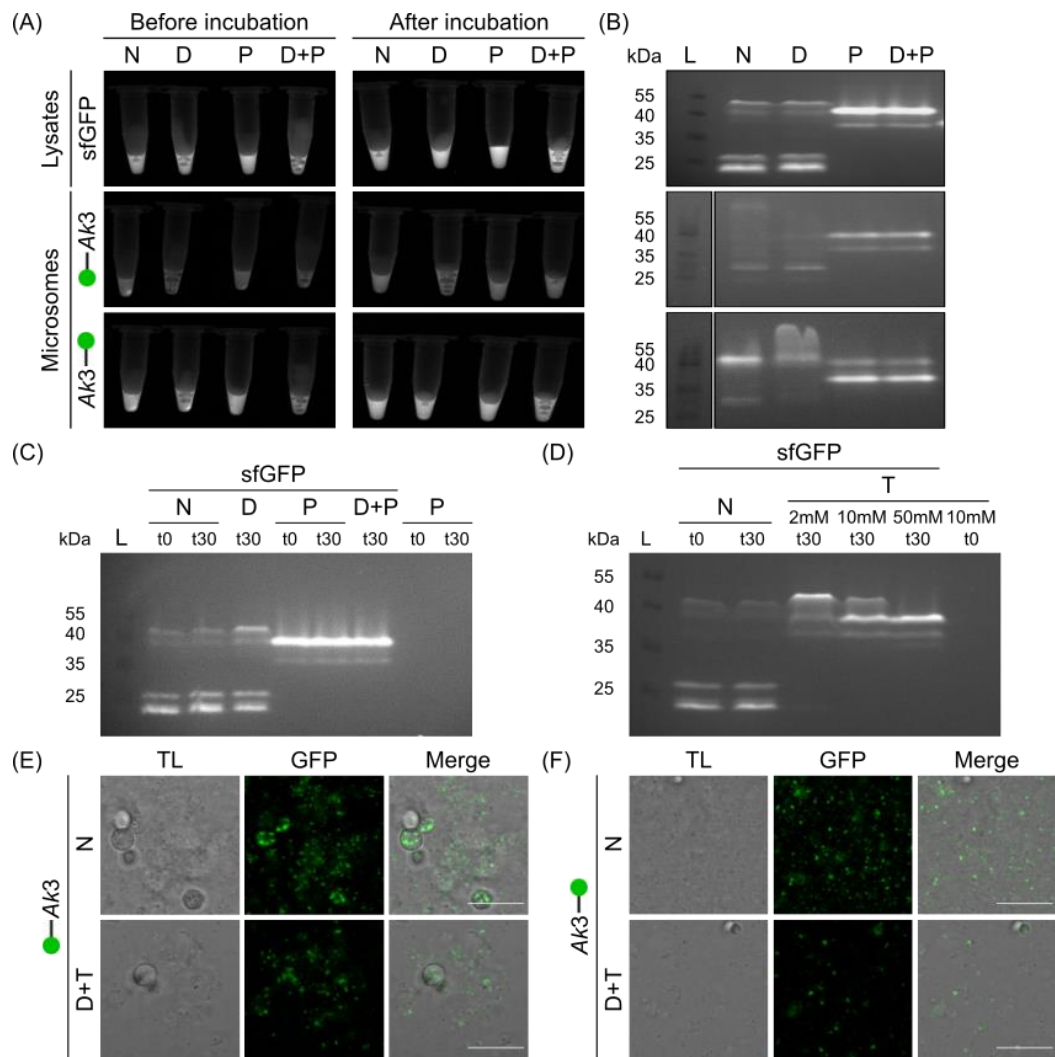


Figure 24: Fluorescent protease protection assay of sfGFP and AkCSLA3 fused with sfGFP on its N- and C-terminus with proteinase K and trypsin. (A) 300 μ g protein from sfGFP lysate and AkCSLA3 tagged microsomal fractions were incubated for 60 min at 1000 rpm and 30°C in a thermocycler with TNG buffer (negative control (N)), detergent (1% DDM; D), proteinase K (P; 3U) or both (D+T). Fluorescent images in front and after the treatment with each chemical and SDS gel with in-gel fluorescence of 20 μ g proteins from each fraction. (B) Proteinase K or (C) trypsin (T) treatment of 300 μ g protein from sfGFP lysates were not incubated (t0) or incubated for 30 min (t30) at 30°C and 1000 rpm in a thermocycler. sfGFP: 26.8 kDa; sfGFP-AkCSLA3: 86.8 kDa; AkCSLA3-sfGFP: 87.1 kDa; Proteinase K: 28.9 kDa. (D) Microscope images of AkCSLA3 tagged microsomes in TNG buffer (N) or treated with detergent and trypsin (D+T). Green dots symbolizing sfGFP fusions on the C- or N-terminus of AkCSLA3. Ak3: *Amorphophallus konjac* CSLA3; GFP: green fluorescent protein; Scale bar: 10 μ m.

3.6 Detailed analysis of fluorescently tagged *Arabidopsis thaliana* CSLAs

Arabidopsis thaliana have nine different isoforms of CSLAs (*At*CSLA1–3, 7, 9–11, 14 and 15) [Richmond and Somerville, 2000]. While a function in (gluco)mannan synthesis was examined for *At*CSLA2, 3, 7 and 9 [Liepman et al., 2005; Goubet et al., 2009, Liepman and Cavalier, 2012], the role of the other CSLA members from *Arabidopsis thaliana* is still unknown. The aim of this study was to characterize their function in HM synthesis and to determine their

structural diversity. In order to observe their gene expression and analyse their localization, AtCSLAs were fused with Venus on their C-terminal end.

AtCSLA-Venus transgenes were transformed into *Pichia* and were screened by growth in 2 ml BMGY for 24 h, followed by 24 h growth in BMMY. Based on the fold change of normalized fluorescence all AtCSLA-Venus isoforms (except AtCSLA1- and AtCSLA15-Venus) were expressed and have shown a significant increase in fluorescence (Figure 25A). AtCSLA1- and AtCSLA15-Venus were examined with a confocal microscope. AtCSLA1-Venus have shown relative less fluorescence that was visible in small puncta, no fluorescent signal was detected for *Pichia* expressing AtCSLA15-Venus (Figure 25B). The low fluorescence of AtCSLA1-Venus is likely related to the fact that only a partial coding sequence with a length of 401 AAs (KJ138750.1) was selected using the Aramemnon database. The full-length coding sequence of AtCSLA1 (AT4G16590.1) obtain comparatively 520 AAs, leading in a truncated protein. The truncated AtCSLA1 is starting 47 AAs after the sequence of the first TMD. This explains the low expression level of AtCSLA1 (Figure 25B). AtCSLA15-Venus DNA from *E. coli* was examined by digestion with different restriction enzymes and verified by sequencing to explain why this strain have shown no fluorescence. As a result, an *E. coli* sequence of 777 bp was incorporated into Venus, leading to a frame shift and a truncated Venus protein (Appendix Figure S1). Therefore, AtCSLA15-Venus was excluded for monosaccharide analysis of the alkaline-insoluble polymers. Representative images of the localization in fluorescent puncta of the other AtCSLA isoforms made with confocal microscope are presented in Appendix Figure S2. Only AkCSLA3-, AtCSLA3- and AtCSLA9-Venus strains have shown significantly increased Glc and Man levels with different amounts in AKI polymers (Figure 25C), suggesting that different CSLAs form distinct HM structures. For all other AtCSLA-Venus enzymes no difference in Glc or Man was detected compared to the empty vector control (Figure 25C). Normally, AtCSLA2 expressed in *Pichia* produce pure mannan [Voiniciuc et al., 2019], leading to the conclusion that the Venus fusion at the C-terminal end of AtCSLA2 impair its activity.

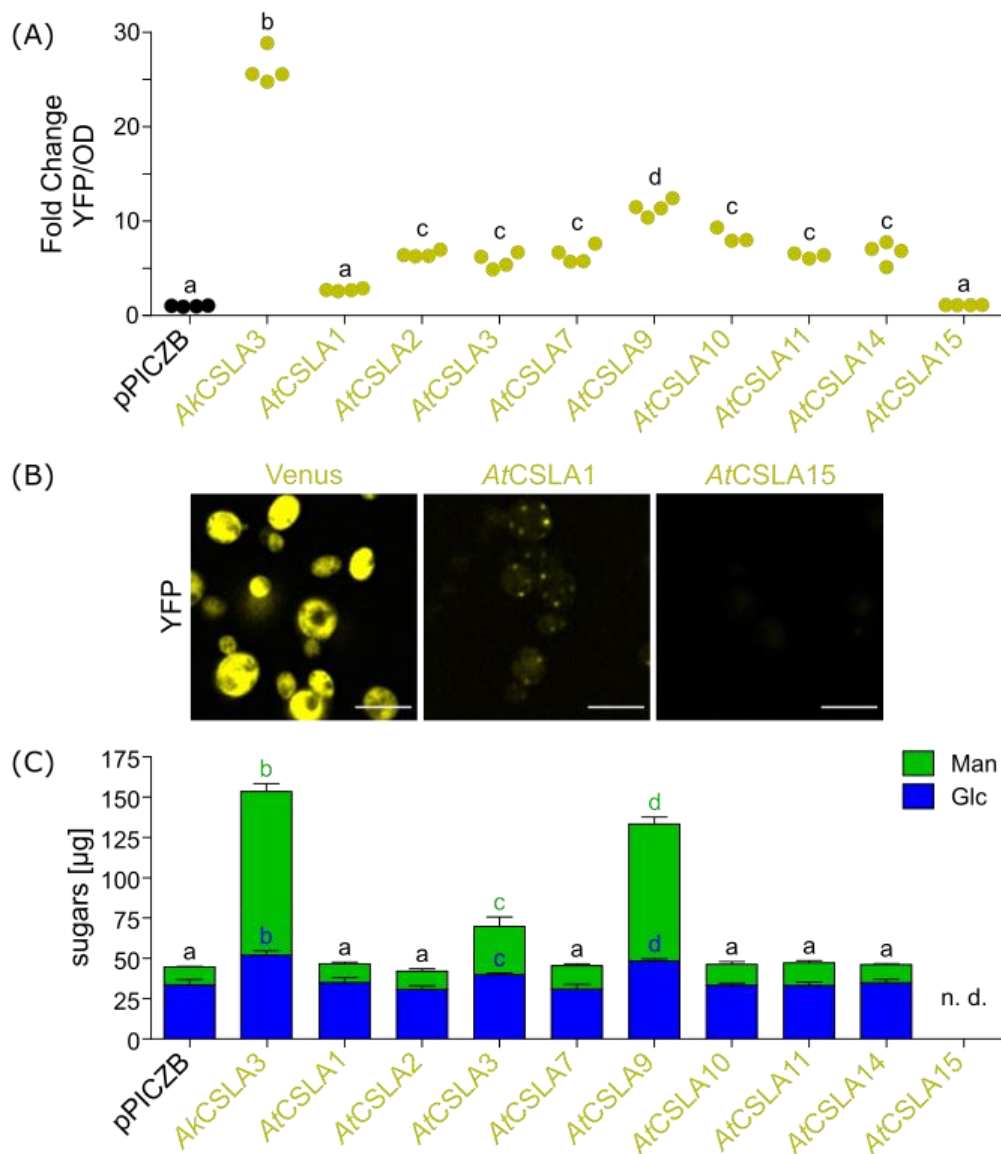


Figure 25: Screening of AtCSLA-Venus isoforms expressed in *Pichia*. Yeast strains were cultivated for 64 h in BMGY followed by 24 h growth in BMMY. (A) Fold change of normalized yellow fluorescence (YFP/OD) relative to the pPICZB control. (B) Microscope images of Venus, AtCSLA1 and AtCSLA15 made with a confocal microscope (LSM880). Intensity of the laser was increased for AtCSLA1 and AtCSLA15 to a gain of 1000, while the gain for Venus was 880. Scale bar: 5 µm. (C) Monosaccharides of AKI wall polymers. Polymers were hydrolysed with sulfuric acid. Data show the mean + SD of four biological replicates, and different letters denote significant changes (one-way ANOVA with Tukey test, $p < 0.05$). *Ak*: *Amorphophallus konjac*; *At*: *Arabidopsis thaliana*.

Because Venus impairs the activity of AtCSLA2 it is possible that the tag also influences the activity of other AtCSLAs. Therefore, other strategies to fluorescently tag AtCSLAs were pursued. Instead using Venus as a fluorescent tag that is able to form weak dimers, a monomeric yellow fluorescent protein (mCitrine) was selected. Furthermore, a self-cleaving 2A peptide from porcine teschovirus 1 (P2A) was cloned as a linker in between the sequence of the specific AtCSLA and sfGFP. During the translation on the ribosomal complex, AtCSLA and sfGFP are “cleaved” by skipping some amino acids on the C-terminal end of the P2A sequence

[Kim et al., 2012]. The P2A enables the visualization of the protein expression by fluorescence in the cytosol of *Pichia* without having the possible negative effects of a fused fluorescent protein. To test these additional fluorescent tags, *AtCSLA2*, *AtCSLA3* and *AtCSLA14* were selected and transformed into *Pichia* (Figure 26).

Based on the fluorescence all *Pichia* strains have expressed the respective protein with the fluorescent tag except *AtCSLA14-2A-sfGFP* that show no fluorescent signal similar to the empty vector control and *AtCSLA2* (Figure 26A). None of the fluorescently labelled *AtCSLA2* expressing *Pichia* variants have shown an outstanding increase in mannose in AKI polymers compared to untagged *AtCSLA2* (Figure 26C), which is probably caused by mislocalization of the fluorescently tagged *AtCSLA2* (visible as double ring pattern in microscope images (Figure 26B)). *AtCSLA3* fused with Venus and mCitrine were detected in puncta (Figure 26B), but only *AtCSLA3-Venus* was able to produce more mannose. Compared to *AtCSLA3-2A-sfGFP* strain, the increase in mannose was negligibly small (Figure 26C) and indicates a negative influence of Venus on *AtCSLA3* activity. *AtCSLA14-Venus* and -mCitrine was localized in a few puncta, but additionally have shown a lot of round-shaped background fluorescence (Figure 26B) and none of these variants were functional in producing HMs (Figure 26C). In summary, the fused fluorescent proteins have severe effects on the localization and function of some *AtCSLAs* and there is no single fluorescent tag that could guarantee HM activity for all *AtCSLA* proteins. In order to check if *CSLA* have a function in HM biosynthesis, it would be better to examine first the untagged proteins first. The AKI of *Pichia* expressing untagged *AtCSLA10*, *AtCSLA11* and *AtCSLA15* was examined and none of these three proteins produce any significant amounts of 4-Man (Appendix Figure S3 for *AtCSLA10* and *AtCSLA15*; data for *AtCSLA11* not shown).

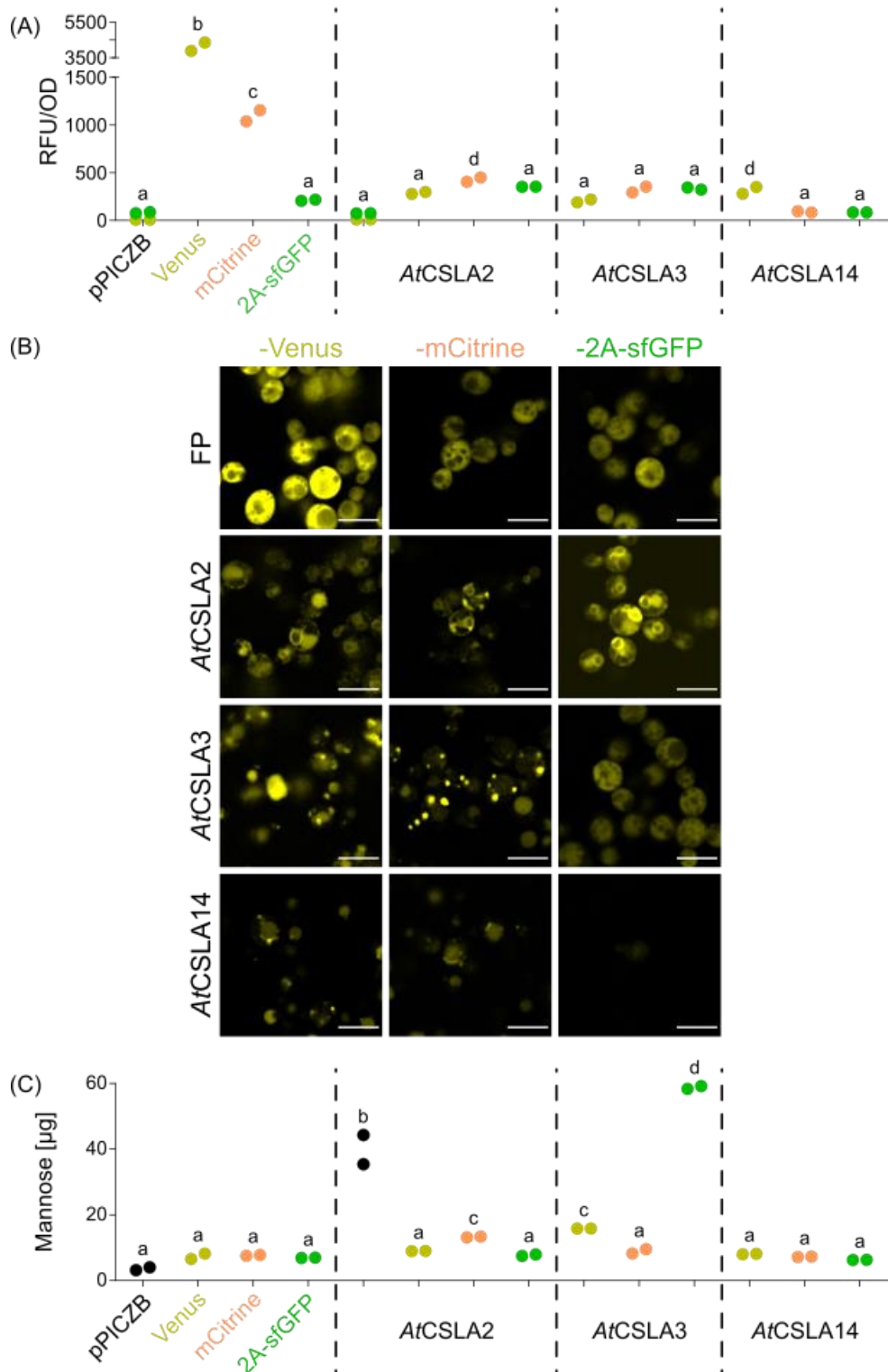


Figure 26: *AtCSLA* proteins fluorescently tagged to Venus, mCitrine and 2A-sfGFP. *Pichia* were grown in 2 ml BMGY for 24 h, followed by 24 h growth in 2 ml BMMY for protein expression. (A) normalized fluorescence (RFU/OD) of *Pichia* expressing *AtCSLA*2, 3 and 14 with different variants of fluorescent tags. (B) Localization of *AtCSLA* tagged enzymes in *Pichia*. Microscope images were made with a confocal microscope (LSM880). Scale bar: 5 μm (C) Monosaccharides of AKI wall polymers. Polymers were hydrolysed with TFA. Data show two biological replicates, and different letters denote significant changes (one-way ANOVA with Tukey test, $p < 0.05$). Yellow dots symbolizing Venus, orange dots mCitrine and green dots 2A-sfGFP fusions on the C-terminal end of the respective *AtCSLA* proteins. *At*: *Arabidopsis thaliana*; FP: fluorescent protein.

Nevertheless, the product specificity of HM producing *AtCSLA*-Venus *Pichia* strains (*AtCSLA3* and *AtCSLA9*) could be verified. Additionally, *AkCSLA3*-Venus (positive control for glucomannan), *AtCSLA2*-Venus (negative control) and *AtCSLA7*-Venus were included, because *AtCSLA7* expression in *Pichia* have shown to significantly increase 4-Man linkages into AKI polysaccharides [Voiniciuc et al., 2019]. The *Pichia* strains were cultivated in YPMG for 48 h instead of using BMGY, because this “co-feeding” strategy minimizes the handling steps and the interruption of growth during the media change.

All *AtCSLA*-Venus genes were expressed in *Pichia* related to the fold change in fluorescence, while *AkCSLA3*-Venus have shown the highest fluorescence (Figure 27A). *AtCSLA2*-Venus have shown fluorescence in a double-ring pattern, whereas the fluorescence of the other *AtCSLA*-Venus expressing strains was localized in punctae, but have shown also a high background fluorescence (Figure 27B).

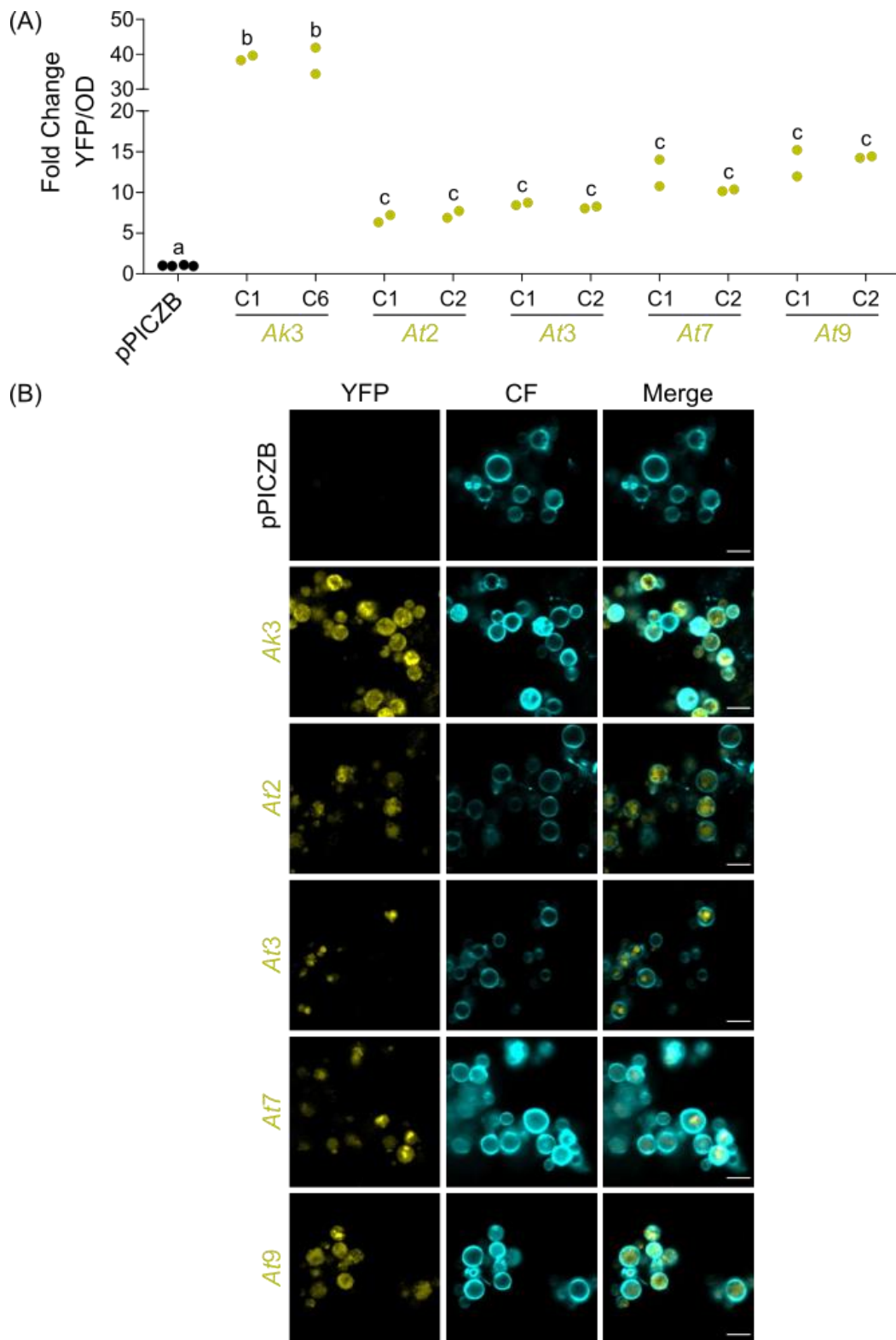


Figure 27: *Pichia* expressing *Ak*CSLA3- and HM-producing *At*CSLA-Venus isoforms. *Pichia* cultures were grown in 2 ml YPMG for 48h. (A) Fold Change of normalized yellow fluorescent proteins (YFP/OD) of two different colonies and two biological replicates for each CSLA expressing strain. (B) Localization of expressed CSLA-Venus using a confocal microscope (LSM900). The plasma membrane of *Pichia* cells was stained with calcofluor white (CF). Scale bar: 5 μ m. *Ak3*: *Amorphophallus konjac* CSLA3; *At*: *Arabidopsis thaliana* CSLA; C: colony.

The AKI of CSLA expressing *Pichia* were digested with endo-1,4- β -mannanase to detect only Glc and Man that was produced by the specific CSLA (Figure 28A). Additionally,

linkage analysis of the AIR of *Pichia* was performed and 4-Man and 4-Glc values represent the HM structure that was produced by the respective CSLA (Figure 28B). *At*CSLA9-Venus have released increased amounts of Glc and Man after mannanase digestion (Figure 28A) and increased 4-Man and 4-Glc linkages were detected (Figure 28B), meaning *At*CSLA9 produce glucomannan. In contrast, *At*CSLA3- and *At*CSLA7-Venus have release only Man into the supernatant after the enzyme treatment (Figure 28A) and only 4-Man linkages could be detected (Figure 28B), indicating pure mannans are produced in yeast.

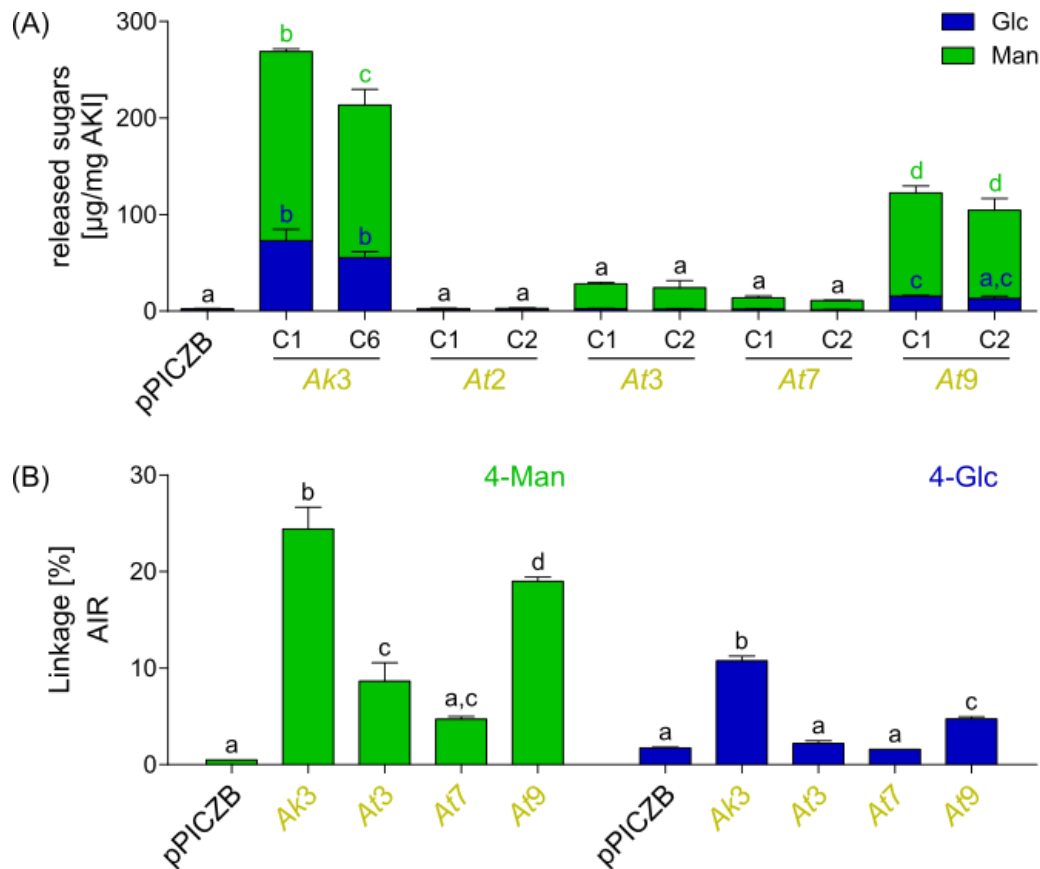


Figure 28: Verification of Heteromannans produced by *At*CSLA-Venus expressing *Pichia* strains. *Pichia* cultures were grown for 48 h in 2 ml YPMG. (A) Alkaline insoluble (AKI) polymers from *Pichia* strains were digested with endo- β -mannanase and released sugars that are specific for mannan and glucomannan produced by each CSLA, were hydrolysed with TFA and determined via HPAEC-PAD. Data show the mean + SD of two different colonies and two biological replicates for each strain. Different letters denote for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). (B) Glycosidic linkage analysis of *Pichia* alcohol-insoluble residue (AIR) polymers. Data show the mean + SD of two biological replicates for each strain. Letter denotes for significant changes (one-way ANOVA with Tukey test, $p < 0.05$). Ak3: *Amorphophallus konjac* CSLA3; At: *Arabidopsis thaliana* CSLA; C: colony.

3.7 Structural analysis of *Arabidopsis thaliana* CSLAs

In order to discover conserved sequence motifs of all nine CSLA isoforms from *Arabidopsis*, the motif discovery tool of the MEME Suite web server (<https://meme-suite.org/meme/>; version 5.5.3) was used [Bailey, 2007; Bailey et al., 2009]. In total ten sequence motifs were searched in all *At*CSLA isoforms. Interestingly, all *At*CSLA proteins that

have shown a function in mannan or glucomannan synthesis (*AtCSLA2*, 3, 7, and 9) on its own share all ten sequence motifs, while *AtCSLA1*, 11 and 14 miss some sequence motifs (Figure 29). Especially, *AtCSLA1* and 14 are interesting, because they miss motifs in the region of their catalytic region that could be necessary for their function.

Furthermore, the amino acid sequences of all *AtCSLA* isoforms were aligned with other *Arabidopsis* CESA and CSLD family members using the Molecular Evolutionary Genetics Analysis (MEGA) software (version 11) to identify amino acid residues participated in binding of nucleotide sugar substrates [Yang et al., 2020]. They pointed out that sequence motifs that distinguish or coordinate nucleotide binding of CESA and CSLD members are mostly different from those of CSLAs and associated with a different specificity for the binding nucleotides. For example, CSLA sequences contain unique aromatic residues (F, Y, W) in the sequence that distinguish between UDP and GDP sugars, and these are necessary to interact with guanine and the proposed catalytic base with the VED motif to bind to GDP-Man [Yang et al., 2020]. The VED motif is represented in all *AtCSLA* isoforms. While the aromatic residue Y is represented in motif 7 (dark blue) for mostly all *AtCSLA* isoforms, it is changed to F for *AtCSLA14*. Because *AtCSLA1* lacks the motif 5 it is the only isoform that misses the aromatic residue F or Y in motif 5 (light orange, Figure 29). Furthermore, the sequence motifs DD or DxD (coordinating UDP or GDP binding) and Q/RxxRW (coordinating Glc binding) were compared in *AtCSLA* isoforms because they are critical for the catalytic activity of the GT [Saxena et al., 1995; Charnock et al., 2001]. For all *AtCSLA* isoform these motifs were present and conserved as DD (motif 2, light blue; motif 9, dark orange), DAD (motif 3, light green) and QQHRWT (motif 2, light blue, Figure 29). Other motifs, such as VQxxxxF and VFFD, are integral in coordinating the glucan chain in CESA and CSLD proteins [Yang et al., 2020]. Notably, the VQxxxxF motif is found in almost all *AtCSLA* isoforms, with the exception of *AtCSLA14* where the sequence is altered to VQxxxxY (motif 1, red, Figure 29, Appendix Figure S4). Regarding the VFFD motif, most *AtCSLA* isoforms feature the NYHF sequence instead. *AtCSLA2* and *AtCSLA9* stand out as the only members of the family that possess the DYHF sequence. Furthermore, *AtCSLA7* diverges from other CSLA isoforms, exhibiting the unique SYHF sequence (Appendix Figure S4).

Overall, the motifs in *AtCSLA* isoforms seem to be very conserved and are different to the ones of CESA and CSLD members [Yang et al., 2020]. Because *AtCSLA1* and *AtCSLA14* are the only isoforms that lack UDP/GDP coordinating motifs in their catalytic region with roles in nucleotide sugar binding, they are especially interesting from the non-functional *AtCSLAs*. Additionally, *AtCSLA11* is missing a motif responsible for coordination of UDP/GDP binding and might explain why *AtCSLA11* is not producing HM alone. Nevertheless, *AtCSLA10* and *AtCSLA15* have all motifs as *AtCSLAs* functional in HM production (*AtCSLA2*, 3, 7, 9;

Figure 28) and still were not able in *Pichia* to produce HMs. Therefore, small changes in one amino acid sequence or other factors as interaction with other proteins and cofactors could affect the activity of the specific CSLA protein.

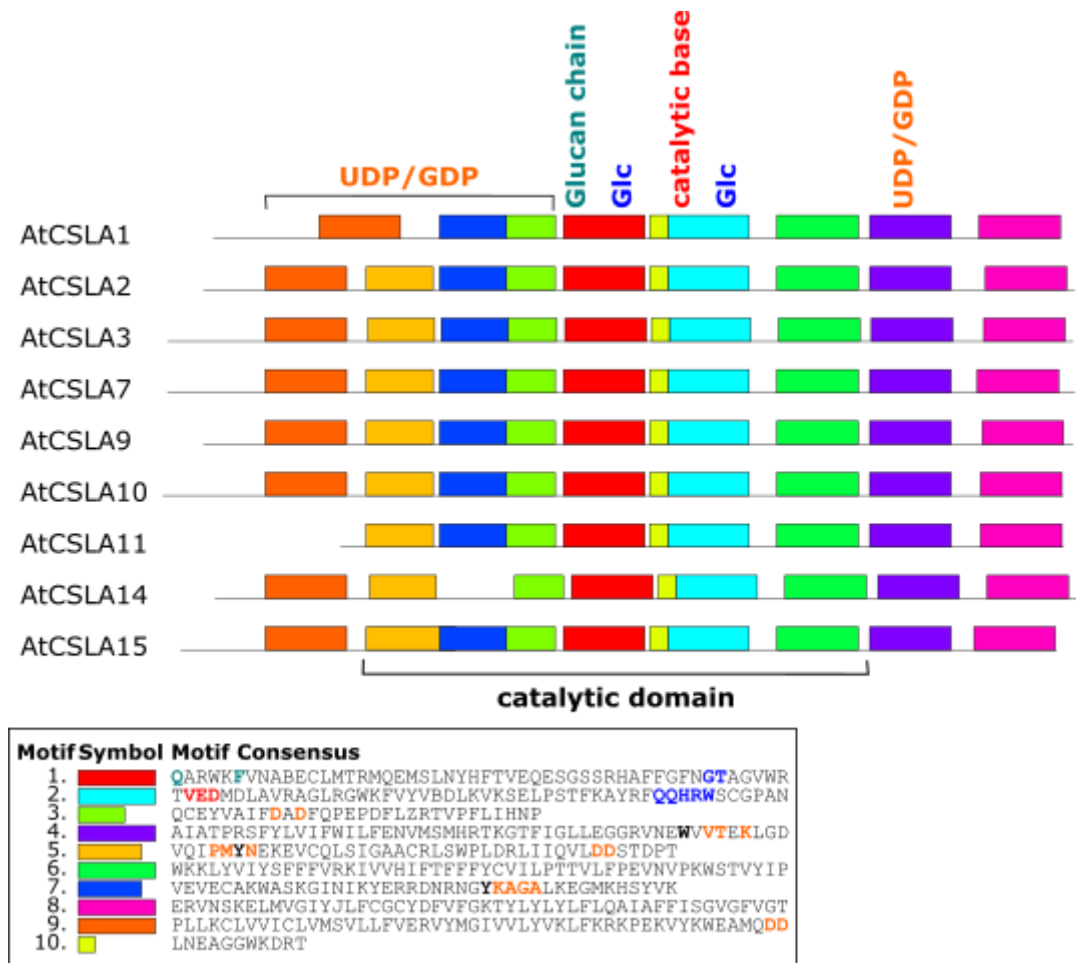


Figure 29: Conserved motif domains of *AtCSLA* isoforms and *AkCSLA3*. Conserved motifs were discovered and illustrated with the MEME web server. The amino acid sequences of *AtCSLAs*, *AkCSLA3* were aligned and compared to each other. Sequences participated in nucleotide binding (published by Yang et al., (2020)) were highlighted in the motif consensus sequence. Amino acids in bold represent identical sequences of CSLA members that are probably coordinating UDP/GDP (DD; DxD; orange), coordinating glucose (Glc; QxxRW; dark blue), coordinating glucan chains (VQxxxxF; light blue), distinguish UDP/GDP (WxVTxK, YKAGA; orange) and catalytic bases (VED).

3.8 *Pichia* strains expressing multiple *AtCSLA* isoforms

Because *csla2* single mutants in *Arabidopsis* have no significant impact in oligosaccharide composition, but *csla2csla9* double mutants have shown tremendously less mannose content in the stem related to Col-0 and *csla9* single mutants [Goubet et al., 2009], an interaction of *AtCSLA2* and *AtCSLA9* would be possible and may enhance glucomannan synthesis. Therefore, multigene constructs of *AtCSLA2* together with *AtCSLA9*, 11 and 14 were cloned in *Pichia*. However, only the single gene constructs of *AtCSLA2* and *AtCSLA9* were functional (Appendix Figure S5). The mannose content of the multigene constructs of

AtCSLA2 and *AtCSLA9* have shown the same level as sfGFP negative controls, meaning that the expression of multigene constructs in *Pichia* had problems in their expression in general.

To circumvent the expression problem of multigene constructs, another cloning strategy was pursued to create *AtCSLA2* and *AtCSLA9* co-expression strains. Instead of creating multigene constructs, *AtCSLA9* was transformed into *AtCSLA2* expressing *Pichia* strains. In order to guarantee *AtCSLA9* expression, *AtCSLA9-Venus* was selected for transformation into *AtCSLA2* expressing *Pichia* strains because its expression could be detected by the fluorescent tag. Both genes were expressed under the control of the strong methanol-inducible promoter pAOX. Based on the fold change of the normalized fluorescence, the protein expression of *AtCSLA2* and *AtCSLA9-Venus* together was mostly similar/ not much less compared to the original *AtCSLA9-Venus* strain for colony 4 and 11, while colony 3 was expressed significantly less (Figure 30A). Although the mannose content of the AKI material was significantly increased for all colonies co-expressing *AtCSLA2* and *AtCSLA9-Venus* compared to the single transgene strains (Figure 30B), the content was not increased in a tremendous manner as it would be expected if both proteins together boost the Man incorporation into the HM backbone. Of course, it could also be that the glucomannan production in *Pichia* cells have reached a maximal saturation point. Based on other experiments in this study where the sugar content of the AKI was calculated, *Pichia* is in general able to produce much more glucomannan *in vivo* and, therefore would be very unlikely. In order to check if the (gluco-)mannan production was really saturated, *Pichia* strains co-expressing *AtCSLA2* and *AtMSR1* could be included as a control in the same growth batch. This strain does not only increase the Man content, it also incorporates Glc in the HM backbone because *AtMSR1* is an important cofactor for *AtCSLA2* [Voiniciuc et al., 2019]. Furthermore, *csla3csla9* double mutants have shown also significantly less mannose content in *Arabidopsis* stems like *csla2csla9* mutants, but this reduction was even surpassed in a tremendous manner by *csla239* triple mutants [Goubet et al., 2009]. Maybe all three CSLAs (CSLA2, 3 and 9) or other co-factors as *AtMSR1* are needed to increase glucomannan production in a higher manner as it was shown in planta.

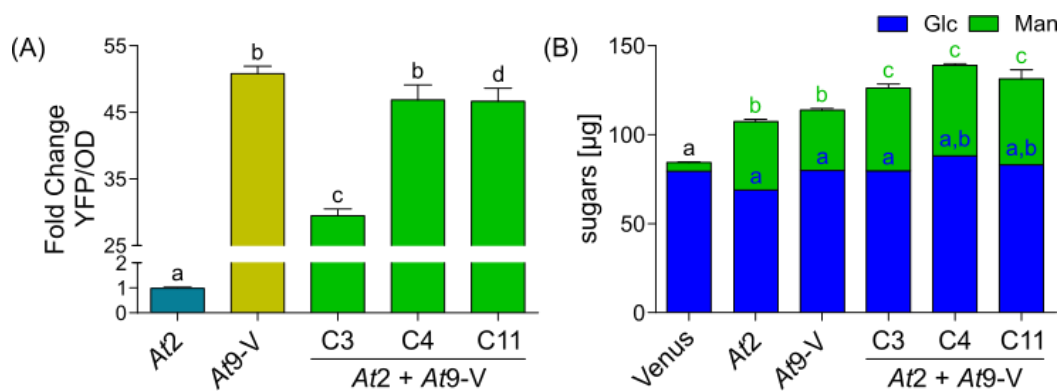


Figure 30: *AtCSLA2* and *AtCSLA9* co-expression in *Pichia*. *Pichia* were cultivated in 2 ml YPMG for 48 h. (A) Fold change of normalized yellow fluorescence (YFP/OD) relative to the non-fluorescent *AtCSLA2* control. (B) Mannose content of AKI wall polymers. Polymers were hydrolysed with TFA. Data show the mean + SD of three biological replicates and letters denote for significant changes (one-way ANOVA with Tukey test; $p < 0.05$). *At*: *Arabidopsis thaliana*; V: Venus.

3.9 Characterization of additional CSLK/As across the plant kingdom

In order to find other CSLK/A enzymes involved in HM biosynthesis that could give insights into their structural diversity during evolution throughout the plant kingdom, the protein sequence of *AtCSLA2* was compared to other biological sequences using the protein BLAST search from the Phytozome database (<https://phytozome-next.jgi.doe.gov/>), the National Center for Biotechnology (NCBI; <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and PhycoCosm database (<https://phycocosm.jgi.doe.gov/phycocosm/home>). The table with the identity and query coverage of targeted sequences is listed in Table 4.

Coding sequences (CDSs) of the selected CSLK/A enzymes were whether synthesized by Genart/GenScript or if plant material was available amplified from cDNA by Phusion PCR. The amplified CDS was then first cloned into *E. coli* and finally integrated into the *Pichia* genome. While the cDNA ordered from GenScript was not codon optimized, the CDSs synthesized by Genart were codon optimized for *Pichia pastoris* using the Codon Usage Frequency Table from GenScript (<https://www.genscript.com/tools/codon-frequency-table>) to increase their translation efficiency in *Pichia* (Table 4). All of the selected CSLK/As were expressed in *Pichia* under the control of the strong methanol-inducible promoter pAOX1. To find out, if the different CSLK/A enzymes produce HMs in *Pichia* expressing strains several colonies were tested first. All of the *Pichia* expressing strains were firstly grown in media using glycerol to accumulate their biomass production for 24 h followed by further 24 h growth using methanol to induce the protein expression of the respective CSLK/A. After the proteins were expressed, *Pichia* cultures were harvested, the AKI was extracted and the sugar content of the AKI was whether directly tested or digested with endo- β -1,4-mannanase.

Table 4: Putative (gluco)mannan synthases from species throughout the plant kingdom homologous to AtCSLA2. For each species, the closest homolog to AtCSLA2 was identified using the protein BLAST of different databases. The table shows the full species name, the abbreviation (Abbrev.), coverage (Cov) and the identity (Ident) in percent of the target sequences. Some coding sequences were codon optimized (Cod. opt.) for *Pichia pastoris* to increase their translational efficiency and protein expression.

Species	Abbrev.	Cov [%]	Ident [%]	Accession	Database	Source	Cod. opt.
<i>Emiliania huxleyi</i>	EhCSLK	48	41	Emihu1I452714	PhycoCosm	Geneart	Yes
<i>Porphyra umbilicalis</i>	PumCSLK1	54	36	Pum0231s0012.1.p	Phytozome	Geneart	Yes
	PumCSLK2	54	36	Pum0372s0002.1.p	Phytozome	Geneart	Yes
<i>Chlamydomonas reinhardtii</i>	CrCSLK	82	36	Cre06.g268550	Phytozome	GenScript	No
<i>Micromonas pusilla</i> CCMP1545	MicpuCSLK	81	44	MicpuC2.e_gw1.16.71.1	Phytozome	GenScript	No
<i>Mesostigma viride</i>	MvCSLK	93	45	Mesvi205S04101	PhycoCosm	Geneart	Yes
<i>Ostreococcus lucimarinus</i>	OICSLK	91	38	O.lucimarinus v2.0 18489	Phytozome	Geneart	Yes
<i>Chara braunii</i>	CbCSLK	67	44	GBG73110.1	NCBI	Geneart	Yes
<i>Physcomitrium patens</i>	PpCSLA1	96	63	Pp3c17_19630v3	Phytozome	cDNA	No
	PpCSLA2	96	62	Pp3c14_13750v3	Phytozome	cDNA	No
	PpCSLA3	96	62	Pp3c14_12270v3	Phytozome	cDNA	No
<i>Marchantia polymorpha</i>	MpoCSLA	96	63	Mapoly0065s0098	Phytozome	cDNA	No
<i>Selaginella moellendorffii</i>	SmCSLA	97	63	S.moellendorffii v1.0 230176	Phytozome	GenScript	No
<i>Amborella trichopoda</i>	AtrCSLA	99	72	XP_006851749.1	NCBI	Geneart	Yes
<i>Amorphophallus konjac</i>	AkCSLA	94	61	ADW7764.1	NCBI	DNA	No
<i>Dendrobium officinale</i>	DoCSLA6	98	72	XP_020673043.1	NCBI	Geneart	Yes

First of all, *Pichia* expressing CSLK/A with CDSs synthesized by Geneart were tested (Figure 31). The *Pichia* CSLK/A expressing strains in which the CDS was not previously codon optimized for gene translation in *Pichia* were the CSLK of the unicellular green alga *Chlamydomonas reinhardtii* (Cr) and *Micromonas pusilla* (Micpu) as well as the ancient vascular plant belonging to the lycophytes *Selaginella moellendorffii* (Sm). All of the tested colonies for these CSLK/A have not produced any HM because they show a similar level of released Man after β -mannanase digestion as Venus control strains (Figure 31A). Of course, this could be due to the fact that the amino acid sequence that is naturally present in plant genes cannot be translated at the ribosomal complex of *Pichia*. Other experiments were performed in this study with CrCSLK, MicpuCSLK and SmCSLA fused to Venus on the C-terminus of the respective CSLK/A. After protein expression, the fluorescence of the *Pichia* strains was examined. CrCSLK-Venus have shown almost no yellow fluorescence, while the

yellow fluorescence of *Micpu*CSLK and *Sm*CSLA was measured in a detectable amount and was visible in small puncta (Appendix Figure S5). This indicate genes of *Micpu*CSLK and *Sm*CSLA were translated and expressed at least in minor amounts. The fluorescent tag is able to impair the activity of the specific CSLK/A, but if these proteins were involved in HM biosynthesis, the increase in Man would be detectable after β -mannanase digestion of the AKI for the untagged constructs. To avoid the problem of a possible loss in gene translation efficiency, other CDSs of CSLK/A that were synthesized by Geneart were codon optimized for *Pichia* afterwards. *Pichia* expressing CSLK of other unicellular green algae – *Ostreococcus lucimarinus* (*O*CSLK) and *Mesostigma viride* (*M*VCSLK) – and a higher developed green algae with leafy and root-like tissue – *Chara braunii* (*C*bCSLK) – were analysed. In case of *O*CSLK and *M*VCSLK expression in *Pichia*, the AKI polymers were treated with β -mannanase and highlighted that no HMs were heterologously produced (Figure 31B). *C*bCSLK expression in *Pichia* have led to a minor increase of Man compared to the sfGFP control (Figure 31C), but it has to be mentioned that only the monosaccharides from the extracted AKI were analysed and it has to be verified if this small increase is based on synthesis of HM. A separated clade of algae are marine haptophytes belonging to the group of phytoplankton. The best known representative of this class is *Emiliana huxleyi* (*E*h). *E*hCSLK expressed in *Pichia* have resulted in a negligible increase of Man content in the AKI compared to the negative control sfGFP (Figure 31C). Surprisingly, the AKI extracted from *Pichia* expressing red algal CSLK1 and 2 from *Porphyra umbilicalis* (*P*umCSLK1 and 2) have shown a remarkable increase in Man release after digestion with β -mannanase (Figure 31B) with no Glc release, indicating a function in pure mannan synthesis. Two other CSLA members from land plants were tested in *Pichia* – CSLA from *Amborella trichopoda* (*A*trCSLA) belonging to the class of angiosperms and a monocot from *Dendrobium officinale* (*D*oCSLA6). The Man content of the extracted AKI have shown to be increased for *A*trCSLA and *D*oCSLA6 compared to sfGFP as a negative control (Figure 31C).

(Figure 32A). Also, the Man content of all three isoforms of the moss *Physcomitrium patens* (*PpCSLA1–3*) was increased in the *Pichia* AKI compared to the empty vector control pPICZB (Figure 32B). The increase of Man of *PpCSLA1* was lower compared to *PpCSLA2* and *PpCSLA3* (Figure 32B).

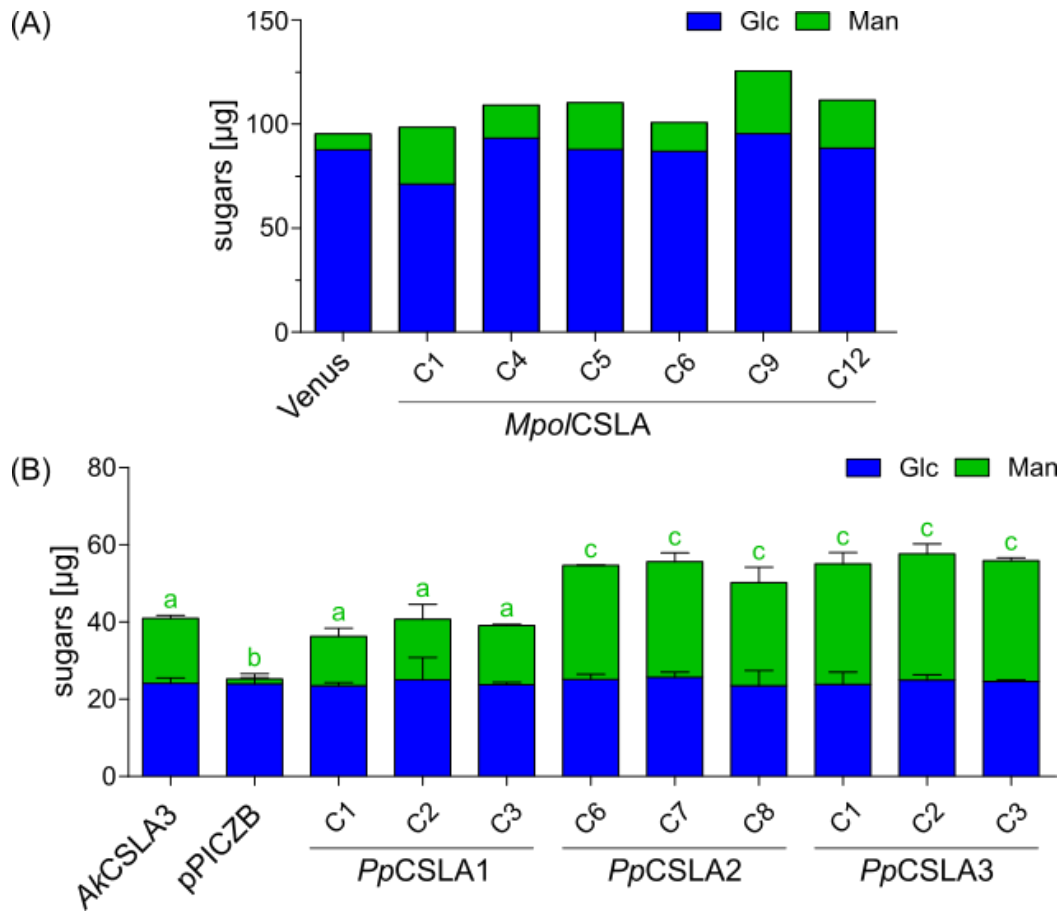


Figure 32: Screening of native CSLA enzymes with coding sequences extracted from plant samples expressed in *Pichia*. *Pichia* colonies expressing *Marchantia polymorpha* CSLA (*MpoCSLA*) were grown in 1 ml YPMG for 48 h in a 24-well EnzyScreen plate (A). *Pichia* colonies expressing *Physcomitrium patens* CSLA isoforms 1–3 (*PpCSLA1–3*) were grown in 600 µl BMGY for 72 h, followed by adding 600 µl BMMY (final concentration 0.5% (v/v) methanol) and further growth for 24 h to express the heterologous proteins (B). Alkaline insoluble polymers from all *Pichia* strains were extracted. Polymers were hydrolysed with sulfuric acid (A) or TFA (B). Data show the mean + SD of three biological replicates and letters denote for significant changes (one-way ANOVA with Tukey test; $p < 0.05$; B). C: colony; Glc: glucose; Man: mannose.

These first tests revealed that most of the ancestral CSLKs are not (green algae: *CrCSLK*, *MicpuCSLK*, *OICSLK*, *MvCSLK*) or may be only negligible (phytoplankton: *EhCSLK*; green alga: *CbCSLK*) able to produce HMs in *Pichia*, except of the algal isoforms *PumCSLK1* and 2. Most of the tested CSLA enzymes that were expressed in *Pichia* were able to increase the Man content, except the spikemoss CSLA (*SmCSLA*). The non-functionality in producing HMs does not mean that the selected CSLK/A enzymes are not involved in HM production. It could be that the genes were not optimal expressed in *Pichia* or other proteins and cofactors that occur naturally *in planta* and normally interact together are missing in our expression host.

3.10 Detailed characterization of CSLK/As in the plant kingdom

To determine the polysaccharide composition of HMs produced by CSLK/A expressing *Pichia* strains from section 3.9 further analysis as enrichment of the AKI with zymolyase or a combination of β -mannanase digestion of the AKI and linkage analysis of the AIR was performed.

The AKI material from three replicates of *Pichia* expressing *PpCSLAs* was pooled and treated with zymolyase. Zymolyase digests β -1,3-glucan [Magnelli et al., 2002] and, therefore enrich specifically the CSLA product in the material. The Glc and Man increase in the enriched material (EM) of *Pichia* strains expressing *PpCSLA* isoforms (Figure 33A) meaning that glucomannans were produced by the *Pichia* strains. *PpCSLA1* expressing *Pichia* strains have produced a glucomannan with the lowest amount of Man and minor amounts of Glc. *Pichia* strains expressing *PpCSLA2* and *PpCSLA3* have very high and similar amounts of Man in the EM, whereas the Glc content in *PpCSLA2* strains was lower compared to *PpCSLA3* (Figure 33A). Because all three isoforms of *PpCSLA* incorporate different amounts of Glc and Man into the produced glucomannan, their backbone structure seem to be highly dependent on the specificity of the respective CSLA protein.

For other CSLK/A enzymes one representative strain or the strain with the highest amount in Man content of the AKI was selected and four biological replicates were re-grown in one growth Batch to better compare the yield in HM production with each other. The AIR of the *Pichia* strains containing all yeast polysaccharides including the produced HMs was extracted and used for linkage analysis, while the AKI was isolated and digested with β -mannanase to test if the respective CSLK/A produce mannans or glucomannans. As already shown in Figure 32, *Pichia* strains expressing CSLK from green algae (*MicpuCSLK*, *OCSLK*) were not able to produce HMs because neither the Glc or Man content of the digested AKI nor the 4-Man and 4-Glc content of the AIR was elevated compared to the Venus control (Figure 33B and 33C). The *Pichia* strains expressing *EhCSLK* and *CbCSLK* that have shown previously a minor increase in Man of the AKI fraction compared to the sfGFP control (Figure 32C), have not released more Man and Glc as the Venus control after the AKI was digested with β -mannanase (Figure 33B) and also the linkages of the analysed AIR from these *Pichia* strains have not shown an elevated amount of 4-Man and 4-Glc (Figure 33C). Therefore, the minor increase in Man for the mentioned strains was not related with HM production. Even other linkages typical for polysaccharides were compared, but none of them were significantly increased in the green algal expressing strains (Appendix Table S11). The significant release of only Man sugars was detected for *PumCSLK1*, *MpoCSLA* and *DoCSLA* expressing *Pichia* strains (Figure 33B). *PumCSLK1* expressing strains were only rich in 4-Man linkages with similar levels of 4-Glc linkages compared to the Venus control, revealing *PumCSLK1* produce

pure mannans (Figure 33C). Although the increase in Glc of the digested AKI was not significantly different for *Mpo*/CSLA and Venus expressing strains, it seems that trace amounts of Glc were incorporated into the produced HM backbone and has to be evaluated. Compared to *Mpo*/CSLA and Venus, the Glc content of *Do*CSLA6 was negligible, indicating the HM backbones consist primarily of Man residues (Figure 33B). *Atr*CSLA expressing *Pichia* strains have released significantly amounts of Glc and Man in their AKI compared to the Venus control (Figure 33B) indicating glucomannan production of these strains.

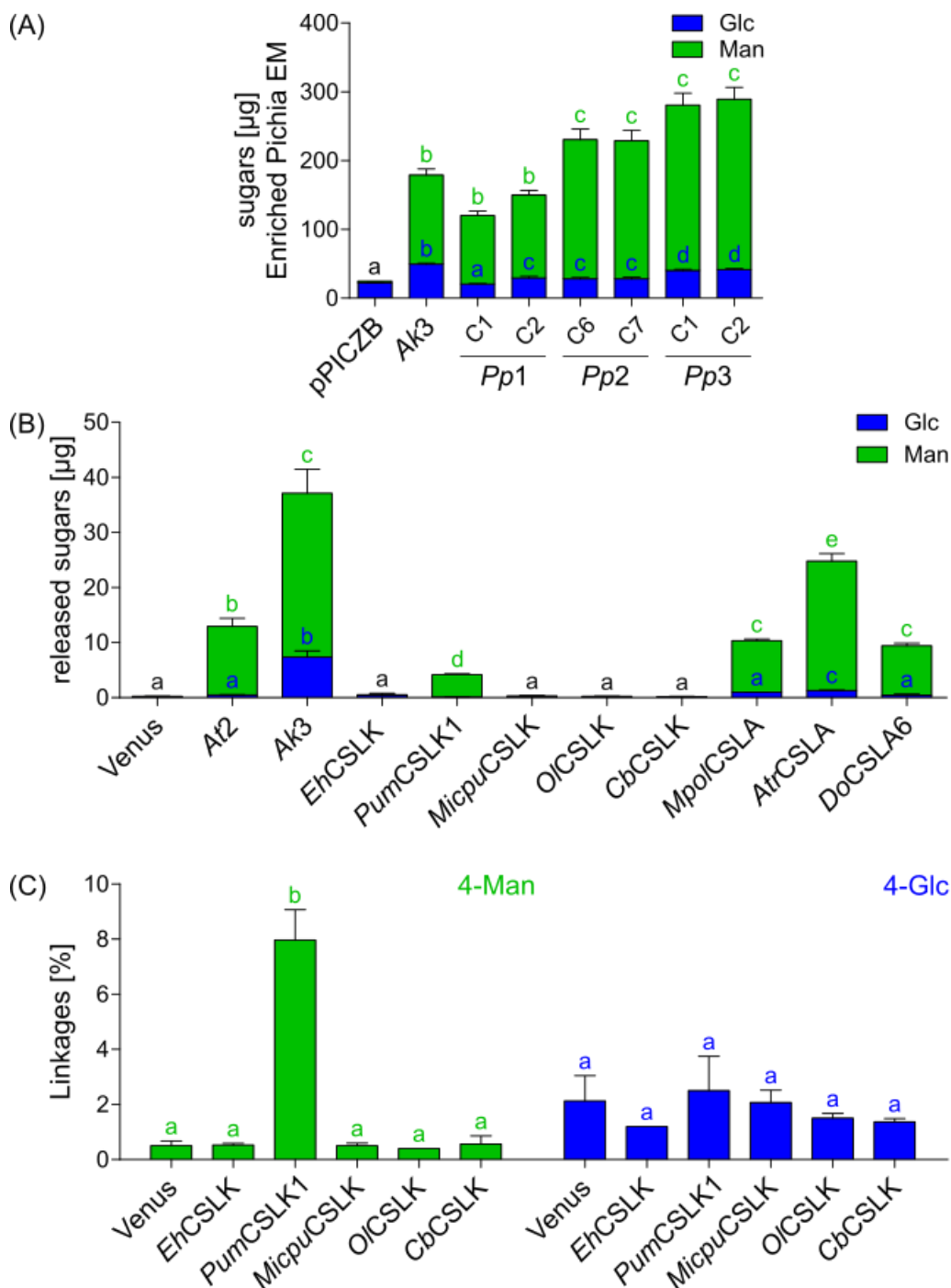


Figure 33: Detailed analysis of *Pichia* expressing different CSLK/A enzymes in HM production. AKI polymers of *Physcomitrium patens* CSLA (*PpCSLA*) expressing strains were pooled and treated with zymolyase to enrich the HMs of the AKI fraction (A). *Pichia* expressing various CSLK/A enzymes were grown in 2 ml YPMG for 48 h. AKI polymers were treated with β -mannanase to release mannose (Man) and glucose (Glc) specifically from the produced (gluco-)mannan into the supernatant (B). EM polysaccharides and AKI supernatants were hydrolysed with TFA and determined via HPAEC-PAD (A and B). Glycosidic linkage analysis of *Pichia* wall AIR (C). Data show the mean + SD of two (C), three (A) or four (B) biological replicates and letters denote for significant changes (one-way ANOVA with Tukey test; $p < 0.05$). *Atr.*: *Amborella trichopoda*; *Cb.*: *Chara braunii*; *Do.*: *Dendrobium officinale*; *Eh.*: *Emiliana huxleyi*; *Micpu.*: *Micromonas pusilla*; *Pum.*: *Porphyra umbilicalis*; *OI.*: *Ostreococcus lucimarinus*;

To gain more detailed structural information of the HMs which were produced in active CSLA expressing *Pichia* strains, a polysaccharide analysis by carbohydrate gel

electrophoresis (PACE) of β -mannanase digested AKI and compared to those from mannose, konjac glucomannan and ivory nut mannan (Figure 34). *AtCSLA2*, *AtCSLA3-2A-sfGFP* and *PumCSLK1* expressing *Pichia* strains released mannose oligosaccharides similar to the bands of INM, whereas *AkCSLA3*, *AtCSLA9-Venus* and *MpolCSLA* expressing *Pichia* released glucomannan oligosaccharides. The different intensity of the bands of the mannan or glucomannan product show the difference in the proportion of Man or Glc incorporated into the mannan or glucomannan chain with several degree of polymerizations. Man₂ and Man₃ is the main sugar in all CSLA expressing with minor amounts of Man₄. Based on the band intensity of the glucosyl residues, *AkCSLA3* expressing *Pichia* strains have incorporated the most Glc in its glucomannan backbone, whereas *AtCSLA9-Venus* and *MpolCSLA* only trace amounts of Glc in their backbones (Figure 34).

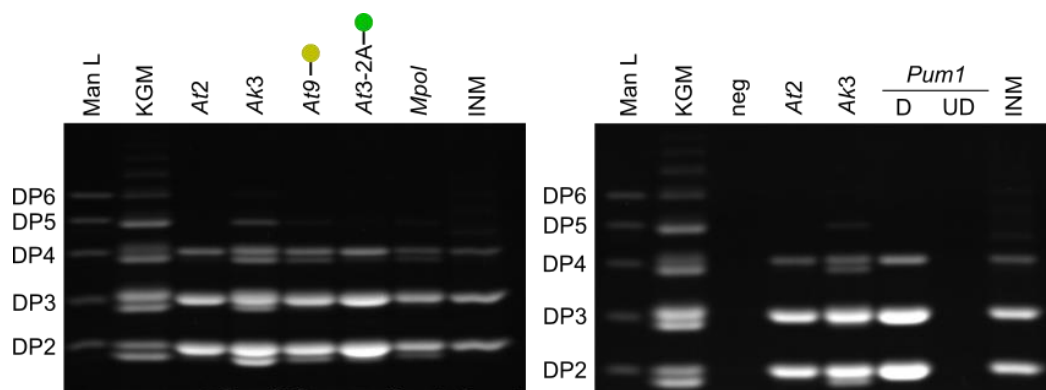


Figure 34: Polysaccharide analysis by carbohydrate gel electrophoresis of active CSLA strains in *Pichia*. Oligosaccharides of the *Pichia* AKI and different ladders (mannooligosaccharides (Man L; DP2 to DP6), konjac glucomannan (KGM), ivory nut mannan (INM)) were released with β -mannanase, derivatised with 2-aminonaphtalene trisulphone (ANTS) and separated by gel electrophoresis. DP: degree of polymerization; *At Arabidoopsis thaliana* CSLAs; *Ak3: Amorphophallus konjac* CSLA; *Mpol: Marchantia polymorpha* CSLA; *Pum1: Porphyra umbilicalis* CSLK1; neg: negative control; D: digested; UD: undigested. Yellow dots indicate fused Venus and green dots illustrate fused sfGFP to the respective CSLA.

So far, the type of HM produced in *Pichia* by the corresponding CSLK/A was determined, however the final polysaccharide may be altered by proteins and cofactors naturally occurring in the respective plant. Furthermore, it is essential to know the overall sugar composition *in planta* because HMs interact with other oligosaccharides and can alter the physical property of the plant cell wall.

First, the AIR material of two red algal NORI leaves (commonly known as Porphyra) from different branches (Nori 1: Obento; Nori 2: Algamar) was extracted, digested with β -mannanase and the oligosaccharides were determined by PACE. Oligosaccharides that were detected in the digested AIR from Nori leaves were only Man₂, Man₃ and Man₄ oligosaccharides, revealing that red alga exhibit pure mannans in their cell wall (Figure 35A). The protocol of Yeats et al. (2016) was performed to determine the matrix and crystalline polymers of the AIR material. Therefore, one set of AIR samples was treated one time with 4%

sulfuric acid to hydrolyse the matrix polysaccharides, while the other set of AIR samples was pre-treated with 72% sulfuric acid followed by 4% sulfuric acid hydrolysis to yield in crystalline polymers. The main matrix polymers found in Nori 1 and Nori 2 AIR was Gal and is comparable to the literature that refer porphyran as the main polysaccharide of the red algal *Porphyra* species consisting of an alternating backbone in 3-linked β -D-galactospyranosyl and 4-linked α -L-galactopyranosyl residues [Peat and Rees, 1961; Rees and Conway, 1962; Frei and Preston, 1964; Hemmingson and Nelson, 2002]. Mannose was the major monosaccharide that was detected in the crystalline polymers of the Nori 1 and Nori 2 AIR (Figure 35B) and form the skeletal component of the plant cell wall of *Porphyra* together with xylan [Frei and Preston, 1964; Usov and Zelinsky, 2013; Brawley et al., 2017]. Taken together, the finding that *PumCSLK* produce pure mannan in *Pichia* and mannooligosaccharides that occur mainly as crystalline polymers in the cell wall support that *PumCSLK* is involved in their synthesis.

Second, the monosaccharide composition of the bryophytes *Physcomitrium patens* and *Marchantia polymorpha* was examined from extracted and destarched AIR material. Major hemicelluloses found in the AIR of bryophyte species are cellulose, xylan, mannan, XyG and RG II, but also arabinogalactan proteins [Popper and Tuohy, 2010, Hornblad et al., 2013]. The major monosaccharide in the *Physcomitrium patens* AIR was Gal (Figure 35C) and can be explained in combination with the detected Ara with a high abundance of arabinogalactan proteins in the gametophore cell wall [Lee et al., 2005; Berry et al., 2016]. Furthermore, Glc, GalA and Xyl were detected in higher amounts in the destarched AIR of the gametophores (Figure 35C). These sugars should be part of xylan (low substituted with α -D-GlcA), HG (unesterified and methyl-esterified) or XyG (non-fucosylated), because several studies have shown binding of antibodies specific for these polymers in the cell wall of gametophores in *Physcomitrium patens* [a nice overview table listed in: Berry et al., 2016]. The Man content in the gametophore was only very low (tested several times; ~3 to 10 mol %) and is controversy to the strong binding of a β -1,4-mannan antibody (BS 400-4) in gametophores [Berry et al., 2016] and the high levels of 4-linked mannans detected using a comprehensive microarray polymer profiling technique [Moller et al., 2007]. Moller et al. treated the AIR with 4 M NaOH and 1% (v/v) NaBH₄ and therefore enrich the hemicellulose fraction what could explain the higher Man content in the gametophytes in their study. Dominant monosaccharides of the destarched AIR of a male (Takaragaike 1 (Tak1)) and female (Tak2) *Marchantia polymorpha* thalli with rhizoids (but without gemmae) were Gal, Ara, Glc and GalA followed by Man, Rha and Xyl (Figure 35D) and could be explained by the presence of arabinogalactan proteins, HG, RG and XyG that were reported [Happ and Classen, 2019]. Interestingly, the Man (9%) and Glc (14%) content of Tak1 plants was slightly but significantly lower compared to Tak2 plants (12% Man and 15% Glc), while the opposite case was shown for the other monosaccharides, except for Gal (Figure 35D). Although the difference in Man and Glc is only minimal between

the two genders, it would be interesting to know if the the Glc and Man is related to glucomannan and if so, the slight increase in glucomannan of Tak2 can compensate other hemicelluloses in their wall.

In summary, the red alga *Porphyra umbilicalis* contain extremely high amount of Man as a crystalline polymer in its wall. *PumCSLK* is most likely involved in its synthesis because this enzyme synthesizes mannans in the *Pichia* expression host. On the other hand, higher amounts of other hemicelluloses than mannan as XyG and xylan is represented beside pectin in the class of bryophytes.

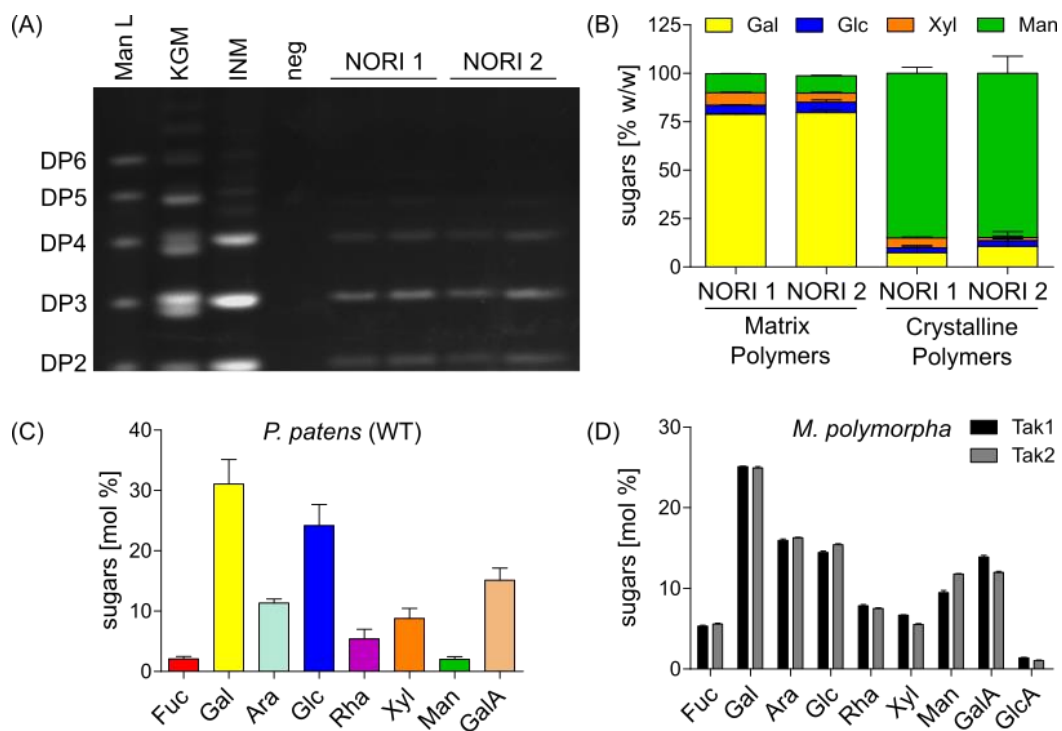


Figure 35: Detailed analysis of polysaccharides in plant material. Oligosaccharides of the extracted AIR material of two red algal NORI leaves from two branches (NORI1: Obento; NORI 2: Algamar) and different ladders (mannooligosaccharides (Man L DP2 to DP6), konjac glucomannan (KGM), ivory nut mannan (INM) were released with β -mannanase, derivatised with 2-aminonaphtalene trisulphone (ANTS) and separated by gel electrophoresis. DP: degree of polymerization; neg: negative control was the undigested AIR of NORI leaf 1 (A). Monosaccharide composition from matrix polymers and crystalline polymers of two NORI leaves from red algae (B). Monosaccharide composition of destarched AIR from leafy gametophores of *Physcomitrium patens* wildtype (WT) Gransden (C) and *Marchantia polymorpha* Takaragaike 1 and 2 (Tak1: male; Tak2: female) (D) Data show the Mean + SD of four (B and D) or six (C) biological replicates.

3.11 CSLA expression *in planta*

In order to investigate CSLK/A expression *in planta*, *AtCSLA9* and *PumCSLK1* (fluorescently tagged to Venus on its C-terminal end and untagged) were transiently transformed under the control of the p35S promoter with Venus as a negative control in *Nicotiana benthamiana* (Tobacco) leaves by using *Agrobacterium tumefaciens* for the mediated gene transfer [Grefen et al., 2010]. The transient expression in planta has the

advantage that proteins and cofactors that might be needed for the activity of the specific CSLK/A or needed to synthesize for the synthesis and modelling of HM in general. Three days after the infiltration with the agrobacteria suspension, the fluorescence was examined and the leaf material was harvested for monosaccharide analysis of the destarched AIR.

Venus expression in Tobacco have shown a very strong fluorescent signal that was most likely uniformly pressed to the plasma membrane (Figure 36A). The fluorescent signal of *AtCSLA9*- and *PumCSLK1*-Venus was much lower than Venus but their expression pattern looks completely different in planta. While the fluorescence of *AtCSLA9* was visible as many puncta pressed to the plasma membrane, *PumCSLK1*-Venus have shown a more uniform fluorescent signal (Figure 36A). In case of Venus and *PumCSLK1*-Venus there were bigger fluorescent aggregates visible that were most likely caused by the overexpression (Figure 36A). But to evaluate the exact localization of CSLK/A in planta a plasmolysis with organelle markers for Golgi, plasma membrane and ER localization should be performed in future.

The leaf spots that were infiltrated with the agrobacterial suspension were harvested and the content of Glc and Man of the extracted and destarched AIR was determined. While *AtCSLA9*-Venus and *AtCSLA2* expression have led to an increase in Glc and Man compared to the Venus control in Tobacco leaves, for *PumCSLK1*(-Venus) only Man was increased after protein expression (Figure 36B), revealing a function in glucomannan synthesis for *AtCSLA2* and *AtCSLA9* whereas *PumCSLK1* synthesize pure mannan. Furthermore, the increase in Man of *PumCSLK1*-Venus expression was lower compared to *PumCSLK1* expression in Tobacco (Figure 36B) pointing out that the fusion to Venus impair the protein activity and synthesis in mannan. The negative influence of the Venus tag can also explain the very low Glc content in Tobacco leaves with *AtCSLA9*-Venus expression, which could only be detected once from destarched AIR material (data of other batches not shown). Because the destarched AIR contains much more polysaccharides than HMs, the material was treated with β -mannanase to evaluate the HM content that was produced by *AtCSLA2* and *PumCSLK1* in Tobacco. Indeed, *AtCSLA2* expressed in Tobacco have increased amounts in Glc and Man compared to the Venus control, meaning the protein synthesize glucomannans, whereas *PumCSLK1* expression have led only to higher Man content in the AIR and underlines their role in synthesizing pure mannans (Figure 36C).

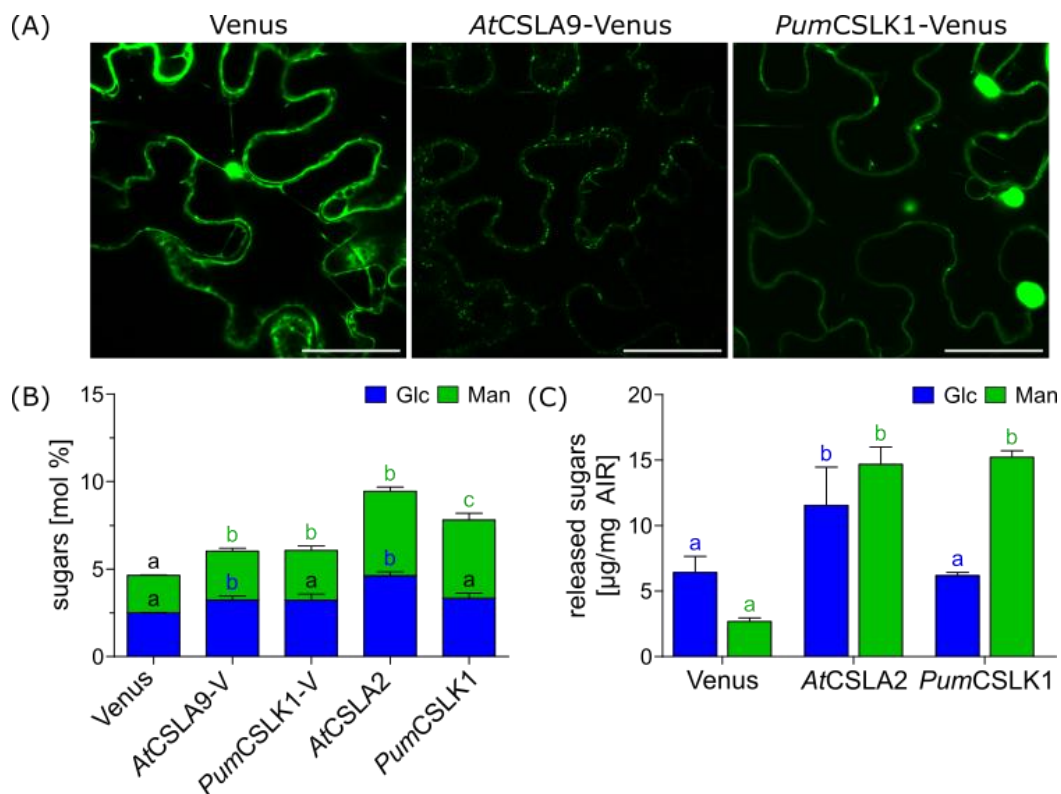


Figure 36: Transient overexpression of CSLK/A in *Nicotiana benthamiana* leaves. Five to six weeks old Tobacco leaves were infiltrated with *Agrobacterium tumefaciens* carrying the fluorescent protein Venus or CSLK/A (with or without Venus tagged to the C-terminal end of the respective CSLK/A). The fluorescence in the epidermal cell of Tobacco leaves was investigated using a confocal microscope (LSM700). Scale bar: 100 µm (A). Tobacco leaves were harvested and the AIR was extracted and destarched prior the monosaccharides of four biological replicates were determined (B). The destarched AIR was digested with β -mannanase to evaluate the glucose (Glc) and mannose (Man) content of each biological replicate (C). Letters denote for significant differences (one-way ANOVA with Tukey test; $p < 0.05$). *At*: *Arabidopsis thaliana*; *Pum*: *Porphyra umbilicalis*; V: Venus.

3.12 Characterization of *csla* mutants in *Arabidopsis* and moss

AkCSLA3 was stably transformed into *Arabidopsis thaliana* plants to evaluate its biological role in planta. Therefore, flowers of Col-0 and *csla239* mutants were transformed by dipping the flowers into *Agrobacterium tumefaciens* suspension carrying *AkCSLA3* or Venus driven by the Cauliflower mosaic virus p35S promoter. Transformed seeds were harvested, re-grown and selected by Basta resistance. Afterwards, the successful transformation was verified by PCR. Although all plants were positively genotyped for Venus or *AkCSLA3*, no fluorescence was detected in *Arabidopsis thaliana* plants carrying Venus (data not shown). The seed mucilage of Col-0 and *csla239* mutant plants non-transformed and transformed was visualized by staining the mucilage with 0.001% ruthenium red. *csla239* mutants have resulted in a more compact mucilage capsule compared to Col-0 (Figure 37A). Because *AkCSLA3* is involved in the synthesis of glucomannan, it is possible that *AkCSLA3* expression in *Arabidopsis thaliana* seeds can lead to more glucomannan incorporation in the mucilage of Col-0 or complement the mucilage defect of *csla239* mutants. Unfortunately, the expression of *AkCSLA3* have not resulted in any changes of the mucilage, neither in Col-0 nor in *csla239*

mutants (Figure 37A). Furthermore, the total mucilage of the four different lines was extracted and their Glc and Man content was determined. The Glc and Man content of *AkCSLA3* transgene lines was similar to the associated original Col-0 and *csla239* mucilages, meaning that the complementation of glucomannan in the mucilage was not successful (Figure 37B). The loss in mannan is not only visible in the seed mucilage of *csla239* mutants, it was also detected in *Arabidopsis* stems [Goubet et al., 2009]. Therefore, the monosaccharide composition of the destarched stem AIR was analysed in the next step. However, the *AkCSLA3* transgene was also not able to complement the lack in mannose of *csla239* mutants in the stem material (data not shown). Even though the *AkCSLA3* transgene was verified by PCR in the transformed plants, the protein level could not be sufficient *in planta* to compensate the lack of mannan in the mucilage and stem.

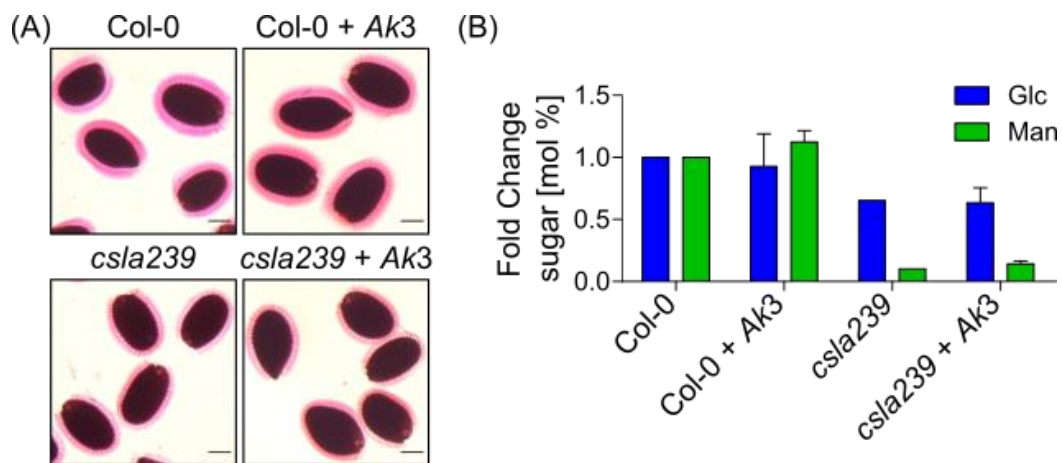


Figure 37: *AkCSLA3* expression in Col-0 and *csla239* mutant *Arabidopsis thaliana*. (A) Seed mucilage was stained with 0.001% ruthenium red. Scale bar: 100 μ m. (B) Fold change of glucose (Glc) and mannose (Man) content (mol %) of Col-0 and *csla239* transformed with *Ak3*. Data show the mean + SD of six biological replicates. *Ak3*: *Amorphophallus konjac* CSLA3.

As a simpler alternative to *Arabidopsis thaliana*, *Physcomitrium patens* which contain only three CSLA isoforms was used to evaluate their function *in planta*. In *Pichia*, all three isoforms of *Physcomitrium patens* CSLA produce glucomannan with different amounts of Man and Glc. Currently, not much is known about their biological role in planta and how the Glc:Man ratio that seem to be specific for *PpCSLA1–3* influences the architecture of the plant cell wall. To investigate the function of each CSLA isoform, single, double and triple mutants of *PpCSLA1–3* were created by our collaboration partners from the Bezanilla Lab in Dartmouth College. In this study, the Man content of the extracted and de-starched AIR from gametophores of *csla2-3csla3-1* double mutant and *csla1-1csla2-2csla3-2* triple mutant was analysed. *csla2-3csla3-1* have exhibited a gain in Man content compared to the WT, whereas *csla1-1csla2-2csla3-2* mutant gametophores have shown decreased (Figure 38). Of course, it has to be evaluated, if the deletion of 6 bp of CSLA2 in the double mutant and of CSLA1 in the triple mutant is enough to destroy their function. These intriguing genotypes can be examined in *Pichia* by introducing the same mutations into the specific CSLA and examining their

functionality in producing HMs. Further investigations of the moss *csla* single, double and triple mutants are necessary and of importance to discover their function in more detail.

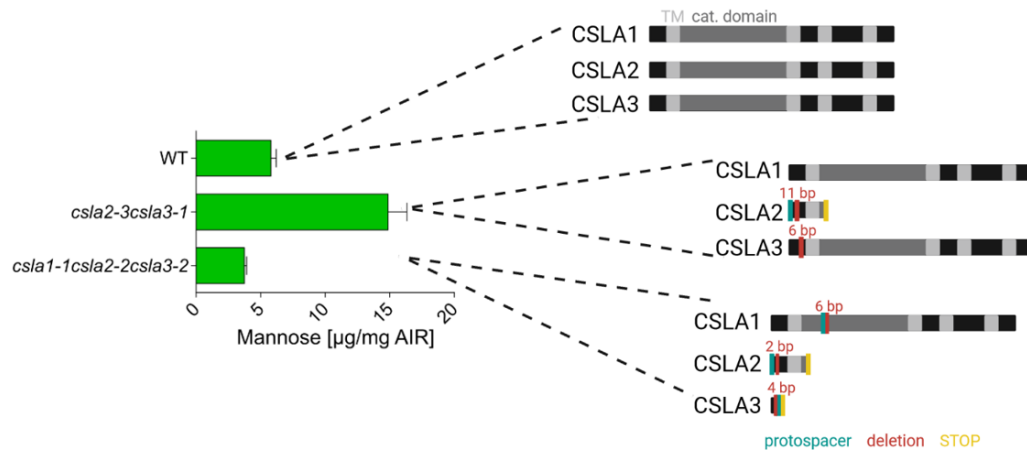


Figure 38: Influence of *csla* mutants of *Physcomitrium patens* mutants on Mannose abundance in gametophores. Mannose content of de-starched AIR material from *Physcomitrium patens* *csla* mutants was determined by HPAEC-PAD. Data show the mean + SD of two biological replicates. The schematic overview shows the position of the protospacer (cyan), the position and number of deleted bases (red) and the position where the protein translation has stopped (yellow).

4. Discussion

Although many genes encoding for GTs involved in plant cell wall polysaccharide synthesis have been identified through sequencing technologies and sequence similarities of these genes, the functions of the translated proteins are best-guesses and frequently lack biochemical evidence. As I extensively described in the introductory chapter 1.7, it is quite challenging to study the biochemical activity and biological role of the respective proteins directly in the plant system. Therefore, the synthetic approach to express candidate genes that are likely involved in HM synthesis in *Pichia* as a heterologous host can help to elucidate their function and specificities.

4.1 Elucidating the Mechanistic Function of AkCSLA3

The fusion of fluorescent proteins to CSLA candidates in *Pichia* cells cannot only help to select the best expressing colony by measuring their fluorescence as a high throughput screening, but additionally helps to uncover the proteins localization, trafficking, interaction with other proteins or degradation processes [Chudakow et al., 2009]. In this thesis, I successfully expressed fully active AkCSLA3 proteins with fluorescent proteins attached at the C-terminal end of the protein (Figure 9). This helped to overcome the issue of lower HM production of sfGFP-AkCSLA3 expression in *Pichia* compared to AkCSLA3 alone that was already shown by Voiniciuc et al. (2019) and was reproduced in this work (Figure 9). While AkCSLA3-Venus exhibited the same activity as AkCSLA3 without any tag, AkCSLA3-sfGFP even exceeded the glucomannan production of AkCSLA3 alone (Figure 9). Therefore, AkCSLA3-sfGFP was selected to characterize the localization, sugar specificity and topology of AkCSLA3 in more detail.

AkCSLA3-sfGFP is visible in small and puncta-like structures in *Pichia* cells (Figure 8C) consistent with Voiniciuc et al. (2019). This punctuate pattern of fluorescence is in agreement with other glycosyltransferases localized in the Golgi apparatus [Chou et al., 2014; Verhertbruggen et al., 2021]. The co-localization of co-expressed AkCSLA3-sfGFP with the Golgi marker ScGOS1-mRuby2 (Figure 14) strongly suggests that AkCSLA3 is localized in the Golgi apparatus. This finding can be further verified using the chemical Brefeldin A (BFA) which blocks the vesicle transport from the ER to the Golgi (conventional secretion pathway), leading to the disassembly of the Golgi apparatus. This would result in the diffuse fluorescence of both the Golgi marker ScGOS1-mRuby2 and AkCSLA3-sfGFP, as these proteins would be recruited back to the ER [Domozych, 1999; Fangel et al., 2011].

Furthermore, BFA can be employed as a tool to investigate whether AkCSLA3 is capable being transported through unconventional trafficking pathways directly to the plasma membrane. Unconventional protein secretion is characterized by bypassing the Golgi

apparatus [Ding et al., 2014]. This is noteworthy because the Golgi is typically involved in the modification and sorting of proteins. If *AkCSLA3* is transported via the unconventional secretion pathway, its localization would be (at least partially) unaffected by BFA. Understanding if *AkCSLA3* is using this alternative pathway is crucial, particularly in the context of the plant's ability to modulate cell wall composition in response to stress or environmental challenges. The proximity of *AkCSLA3* to the plasma membrane could be essential, as it positions the enzyme near the site where the synthesis and incorporation of polysaccharides into the cell wall occur. For this purpose, I tried two different approaches in this thesis to mis-localize *AkCSLA3* to the plasma membrane of *Pichia* cells and examine its activity in producing glucomannans: (i) fusion of N-terminal 33 AAs of Rat RGS4 on *AkCSLA3*-pmVenus and (ii) deletion of the N-terminal sequence and the first transmembrane-spanning domain (TMD1) of *AkCSLA3*-sfGFP. Unfortunately, both approaches were not able to mis-localize *AkCSLA3* to the plasma membrane in *Pichia*.

The formation of protein aggregates (Figure 16, 18 and 20C) suggests improper protein folding, potentially leading to functional issues (Figure 20B) or probably vacuolar degradation (Figure 18) [Heiss et al., 2015; Puxbaum et al., 2015]. The formation of aggregates by *AkCSLA3*-pmVenus (Figure 19), implies that the fluorescent tag could be problematic. *AkCSLA3*-pmVenus are likely degraded and have a vacuolar localization that could be verified by FM4-64 staining [Vida and Emr, 1995; Ito et al., 2022]. Nevertheless, RGS4:pmVenus still show a cytosolic localization (Figure 17), suggesting that the RGS4 sequence is not working the same way as it was shown in *Saccharomyces cerevisiae* [Srinavasa et al., 1998]. Whether pmVenus affect the properties of RGS4 and its final localization or if the cellular machinery of *Pichia* and *Saccharomyces* recognize RGS4 differently was not verified in this study, but could be tested by repeating the experiment with sfGFP and *AkCSLA3*-sfGFP.

Deleting the N-terminal sequence up to the first TMD of *AkCSLA3* resulted in a total loss of function in glucomannan synthesis within *Pichia* cells expressing *AkCSLA3*-sfGFP (w/o TMD1) (Figure 20B). Despite reduced fluorescence, the fluorescence remained localized in punctate structures. No signal peptide sequence is predicted in the N-terminal sequence of *AkCSLA3* which altogether contains five TMDs (Appendix Table S12). Even after deleting the first TMD, the remaining four TMDs might be sufficient to anchor *AkCSLA3*-sfGFP (w/o TMD1) to the Golgi apparatus which could explain the observed punctate localization. To confirm whether the protein indeed localizes within the Golgi, co-localization studies can be conducted using *AkCSLA3*-sfGFP (w/o TMD1) and *ScGOS1*-mRuby2 expressing *Pichia* cells. Overall, the N-terminal sequence of *AkCSLA3* appears to be crucial for glucomannan synthesis. The N-terminal tagged sfGFP (Figure 9C) and domain swaps with the N-terminal sequence of *AtCSLA2* [Robert et al., 2021] both seem to have a critical effect on the enzymes function. The

TMDs of CESA [Purushotham et al., 2020] and CSLC [Davis et al., 2010] proteins have been shown or are hypothesized to form a pore for translocating the sugar molecules. The N-terminal domain of CESA proteins seem to play a special role in this translocation process [Purushotham et al., 2020], why future investigations on the N-terminal sequence of *AkCSLA3* including the first TMD are especially important to analyse.

It is relevant to consider that the deletion of the N-terminal sequence of *AkCSLA3* might led to an altered orientation of *AkCSLA3* within the Golgi membrane, which could account for its non-functionality. The presence of an odd number of TMDs in *AkCSLA3* indicates that the N- and C-terminal ends of the protein are situated on opposite sides. Consequently, the catalytic active domain could be orientated either towards the cytosolic or luminal face of the Golgi. The orientation is critical as it determines whether the cytosolic or luminal nucleotide sugars are utilized for polysaccharide synthesis [Davis et al., 2010]. To investigate the orientation of *AkCSLA3*, I employed sfGFP-*AkCSLA3* and *AkCSLA3*-sfGFP constructs and conducted a FPP assay with extracted microsomes of both strains. The underlying hypothesis was that if the fluorescence of the protein facing the cytosolic side could be quenched through the application of proteinase K or trypsin, it would provide insights into the orientation. However, neither of these enzymes succeeded in diminishing the fluorescence of sfGFP when attached to either the C- or N-terminal end of *AkCSLA3*.

Although the determination of the topology of fluorescently tagged sfGFP on *AkCSLA3* microsomes in the Golgi was unsuccessful (Figure 17B), sfGFP seems to highly stabilize *AkCSLA3* because even after multiple times of thawing and freezing as well as treatment with proteinase K or trypsin (Figure 19), the fluorescence was visible in puncta. The higher stability of *AkCSLA3*-sfGFP might also be the reason why *AkCSLA3*-sfGFP exceeds the HM content compared to *AkCSLA3* expression alone (Figure 9C). This makes *AkCSLA3*-sfGFP the perfect candidate to perform many *in vitro* assays because the limiting factor for *in vitro* assays is generally known their high instability [Brown et al., 2012].

Amorphophallus konjac obtain glucomannans with a Glc:Man ratio of 1:1.4–1.8 *in planta* [Gille et al., 2011] and is used in various applications ranging from food industry to pharmaceuticals [Behera et al., 2016; Yang et al., 2017]. Glucomannan containing a higher molecular weight and higher ratio of Glc:Man was positively associated with antidiabetic effects (lower lipoproteins and sugar in the blood serum) in type 2 diabetic rats [Chen et al., 2021], possibly due to the positive influence on glycaemic response, lipid regulation and intestinal enzyme activity [Chen et al., 2021]. Therefore, it is quite of interest to modulate the glucomannan structure as it is needed for different purposes in the industry. To examine if the Glc:Man ratio of glucomannan produced by *AkCSLA3* can be altered and also to determine the specificity of *AkCSLA3* to synthesize a specific glucomannan structure two different

approaches were tested. In the first approach *Pichia* strains were co-fed with several carbon sources, while the other approach used a glycoengineered strain (SuperMan5) which reduce hypermannosylation and, therefore have in theory an excess of mannose in the cell that might be redirected to GDP-Man formation in the cell that could be built in into the glucomannan backbone. Both approaches were not able to change the glucose and mannan content in the produced glucomannan structure of *AkCSLA3* expressing *Pichia* strains (Figure 21B, 23B and 23D).

There is no confirmation that the supplementation of Man and Glc is effectively converted into GDP-Man or GDP-Glc, which could be utilized for glucomannan production. Employing radioactively labelled Glc or Man combined with nucleotide sugar extraction followed by separation through chromatography or mass spectrometry techniques could provide insights into whether these precursors contribute to the formation of activated nucleotides. However, an alternate possibility is that these precursors are converted into Fru6P, a key intermediate that serves as a branching point for biomass production and the pentose phosphate pathway in *Pichia* [Zepeda et al., 2018; Gündüz Ergün et al., 2022]. The observed increase in glucose concentration during the supplementation of glycerol, Glc and Man in Venus expressing *Pichia* strains (Figure 21A) indicates the incorporation of these carbon sources into glucans specific for *Pichia*. This observation can be attributed to *Pichia*'s preference to utilize carbon sources for biomass production initially, and resort to methanol utilization once these sources are depleted. It has been documented that *AkCSLA3* expression exhibits toxicity to *Pichia* cells [Robert et al., 2021]. The combination of carbon sources together with methanol appears to alleviate stress responses induced by the excessive accumulation of the recombinant protein [Jordà et al., 2013] as evidenced by the reduced abundance of glucomannan content in YPM-cultivated *Pichia AkCSLA3*-Venus cells.

Although the reduced hypermannosylation capacity of the SuperMan5 *Pichia* strain theoretically result in a higher availability of Man compared to X-33 *Pichia* strains, the expression of *AkCSLA3*-sfGFP led to reduced production of glucomannans without any alteration in the Glc:Man ratio (Figure 23). While the reduced hypermannosylation capacity of SuperMan5 might be expected to affect the glycosylation accuracy or pattern of *AkCSLA3*, the reduced expression (lower fluorescence but similar localization pattern; Figure 22A) and reduced viability of SuperMan5 (propidium iodide uptake in *Pichia* cells; Figure 22B) [Robert et al., 2021], are more likely to contribute to the observed reduction in glucomannan production. Furthermore, a higher abundance of Man availability does not necessarily mean that Man is redirected to GDP-Man and GDP-Man is incorporated in glucomannan. Therefore, whether the specificity of *AkCSLA3* for a specific glucomannan structure is highly determined by the enzyme or the change in Glc or Man content is not used for formation of activated sugar

nucleotides and glucomannan production. Further studies are needed to elucidate the underlying mechanism and the role of *AkCSLA3* specificity in glucomannan synthesis.

In order to analyse the biological role of *AkCSLA3* in planta, I stably transformed *AkCSLA3* under the control of the p35S promoter in *Arabidopsis thaliana csla239* mutant and Col-0 plants. Although, the transformation was successful, I was never able to detect fluorescence in the leaf material of transformed Venus plants, nor overexpression of *AkCSLA3* was able to complement the compact mucilage phenotype of *csla239* mutants (Figure 37A) or increase the glucomannan content. Nevertheless, my effort in seeing fluorescence for the Venus control in the leaf material of the plants and the compensation of the compact mucilage phenotype of *csla239* mutants (Figure 37) have been unsuccessful. Zhu et al (2003) identified an *Arabidopsis* mutant *rat4* where the T-DNA was inserted into the region of the 3'-untranslated region of *CSLA9* gene. This disruption led to a significant reduction in *CSLA9* transcripts in the mutant plant and result in resistance to *Agrobacterium tumefaciens* transformation due to the reduced numbers and growth rate of lateral roots. The alterations in the root surface probably affects the binding of *Agrobacterium* [Zhu et al., 2003]. *CSLA9* is expressed in various tissues (from the germinating seedling until the stage of flowering and developed siliques [Voiniciuc, 2022] the root, rather it expressed in most stages from the germinating seed until the flowers and siliques were developed. Therefore, the disruption of *CSLA9* can have similar effects on the structure and function of other tissues as well, why they are resistant to *Agrobacteria tumefaciens* in *csla239* mutant plants [Zhu et al., 2003].

4.2 Specificity of several CSLA isoforms of Arabidopsis

Although nine CSLA isoforms exist in *Arabidopsis thaliana*, only four isoforms (*AtCSLA2*, 3, 7 and 9) were identified as (gluco)mannan synthases involved in HM production [Liepman et al., 2005 and 2007; Goubet et al., 2009; Voiniciuc et al., 2019]. To investigate the function of these enzymes and their localization in *Pichia*, each was fused with a fluorescent protein. Although C-terminal tagged *AkCSLA3* was fully functional and actively produced glucomannans (Figure 9C), the activity of *AtCSLA* proteins was dramatically affected by the fluorescent tag. This effect was particularly visible for *AtCSLA2* which lost its mannan synthase activity (Figure 26C). The loss of activity can potentially be explained by mis-localization (visible as double-ring patterns indicating ER localization [Soderholm et al., 2004]). Mis-localization could occur due to improper protein folding which lead to steric hindrance of binding sites and impede appropriate interactions or trafficking [Heymann et al., 2015]. Therefore, fusion of proteins with fluorescent tags has to carefully tested for each protein on its own to determine their correct function and subcellular localization.

Nevertheless, *AtCSLA3*-, 7- and 9-Venus constructs were active in HM synthesis in *Pichia*. I was able to identify *AtCSLA9* as a glucomannan synthase, whereas *AtCSLA2*, 3 and 7 yielded in pure mannan production (Figure 26C, 28, 34). The production of mannan in *AtCSLA2* and *AtCSLA7* expressing *Pichia* cells was comparable to Voiniciuc et al. (2019), although *AtCSLA2* expression in *Drosophila* S2 cells [Liepman et al., 2005] or its function *in planta* [Goubet et al., 2009; Yu et al., 2014; Voiniciuc et al., 2015b] yielded in glucomannan. Voiniciuc et al. (2019) suggested that the difference between these organisms lay in their post-translational modification or accessory proteins. *AtMSR1* is such an accessory protein *in planta* and is able to enhance the HM production of *AkCSLA3* and *AtCSLA2* in *Pichia*, but also incorporate Glc into the mannan backbone of some CSLA (*AtCSLA2*, but not *AtCSLA7*). Therefore, MSR proteins modulate the activity of CSLAs [Voiniciuc et al., 2019].

AtCSLA7 is also suggested to be involved in glucomannan production *in planta*. One reason for this is that *AtCSLA7* overexpression led to an increased glucomannan content in stems. Additionally, the embryo lethal phenotype of *atcsla7* mutants can be complemented by *AtCSLA9* overexpression, indicating that the products are similar to each other [Goubet et al., 2009]. But up to date, heterologous expression of *AtCSLA7* only showed pure mannan production in insect [Liepman et al., 2007] or *Pichia* cells [Voiniciuc et al., 2019] and is comparable with my data. The fact that *AtMSR1* is not able to incorporate Glc into the mannan backbone produced by recombinant *AtCSLA7* in *Pichia* [Voiniciuc et al., 2019] indicate that other regulatory elements, key components or interaction partners than *AtMSR1* naturally occur *in planta* that are necessary for glucomannan synthesis of *AtCSLA7* and are missing in *Pichia*. To uncover the specific mechanism behind that more research regarding the enzyme specificities, structure and interacting proteins *in planta* and yeast is needed.

In general, CSLA share highly conserved sequences for nucleotide sugar binding [Yang et al., 2020] that can be found for most *AtCSLA* isoforms. Phylogenetically, *AtCSLA2* and *AtCSLA9* are part of the same clade, whereas *AtCSLA7* is part of another clade with all other *AtCSLA* isoforms [Liepman et al., 2007]. This, potentially indicate for diverged functions of those proteins. A glucose sugar coordinating motif DYHF occurs in *AtCSLA2*, *AtCSLA9* as well in *AkCSLA3*, whereas mostly all other *AtCSLA* proteins have NYHF (Appendix Figure S6). *AtCSLA7* is categorized with the other *AtCSLAs* in a second clade [Liepman et al., 2007] and has SYHF as a unique glucose sugar coordinating motif (Appendix Figure S6). The DYHF and SYHF motif could indicate for their ability to incorporate Glc into the mannan backbone or at least demonstrates the proteins phylogenetic difference to each other.

AtCSLA2 and *AtCSLA9*-Venus expressing *Pichia* cells showed the highest amounts of produced mannan or glucomannan (Figure 30B). This is comparable to the high expression levels of *AtCsla2* and *AtCsla9* genes in vegetative and reproductive tissues [Liepman et al.,

2007]. In *Pichia*, AtCSLA3- and AtCSLA7-Venus expression resulted in lower levels of pure mannan synthesis compared to AtCSLA2 and AtCSLA9 (Figure 28). The lower levels of mannan produced by AtCSLA3 and AtCSLA7 are comparable to the amount of gene expression in *Drosophila* expressing cells [Liepman et al., 2007] and immunofluorescence labelling of stem sections from several *csla* mutants in *Arabidopsis* [Goubet et al., 2009] and is comparable to AtCSLA7 expression in *Pichia* by Voiniciuc et al. (2019). Therefore, AtCSLA2 and AtCSLA9 are likely to be the main players in glucomannan synthesis.

All other AtCSLA isoforms (AtCSLA1, 10, 11, 14 and 15) failed to produce HMs in *Pichia*. The low expression and its non-functionality of AtCSLA1-Venus could potentially be explained by the fact that only a partial coding sequence missing 119 AAs (KJ138750.1) compared to the full-length cDNA (AT4G16590.1) was cloned in this study (Figure 25B and 25C). Liepman et al. (2005) expressed full length cDNA of AtCSLA1 in *Drosophila* S2 cells, and detected GDP-Glc and GDP-Man activity. Therefore, the full-length cDNA of AtCSLA1 has to be re-transformed in *Pichia* for future localization studies and *in vitro* assays.

AtCSLA10-, 11-, 14- and 15-Venus had no function in HM synthesis on its own. All of these constructs, except AtCSLA15, showed fluorescence. The fluorescence was visible in one to two punctate per cell (AtCSLA10 and AtCSLA11) or was visible as a bigger aggregate (AtCSLA14). To exclude that this effect was detected in the expressing *Pichia* strains by impairments of the fluorescent protein or in case of AtCSLA15 due to the attached contamination of the *E. coli* sequence, AtCSLA10, 11 and 15 were re-cloned and expressed without a fluorescent protein but still remain unfunctional (Appendix Figure S3; AtCSLA11 data not shown). Therefore, the non-functionality of AtCSLA10, 11 and 15 was not triggered by the fluorescent protein. The non-functionality of AtCSLA15 was surprising because compared to AtCSLA1, 11 and 14 it is the only protein of these isoforms containing all of the conserved motifs – and all of them were involved in coordination nucleotide sugar binding – similar to the functional AtCSLAs (Figure 29). In future experiments, the untagged AtCSLA14 still needs to be verified to determine if this recombinant protein remain unfunctional in (gluco)mannan synthesis.

Reasons why most of the AtCSLAs are not functional in *Pichia* could be that other proteins that naturally occur *in planta* as AtMSR1/2 are missing in *Pichia* and necessary for some CSLA to produce (gluco)mannans. AtMSR1 shows sequence similarity to human GDP-Fucose Protein Fucosyltransferase 1 (POFUT1) and even has motifs that are responsible for GDP-sugar binding [Voiniciuc et al., 2019]. Therefore, Voiniciuc et al. (2019) suggested that MSR proteins may glycosylate CSLA enzymes or engage in physical interaction by forming a complex has been suggested by Voiniciuc et al., (2019). The observation of Glc incorporation into the mannan backbone of AtCSLA2 upon co-expression with AtMSR1 suggest a potential

affinity of AtMSR1 for GDP-Glc binding. This specificity holds significant implications, particularly considering that, to date, no GDP-Glc transporter to the Golgi apparatus has been identified. It raises the possibility that AtMSR1 might play a role in this context. Further investigations are required to elucidate the specificities and mechanistic underpinnings of these interactions.

Additionally, it is not well understood why *Arabidopsis* but also other plant species as *Physcomitrium patens* have more than one isoform. One option might be that all of the CSLA members from *Arabidopsis* are involved in glucomannan synthesis, but expressed during several stages of development [Liepman et al., 2007; Goubet et al., 2009]. The other option is that CSLA can interact with each other by forming multimeric complexes as known for CESA isoforms and cellulose synthesis [McFarlane et al., 2014]. In this thesis, I co-expressed AtCSLA2 and AtCSLA9-Venus in *Pichia* (Figure 30). This resulted in a significant but only slight increase in glucomannan production. If these proteins would interact with each other a more tremendous amount of glucomannan would be expected as shown for the co-expression of AtCSLA2 with AtMSR1 [Voiniciuc et al., 2019]. As learned from this thesis, the use of fluorescent tags offers numerous advantages, but can cause sometimes complications. Consequently, it is within the realm of possibility that the attachment of the Venus tag to AtCSLA9 might create steric hindrance, impeding its ability to effectively bind to AtCSLA2. The observation that *csla239* triple mutants showed no detectable glucomannan in the stem of *Arabidopsis*, while a small amount of glucomannan was detectable in *csla2csla9* double mutants, could suggest that a complex is formed involving all three AtCSLA proteins. For further purposes, the expression level in different tissues as well as developmental stages of several CSLA isoforms as well as their structure predictions need to be examined in detail to select the most promising candidates for protein interaction.

4.3 Function of non-characterized CSLA and CSLK

Compared with *Arabidopsis*, other plant species encode for significant fewer CSLA or only single CSLA proteins. This reduces the complexity and make it much easier to examine the biological role and specificity of the respective CSLA *in planta*. The three isoforms of *Physcomitrium patens* CSLA produce all glucomannans, but with a different Glc and Man ratio in *Pichia* (Figure 33A), indicating that all three isoforms are highly specific for how much Glc and Man is built into the HM structure. The biological role of *PpCSLA1–3* can be elucidated in the *csla* single, double and triple mutants that were created with CRISPR Cas9. The first insights look promising. While the gametophores of *csla1-1csla2-2csla3-2* triple mutant resulted in less Man compared to the WT plant, the *csla2-3csla3-1* double mutant intriguingly increased Man (Figure 38). The deletion of only 6 bp in the CSLA1 protein of the triple mutant could be not enough to disrupt the whole function of the protein. For further analysis, the

mutations of the CRISPR Cas9 mutants can be introduced in the sequences of *Pichia* to analyse the effect of each specific mutation on its own. Thanks to their haploid-dominant life cycle, these *Physcomitrium patens* lines should enable us to overcome the limitations of *Arabidopsis csla* mutants, including genetic redundancy and embryo-lethality.

The recombinant expression of *Mpo*CSLA resulted in mostly pure mannan production with a very small amount of Glc incorporated into the backbone in *Pichia* (Figure 33B and 34) which is comparable with the high amounts of Man detected in the in the extracted and enriched cell wall of *Marchantia* but also to other bryophyte species [Happ and Classen, 2019; Kolkas et al., 2022]. Even though HMs of non-vascular plants (*P. patens*, *M. polymorpha*) have been assumed to be more abundant than in higher plants [Sørensen et al., 2011; Moller et al., 2007] in this study the Man content was only minimal (~3% for *P.patens* and ~10% for *M. polymorpha*; Figure 35C and 35D). This indicate that stainings with antibodies overestimate the abundance of glucomannans. *Atr*CSLA made glucomannan with only low amounts of incorporated Glc, whereas *Do*CSLA only showed increased Man content in *Pichia* cells (Figure 33B). All of the land plants contain a MSR genes included in the class GT106 [Wang et al., 2013], suggesting that Glc can be incorporated into the mannan backbones (especially for *Dendrobium* which exhibit acetylated glucomannan *in planta* [Xing et al., 2015])

During the evolution from aqueous to terrestrial environments, modifications of the cell wall and changes in their structure and polysaccharide composition were necessary for land plants to adapt to terrestrial environments. Heteromannans has been regarded as the most ancestral hemicelluloses. Red algae and chlorophytic green algae contain mannan in their cell wall but have no CSLA genes. The observation of a single copy CSL (later renamed as CSLK) was found first in chlorophytic green algae [Yin et al., 2009 and 2014] and later two CSLK proteins were also discovered in *Porphyra umbilicalis* [Brawley et al., 2017]. Their close relation to CSLA and CSLC proteins from land plants indicates that these proteins are responsible for mannan synthesis [Yin et al., 2009 and 2014; Brawley et al., 2017]. If this is the case, the CSLA gene family and mannan synthesis likely diverged in the last-common ancestor of rhodophytes and chlorophytes [Brawley et al., 2017]. The functional characterization of CSLK proteins will provide valuable insights into the diversity and complexity of cell wall polysaccharides and the specificity of the associated protein. The expression of the two red algal CSLK from *Porphyra umbilicalis* (*Pum*CSLK1 and *Pum*CSLK2) in *Pichia* resulted in the production of pure mannans (Figure 33B, 33C, and 34). To elucidate the structure of mannan within the plant, the AIR from two red algal NORI leaves, originating from different branches, was extracted and analysed. The PACE gel analysis yielded pure mannoooligosaccharides (Figure 35A), while the monosaccharide composition analysis revealed the presence of Man in the form of crystalline polymers (Figure 35B). This is akin to the crystalline β -1,4-mannan in the outer cell wall of the

skeletal component within the cell walls of blade cells [Brawley et al., 2017]. The transient transformation in *Nicotiana tumefaciens* leaves revealed that *PumCSLK1* is able to produce pure mannan in planta (Figure 36B). Interestingly, the localization pattern of the fluorescence of *AtCSLA9-Venus* (puncta-like structure) and *PumCSLK1-Venus* (continuous fluorescence) behaves differently (Figure 33A). The verification of the fluorescent localization of both proteins can give important insights about their functional mechanism. The punctate localization of *AtCSLA9* indicates for Golgi localization, whereas I would hypothesize that *PumCSLK1* is located at the plasma membrane.

While all of the CSLK containing green algae were hypothesized to function as mannan synthases, *Pichia* cells expressing CSLKs from chlorophytes (*Chlamydomonas reinhardtii*, *Micromonas pusilla*, *Ostreococcus lucimarinus*) as well as from charophytes (*Mesostigma viride* and *Chara braunii*) did not produce mannans (Figure 31 and 33). It's important to note that the CDS of *Chlamydomonas reinhardtii* and *Micromonas pusilla* was not codon optimized for *Pichia* (Table 4) which could potentially cause issues in the translation of the respective protein in *Pichia*. Because also the codon optimized algal CSLKs have no function in mannan synthesis on their own, it would be possible that CSLKs need other additional co-factors for their expression. But Wang et al. (2013) have shown that multicellular green algae have no homologous MSR genes, indicating that CSLKs are likely not involved in mannan synthesis, or could be essential for the biosynthesis of other glycan structures. Important to note is also that CSLDs were thought to make mannan. While the tested chlorophytes in this study had no CSLD genes, charophytic green algae have at least one or even several CSLD genes [Mikkelsen et al., 2014]. Recent evidence from the last two years [Yang et al., 2020], unequivocally demonstrates that this group of enzymes is involved in glucan-like structures. Therefore, it seems improbable that CSLD enzymes play a role in mannan synthesis within green algae. For a more comprehensive understanding of the function of both CSLDs and CSLKs in charophytic green algae, additional research on these enzymes is essential.

4.4 Conclusion

I have found many CSLA and even one CSLK candidate(s) which produced mannan or glucomannan when they were expressed in *Pichia*. The usage of fluorescent tags is very beneficial for analyzing the CSLK/A specificities, localization pattern, trafficking routes and also protein-protein co-localization or interaction. Unfortunately, some CSLA enzymes (ie. *AtCSLA2* and likely *AtCSLA14*) are very sensitive to their structural behaviour or stability and were misfolded or not able to locate right. Due to this bottleneck, most of the CSLK/A enzymes were characterized in *Pichia* without any fluorescent tag. Nevertheless, several CSLA enzymes were able to produce heteromannans with distinguishable amounts of Man and with or without Glc in *Pichia*. In future experiments, the distribution of Glc and Man in the backbone and the

heteromannans length remain to be determined. All these factors play an important role in the physicochemical properties of the formed heteromannans. In an industrial context, the importance of these enzymes extends beyond basic plant biology. There are several applications for which heteromannans can be used such as ingredients or thickener agents for food or cosmetics [López-Hortas et al., 2021], coating pharmaceuticals [Aziz et al., 2020] or as an additive to create biodegradable plastic [Harsojuwono et al., 2020]. All these applications depend on their HMs physicochemical properties. Additionally, the HM products produced by CSLK/A proteins can be re-designed and controlled through genetical modifications of the enzymes. This offers more possibilities in changing the polymers structure to the specific properties needed for the respective application. The capability to re-design and control the HM products open up an innovative platform for developing sustainable and biodegradable applications.

5. References

- Aanisah N, Wardhana YW, Chaerunisaa AY, Budiman A. Review on Modification of Glucomannan as an Excipient in Solid Dosage Forms. *Polymers*. 2022;14(13):2550. doi:10.3390/polym14132550
- Albersheim P, ed. *Plant Cell Walls: From Chemistry to Biology*. Garland Science; 2011.
- Amos RA, Mohnen D. Critical Review of Plant Cell Wall Matrix Polysaccharide Glycosyltransferase Activities Verified by Heterologous Protein Expression. *Front Plant Sci*. 2019;10:915. doi:10.3389/fpls.2019.00915
- Arnling Bååth J, Martínez-Abad A, Berglund J, Larsbrink J, Vilaplana F, Olsson L. Mannanase hydrolysis of spruce galactoglucomannan focusing on the influence of acetylation on enzymatic mannan degradation. *Biotechnol Biofuels*. 2018;11(1):114. doi:10.1186/s13068-018-1115-y
- Aspinall GO, Hirst EL, Percival EGV, Williamson IR. 635. The mannans of ivory nut (*Phytelephas macrocarpa*). Part I. The methylation of mannan A and mannan B. *J Chem Soc*. Published online 1953:3184. doi:10.1039/jr9530003184
- Bailey TL, Boden M, Buske FA, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*. 2009;37(Web Server):W202-W208. doi:10.1093/nar/gkp335
- Bakker H. Molecular cloning of two Arabidopsis UDP-galactose transporters by complementation of a deficient Chinese hamster ovary cell line. *Glycobiology*. 2004;15(2):193-201. doi:10.1093/glycob/cwh159
- Baldwin TC, Handford MG, Yuseff MI, Orellana A, Dupree P. Identification and Characterization of GONST1, a Golgi-Localized GDP-Mannose Transporter in Arabidopsis. *Plant Cell*. 2001;13(10):2283-2295. doi:10.1105/tpc.010247
- Bar-Peled M, O'Neill MA. Plant Nucleotide Sugar Formation, Interconversion, and Salvage by Sugar Recycling*. *Annu Rev Plant Biol*. 2011;62(1):127-155. doi:10.1146/annurev-arplant-042110-103918
- Basa S, Nampally M, Honorato T, et al. The Pattern of Acetylation Defines the Priming Activity of Chitosan Tetramers. *J Am Chem Soc*. 2020;142(4):1975-1986. doi:10.1021/jacs.9b11466
- Behera SS, Ray RC. Konjac glucomannan, a promising polysaccharide of *Amorphophallus konjac* K. Koch in health care. *International Journal of Biological Macromolecules*. 2016;92:942-956. doi:10.1016/j.ijbiomac.2016.07.098
- Berglund J, Angles d'Ortoli T, Vilaplana F, et al. A molecular dynamics study of the effect of glycosidic linkage type in the hemicellulose backbone on the molecular chain flexibility. *Plant J*. 2016;88(1):56-70. doi:10.1111/tpj.13259

- Berglund J, Azhar S, Lawoko M, et al. The structure of galactoglucomannan impacts the degradation under alkaline conditions. *Cellulose*. 2019;26(3):2155-2175. doi:10.1007/s10570-018-1737-z
- Bernal AJ, Yoo CM, Mutwil M, et al. Functional Analysis of the Cellulose Synthase-Like Genes CSLD1 , CSLD2 , and CSLD4 in Tip-Growing Arabidopsis Cells. *Plant Physiology*. 2008;148(3):1238-1253. doi:10.1104/pp.108.121939
- Berry EA, Tran ML, Dimos CS, Budziszek MJ, Scavuzzo-Duggan TR, Roberts AW. Immuno and Affinity Cytochemical Analysis of Cell Wall Composition in the Moss *Physcomitrella patens*. *Front Plant Sci*. 2016;7. doi:10.3389/fpls.2016.00248
- Bouzouita N, Khaldi A, Zgoulli S, et al. The analysis of crude and purified locust bean gum: A comparison of samples from different carob tree populations in Tunisia. *Food Chemistry*. 2007;101(4):1508-1515. doi:10.1016/j.foodchem.2006.03.056
- Brawley SH, Blouin NA, Ficko-Blean E, et al. Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proc Natl Acad Sci USA*. 2017;114(31). doi:10.1073/pnas.1703088114
- Bromley JR, Busse-Wicher M, Tryfona T, et al. GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J*. 2013;74(3):423-434. doi:10.1111/tpj.12135
- Brown C, Leijon F, Bulone V. Radiometric and spectrophotometric in vitro assays of glycosyltransferases involved in plant cell wall carbohydrate biosynthesis. *Nat Protoc*. 2012;7(9):1634-1650. doi:10.1038/nprot.2012.089
- Brown DM, Goubet F, Wong VW, et al. Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J*. 2007;52(6):1154-1168. doi:10.1111/j.1365-313X.2007.03307.x
- Broxterman SE, Schols HA. Interactions between pectin and cellulose in primary plant cell walls. *Carbohydrate Polymers*. 2018;192:263-272. doi:10.1016/j.carbpol.2018.03.070
- Broxterman SE, Schols HA. Characterisation of pectin-xylan complexes in tomato primary plant cell walls. *Carbohydrate Polymers*. 2018;197:269-276. doi:10.1016/j.carbpol.2018.06.003
- Brummer Y, Cui W, Wang Q. Extraction, purification and physicochemical characterization of fenugreek gum. *Food Hydrocolloids*. 2003;17(3):229-236. doi:10.1016/S0268-005X(02)00054-1
- Buckeridge MS. Seed Cell Wall Storage Polysaccharides: Models to Understand Cell Wall Biosynthesis and Degradation. *Plant Physiology*. 2010;154(3):1017-1023. doi:10.1104/pp.110.158642
- Burton RA, Fincher GB. (1,3;1,4)-beta-D-glucans in cell walls of the poaceae, lower plants, and fungi: a tale of two linkages. *Mol Plant*. 2009;2(5):873-882. doi:10.1093/mp/ssp063

- Burton RA, Wilson SM, Hrmova M, et al. Cellulose Synthase-Like CslF Genes Mediate the Synthesis of Cell Wall (1,3;1,4)- β -D-Glucans. *Science*. 2006;311(5769):1940-1942. doi:10.1126/science.1122975
- Campestrini LH, Silveira JLM, Duarte MER, Koop HS, Nosedá MD. NMR and rheological study of *Aloe barbadensis* partially acetylated glucomannan. *Carbohydrate Polymers*. 2013;94(1):511-519. doi:10.1016/j.carbpol.2013.01.020
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*. 2009;37(Database):D233-D238. doi:10.1093/nar/gkn663
- Capek P, Alföldi J, Lišková D. An acetylated galactoglucomannan from *Picea abies* L. Karst. *Carbohydrate Research*. 2002;337(11):1033-1037. doi:10.1016/S0008-6215(02)00090-3
- Cavalier DM, Lerouxel O, Neumetzler L, et al. Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell*. 2008;20(6):1519-1537. doi:10.1105/tpc.108.059873
- Cescutti P, Campa C, Delben F, Rizzo R. Structure of the oligomers obtained by enzymatic hydrolysis of the glucomannan produced by the plant *Amorphophallus konjac*. *Carbohydrate Research*. 2002;337(24):2505-2511. doi:10.1016/S0008-6215(02)00332-4
- Chanzy HD, Grosrenaud A, Vuong R, Mackie W. The crystalline polymorphism of mannan in plant cell walls and after recrystallisation. *Planta*. 1984;161(4):320-329. doi:10.1007/BF00398722
- Charnock SJ, Henrissat B, Davies GJ. Three-Dimensional Structures of UDP-Sugar Glycosyltransferases Illuminate the Biosynthesis of Plant Polysaccharides. *Plant Physiology*. 2001;125(2):527-531. doi:10.1104/pp.125.2.527
- Chebli Y, Geitmann A. Cellular growth in plants requires regulation of cell wall biochemistry. *Current Opinion in Cell Biology*. 2017;44:28-35. doi:10.1016/j.ceb.2017.01.002
- Chen H, Nie Q, Hu J, Huang X, Yin J, Nie S. Multiomics Approach to Explore the Amelioration Mechanisms of Glucomannans on the Metabolic Disorder of Type 2 Diabetic Rats. *J Agric Food Chem*. 2021;69(8):2632-2645. doi:10.1021/acs.jafc.0c07871
- Chokboribal J, Tachaboonyakiat W, Sangvanich P, Ruangpornvisuti V, Jettanacheawchankit S, Thunyakitpisal P. Deacetylation affects the physical properties and bioactivity of acemannan, an extracted polysaccharide from *Aloe vera*. *Carbohydrate Polymers*. 2015;133:556-566. doi:10.1016/j.carbpol.2015.07.039
- Chou YH, Pogorelko G, Zabolina OA. Xyloglucan Xylosyltransferases XXT1, XXT2, and XXT5 and the Glucan Synthase CSLC4 Form Golgi-Localized Multiprotein Complexes. *Plant Physiology*. 2012;159(4):1355-1366. doi:10.1104/pp.112.199356
- Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews*. 2010;90(3):1103-1163. doi:10.1152/physrev.00038.2009

- Ciancia M, Fernández PV, Leliaert F. Diversity of Sulfated Polysaccharides From Cell Walls of Coenocytic Green Algae and Their Structural Relationships in View of Green Algal Evolution. *Front Plant Sci.* 2020;11:554585. doi:10.3389/fpls.2020.554585
- Ciancia M, Quintana I, Vizcargüénaga MI, et al. Polysaccharides from the green seaweeds *Codium fragile* and *C. vermilara* with controversial effects on hemostasis. *International Journal of Biological Macromolecules.* 2007;41(5):641-649. doi:10.1016/j.ijbiomac.2007.08.007
- Cocuron JC, Lerouxel O, Drakakaki G, et al. A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase. *Proc Natl Acad Sci U S A.* 2007;104(20):8550-8555. doi:10.1073/pnas.0703133104
- Cosgrove DJ. Growth of the plant cell wall. *Nat Rev Mol Cell Biol.* 2005;6(11):850-861. doi:10.1038/nrm1746
- Cosson P, Perrin J, Bonifacino JS. Anchors aweigh: protein localization and transport mediated by transmembrane domains. *Trends in Cell Biology.* 2013;23(10):511-517. doi:10.1016/j.tcb.2013.05.005
- Cronshaw J, Myers A, Preston RD. A chemical and physical investigation of the cell walls of some marine algae. *Biochimica et Biophysica Acta.* 1958;27:89-103. doi:10.1016/0006-3002(58)90295-6
- Culbertson AT, Smith AL, Cook MD, Zabolina OA. Truncations of xyloglucan xylosyltransferase 2 provide insights into the roles of the N- and C-terminus. *Phytochemistry.* 2016;128:12-19. doi:10.1016/j.phytochem.2016.03.016
- Daas PJH, Grolle K, Van Vliet T, Schols HA, De Jongh HHJ. Toward the Recognition of Structure–Function Relationships in Galactomannans. *J Agric Food Chem.* 2002;50(15):4282-4289. doi:10.1021/jf011399t
- Daras G, Templalexis D, Avgeri F, Tsitsekian D, Karamanou K, Rigas S. Updating Insights into the Catalytic Domain Properties of Plant Cellulose synthase (CesA) and Cellulose synthase-like (Csl) Proteins. *Molecules.* 2021;26(14):4335. doi:10.3390/molecules26144335
- Darby RAJ, Cartwright SP, Dilworth MV, Bill RM. Which Yeast Species Shall I Choose? *Saccharomyces cerevisiae* Versus *Pichia pastoris* (Review). In: Bill RM, ed. *Recombinant Protein Production in Yeast.* Vol 866. *Methods in Molecular Biology.* Humana Press; 2012:11-23. doi:10.1007/978-1-61779-770-5_2
- Davis J, Brandizzi F, Liepman AH, Keegstra K. Arabidopsis mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane: Hemicellulosic glycan synthase topology. *The Plant Journal.* 2010;64(6):1028-1037. doi:10.1111/j.1365-313X.2010.04392.x
- Dea ICM, Morrison A. Chemistry and Interactions of Seed Galactomannans. In: *Advances in Carbohydrate Chemistry and Biochemistry.* Vol 31. Elsevier; 1975:241-312. doi:10.1016/S0065-2318(08)60298-X

- Dean GH, Zheng H, Tewari J, et al. The Arabidopsis MUM2 Gene Encodes a β -Galactosidase Required for the Production of Seed Coat Mucilage with Correct Hydration Properties. *The Plant Cell*. 2008;19(12):4007-4021. doi:10.1105/tpc.107.050609
- Dhugga KS, Barreiro R, Whitten B, et al. Guar seed beta-mannan synthase is a member of the cellulose synthase super gene family. *Science*. 2004;303(5656):363-366. doi:10.1126/science.1090908
- Dhugga KS, Barreiro R, Whitten B, et al. Guar Seed β -Mannan Synthase Is a Member of the Cellulose Synthase Super Gene Family. *Science*. 2004;303(5656):363-366. doi:10.1126/science.1090908
- Ding Y, Robinson DG, Jiang L. Unconventional protein secretion (UPS) pathways in plants. *Current Opinion in Cell Biology*. 2014;29:107-115. doi:10.1016/j.ceb.2014.05.008
- Doblin MS, Pettolino FA, Wilson SM, et al. A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-beta-D-glucan synthesis in transgenic Arabidopsis. *Proc Natl Acad Sci U S A*. 2009;106(14):5996-6001. doi:10.1073/pnas.0902019106
- Domozych DS, Ciancia M, Fangel JU, Mikkelsen MD, Ulvskov P, Willats WGT. The Cell Walls of Green Algae: A Journey through Evolution and Diversity. *Front Plant Sci*. 2012;3. doi:10.3389/fpls.2012.00082
- Du X, Li J, Chen J, Li B. Effect of degree of deacetylation on physicochemical and gelation properties of konjac glucomannan. *Food Research International*. 2012;46(1):270-278. doi:10.1016/j.foodres.2011.12.015
- Dwivany FM, Yulia D, Burton RA, et al. The CELLULOSE-SYNTHASE LIKE C (CSLC) family of barley includes members that are integral membrane proteins targeted to the plasma membrane. *Mol Plant*. 2009;2(5):1025-1039. doi:10.1093/mp/ssp064
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 2004;32(5):1792-1797. doi:10.1093/nar/gkh340
- Edwards ME, Dickson CA, Chengappa S, Sidebottom C, Gidley MJ, Reid JSG. Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *Plant J*. 1999;19(6):691-697. doi:10.1046/j.1365-313x.1999.00566.x
- Edwards ME, Marshall E, Gidley MJ, Reid JSG. Transfer Specificity of Detergent-Solubilized Fenugreek Galactomannan Galactosyltransferase. *Plant Physiology*. 2002;129(3):1391-1397. doi:10.1104/pp.002592
- Elbein AD. Biosynthesis of a Cell Wall Glucomannan in Mung Bean Seedlings. *Journal of Biological Chemistry*. 1969;244(6):1608-1616. doi:10.1016/S0021-9258(18)91803-X
- Ergün BG, Laçın K, Çaloğlu B, Binay B. Second generation *Pichia pastoris* strain and bioprocess designs. *Biotechnol Biofuels*. 2022;15(1):150. doi:10.1186/s13068-022-02234-7

- Faik A, Price NJ, Raikhel NV, Keegstra K. An Arabidopsis gene encoding an alpha-xylosyltransferase involved in xyloglucan biosynthesis. *Proc Natl Acad Sci U S A*. 2002;99(11):7797-7802. doi:10.1073/pnas.102644799
- Fangel JU, Petersen BL, Jensen NB, Willats WGT, Bacic A, Egelund J. A putative Arabidopsis thaliana glycosyltransferase, At4g01220, which is closely related to three plant cell wall-specific xylosyltransferases, is differentially expressed spatially and temporally. *Plant Science*. 2011;180(3):470-479. doi:10.1016/j.plantsci.2010.11.002
- Fernández PV, Arata PX, Ciancia M. Polysaccharides from Codium Species. In: *Advances in Botanical Research*. Vol 71. Elsevier; 2014:253-278. doi:10.1016/B978-0-12-408062-1.00009-3
- Fernández PV, Estevez JM, Cerezo AS, Ciancia M. Sulfated β -d-mannan from green seaweed Codium vermilara. *Carbohydrate Polymers*. 2012;87(1):916-919. doi:10.1016/j.carbpol.2011.06.063
- Fincher GB. Revolutionary Times in Our Understanding of Cell Wall Biosynthesis and Remodeling in the Grasses. *Plant Physiology*. 2009;149(1):27-37. doi:10.1104/pp.108.130096
- Frei, Eva, Preston, Reginald Dawson. Non-cellulosic structural polysaccharides in algal cell walls - II. Association of xylan and mannan in Porphyra umbilicalis. *Proc R Soc Lond B*. 1964;160(980):314-327. doi:10.1098/rspb.1964.0042
- Gadkari PV, Reaney MJT, Ghosh S. Assessment of gelation behaviour of fenugreek gum and other galactomannans by dynamic viscoelasticity, fractal analysis and temperature cycle. *International Journal of Biological Macromolecules*. 2019;126:337-344. doi:10.1016/j.ijbiomac.2018.12.132
- Gadkari PV, Tu S, Chiyarda K, Reaney MJT, Ghosh S. Rheological characterization of fenugreek gum and comparison with other galactomannans. *International Journal of Biological Macromolecules*. 2018;119:486-495. doi:10.1016/j.ijbiomac.2018.07.108
- Gao S, Nishinari K. Effect of Degree of Acetylation on Gelation of Konjac Glucomannan. *Biomacromolecules*. 2004;5(1):175-185. doi:10.1021/bm034302f
- Gellissen G. Production of recombinant proteins: novel microbial and eukaryotic expression systems. Wiley-VCH; 2005.
- Gendre D, Jonsson K, Boutté Y, Bhalerao RP. Journey to the cell surface—the central role of the trans-Golgi network in plants. *Protoplasma*. 2015;252(2):385-398. doi:10.1007/s00709-014-0693-1
- Gibeaut DM, Carpita NC. Synthesis of (1 \rightarrow 3), (1 \rightarrow 4)-beta-D-glucan in the Golgi apparatus of maize coleoptiles. *Proc Natl Acad Sci USA*. 1993;90(9):3850-3854. doi:10.1073/pnas.90.9.3850
- Gille S, Cheng K, Skinner ME, Liepman AH, Wilkerson CG, Pauly M. Deep sequencing of voodoo lily (Amorphophallus konjac): an approach to identify relevant genes involved in

- the synthesis of the hemicellulose glucomannan. *Planta*. 2011;234(3):515-526. doi:10.1007/s00425-011-1422-z
- Goubet F, Barton CJ, Mortimer JC, et al. Cell wall glucomannan in *Arabidopsis* is synthesised by CSLA glycosyltransferases, and influences the progression of embryogenesis. *The Plant Journal*. 2009;60(3):527-538. doi:10.1111/j.1365-313X.2009.03977.x
- Grefen C, Donald N, Hashimoto K, Kudla J, Schumacher K, Blatt MR. A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies: Fluorescence tagging and expression in *Arabidopsis*. *The Plant Journal*. 2010;64(2):355-365. doi:10.1111/j.1365-313X.2010.04322.x
- Grieß-Osowski A, Voiniciuc C. Branched mannan and xyloglucan as a dynamic duo in plant cell walls. *The Cell Surface*. 2023;9:100098. doi:10.1016/j.tcs.2023.100098
- Gu F, Bringmann M, Combs J, Yang J, Bergmann D, Nielsen E. The *Arabidopsis* CSLD5 functions in cell plate formation in a cell cycle dependent manner. *Plant Cell*. Published online June 27, 2016:tpc.00203.2016. doi:10.1105/tpc.16.00203
- Guo Q, Ai L, Cui S. *Methodology for Structural Analysis of Polysaccharides*. Springer International Publishing; 2018. doi:10.1007/978-3-319-96370-9
- Handford MG, Sicilia F, Brandizzi F, Chung JH, Dupree P. *Arabidopsis thaliana* expresses multiple Golgi-localised nucleotide-sugar transporters related to GONST1. *Mol Genet Genomics*. 2004;272(4):397-410. doi:10.1007/s00438-004-1071-z
- Handford M. Biosynthesis of plant cell walls. *Ciencia e Investigación Agraria*. 2006;33(3):149-166. doi:10.7764/rcia.v33i3.345
- Handford MG, Baldwin TC, Goubet F, et al. Localisation and characterisation of cell wall mannan polysaccharides in *Arabidopsis thaliana*. *Planta*. 2003;218(1):27-36. doi:10.1007/s00425-003-1073-9
- Handford M, Rodríguez-Furlán C, Marchant L, et al. *Arabidopsis thaliana* AtUTr7 Encodes a Golgi-Localized UDP-Glucose/UDP-Galactose Transporter that Affects Lateral Root Emergence. *Molecular Plant*. 2012;5(6):1263-1280. doi:10.1093/mp/sss074
- Hannuksela T, Hervé Du Penhoat C. NMR structural determination of dissolved O-acetylated galactoglucomannan isolated from spruce thermomechanical pulp. *Carbohydrate Research*. 2004;339(2):301-312. doi:10.1016/j.carres.2003.10.025
- Happ, Classen. Arabinogalactan-Proteins from the Liverwort *Marchantia polymorpha* L., a Member of a Basal Land Plant Lineage, Are Structurally Different to Those of Angiosperms. *Plants*. 2019;8(11):460. doi:10.3390/plants8110460
- Harholt J, Sørensen I, Fangel J, et al. The Glycosyltransferase Repertoire of the Spikemoss *Selaginella moellendorffii* and a Comparative Study of Its Cell Wall. Newbigin E, ed. *PLoS ONE*. 2012;7(5):e35846. doi:10.1371/journal.pone.0035846

- Harris PJ, Stone BA. Chemistry and Molecular Organization of Plant Cell Walls. In: Himmel ME, ed. Biomass Recalcitrance. Blackwell Publishing Ltd.; 2008:61-93. doi:10.1002/9781444305418.ch4
- Harris PJ, Stone BA. Chemistry and Molecular Organization of Plant Cell Walls. In: Himmel ME, ed. Biomass Recalcitrance. Blackwell Publishing Ltd.; 2008:61-93. doi:10.1002/9781444305418.ch4
- He C, Wu K, Zhang J, et al. Cytochemical Localization of Polysaccharides in *Dendrobium officinale* and the Involvement of DoCSLA6 in the Synthesis of Mannan Polysaccharides. *Front Plant Sci.* 2017;8. doi:10.3389/fpls.2017.00173
- Hemmingson JA, Nelson WA. [No title found]. *Journal of Applied Phycology.* 2002;14(5):357-364. doi:10.1023/A:1022110722705
- Heppert JK, Dickinson DJ, Pani AM, et al. Comparative assessment of fluorescent proteins for in vivo imaging in an animal model system. Strome S, ed. *MBoC.* 2016;27(22):3385-3394. doi:10.1091/mbc.e16-01-0063
- Heymann MC, Rabe S, Ruß S, et al. Fluorescent tags influence the enzymatic activity and subcellular localization of procaspase-1. *Clinical Immunology.* 2015;160(2):172-179. doi:10.1016/j.clim.2015.05.011
- Hoffmann N, King S, Samuels AL, McFarlane HE. Subcellular coordination of plant cell wall synthesis. *Developmental Cell.* 2021;56(7):933-948. doi:10.1016/j.devcel.2021.03.004
- Hohenblum H, Borth N, Mattanovich D. Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *Journal of Biotechnology.* 2003;102(3):281-290. doi:10.1016/S0168-1656(03)00049-X
- Hooks SB, Martemyanov K, Zachariou V. A role of RGS proteins in drug addiction. *Biochemical Pharmacology.* 2008;75(1):76-84. doi:10.1016/j.bcp.2007.07.045
- Houston K, Tucker MR, Chowdhury J, Shirley N, Little A. The Plant Cell Wall: A Complex and Dynamic Structure As Revealed by the Responses of Genes under Stress Conditions. *Front Plant Sci.* 2016;7. doi:10.3389/fpls.2016.00984
- Ishii T. O-acetylated oligosaccharides from pectins of potato tuber cell walls. *Plant Physiol.* 1997;113(4):1265-1272. doi:10.1104/pp.113.4.1265
- Ishii T. O-Acetylated Oligosaccharides from Pectins of Potato Tuber Cell Walls. *Plant Physiology.* 1997;113(4):1265-1272. doi:10.1104/pp.113.4.1265
- Ishrud O, Zahid M, Ahmad VU, Pan Y. Isolation and Structure Analysis of a Glucomannan from the Seeds of Libyan Dates. *J Agric Food Chem.* 2001;49(8):3772-3774. doi:10.1021/jf0103976
- Ito Y, Ishigami M, Hashiba N, et al. Avoiding entry into intracellular protein degradation pathways by signal mutations increases protein secretion in *Pichia pastoris*. *Microbial Biotechnology.* 2022;15(9):2364-2378. doi:10.1111/1751-7915.14061

- Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N. Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nat Protoc.* 2009;4(1):58-70. doi:10.1038/nprot.2008.213
- Jing B, Ishikawa T, Soltis N, et al. The *Arabidopsis thaliana* nucleotide sugar transporter GONST2 is a functional homolog of GONST1. *Plant Direct.* 2021;5(3). doi:10.1002/pld3.309
- Johnson WL, Straight AF. Fluorescent Protein Applications in Microscopy. In: *Methods in Cell Biology.* Vol 114. Elsevier; 2013:99-123. doi:10.1016/B978-0-12-407761-4.00005-1
- Johnson WL, Straight AF. Fluorescent Protein Applications in Microscopy. In: *Methods in Cell Biology.* Vol 114. Elsevier; 2013:99-123. doi:10.1016/B978-0-12-407761-4.00005-1
- Jordà J, De Jesus SS, Peltier S, Ferrer P, Albiol J. Metabolic flux analysis of recombinant *Pichia pastoris* growing on different glycerol/methanol mixtures by iterative fitting of NMR-derived ¹³C-labelling data from proteinogenic amino acids. *New Biotechnology.* 2014;31(1):120-132. doi:10.1016/j.nbt.2013.06.007
- Julian JD, Zabolina OA. Xyloglucan Biosynthesis: From Genes to Proteins and Their Functions. *Front Plant Sci.* 2022;13:920494. doi:10.3389/fpls.2022.920494
- Kang BH, Nielsen E, Preuss ML, Mastronarde D, Staehelin LA. Electron Tomography of RabA4b- and PI-4K β 1-Labeled Trans Golgi Network Compartments in *Arabidopsis*. *Traffic.* 2011;12(3):313-329. doi:10.1111/j.1600-0854.2010.01146.x
- Katuraya K, Okuyama K, Hatanaka K, Oshima R, Sato T, Matsuzaki K. Constitution of konjac glucomannan: chemical analysis and ¹³C NMR spectroscopy. *Carbohydrate Polymers.* 2003;53(2):183-189. doi:10.1016/S0144-8617(03)00039-0
- Keegstra K, Cavalier D. Glycosyltransferases of the GT34 and GT37 Families. In: Ulvskov P, ed. *Annual Plant Reviews.* Wiley-Blackwell; 2010:235-249. doi:10.1002/9781444391015.ch8
- Keegstra K, Raikhel N. Plant glycosyltransferases. *Current Opinion in Plant Biology.* 2001;4(3):219-224. doi:10.1016/S1369-5266(00)00164-3
- Kim SJ, Brandizzi F. The Plant Secretory Pathway: An Essential Factory for Building the Plant Cell Wall. *Plant and Cell Physiology.* 2014;55(4):687-693. doi:10.1093/pcp/pct197
- Kim SJ, Chandrasekar B, Rea AC, et al. The synthesis of xyloglucan, an abundant plant cell wall polysaccharide, requires CSLC function. *Proc Natl Acad Sci U S A.* 2020;117(33):20316-20324. doi:10.1073/pnas.2007245117
- Kong Z, Ioki M, Braybrook S, et al. Kinesin-4 Functions in Vesicular Transport on Cortical Microtubules and Regulates Cell Wall Mechanics during Cell Elongation in Plants. *Molecular Plant.* 2015;8(7):1011-1023. doi:10.1016/j.molp.2015.01.004

- Koroskenyi B, McCarthy SP. Synthesis of Acetylated Konjac Glucomannan and Effect of Degree of Acetylation on Water Absorbency. *Biomacromolecules*. 2001;2(3):824-826. doi:10.1021/bm010014c
- Kotake T, Hojo S, Yamaguchi D, Aohara T, Konishi T, Tsumuraya Y. Properties and Physiological Functions of UDP-Sugar Pyrophosphorylase in Arabidopsis. *Bioscience, Biotechnology, and Biochemistry*. 2007;71(3):761-771. doi:10.1271/bbb.60605
- Kraemer FJ, Lunde C, Koch M, et al. A mixed-linkage (1,3;1,4)- β -D-glucan specific hydrolase mediates dark-triggered degradation of this plant cell wall polysaccharide. *Plant Physiol*. 2021;185(4):1559-1573. doi:10.1093/plphys/kiab009
- Kumar M, Turner S. Plant cellulose synthesis: CESA proteins crossing kingdoms. *Phytochemistry*. 2015;112:91-99. doi:10.1016/j.phytochem.2014.07.009
- Kuravadi NA, Verma S, Pareek S, et al. Guar. In: *Agricultural Sustainability*. Elsevier; 2013:47-60. doi:10.1016/B978-0-12-404560-6.00003-4
- Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu Rev Biochem*. 2008;77(1):521-555. doi:10.1146/annurev.biochem.76.061005.092322
- Lee KJD, Sakata Y, Mau SL, et al. Arabinogalactan Proteins Are Required for Apical Cell Extension in the Moss *Physcomitrella patens*. *Plant Cell*. 2005;17(11):3051-3065. doi:10.1105/tpc.105.034413
- Levesque-Tremblay G, Pelloux J, Braybrook SA, Müller K. Tuning of pectin methylesterification: consequences for cell wall biomechanics and development. *Planta*. 2015;242(4):791-811. doi:10.1007/s00425-015-2358-5
- Liepman AH, Cavalier DM. The CELLULOSE SYNTHASE-LIKE A and CELLULOSE SYNTHASE-LIKE C families: recent advances and future perspectives. *Front Plant Sci*. 2012;3. doi:10.3389/fpls.2012.00109
- Liepman AH, Nairn CJ, Willats WGT, Sørensen I, Roberts AW, Keegstra K. Functional Genomic Analysis Supports Conservation of Function Among Cellulose Synthase-Like A Gene Family Members and Suggests Diverse Roles of Mannans in Plants. *Plant Physiology*. 2007;143(4):1881-1893. doi:10.1104/pp.106.093989
- Liepman AH, Wightman R, Geshi N, Turner SR, Scheller HV. Arabidopsis - a powerful model system for plant cell wall research. *The Plant Journal*. 2010;61(6):1107-1121. doi:10.1111/j.1365-313X.2010.04161.x
- Liepman AH, Wilkerson CG, Keegstra K. Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proc Natl Acad Sci USA*. 2005;102(6):2221-2226. doi:10.1073/pnas.0409179102
- Litterer LA, Schnurr JA, Plaisance KL, Storey KK, Gronwald JW, Somers DA. Characterization and expression of Arabidopsis UDP-sugar pyrophosphorylase. *Plant Physiology and Biochemistry*. 2006;44(4):171-180. doi:10.1016/j.plaphy.2006.04.004

- Little A, Schwerdt JG, Shirley NJ, et al. Revised Phylogeny of the Cellulose Synthase Gene Superfamily: Insights into Cell Wall Evolution. *Plant Physiol.* 2018;177(3):1124-1141. doi:10.1104/pp.17.01718
- Liu Q, Luo L, Zheng L. Lignins: Biosynthesis and Biological Functions in Plants. *IJMS.* 2018;19(2):335. doi:10.3390/ijms19020335
- López-Hortas L, Flórez-Fernández N, Torres MD, et al. Applying Seaweed Compounds in Cosmetics, Cosmeceuticals and Nutricosmetics. *Marine Drugs.* 2021;19(10):552. doi:10.3390/md19100552
- Lund CH, Bromley JR, Stenbæk A, Rasmussen RE, Scheller HV, Sakuragi Y. A reversible Renilla luciferase protein complementation assay for rapid identification of protein–protein interactions reveals the existence of an interaction network involved in xyloglucan biosynthesis in the plant Golgi apparatus. *Journal of Experimental Botany.* 2015;66(1):85-97. doi:10.1093/jxb/eru401
- Lundqvist J, Teleman A, Junell L, et al. Isolation and characterization of galactoglucomannan from spruce (*Picea abies*). *Carbohydrate Polymers.* 2002;48(1):29-39. doi:10.1016/S0144-8617(01)00210-7
- Mackie W, Sellen DB. The degree of polymerization and polydispersity of mannan from the cell wall of the green seaweed *Codium fragile*. *Polymer.* 1969;10:621-632. doi:10.1016/0032-3861(69)90082-2
- Mackie W, Preston RD. The occurrence of mannan microfibrils in the green algae *Codium fragile* and *Acetabularia crenulata*. *Planta.* 1968;79(3):249-253. doi:10.1007/BF00396031
- Magnelli P, Cipollo JF, Abeijon C. A Refined Method for the Determination of *Saccharomyces cerevisiae* Cell Wall Composition and β -1,6-Glucan Fine Structure. *Analytical Biochemistry.* 2002;301(1):136-150. doi:10.1006/abio.2001.5473
- Manabe Y, Verhertbruggen Y, Gille S, et al. Reduced Wall Acetylation Proteins Play Vital and Distinct Roles in Cell Wall O -Acetylation in *Arabidopsis*. *Plant Physiology.* 2013;163(3):1107-1117. doi:10.1104/pp.113.225193
- Matheson NK. Mannose-based Polysaccharides. In: *Methods in Plant Biochemistry.* Vol 2. Elsevier; 1990:371-413. doi:10.1016/B978-0-12-461012-5.50017-3
- Mathur NK. *Industrial Galactomannan Polysaccharides.* Taylor & Francis; 2012.
- McFarlane HE, Döring A, Persson S. The Cell Biology of Cellulose Synthesis. *Annu Rev Plant Biol.* 2014;65(1):69-94. doi:10.1146/annurev-arplant-050213-040240
- Meier H. On the structure of cell walls and cell wall mannans from ivory nuts and from dates. *Biochimica et Biophysica Acta.* 1958;28:229-240. doi:10.1016/0006-3002(58)90468-2
- Mikkelsen MD, Harholt J, Ulvskov P, et al. Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae. *Annals of Botany.* 2014;114(6):1217-1236. doi:10.1093/aob/mcu171

- Mikkonen KS, Tenkanen M. Sustainable food-packaging materials based on future biorefinery products: Xylans and mannans. *Trends in Food Science & Technology*. 2012;28(2):90-102. doi:10.1016/j.tifs.2012.06.012
- Minjares-Fuentes R, Femenia A, Comas-Serra F, Rodríguez-González VM. Compositional and Structural Features of the Main Bioactive Polysaccharides Present in the Aloe vera Plant. *Journal of AOAC INTERNATIONAL*. 2018;101(6):1711-1719. doi:10.5740/jaoacint.18-0119
- Mohnen D. Pectin structure and biosynthesis. *Curr Opin Plant Biol*. 2008;11(3):266-277. doi:10.1016/j.pbi.2008.03.006
- Moller I, Sørensen I, Bernal AJ, et al. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays: Glycan microarrays for plant cell-wall analysis. *The Plant Journal*. 2007;50(6):1118-1128. doi:10.1111/j.1365-313X.2007.03114.x
- Moreira LRS, Filho EXF. An overview of mannan structure and mannan-degrading enzyme systems. *Appl Microbiol Biotechnol*. 2008;79(2):165-178. doi:10.1007/s00253-008-1423-4
- Morrison IM. Polysaccharides: Plant Noncellulosic. In: John Wiley & Sons, Ltd, ed. ELS. 1st ed. Wiley; 2001. doi:10.1038/npg.els.0000696
- Mortimer JC, Yu X, Albrecht S, et al. Abnormal Glycosphingolipid Mannosylation Triggers Salicylic Acid-Mediated Responses in Arabidopsis. *The Plant Cell*. 2013;25(5):1881-1894. doi:10.1105/tpc.113.111500
- Mudgil D, Barak S, Khatkar BS. Guar gum: processing, properties and food applications—A Review. *J Food Sci Technol*. 2014;51(3):409-418. doi:10.1007/s13197-011-0522-x
- Munro S. Localization of proteins to the Golgi apparatus. *Trends in Cell Biology*. 1998;8(1):11-15. doi:10.1016/S0962-8924(97)01197-5
- Nagai T, Iyata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol*. 2002;20(1):87-90. doi:10.1038/nbt0102-87
- Nakano Y, Yamaguchi M, Endo H, Rejab NA, Ohtani M. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front Plant Sci*. 2015;6. doi:10.3389/fpls.2015.00288
- Nakayama K, Nagasu T, Shimma Y, Kuromitsu J, Jigami Y. OCH1 encodes a novel membrane bound mannosyltransferase: outer chain elongation of asparagine-linked oligosaccharides. *The EMBO Journal*. 1992;11(7):2511-2519. doi:10.1002/j.1460-2075.1992.tb05316.x
- Nakayama K ichi, Nakanishi-Shindo Y, Tanaka A, Haga-Toda Y, Jigami Y. Substrate specificity of α -1,6-mannosyltransferase that initiates N-linked mannose outer chain elongation in *Saccharomyces cerevisiae*. *FEBS Letters*. 1997;412(3):547-550. doi:10.1016/S0014-5793(97)00634-0

- Nguema-Ona E, VicrÃ©-Gibouin M, GottÃ© M, et al. Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. *Front Plant Sci.* 2014;5. doi:10.3389/fpls.2014.00499
- Nishinari K, Takemasa M, Zhang H, Takahashi R. Storage Plant Polysaccharides: Xyloglucans, Galactomannans, Glucomannans. In: *Comprehensive Glycoscience.* Elsevier; 2007:613-652. doi:10.1016/B978-044451967-2/00146-X
- Norambuena L, Marchant L, Berninsone P, Hirschberg CB, Silva H, Orellana A. Transport of UDP-galactose in Plants. *Journal of Biological Chemistry.* 2002;277(36):32923-32929. doi:10.1074/jbc.M204081200
- Norambuena L, Nilo R, Handford M, et al. AtUTr2 is an Arabidopsis thaliana nucleotide sugar transporter located in the Golgi apparatus capable of transporting UDP-galactose. *Planta.* 2005;222(3):521-529. doi:10.1007/s00425-005-1557-x
- Nunes FM, Coimbra MA. Chemical Characterization of Galactomannans and Arabinogalactans from Two Arabica Coffee Infusions As Affected by the Degree of Roast. *J Agric Food Chem.* 2002;50(6):1429-1434. doi:10.1021/jf0109625
- Nunes FM, Domingues MR, Coimbra MA. Arabinosyl and glucosyl residues as structural features of acetylated galactomannans from green and roasted coffee infusions. *Carbohydrate Research.* 2005;340(10):1689-1698. doi:10.1016/j.carres.2005.05.002
- Oikawa A, Lund CH, Sakuragi Y, Scheller HV. Golgi-localized enzyme complexes for plant cell wall biosynthesis. *Trends in Plant Science.* 2013;18(1):49-58. doi:10.1016/j.tplants.2012.07.002
- Pagny S, Bouissonnie F, Sarkar M, et al. Structural requirements for Arabidopsis beta1,2-xylosyltransferase activity and targeting to the Golgi. *Plant J.* 2003;33(1):189-203. doi:10.1046/j.0960-7412.2002.01604.x
- Painter TJ. Structural evolution of glycans in algae. *Pure and Applied Chemistry.* 1983;55(4):677-694. doi:10.1351/pac198855040677
- Pancaldi F, Van Loo EN, Schranz ME, Trindade LM. Genomic Architecture and Evolution of the Cellulose synthase Gene Superfamily as Revealed by Phylogenomic Analysis. *Front Plant Sci.* 2022;13:870818. doi:10.3389/fpls.2022.870818
- Park S, Szumlanski AL, Gu F, Guo F, Nielsen E. A role for CSLD3 during cell-wall synthesis in apical plasma membranes of tip-growing root-hair cells. *Nat Cell Biol.* 2011;13(8):973-980. doi:10.1038/ncb2294
- Parsons HT, Stevens TJ, McFarlane HE, et al. Separating Golgi Proteins from Cis to Trans Reveals Underlying Properties of Cisternal Localization. *Plant Cell.* 2019;31(9):2010-2034. doi:10.1105/tpc.19.00081
- Pauly M, Gawenda N, Wagner C, et al. The Suitability of Orthogonal Hosts to Study Plant Cell Wall Biosynthesis. *Plants.* 2019;8(11):516. doi:10.3390/plants8110516

- Pauly M, Gille S, Liu L, et al. Hemicellulose biosynthesis. *Planta*. 2013;238(4):627-642. doi:10.1007/s00425-013-1921-1
- Pauly M, Keegstra K. Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan. *Annu Rev Plant Biol*. 2016;67(1):235-259. doi:10.1146/annurev-arplant-043015-112222
- Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM. Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. *Proc Natl Acad Sci USA*. 1996;93(22):12637-12642. doi:10.1073/pnas.93.22.12637
- Peat S, Rees DA. Carbohydrase and sulphatase activities of *Porphyra umbilicalis*. *Biochemical Journal*. 1961;79(1):7-12. doi:10.1042/bj0790007
- Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*. 2006;24(1):79-88. doi:10.1038/nbt1172
- Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA. Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology*. 2008;18(11):891-904. doi:10.1093/glycob/cwn078
- Petrovic B, Pokluda R, Richard C, Fonteno K. Production cost analysis of organic and conventional vegetable seedlings. *JAH*. 2021;23(2):106-110. doi:10.37855/jah.2020.v22i03.32
- Pettolino FA, Walsh C, Fincher GB, Bacic A. Determining the polysaccharide composition of plant cell walls. *Nat Protoc*. 2012;7(9):1590-1607. doi:10.1038/nprot.2012.081
- Piirainen MA, Salminen H, Frey AD. Production of galactosylated complex-type N-glycans in glycoengineered *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol*. 2022;106(1):301-315. doi:10.1007/s00253-021-11727-8
- Piro G, Zuppa A, Dalessandro G, Northcote DH. Glucomannan synthesis in pea epicotyls: The mannose and glucose transferases. *Planta*. 1993;190(2). doi:10.1007/BF00196613
- Pitkänen L, Tuomainen P, Mikkonen KS, Tenkanen M. The effect of galactose side units and mannan chain length on the macromolecular characteristics of galactomannans. *Carbohydrate Polymers*. 2011;86(3):1230-1235. doi:10.1016/j.carbpol.2011.06.018
- Podzimek S. Characterization of Branched Polymers. In: *Light Scattering, Size Exclusion Chromatography and Asymmetric Flow Field Flow Fractionation*. John Wiley & Sons, Inc.; 2011:307-345. doi:10.1002/9780470877975.ch6
- Popper ZA, Michel G, Hervé C, et al. Evolution and Diversity of Plant Cell Walls: From Algae to Flowering Plants. *Annu Rev Plant Biol*. 2011;62(1):567-590. doi:10.1146/annurev-arplant-042110-103809

- Popper ZA, Tuohy MG. Beyond the Green: Understanding the Evolutionary Puzzle of Plant and Algal Cell Walls. *Plant Physiology*. 2010;153(2):373-383. doi:10.1104/pp.110.158055
- Poulsen CP, Dilokpimol A, Mouille G, Burow M, Geshi N. Arabinogalactan Glycosyltransferases Target to a Unique Subcellular Compartment That May Function in Unconventional Secretion in Plants. *Traffic*. 2014;15(11):1219-1234. doi:10.1111/tra.12203
- Prakash R, Johnston SL, Bolding HL, et al. Mannans in tomato fruit are not depolymerized during ripening despite the presence of endo- β -mannanase. *J Plant Physiol*. 2012;169(12):1125-1133. doi:10.1016/j.jplph.2012.03.017
- Prielhofer R, Barrero JJ, Steuer S, et al. GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. *BMC Syst Biol*. 2017;11(1):123. doi:10.1186/s12918-017-0492-3
- Prielhofer R, Cartwright SP, Graf AB, et al. *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. *BMC Genomics*. 2015;16(1):167. doi:10.1186/s12864-015-1393-8
- Purushotham P, Ho R, Zimmer J. Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science*. 2020;369(6507):1089-1094. doi:10.1126/science.abb2978
- Puxbaum V, Mattanovich D, Gasser B. Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. *Appl Microbiol Biotechnol*. 2015;99(7):2925-2938. doi:10.1007/s00253-015-6470-z
- Quezada MP, Salinas C, Gotteland M, Cardemil L. Acemannan and Fructans from *Aloe vera* (*Aloe barbadensis* Miller) Plants as Novel Prebiotics. *J Agric Food Chem*. 2017;65(46):10029-10039. doi:10.1021/acs.jafc.7b04100
- Rácz B, Blanpied TA, Ehlers MD, Weinberg RJ. Lateral organization of endocytic machinery in dendritic spines. *Nat Neurosci*. 2004;7(9):917-918. doi:10.1038/nn1303
- Ralet MC, Crépeau MJ, Vigouroux J, et al. Xylans Provide the Structural Driving Force for Mucilage Adhesion to the Arabidopsis Seed Coat. *Plant Physiol*. 2016;171(1):165-178. doi:10.1104/pp.16.00211
- Rautengarten C, Ebert B, Liu L, et al. The Arabidopsis Golgi-localized GDP-L-fucose transporter is required for plant development. *Nat Commun*. 2016;7(1):12119. doi:10.1038/ncomms12119
- Rautengarten C, Ebert B, Moreno I, et al. The Golgi localized bifunctional UDP-rhamnose/UDP-galactose transporter family of Arabidopsis. *Proc Natl Acad Sci USA*. 2014;111(31):11563-11568. doi:10.1073/pnas.1406073111
- Rayner JC, Munro S. Identification of the MNN2 and MNN5 Mannosyltransferases Required for Forming and Extending the Mannose Branches of the Outer Chain Mannans of

- Saccharomyces cerevisiae*. *Journal of Biological Chemistry*. 1998;273(41):26836-26843. doi:10.1074/jbc.273.41.26836
- Redgwell RJ, Curti D, Rogers J, Nicolas P, Fischer M. Changes to the galactose/mannose ratio in galactomannans during coffee bean (*Coffea arabica* L.) development: implications for in vivo modification of galactomannan synthesis. *Planta*. 2003;217(2):316-326. doi:10.1007/s00425-003-1003-x
- Rees D, Conway E. The structure and biosynthesis of porphyran: a comparison of some samples. *Biochemical Journal*. 1962;84(2):411-416. doi:10.1042/bj0840411
- Rennie EA, Scheller HV. Xylan biosynthesis. *Curr Opin Biotechnol*. 2014;26:100-107. doi:10.1016/j.copbio.2013.11.013
- Reyes F, León G, Donoso M, Brandizzí F, Weber APM, Orellana A. The nucleotide sugar transporters AtUTr1 and AtUTr3 are required for the incorporation of UDP-glucose into the endoplasmic reticulum, are essential for pollen development and are needed for embryo sac progress in *Arabidopsis thaliana*. *The Plant Journal*. 2010;61(3):423-435. doi:10.1111/j.1365-313X.2009.04066.x
- Reyes F, Marchant L, Norambuena L, Nilo R, Silva H, Orellana A. AtUTr1, a UDP-glucose/UDP-galactose Transporter from *Arabidopsis thaliana*, Is Located in the Endoplasmic Reticulum and Up-regulated by the Unfolded Protein Response. *Journal of Biological Chemistry*. 2006;281(14):9145-9151. doi:10.1074/jbc.M512210200
- Reynolds T, Dweck AC. Aloe vera leaf gel: a review update. *Journal of Ethnopharmacology*. 1999;68(1-3):3-37. doi:10.1016/S0378-8741(99)00085-9
- Richmond T. Higher plant cellulose synthases. *Genome Biol*. 2000;1(4):reviews3001.1. doi:10.1186/gb-2000-1-4-reviews3001
- Richmond TA, Somerville CR. The Cellulose Synthase Superfamily. *Plant Physiology*. 2000;124(2):495-498. doi:10.1104/pp.124.2.495
- Ridley BL, O'Neill MA, Mohnen D. Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry*. 2001;57(6):929-967. doi:10.1016/S0031-9422(01)00113-3
- Robert M, Waldhauer J, Stritt F, Yang B, Pauly M, Voiniciuc C. Modular biosynthesis of plant hemicellulose and its impact on yeast cells. *Biotechnol Biofuels*. 2021;14(1):140. doi:10.1186/s13068-021-01985-z
- Rollwitz I, Santaella M, Hille D, Flügge UI, Fischer K. Characterization of AtNST-KT1, a novel UDP-galactose transporter from *Arabidopsis thaliana*. *FEBS Letters*. 2006;580(17):4246-4251. doi:10.1016/j.febslet.2006.06.082
- Rossanese OW, Soderholm J, Bevis BJ, et al. Golgi Structure Correlates with Transitional Endoplasmic Reticulum Organization in *Pichia pastoris* and *Saccharomyces cerevisiae*. *Journal of Cell Biology*. 1999;145(1):69-81. doi:10.1083/jcb.145.1.69

- Saez-Aguayo S, Parra-Rojas JP, Sepúlveda-Orellana P, et al. Transport of UDP-rhamnose by URG2, URG4, and URG6 modulates rhamnogalacturonan-I length. *Plant Physiology*. 2021;185(3):914-933. doi:10.1093/plphys/kiaa070
- Salinas P, Salinas C, Contreras RA, Zuñiga GE, Dupree P, Cardemil L. Water deficit and abscisic acid treatments increase the expression of a glucomannan mannosyltransferase gene (GMMT) in *Aloe vera* Burm. F. *Phytochemistry*. 2019;159:90-101. doi:10.1016/j.phytochem.2018.12.009
- Saxena IM, Brown RM, Fevre M, Geremia RA, Henrissat B. Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. *J Bacteriol*. 1995;177(6):1419-1424. doi:10.1128/jb.177.6.1419-1424.1995
- Scheller HV, Ulvskov P. Hemicelluloses. *Annu Rev Plant Biol*. 2010;61:263-289. doi:10.1146/annurev-arplant-042809-112315
- Schröder R, Nicolas P, Vincent SJ, Fischer M, Reymond S, Redgwell RJ. Purification and characterisation of a galactoglucomannan from kiwifruit (*Actinidia deliciosa*). *Carbohydr Res*. 2001;331(3):291-306. doi:10.1016/s0008-6215(01)00046-5
- Schröder R, Atkinson RG, Redgwell RJ. Re-interpreting the role of endo- β -mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. *Annals of Botany*. 2009;104(2):197-204. doi:10.1093/aob/mcp120
- Sechet J, Htwe S, Urbanowicz B, et al. Suppression of *Arabidopsis* GGLT 1 affects growth by reducing the L-galactose content and borate cross-linking of rhamnogalacturonan- II. *Plant J*. 2018;96(5):1036-1050. doi:10.1111/tpj.14088
- Seymour GB, Colquhoun IJ, Dupont MS, Parsley KR, R. Selvendran R. Composition and structural features of cell wall polysaccharides from tomato fruits. *Phytochemistry*. 1990;29(3):725-731. doi:10.1016/0031-9422(90)80008-5
- Shi XD, Yin JY, Cui SW, Wang Q, Wang SY, Nie SP. Comparative study on glucomannans with different structural characteristics: Functional properties and intestinal production of short chain fatty acids. *International Journal of Biological Macromolecules*. 2020;164:826-835. doi:10.1016/j.ijbiomac.2020.07.186
- Shobha M, Vishukumar A, Tharanathan R, Koka R, Gaonkar A. Modification of guar galactomannan with the aid of pectinase. *Carbohydrate Polymers*. 2005;62(3):267-273. doi:10.1016/j.carbpol.2005.07.031
- Simões J, Nunes FM, Domingues P, Coimbra MA, Domingues MR. Mass spectrometry characterization of an *Aloe vera* mannan presenting immunostimulatory activity. *Carbohydrate Polymers*. 2012;90(1):229-236. doi:10.1016/j.carbpol.2012.05.029
- Sims IM, Craik DJ, Bacic A. Structural characterisation of galactoglucomannan secreted by suspension-cultured cells of *Nicotiana glauca*. *Carbohydrate Research*. 1997;303(1):79-92. doi:10.1016/S0008-6215(97)00144-4
- Sinclair R, Rosquete MR, Drakakaki G. Post-Golgi Trafficking and Transport of Cell Wall Components. *Front Plant Sci*. 2018;9:1784. doi:10.3389/fpls.2018.01784

- Singh S, Singh G, Arya SK. Mannans: An overview of properties and application in food products. *International Journal of Biological Macromolecules*. 2018;119:79-95. doi:10.1016/j.ijbiomac.2018.07.130
- Smith SM. Strategies for the Purification of Membrane Proteins. In: Walls D, Loughran ST, eds. *Protein Chromatography*. Vol 1485. *Methods in Molecular Biology*. Springer New York; 2017:389-400. doi:10.1007/978-1-4939-6412-3_21
- Soderholm J, Bhattacharyya D, Strongin D, et al. The Transitional ER Localization Mechanism of *Pichia pastoris* Sec12. *Developmental Cell*. 2004;6(5):649-659. doi:10.1016/S1534-5807(04)00129-7
- Somerville C, Bauer S, Brininstool G, et al. Toward a systems approach to understanding plant cell walls. *Science*. 2004;306(5705):2206-2211. doi:10.1126/science.1102765
- Song T, Pranovich A, Holmbom B. Separation of polymeric galactoglucomannans from hot-water extract of spruce wood. *Bioresource Technology*. 2013;130:198-203. doi:10.1016/j.biortech.2012.11.149
- Sørensen I, Pettolino FA, Bacic A, et al. The charophycean green algae provide insights into the early origins of plant cell walls: Cell-wall evolution and the Charophycean green algae. *The Plant Journal*. 2011;68(2):201-211. doi:10.1111/j.1365-313X.2011.04686.x
- Sowinski EE, Westman BM, Redmond CR, et al. Lack of xyloglucan in the cell walls of the *Arabidopsis* xxt1/xxt2 mutant results in specific increases in homogalacturonan and glucomannan. *The Plant Journal*. 2022;110(1):212-227. doi:10.1111/tpj.15666
- Srinivasa SP, Bernstein LS, Blumer KJ, Linder ME. Plasma membrane localization is required for RGS4 function in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA*. 1998;95(10):5584-5589. doi:10.1073/pnas.95.10.5584
- Suzuki S, Li L, Sun YH, Chiang VL. The Cellulose Synthase Gene Superfamily and Biochemical Functions of Xylem-Specific Cellulose Synthase-Like Genes in *Populus trichocarpa*. *Plant Physiology*. 2006;142(3):1233-1245. doi:10.1104/pp.106.086678
- Suzuki T, Fujikura K, Higashiyama T, Takata K. DNA Staining for Fluorescence and Laser Confocal Microscopy. *J Histochem Cytochem*. 1997;45(1):49-53. doi:10.1177/002215549704500107
- Tai-Nin Chow J, Williamson DA, Yates KM, Goux WJ. Chemical characterization of the immunomodulating polysaccharide of *Aloe vera* L. *Carbohydrate Research*. 2005;340(6):1131-1142. doi:10.1016/j.carres.2005.02.016
- Talmadge J, Chavez J, Jacobs L, et al. Fractionation of *Aloe vera* L. inner gel, purification and molecular profiling of activity. *International Immunopharmacology*. 2004;4(14):1757-1773. doi:10.1016/j.intimp.2004.07.013
- Tamura K, Shimada T, Kondo M, Nishimura M, Hara-Nishimura I. KATAMARI1/MURUS3 Is a Novel Golgi Membrane Protein That Is Required for Endomembrane Organization in *Arabidopsis*. *The Plant Cell*. 2005;17(6):1764-1776. doi:10.1105/tpc.105.031930

- Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Battistuzzi FU, ed. *Molecular Biology and Evolution*. 2021;38(7):3022-3027. doi:10.1093/molbev/msab120
- Temple H, Saez-Aguayo S, Reyes FC, Orellana A. The inside and outside: topological issues in plant cell wall biosynthesis and the roles of nucleotide sugar transporters. *Glycobiology*. 2016;26(9):913-925. doi:10.1093/glycob/cww054
- Terrett OM, Dupree P. Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Curr Opin Biotechnol*. 2019;56:97-104. doi:10.1016/j.copbio.2018.10.010
- Terrett OM, Lyczakowski JJ, Yu L, et al. Molecular architecture of softwood revealed by solid-state NMR. *Nat Commun*. 2019;10(1):4978. doi:10.1038/s41467-019-12979-9
- Toyooka K, Goto Y, Asatsuma S, Koizumi M, Mitsui T, Matsuoka K. A Mobile Secretory Vesicle Cluster Involved in Mass Transport from the Golgi to the Plant Cell Exterior. *The Plant Cell*. 2009;21(4):1212-1229. doi:10.1105/tpc.108.058933
- Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res*. 2015;43(W1):W401-W407. doi:10.1093/nar/gkv485
- Tu L, Banfield DK. Localization of Golgi-resident glycosyltransferases. *Cell Mol Life Sci*. 2010;67(1):29-41. doi:10.1007/s00018-009-0126-z
- Usov AI, Zelinsky ND. Chemical structures of algal polysaccharides. In: *Functional Ingredients from Algae for Foods and Nutraceuticals*. Elsevier; 2013:23-86. doi:10.1533/9780857098689.1.23
- Usov AI. Polysaccharides of the red algae. In: *Advances in Carbohydrate Chemistry and Biochemistry*. Vol 65. Elsevier; 2011:115-217. doi:10.1016/B978-0-12-385520-6.00004-2
- Valli M, Grillitsch K, Grünwald-Gruber C, et al. A subcellular proteome atlas of the yeast *Komagataella phaffii*. *FEMS Yeast Research*. 2020;20(1):foaa001. doi:10.1093/femsyr/foaa001
- Van De Meene AML, Doblin MS, Bacic A. The plant secretory pathway seen through the lens of the cell wall. *Protoplasma*. 2017;254(1):75-94. doi:10.1007/s00709-016-0952-4
- Varki A, Cummings RD, Esko JD, et al., eds. *Essentials of Glycobiology*. 4th ed. Cold Spring Harbor Laboratory Press; 2022. Accessed June 28, 2023. <http://www.ncbi.nlm.nih.gov/books/NBK579918/>
- Verbančič J, Lunn JE, Stitt M, Persson S. Carbon Supply and the Regulation of Cell Wall Synthesis. *Molecular Plant*. 2018;11(1):75-94. doi:10.1016/j.molp.2017.10.004
- Verherbruggen Y, Boudier A, Vigouroux J, et al. The TaCslA12 gene expressed in the wheat grain endosperm synthesizes wheat-like mannan when expressed in yeast and *Arabidopsis*. *Plant Science*. 2021;302:110693. doi:10.1016/j.plantsci.2020.110693

- Verherthbruggen Y, Yin L, Oikawa A, Scheller HV. Mannan synthase activity in the CSLD family. *Plant Signaling & Behavior*. 2011;6(10):1620-1623. doi:10.4161/psb.6.10.17989
- Vida TA, Emr SD. A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast. *Journal of Cell Biology*. 1995;128(5):779-792. doi:10.1083/jcb.128.5.779
- Vieira Gomes A, Souza Carmo T, Silva Carvalho L, Mendonça Bahia F, Parachin N. Comparison of Yeasts as Hosts for Recombinant Protein Production. *Microorganisms*. 2018;6(2):38. doi:10.3390/microorganisms6020038
- Voiniciuc C. Modern mannan: a hemicellulose's journey. *New Phytologist*. 2022;234(4):1175-1184. doi:10.1111/nph.18091
- Voiniciuc C, Dama M, Gawenda N, Stritt F, Pauly M. Mechanistic insights from plant heteromannan synthesis in yeast. *Proc Natl Acad Sci USA*. 2019;116(2):522-527. doi:10.1073/pnas.1814003116
- Voiniciuc C, Dean GH, Griffiths JS, et al. FLYING SAUCER1 Is a Transmembrane RING E3 Ubiquitin Ligase That Regulates the Degree of Pectin Methylesterification in Arabidopsis Seed Mucilage. *The Plant Cell*. 2013;25(3):944-959. doi:10.1105/tpc.112.107888
- Voiniciuc C, Guenl M, Schmidt MHW, Usadel B. Highly Branched Xylan Made by IRX14 and MUCI21 Links Mucilage to Arabidopsis Seeds. *Plant Physiol*. Published online October 19, 2015;pp.01441.2015. doi:10.1104/pp.15.01441
- Voiniciuc C, Yang B, Schmidt M, Günl M, Usadel B. Starting to Gel: How Arabidopsis Seed Coat Epidermal Cells Produce Specialized Secondary Cell Walls. *IJMS*. 2015;16(2):3452-3473. doi:10.3390/ijms16023452
- Von Freiesleben P, Spodsberg N, Stenbæk A, Stålbrand H, Krogh KBRM, Meyer AS. Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases. *Biotechnol Biofuels*. 2018;11(1):194. doi:10.1186/s13068-018-1184-y
- Wang P, Chen X, Goldbeck C, Chung E, Kang BH. A distinct class of vesicles derived from the trans -Golgi mediates secretion of xylogalacturonan in the root border cell. *Plant J*. 2017;92(4):596-610. doi:10.1111/tpj.13704
- Wang Y, Seemann J. Golgi Biogenesis. *Cold Spring Harbor Perspectives in Biology*. 2011;3(10):a005330-a005330. doi:10.1101/cshperspect.a005330
- Wang Y, Mortimer JC, Davis J, Dupree P, Keegstra K. Identification of an additional protein involved in mannan biosynthesis. *Plant J*. 2013;73(1):105-117. doi:10.1111/tpj.12019
- Weber APM, Linka N. Connecting the Plastid: Transporters of the Plastid Envelope and Their Role in Linking Plastidial with Cytosolic Metabolism. *Annu Rev Plant Biol*. 2011;62(1):53-77. doi:10.1146/annurev-arplant-042110-103903

- Weber A, Servaites JC, Geiger DR, et al. Identification, Purification, and Molecular Cloning of a Putative Plastidic Glucose Translocator. *Plant Cell*. 2000;12(5):787-801. doi:10.1105/tpc.12.5.787
- Welner DH, Shin D, Tomaleri GP, et al. Plant cell wall glycosyltransferases: High-throughput recombinant expression screening and general requirements for these challenging enzymes. Hofmann A, ed. *PLoS ONE*. 2017;12(6):e0177591. doi:10.1371/journal.pone.0177591
- Wilkop T, Pattathil S, Ren G, et al. A Hybrid Approach Enabling Large-Scale Glycomic Analysis of Post-Golgi Vesicles Reveals a Transport Route for Polysaccharides. *Plant Cell*. 2019;31(3):627-644. doi:10.1105/tpc.18.00854
- Willför S, Sjöholm R, Laine C, Roslund M, Hemming J, Holmbom B. Characterisation of water-soluble galactoglucomannans from Norway spruce wood and thermomechanical pulp. *Carbohydrate Polymers*. 2003;52(2):175-187. doi:10.1016/S0144-8617(02)00288-6
- Willför S, Sundberg A, Hemming J, Holmbom B. Polysaccharides in some industrially important softwood species. *Wood Sci Technol*. 2005;39(4):245-257. doi:10.1007/s00226-004-0280-2
- Wolf S, Hématy K, Höfte H. Growth Control and Cell Wall Signaling in Plants. *Annu Rev Plant Biol*. 2012;63(1):381-407. doi:10.1146/annurev-arplant-042811-105449
- Wolfrom ML, Laver ML, Patin DL. Carbohydrates of the Coffee Bean. II. Isolation and Characterization of a Mannan 1. *J Org Chem*. 1961;26(11):4533-4535. doi:10.1021/jo01069a080
- Worden N, Park E, Drakakaki G. Trans -Golgi Network-An Intersection of Trafficking Cell Wall Components. *Journal of Integrative Plant Biology*. Published online October 2012:no-no. doi:10.1111/j.1744-7909.2012.01179.x
- Xing X, Cui SW, Nie S, Phillips GO, Goff HD, Wang Q. Study on *Dendrobium officinale* O-acetyl-glucomannan (Dendronan®): Part II. Fine structures of O-acetylated residues. *Carbohydrate Polymers*. 2015;117:422-433. doi:10.1016/j.carbpol.2014.08.121
- Yang B, Hofmann F, Usadel B, Voiniciuc C. Seed hemicelluloses tailor mucilage properties and salt tolerance. *New Phytol*. 2021;229(4):1946-1954. doi:10.1111/nph.17056
- Yang J, Bak G, Burgin T, et al. Biochemical and Genetic Analysis Identify CSLD3 as a beta-1,4-Glucan Synthase That Functions during Plant Cell Wall Synthesis. *Plant Cell*. 2020;32(5):1749-1767. doi:10.1105/tpc.19.00637
- Ye ZH, Zhong R. Cell wall biology of the moss *Physcomitrium patens*. Degola F, ed. *Journal of Experimental Botany*. 2022;73(13):4440-4453. doi:10.1093/jxb/erac122
- Yeats T, Vellosillo T, Sorek N, Ibáñez A, Bauer S. Rapid Determination of Cellulose, Neutral Sugars, and Uronic Acids from Plant Cell Walls by One-step Two-step Hydrolysis and HPAEC-PAD. *BIO-PROTOCOL*. 2016;6(20). doi:10.21769/BioProtoc.1978

- Yin Y, Huang J, Gu X, Bar-Peled M, Xu Y. Evolution of Plant Nucleotide-Sugar Interconversion Enzymes. Moustafa A, ed. PLoS ONE. 2011;6(11):e27995. doi:10.1371/journal.pone.0027995
- Yin Y, Huang J, Xu Y. The cellulose synthase superfamily in fully sequenced plants and algae. BMC Plant Biol. 2009;9(1):99. doi:10.1186/1471-2229-9-99
- Yin Y, Johns MA, Cao H, Rupani M. A survey of plant and algal genomes and transcriptomes reveals new insights into the evolution and function of the cellulose synthase superfamily. BMC Genomics. 2014;15(1):260. doi:10.1186/1471-2164-15-260
- Youngs HL, Hamann T, Osborne E, Somerville C. The Cellulose Synthase Superfamily. In: Brown RM, Saxena IM, eds. Cellulose: Molecular and Structural Biology. Springer Netherlands; 2007:35-48. doi:10.1007/978-1-4020-5380-1_3
- Yu L, Lyczakowski JJ, Pereira CS, et al. The Patterned Structure of Galactoglucomannan Suggests It May Bind to Cellulose in Seed Mucilage. Plant Physiol. 2018;178(3):1011-1026. doi:10.1104/pp.18.00709
- Yu L, Yoshimi Y, Cresswell R, et al. Eudicot primary cell wall glucomannan is related in synthesis, structure, and function to xyloglucan. The Plant Cell. 2022;34(11):4600-4622. doi:10.1093/plcell/koac238
- Zabotina OA, Avci U, Cavalier D, et al. Mutations in multiple XXT genes of Arabidopsis reveal the complexity of xyloglucan biosynthesis. Plant Physiol. 2012;159(4):1367-1384. doi:10.1104/pp.112.198119
- Zabotina OA, Zhang N, Weerts R. Polysaccharide Biosynthesis: Glycosyltransferases and Their Complexes. Front Plant Sci. 2021;12:625307. doi:10.3389/fpls.2021.625307
- Zeng Y, Zhao S, Yang S, Ding SY. Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. Current Opinion in Biotechnology. 2014;27:38-45. doi:10.1016/j.copbio.2013.09.008
- Zepeda AB, Pessoa A, Farias JG. Carbon metabolism influenced for promoters and temperature used in the heterologous protein production using *Pichia pastoris* yeast. Brazilian Journal of Microbiology. 2018;49:119-127. doi:10.1016/j.bjm.2018.03.010
- Zhang GF, Staehelin LA. Functional Compartmentation of the Golgi Apparatus of Plant Cells: Immunocytochemical Analysis of High-Pressure Frozen- and Freeze-Substituted Sycamore Maple Suspension Culture Cells. Plant Physiol. 1992;99(3):1070-1083. doi:10.1104/pp.99.3.1070
- Zhong R, Cui D, Phillips DR, Richardson EA, Ye ZH. A Group of O-Acetyltransferases Catalyze Xyloglucan Backbone Acetylation and Can Alter Xyloglucan Xylosylation Pattern and Plant Growth When Expressed in Arabidopsis. Plant and Cell Physiology. 2020;61(6):1064-1079. doi:10.1093/pcp/pcaa031
- Zhong R, Cui D, Richardson EA, et al. Cytosolic Acetyl-CoA Generated by ATP-Citrate Lyase Is Essential for Acetylation of Cell Wall Polysaccharides. Plant and Cell Physiology. 2020;61(1):64-75. doi:10.1093/pcp/pcz178

- Zhong R, Cui D, Ye ZH. Members of the DUF231 Family are O-Acetyltransferases Catalyzing 2-O- and 3-O-Acetylation of Mannan. *Plant and Cell Physiology*. Published online August 9, 2018. doi:10.1093/pcp/pcy159
- Zhong R, Cui D, Ye ZH. Xyloglucan O-acetyltransferases from *Arabidopsis thaliana* and *Populus trichocarpa* catalyze acetylation of fucosylated galactose residues on xyloglucan side chains. *Planta*. 2018;248(5):1159-1171. doi:10.1007/s00425-018-2972-0
- Zhong R, Cui D, Ye Z. Secondary cell wall biosynthesis. *New Phytol*. 2019;221(4):1703-1723. doi:10.1111/nph.15537
- Zhong R, Cui D, Ye Z. Evolutionary origin of O -acetyltransferases responsible for glucomannan acetylation in land plants. *New Phytol*. 2019;224(1):466-479. doi:10.1111/nph.15988
- Zhu Y, Nam J, Carpita NC, Matthyse AG, Gelvin SB. Agrobacterium-Mediated Root Transformation Is Inhibited by Mutation of an Arabidopsis Cellulose Synthase-Like Gene. *Plant Physiology*. 2003;133(3):1000-1010. doi:10.1104/pp.103.030726

A. Appendices

Table S1: Overview of primers used for Phusion PCR from cDNA and cloning into the pPICZ-based vectors. Bold letters denote for the restriction sites of the enzyme and modified bases are labelled in red. Big letters show gene-specific bases of the sequence, while bases in small letters have been added to either facilitate the cutting of the restriction enzyme or to assemble the sequence in frame. Bases marked in blue denotes for a stop codon (*) at the end of the sequence.

Template (cDNA)	Phusion ID	Primer Target	Direction	Restriction Site	5' to 3' Sequence
<i>A. thaliana</i> siliques	P1	<i>AtCSLA1</i>	Forward	<i>Bsp119I</i>	gcc TTCGAA ATGTCACCATTTTTGAAGTTC
			Reverse	<i>Eco72I</i>	ttat CACGTG AGGTTCCGACAAAACCA
	P2	<i>AtCSLA3</i>	Forward	<i>Bsp119I</i>	gcc TTCGAA ATGTCACCATTTTTGAAGTTC
Reverse			<i>Eco72I</i>	tac CACGTG AACAGTTGGAACAAATGTGC	
<i>M. polymorpha</i> leaf	P3	<i>AtCSLA14</i>	Forward	<i>Bsp119I</i>	aca TTCGAA ATGGCGACATTGTCTGATG
			Reverse	<i>Eco72I</i>	tata CACGTG aCTCGGAAGTCTGGACCG
<i>P. patens</i> protonemata	P4	<i>MpoCSLA</i>	Forward	<i>Eco72I</i>	aca CACGTG ATGGCTCGTCTGGGGTCT
			Reverse	<i>Eco72I</i>	cct CACGTG aATGAGGTACAATAGTACCGA
<i>P. patens</i> protonemata	P5	<i>PpCSLA1*</i>	Forward	<i>Eco72I</i>	att CACGTG ATGGCGGAGAGCGGGA
			Reverse	<i>XhoI</i>	cgataa CTCGAGCTA AGAGTAGTGCGGAACGAATG
	P6	<i>PpCSLA2*</i>	Forward	<i>Bsp119I</i>	atc TTCGAA ATGGCGGAAAGCGCACGC
Reverse			<i>Eco72I</i>	taa CACGTGCTA CGAATAGTGGGGGACGAA	
P7	<i>PpCSLA3*</i>	Forward	<i>Bsp119I</i>	atc TTCGAA ATGGCGGATAGCTTCAGTGAA	
			Reverse	<i>Eco72I</i>	cga CACGTGCTA TGAATATTGTGGGACGAAAGTACC

Table S2: Overview of primers used for cloning into the pPICZ cloning system from plasmid DNA. Bold letters denote for the restriction sites of the enzyme and modified bases are labelled in red. Big letters show gene-specific bases of the sequence, while bases in small letters have been added to either facilitate the cutting of the restriction enzyme or to assemble the sequence in frame. Bases marked in blue denotes for a stop codon at the end of the sequence.

Template (plasmid DNA)	Phusion ID	Primer Target	Direction	Restriction Site	5' to 3' Sequence
pPICZB + <i>AkCSLA3</i> -2A-sfGFP	P8	sfGFP <i>AkCSLA3</i>	Forward	None	ATGCGTAAAGGCGAAGAGCTG
			Reverse	None	CTTTTCACTAGGAACAAAGGTGCCA
pPICZB + <i>AkCSLA3</i>	P9	<i>AkCSLA3</i>	Forward	<i>Bsp119I</i>	a TTCGAA acgatATGGCCATCGACTG
			Reverse	<i>Eco72I</i>	taa CACGTG aTGACTTTTCACTAGGAACAAAGGTG

pPICZB + <i>AtCSLA2</i>	P10	<i>AtCSLA2</i>	Forward Reverse	<i>Bsp119I</i> <i>Eco72I</i>	cca TTCGAA ATGGACGGTGTATCACC tgt CACGTG aACTCGGGACATAAGTCC
pPICZB + <i>AtCSLA7</i>	P11	<i>AtCSLA7</i>	Forward Reverse	<i>Eco72I</i> <i>Eco72I</i>	aag CACGTG ATGTCTCCTCTCCCAATCTTC tcc CACGTG aTGCTAAGTAAGAGGAAGCAG
pPICZB + <i>AtCSLA9</i>	P12	<i>AtCSLA9</i>	Forward Reverse	<i>Eco72I</i> <i>Eco72I</i>	tct CACGTG ATGGAGCTAGGAGATACGAC tca CACGTG AATGGTTAGGCACAATTGTC
pPICZB + <i>AtCSLA10</i>	P13	<i>AtCSLA10</i>	Forward Reverse	<i>Bsp119I</i> <i>Eco72I</i>	ggc TTCGAA ATGACTACGTTTCTGAAGTCACTC tat CACGTG AAGTGCCGACAAAGCCAA
pPICZB + <i>AtCSLA11</i>	P14	<i>AtCSLA11</i>	Forward Reverse	<i>Bsp119I</i> <i>Eco72I</i>	ctc TTCGAA ATGCAAGAGGACTTGGAGCT tgt CACGTG AAGTTCCTATAAAACCAATTCCGG
pPICZB + <i>AtCSLA15</i>	P15	<i>AtCSLA15</i>	Forward Reverse	<i>Eco72I</i> <i>Eco72I</i>	aca CACGTG ATGTTTTGCTTCTGAAGCC tgt CACGTG AAGTTCACAAAGCCAACT
pDONR221 + mCitrine (no STOP)	P16	mCitrine	Forward Reverse	<i>Eco72I</i> <i>XhoI</i>	att CACGTG tcATGGTGAGCAAGGGCGAG cag CTCGAG TTACTTGTACAGCTCGTCCATGC
pPICZX + mRuby2- <i>AtMSR1</i>	P17	mRuby2	Forward Reverse	<i>Eco72I</i> <i>XhoI</i>	caa CACGTG tcATGGTGTCACAAAGGAGAGGA gtg CTCGAG CTACTTATACAATTCATCCATACCACCG
pPICZB + Venus- <i>AtCSLA2</i>	P18	Venus	Forward Reverse	<i>Eco72I</i> <i>XhoI</i>	caa CACGTG tcATGTCTAAAGGTGAAGAATTATTCAGT G gac CTCGAG TCAATTTGTACAATTCATCCATACCATG
pPICZB + <i>AkCSLA3</i> - 2A-sfGFP	P19	2A-sfGFP	Forward Reverse	<i>Eco72I</i> <i>XhoI</i>	aag CACGTG tcGCTACTAACTTCTCTCTGCTCAAG agg CTCGAG CTAATTTGTACAGTTCATCCATACC
none	P20	Overlap extension primers for STOP cassette	Forward Reverse	<i>Eco72I</i> <i>XhoI</i>	cca CACGTG acTGAGTTTGTAGCCTTAGACATG cgt CTCGAG CATGTCTAAGGCTACAAACTCAGT

Table S3: Overview of pPICZ-based plasmids used for transformation into *Pichia X-33*. Vector and insert DNA were digested with restriction enzymes and sticky-ends of the vector and insert DNA were ligated.

Final plasmid	Vector DNA	Insert DNA (Phusion ID)	Resistance
pPICZB + <i>AkCSLA3</i> -sfGFP	pPICZB + <i>AkCSLA3</i> -2A-sfGFP	None (amplified (P8) & re-ligated)	Zeocin
pPICZX + mRuby2	pPICZX	mRuby2 (P17)	Zeocin
pPICZX + <i>AkCSLA3</i> -mRuby2	pPICZX + mRuby2	<i>AkCSLA3</i> (P9)	Zeocin
pPICZB + Venus	pPICZB	Venus (P18)	Zeocin
pPICZB + <i>AkCSLA3</i> -Venus	pPICZB + Venus	<i>AkCSLA3</i> (P9)	Zeocin
pPICZB + <i>AtCSLA1</i> -Venus	pPICZB + Venus	<i>AtCSLA1</i> (P1)	Zeocin
pPICZB + <i>AtCSLA2</i> -Venus	pPICZB + Venus	<i>AtCSLA2</i> (P10)	Zeocin
pPICZB + <i>AtCSLA3</i> -Venus	pPICZB + Venus	<i>AtCSLA3</i> (P2)	Zeocin
pPICZB + <i>AtCSLA7</i> -Venus	pPICZB + Venus	<i>AtCSLA7</i> (P11)	Zeocin
pPICZB + <i>AtCSLA9</i> -Venus	pPICZB + Venus	<i>AtCSLA9</i> (P12)	Zeocin
pPICZB + <i>AtCSLA10</i> -Venus	pPICZB + Venus	<i>AtCSLA10</i> (P13)	Zeocin
pPICZB + <i>AtCSLA11</i> -Venus	pPICZB + Venus	<i>AtCSLA11</i> (P14)	Zeocin
pPICZB + <i>AtCSLA14</i> -Venus	pPICZB + Venus	<i>AtCSLA14</i> (P3)	Zeocin
pPICZB + <i>AtCSLA15</i> -Venus	pPICZB + Venus	<i>AtCSLA15</i> (P15)	Zeocin
pPICZB + mCitrine	pPICZB	mCitrine (P16)	Zeocin
pPICZB + <i>AtCSLA2</i> -mCitrine	pPICZB + mCitrine	<i>AtCSLA2</i> (P10)	Zeocin
pPICZB + <i>AtCSLA3</i> -mCitrine	pPICZB + mCitrine	<i>AtCSLA3</i> (P2)	Zeocin
pPICZB + <i>AtCSLA14</i> -mCitrine	pPICZB + mCitrine	<i>AtCSLA14</i> (P3)	Zeocin
pPICZB + 2A-sfGFP	pPICZB	2A-sfGFP (P19)	Zeocin
pPICZB + <i>AtCSLA2</i> -2A-sfGFP	pPICZB + 2A-sfGFP	<i>AtCSLA2</i> (P10)	Zeocin
pPICZB + <i>AtCSLA3</i> -2A-sfGFP	pPICZB + 2A-sfGFP	<i>AtCSLA3</i> (P2)	Zeocin
pPICZB + <i>AtCSLA14</i> -2A-sfGFP	pPICZB + 2A-sfGFP	<i>AtCSLA14</i> (P3)	Zeocin
pPICZB (STOP cassette)	pPICZB	STOP cassette (P20)	Zeocin
pPICZB + <i>CrCSLK</i> (STOP)	pPICZB (STOP cassette)	<i>CrCSLK</i>	Zeocin
pPICZB + <i>MicpuCSLK</i> (STOP)	pPICZB (STOP cassette)	<i>MicpuCSLK</i>	Zeocin
pPICZB + <i>MpoCSLA</i> (STOP)	pPICZB (STOP cassette)	<i>MpoCSLA</i> (P4)	Zeocin
pPICZB + <i>SmCSLA</i> (STOP)	pPICZB (STOP cassette)	<i>SmCSLA</i>	Zeocin
pPICZB + <i>EhCSLK</i> *	pPICZB	<i>EhCSLK</i> *	Zeocin
pPICZB + <i>PumCSLK2</i> *	pPICZB	<i>PumCSLK2</i> *	Zeocin
pPICZB + <i>MvCSLK</i> *	pPICZB	<i>MvCSLK</i> *	Zeocin
pPICZB + <i>CbCSLK</i> *	pPICZB	<i>CbCSLK</i> *	Zeocin
pPICZB + <i>PpCSLA1</i> *	pPICZB	<i>PpCSLA1</i> * (P5)	Zeocin

pPICZB + <i>PpCSLA2</i> *	pPICZB	<i>PpCSLA2</i> * (P6)	Zeocin
pPICZB + <i>PpCSLA3</i> *	pPICZB	<i>PpCSLA3</i> * (P7)	Zeocin
pPICZB + <i>AtrCSLA</i> *	pPICZB	<i>AtrCSLA</i> *	Zeocin
pPICZB + <i>DoCSLA6</i> *	pPICZB	<i>DoCSLA6</i> *	Zeocin

Table S4: Overview of primers used for Phusion PCR from cDNA and cloning into the Golden PiCs vectors. Bold letters denote for the restriction sites of the enzyme and modified bases are labelled in red. Big letters show gene-specific bases of the sequence, while bases in small letters have been added to either facilitate the cutting of the restriction enzyme or to assemble the sequence in frame. Bases marked in blue denotes for a stop codon at the end of the sequence.

Template (cDNA)	Phusion ID	Primer Target	Description	Direction	Adapter	Fusion Site	Gene-specific Sequence
<i>S. cerevisiae</i> cell pellet	P21	ScGOS1	Fs2-ScGOS1	Forward	acga GGTCTC a	CATG	AGCTCACAACCGTCTTTTCG
			ScGOS1-custom (mRuby2)	Reverse	gctc GGTCTC a	CTCC	CCATGTGAAAAACAAAAACAG
	P22	ScOCH1	Fs2-ScOCH1	Forward	gcgc GGTCTC a	CATG	TCTAGGAAGTTGTCCCAC
			Domestication Site 1	Reverse	gcc GGTCTC g	ATCT	TCTTCAAATCTTGGATCAATC
P23	ScOCH1	Domestication Site 1	Forward	agcc GGTCTC g	AGA	TACAACGTAAACTATAGGCACAAAA GACGTCATGATGAAACCTAC	
		ScOCH1-custom (mRuby2)	Reverse	acgc GGTCTC t	CTCC	TTTATGACCTGCATTTTTTATCAGC	

Table S5: Overview of primers used for amplification and domestication (Dom) of transgenes into the Golden PiCs cloning system. Introduced restriction sites are marked in bold, introduced bases are labelled in red and * denote the native stop codon of the protein.

Template (plasmid DNA)	Phusion ID	Primer Target	Description	Direction	Adapter	Fusion Site	Gene-specific Sequence
pPICZX + ScMnn2-SpUGE1	P24	ScMnn2p	Fs2-ScMNN2p	Forward	atat GGTCTC a	CATG	CTGCTGACCAAGAGGTTTC
			ScMNN2p-custom (mRuby2)	Reverse	acca GGTCTC t	GGCT	CTACCTCTGTCCAGGTATTC
BB1_23 + mRuby2	P25	mRuby2	custom-mRuby2	Forward	acca GGTCTC t	AGCC	GTGTCCAAAGGAGAGGAG

BB1_12 + pAOX1	P26	pAOX1	mRuby2-Fs3	Reverse	acga GGTCTC g	AAGC	T TACTTATACAATTTCATCCATACCACC G
			Fs1-pAOX1	Forward	agcc GGTCTC a	GGAG	AGACGAAAGGTTGAATGAAAC
			pAOX1- custom (RGS4)	Reverse	gcgc GGTCTC c	CAAC	GTTTCGAATAATTAGTTG
pPICZB + AkCSLA3	P27		Fs2-AkCSLA3	Forward	atat GGTCTC a	CATG	GCCATCGACTGGGC
				Reverse	cgac GGTCTC t	ATAC	CTTCACGTACAGGCTCA
	P28		Domestication Site 1	Forward	ataa GGTCTC a	GTAT GTaT	TCCG TCGGCGGGCGCAAC CAA
				Reverse	cggc GGTCTC t	TTTC	ATATTTTATGTTACCCCATCC
	P29	AkCSLA3	Domestication Site 2	Forward	acat GGTCTC a	GAAA	CCAGGGCCAACAGAAATG
				Reverse	atca GGTCTC t	GTCC	GATCTCCGGTGTGGAT
	P30		Domestication Site 3	Forward	atat GGTCTC a	GGAC	TCGCTCAGGCTCGCTGGAA
			AkCSLA3-Fs3	Reverse	cgcg GGTCTC t	AAGC	CTACTTTTCACTAGGAACAAAGGTG
	P31		Domestication Site 3	Forward	atat GGTCTC a	GGAC	TCGCTCAGGCTCGCTGGAA
			AkCSLA3*-Fs3	Reverse	cgcg GGTCTC t	AAGC AAGC	CTACTTTTCACTAGGAACAAAGGT, GGT
BB1_23 + pmVenus	P32	pmVenus	Fs3-pmVenus	Forward	tcac GGTCTC a	GCTT	C TATGGTGTCCAAGGGAGAAG
			pmVenus- custom (RPP1Btt)	Reverse	gcgc GGTCTC a	CTAT	TTATTTATAGAGTTCGTCCATGCC
BB1_34 + RPP1Btt	P33	RPP1Btt	custom- RPP1Btt	Forward	gctc GGTCTC g	ATAG	GGATGATACTTTAATTGATGCATGTG
			RPP1Btt-Fs4	Reverse	tcga GGTCTC c	AGCG	GCTCCTCTAGATACCATCAAG
BB3aZ_14 + pAOX1:AkCSL	P34	AkCSLA3	Fs2-AkCSLA3 (w/o TMD1)	Forward	acta GGTCTC a	CATG	ATGGGTGGTGTGAGCCTGTAC
			AkCSLA3-Fs3	Reverse	cgcg GGTCTC t	AAGC	CTTTTCACTAGGAACAAAGGTG

A3:sfGFP-
RPP1Btt

E3 + sfGFP	P35	sfGFP	Fs3-sfGFP- sfGFP-custom (RPP1Btt)	Forward Reverse	atat GGTCTC a taat GGTCTC a	GCTT CTAT	ccCGTAAAGGCGAAGAGCTG TTGTACAGTTCATCCATACCATGC
BB1_34 + RPP1Btt	P36	RPP1Btt	Custom- RPP1Btt	Forward Reverse	atat GGTCTC a taat GGTCTC a	ATAG AGCG	GGATGATACTTTAATTGATGC GCTCCTCTAGATACCATCAAGAC
pPICZB + AtCSLA2	P37	AtCSLA2	Fs2-AtCSLA2	Forward	acat GGTCTC a	CATG	GACGGTGTATCACCAAAGT
			Domestication Site 1	Reverse	atat GGTCTC t	GAGA	TGCTCGGACGGCGAGAT
	P38	AtCSLA2*-Fs3	Forward	ataa GGTCTC a	TCTC	TTCGCGGCTGGAAAT	
pPICZB + AtCSLA9	P39	AtCSLA9	Fs2-AtCSLA9	Forward	atca GGTCTC	CATG	GAGCTAGGAGATACGACG
			AtCSLA9*-Fs3	Reverse	atca GGTCTC t	AAGC	GTG <u>TCA</u> ATGGTTAGGCACAATT
pPICZB + AtCSLA11	P40	AtCSLA11	Fs2-AtCSLA11	Forward	atca GGTCTC a	CATG	CAAGAGGACTTGGAGCT
			Domestication Site 1	Reverse	atca GGTCTC a	ACCA	AAGACAAGGTTATATGATGC
	P41	AtCSLA11*- Fs3	Forward	gcc GGTCTC t	TGGT	AAGACAGTGTTATATATC	
pPICZB + AtCSLA14	P42	AtCSLA14	Fs2-AtCSLA14	Forward	atca GGTCTC a	CATG	GCGACATTGTCTGATGG
			AtCSLA14*- Fs3	Reverse	atca GGTCTC t	AAGC	TTACTCGGAAGTCTGGACC

Table S6: Overview of Golden Gate Assemblies and created plasmids used for transformation into *Pichia X-33*. * denote the native stop codon of the protein.

Final plasmid	Vector DNA	Insert DNA (Phusion ID)	Resistance
BB3rN_14 + pGAP:mRuby2:ScCYC1tt	BB3rN_14	BB1_12 + pGAP BB1_23 + mRuby2 BB1_34 + ScCYC1tt	Nourseothricin
BB1_23 + ScMnn2p-mRuby2	BB1_23	Fs2-ScMnn2p-custom (mRuby2) (P24) Custom-mRuby2-Fs3 (P25)	Kanamycin
BB3rN_14 + pGAP:ScMnn2p-mRuby2:ScCYC1tt	BB3rN_14	BB1_12 + pGAP BB1_23 + ScMnn2p-mRuby2 BB1_34 + ScCYC1tt	Nourseothricin
BB1_23 + ScGOS1-mRuby2	BB1_23	Fs2-ScGOS1-custom (mRuby2) (P21) Custom-mRuby2-Fs3 (P25)	Kanamycin
BB3rN_14 + pGAP:ScGOS1-mRuby2:ScCYC1tt	BB3rN_14	BB1_12 + pGAP BB1_23 + ScGOS1-mRuby2 BB1_34 + ScCYC1tt	Nourseothricin
BB1_23 + ScOCH1-mRuby2	BB1_23	Fs2-ScOCH1-dom1 (P22) Dom1-ScOCH1-custom (mRuby2) (P23) Custom-mRuby2-Fs3 (P25)	Kanamycin
BB3rN_14 + pGAP:ScOCH1-mRuby2:ScCYC1tt	BB3rN_14	BB1_12 + pGAP BB1_23 + ScOCH1-mRuby2 BB1_34 + ScCYC1tt	Nourseothricin
BB1_12 + pAOX1-RGS4	BB1_12	Fs1-pAOX1-custom (RGS4) (P26) Custom-RGS4-Fs2 (Geneart digested)	Kanamycin
BB1_23 + AkCSLA3*	BB1_23	Fs2-AkCSLA3-Dom1 (P27) Dom1-AkCSLA3-Dom2 (P28) Dom2-AkCSLA3-dom3 (P29) Dom3-AkCSLA3*-Fs3 (P31)	Kanamycin
BB1_23 + AkCSLA3	BB1_23	Fs2-AkCSLA3-Dom1 (P27) Dom1-AkCSLA3-Dom2 (P28) Dom2-AkCSLA3-Dom3 (P29) Dom3-AkCSLA3-Fs3 (P30)	Kanamycin

BB3aZ_14 + pAOX1: <i>AkCSLA3</i> *:RPP1Btt	BB3aZ_14	BB1_12 + pAOX1 BB1_23 + <i>AkCSLA3</i> * BB1_34 + RPP1Btt	Zeocin
BB1_34 + pmVenus-RPP1Btt	BB1_34	Fs3-pmVenus-custom (RPP1Btt) (P32) Custom-RPP1Btt-Fs4 (P33)	Kanamycin
BB3aZ_14 + pAOX1:pmVenus:RPP1Btt	BB3aZ_14	BB1_12 + pAOX1-RGS4 BB1_23 + pmVenus BB1_34 + RPP1Btt	Zeocin
BB3aZ_14 + pAOX1-RGS4:pmVenus:RPP1Btt	BB3aZ_14	BB1_12 + pAOX1-RGS4 BB1_23 + pmVenus BB1_34 + RPP1Btt	Zeocin
BB3aZ_14 + pAOX1: <i>AkCSLA3</i> :pmVenus-RPP1Btt	BB3aZ_14	BB1_12 + pAOX1 BB1_23 + <i>AkCSLA3</i> BB1_34 + pmVenus-RPP1Btt	Zeocin
BB3aZ_14 + pAOX1-RGS4: <i>AkCSLA3</i> :pmVenus-RPP1Btt	BB3aZ_14	BB1_12 + pAOX1-RGS4 BB1_23 + <i>AkCSLA3</i> BB1_34 + pmVenus-RPP1Btt	Zeocin
BB1_23 + <i>AkCSLA3</i> (w/oTMD1)	BB1_23	Fs2- <i>AkCSLA3</i> (w/oTMD1)-Fs3 (P34)	Kanamycin
BB1_34 + sfGFP-RPP1Btt	BB1_34	Fs3-sfGFP-custom (RPP1Btt) (P35) Custom-RPP1Btt (P36)	Kanamycin
BB3aZ_14 + pAOX1: <i>AkCSLA3</i> :sfGFP-RPP1Btt	BB3aZ_14	BB1_12 + pAOX1 BB1_23 + <i>AkCSLA3</i> BB1_34 + sfGFP-RPP1Btt	Zeocin
BB3aZ_14 + pAOX1: <i>AkCSLA3</i> (w/oTMD1):sfGFP-RPP1Btt	BB3aZ_14	BB1_12 + pAOX1 BB1_23 + <i>AkCSLA3</i> (w/oTMD1) BB1_34 + sfGFP-RPP1Btt	Zeocin
BB1_23 + <i>AtCSLA2</i> *	BB1_23	Fs2- <i>AtCSLA2</i> -Dom site 1 (P37) Dom Site 1- <i>AtCSLA2</i> *-Fs3 (P38)	Kanamycin
BB2_AB + pAOX1: <i>AtCSLA2</i> *:RPP1Btt	BB2_AB	BB1_12 + pAOX1 BB1_23 + <i>AtCSLA2</i> * BB1_34 + RPP1Btt	Ampicillin

BB3eH_AC + pAOX1:AtCSLA2*:RPP1Btt	BB3eH_AC	BB2_AB + pAOX1:AtCSLA2*:RPP1Btt BB2_BC (empty)	Hygromycin
BB1_23 + AtCSLA9*	BB1_23	Fs2-AtCSLA9*-Fs3 (P39)	Kanamycin
BB2_BC + pDAS2:AtCSLA9*:ScCYC1tt	BB2_BC	BB1_12 + pDAS2 BB1_23 + AtCSLA9* BB1_34 + ScCYC1tt	Ampicillin
BB3eH_AC + pDAS2:AtCSLA9*:ScCYC1tt	BB3eH_AC	BB2_AB (empty) BB2_BC + pDAS2:AtCSLA9*:ScCYC1tt	Hygromycin
BB3eH_AC + pAOX1:AtCSLA2*:RPP1Btt + pDAS2:AtCSLA9*:ScCYC1tt	BB3eH_AC	BB2_AB + pAOX1:AtCSLA2*:RPP1Btt BB2_BC + pDAS2:AtCSLA9*:ScCYC1tt	Hygromycin
BB1_23 + AtCSLA11*	BB1_23	Fs2-AtCSLA11-Dom site 1 (P40) Dom Site 1-AtCSLA11*-Fs3 (P41)	Kanamycin
BB2_BC + pDAS2:AtCSLA11*:ScCYC1tt	BB2_BC	BB1_12 + pDAS2 BB1_23 + AtCSLA11* BB1_34 + ScCYC1tt	Ampicillin
BB3eH_AC + pDAS2:AtCSLA11*:ScCYC1tt	BB3eH_AC	BB2_AB (empty) BB2_BC + pDAS2:AtCSLA11*:ScCYC1tt	Hygromycin
BB3eH_AC + pAOX1:AtCSLA2*:RPP1Btt + pDAS2:AtCSLA11*:ScCYC1tt	BB3eH_AC	BB2_AB + pAOX1:AtCSLA2*:RPP1Btt BB2_BC + pDAS2:AtCSLA11*:ScCYC1tt	Hygromycin
BB1_23 + AtCSLA14*	BB1_23	Fs2-AtCSLA14*-Fs3 (P42)	Kanamycin
BB2_BC + pDAS2:AtCSLA14*:ScCYC1tt	BB2_BC	BB1_12 + pDAS2 BB1_23 + AtCSLA14* BB1_34 + ScCYC1tt	Ampicillin
BB3eH_AC + pDAS2:AtCSLA14*:ScCYC1tt	BB3eH_AC	BB2_AB (empty) BB2_BC + pDAS2:AtCSLA14*:ScCYC1tt	Hygromycin
BB3eH_AC + pAOX1:AtCSLA2*:RPP1Btt + pDAS2:AtCSLA14*:ScCYC1tt	BB3eH_AC	BB2_AB + pAOX1:AtCSLA2*:RPP1Btt BB2_BC + pDAS2:AtCSLA14*:ScCYC1tt	Hygromycin

Table S7: Overview of primers for Phusion PCRs for expression *in planta*. Bold letters denote for the restriction sites of the enzyme and modified bases are labelled in red. Big letters show gene-specific bases of the sequence, while bases in small letters have been added to either facilitate the cutting of the restriction enzyme or to assemble the sequence in frame. Bases marked in blue denotes for a stop codon at the end of the sequence.

Template (plasmid DNA)	Phusion ID	Primer Target	Description	Direction	Adapter	Fusion Site	Gene-specific Sequence
LIC Cloning							
pPICZB + AkCSLA3	P43	AkCSLA3		Forward	TAGTTGGAATAG	-----	ATGGCCATCGACTGGGCGA
				Reverse	GTTC AGTATGGAGTTG GGTTCAC	-----	CTTTTCACTAGGAACAAAGGTG
MoClo vectors							
pPICZB + Venus	P44	Venus	Fs2-Venus	Forward	cgct GAAGAC gg	AATG	tctaaaggtgaagaatta
			Venus-Fs3	Reverse	acga GAAGAC gg	aagc	TTA tttgtacaattcatcatacc
BB1_23 + AtCSLA2	P45	AtCSLA2	Fs2-AtCSLA2	Forward	agtt GAAGAC gg	AATG	GACGGTGTATCACCAAAG
			AtCSLA2-Fs3	Reverse	acga GAAGAC gg	aagc	CTAACTCGGGACATAAGTC
pPICZB + AtCSLA9-Venus	P46	AtCSLA9	Fs2-AtCSLA9	Forward	aatt GAAGAC aa	AATG	GAGCTAGGAGATACGACGTCGGTG
		Venus	Venus-Fs3	Reverse	gcgc GAAGAC ag	aagc	TTA tttgtacaattcatcatacc
BB1_23 + PumCSLK1	P47	PumCSLK1	Fs2-PumCSLK1	Forward	agtt GAAGAC gg	AATG	GCTCTACCAGAAAGAGCCATTG
			PumCSLK1-Fs3	Reverse	acga GAAGAC gc	aagc	TTAAACCTCAACAACCTGTAGCG
BB1_23 + PumCSLK1-Venus	P48	PumCSLK1	PumCSLK1-custom (Venus)	Reverse	agtt GAAGAC gg	AGAA	AACCTCAACAACCTGTAGCGAGAGGG
		Venus	Custom-Venus	Forward	gcgc GAAGAC gg	TTCT	CT tctaaaggtgaagaatta

Table S8: Overview of created plasmids for overexpression *in planta*.

Final plasmid	Vector DNA	Insert DNA (Phusion ID)	Resistance
pCV01 + p35S: <i>AkCSLA3</i> :YFP-tNOS	pCV01 + p35S-YFP-tNOS	<i>AkCSLA3</i> (P43)	Kanamycin
pAGM4031 + Venus	pAGM4031	Venus (P44)	Spectromycin
pICH75055 + p35S:Venus:tOCS	pICH75055	pICH51277 + p35S pAGM4031 + Venus pICH41432 + tOCS	Carbenicillin
pAGM4031 + <i>AtCSLA2</i>	pAGM4031	<i>AtCSLA2</i> (P45)	Spectromycin
pICH75055 + p35S: <i>AtCSLA2</i> :tOCS	pICH75055	pICH51277 + p35S pAGM4031 + <i>AtCSLA2</i> pICH41432 + tOCS	Carbenicillin
pAGM4031 + <i>AtCSLA9</i> -Venus	pAGM4031	<i>AtCSLA9</i> -Venus (P46)	Spectromycin
pICH75055 + p35S: <i>AtCSLA9</i> - Venus:tOCS	pICH75055	pICH51277 + p35S pAGM4031 + <i>AtCSLA9</i> -Venus pICH41432 + tOCS	Carbenicillin
pAGM4031 + <i>PumCSLK1</i>	pAGM4031	<i>PumCSLK1</i> (P47)	Spectromycin
pICH75055 + p35S: <i>PumCSLK1</i> :tOCS	pICH75055	pICH51277 + p35S pAGM4031 + <i>PumCSLK1</i> pICH41432 + tOCS	Carbenicillin
pAGM4031 + <i>PumCSLK1</i> -Venus	pAGM4031	<i>PumCSLK1</i> -Venus (P48)	Spectromycin
pICH75055 + p35S: <i>PumCSLK1</i> - Venus:tOCS	pICH75055	pICH51277 + p35S pAGM4031 + <i>PumCSLK1</i> -Venus pICH41432 + tOCS	Carbenicillin

Table S9: Overview of primers used for genotyping and sequencing.

	Primer Target	Direction	5' to 3' Sequence
Promoter	pAOX1	Forward	GACTGGTTCCAATTGACAAGC
	pDAS2	Forward	CTTGTATCCTGAGTGACCGTTG
	pGAP	Forward	TGGTTTCTCCTGACCCAAAG
	p35S	Forward	TGACGTAAGGGATGACGCA
Backbone-specific primer	pPICZ	Reverse	GCAAATGGCATTCTGACATCC
	BB1	Forward	TGTA AACGACGGCCAGT
		Reverse	CAGGAAACAGCTATGAC
	BB2	Forward	CTGCGTTATCCCCTGATTCT
		Reverse	GGGTGAGCAAAAACAGGAAG
	BB3	Forward	TAGCTCTTGATCCGGCAAAC
	pAGM4031	Forward	CCTGTGCGGTTTTCGCCACCT
		Reverse	GCCGTTACCACCGCTGCGTT
	pICH75055 (TDNA-left border)	Forward	CTGATGGGCTGCCTGTATCG
	pICH75055 (TDNA-right border)	Reverse	CATGCACATACAAATGGACGAACG
Terminator	RPP1Btt	Reverse	GGACAAGCCGTCACITTTAC
	ScCYC1tt	Reverse	CTTTTCGGTTAGAGCGGATG
	tOCS	Reverse	TCAAGGAATCTCACCCATGC
Gene-specific primer	AkCSLA3	Forward	GAGCAGGAGTCAGGCTCATC
		Reverse	TATTGCCATCTTCCTCACCAG
	ScGOS1	Forward	ATTCTAACCCGGCAATCTCC
	ScMnn2p	Reverse	CCAAGAGGTTCTCCAAGCTG
	ScOCH1	Forward	ATCCGGACAGAGATGATTGG
	RGS4	Reverse	ATCTTTGGCTGAACGCAGAC
	AtCSLA1	Forward	CGAGGATGCAAGAGATGTCTC
	AtCSLA2	Forward	CGCCGGAATATGGAGAATAG
	AtCSLA3	Reverse	TGGCAATTCATTTTTCACCTC
	AtCSLA7	Reverse	ATAAAATGCAGCCCATTTTCG
	AtCSLA9	Reverse	TAGCTGGACCACATGACCAC
	AtCSLA10	Reverse	AATTTCCAGCCGAGAAGACC
	AtCSLA11	Reverse	TTGACCAAGTCGGAATGTTG
AtCSLA14	Reverse	CTCGGAAGCTCGTTTTTCAC	

	<i>AtCSLA15</i>	Forward	TGACGAGGATGCAAGAGATG
		Reverse	AAGATTGGCTGGACCACAAG
	<i>EhCSLK</i>	Forward	GTGCAAGAGGAGTGAGATGC
	<i>PumCSLK1</i>	Reverse	CTCTGTAGCAGCACGTTTGG
	<i>PumCSLK2</i>	Reverse	TTCTGCTTCGCATTTGTGTC
	<i>CrCSLK</i>	Reverse	GCCAGATGTTGGAGAAGGAC
	<i>MicpuCSLK</i>	Reverse	GTGCCGAAGAAGAAGACGAC
	<i>MvCSLK</i>	Reverse	ACTGCATTGGACCTCCTGTC
	<i>OiCSLK</i>	Reverse	GTAGCAGCCATTCTCGTTCC
	<i>CbCSLK</i>	Forward	TTGATGCCGATTTCCAACCC
	<i>MpoCSLA</i>	Reverse	CTTGCGGACGAAGAAGAAAG
	<i>PpCSLA1</i>	Reverse	AGCTCGTTCTTCACCTCCAG
	<i>PpCSLA2</i>	Reverse	GGCAGCTCATTCTTCACCTC
	<i>PpCSLA3</i>	Reverse	CTTCCGCACAAAGAAAAAGG
	<i>SmCSLA</i>	Reverse	GCCAAGTCCATGTCCTCAAC
	<i>AtrCSLA</i>	Forward	GGTCTAGTCCAGGCTCGATG
	<i>DoCSLA6</i>	Forward	CAGGCACGGTGGAAATTCAG
Fluorescent proteins	mCitrine	Reverse	AAGTCGTGCTGCTTCATGTG
	mRuby2	Reverse	CAAATGACCTCCACCATCAA
	pmVenus	Reverse	GTTCGAGGGTGATACGTTGG
	sfGFP	Forward	ATGCGTAAAGGCGAAGAGCTG
		Reverse	GTTCTGTCACATAGCCTTCC
	Venus	Reverse	CAAGACTGGACCATCACCAA

Table S10: Glycosyl linkages in *Pichia* expressing CSLA-Venus. Data show the mean of the relative abundance (%) of each linkage \pm SD of two biological replicates. Linkages with significant differences compared to pPICZB are labelled in green for higher or blue for lower content in *Pichia* (two-tailed t-test, $p < 0.05$). *Ak*: *Amorphophallus konjac* CSLA3; *At*: *Arabidopsis thaliana* CSLA; *V*: Venus tag fused on the C-terminal end of the respective CSLA.

Strain	pPICZB	<i>Ak3-V</i>	<i>At3-V</i>	<i>At7-V</i>	<i>At9-V</i>
t-Man	10.33 \pm 0.38	5.05 \pm 0.08	8.68 \pm 0.21	8.43 \pm 0.83	6.15 \pm 0.60
t-Glc	4.09 \pm 0.18	2.49 \pm 0.34	4.15 \pm 0.04	3.70 \pm 0.38	3.63 \pm 0.25
3-Glc	17.76 \pm 0.44	11.96 \pm 0.82	16.16 \pm 0.68	17.13 \pm 0.43	14.00 \pm 0.73
2-Man	28.67 \pm 2.49	18.62 \pm 1.47	24.49 \pm 3.54	30.40 \pm 2.75	21.73 \pm 0.35
4-Man	0.51 \pm 0.03	24.38 \pm 2.30	8.65 \pm 1.90	4.71 \pm 0.30	18.98 \pm 0.40
6-Man	1.82 \pm 0.04	1.00 \pm 0.03	1.79 \pm 0.20	1.41 \pm 0.12	1.50 \pm 0.09
6-Glc	14.01 \pm 1.11	11.55 \pm 0.12	11.91 \pm 1.11	13.83 \pm 0.30	12.38 \pm 1.24
4-Glc	1.79 \pm 0.08	10.72 \pm 0.51	2.25 \pm 0.28	1.59 \pm 0.05	4.80 \pm 0.22
2,3-Man	1.60 \pm 0.11	0.54 \pm 0.06	1.70 \pm 0.14	1.20 \pm 0.09	0.87 \pm 0.13
2,3-Glc	6.08 \pm 0.19	3.47 \pm 0.95	7.14 \pm 0.96	5.10 \pm 0.36	4.38 \pm 1.29
4,6-Man	0.19 \pm 0.02	1.37 \pm 0.16	0.68 \pm 0.13	0.32 \pm 0.05	1.08 \pm 0.01
2,6-Glc	10.84 \pm 0.42	7.90 \pm 0.88	10.18 \pm 0.12	10.37 \pm 0.27	9.07 \pm 0.65
3,6-Man	2.30 \pm 0.03	0.93 \pm 0.04	2.23 \pm 0.25	1.80 \pm 0.18	1.44 \pm 0.09

Table S11: Glycosyl linkages in *Pichia* expressing CSLK-Venus. Data show the mean of the relative abundance (%) of each linkage \pm SD of two biological replicates. Linkages with significant differences compared to pPICZB are labelled in green for higher or blue for lower content in *Pichia* (two-tailed t-test, $p < 0.05$). *Eh*: *Emiliania huxleyi* CSLK; *Pum*: *Porphyra umbilicalis* CSLK; *Oi*: *Ostreococcus lucimarinus* CSLK; *Micpu*: *Micromonas pusilla* CSLK.

Strain	Venus	<i>EhCSLK</i>	<i>PumCSLK1</i>	<i>OiCSLK</i>	<i>MicpuCSLK</i>
t-Man	8.58 \pm 1.55	8.59 \pm 0.51	5.50 \pm 0.28	7.59 \pm 0.15	6.87 \pm 0.16
t-Glc	3.31 \pm 0.06	3.26 \pm 0.43	1.60 \pm 0.41	2.44 \pm 0.12	1.93 \pm 0.02
3-Glc	14.41 \pm 4.82	16.68 \pm 0.71	14.54 \pm 0.51	16.92 \pm 0.92	17.33 \pm 0.14
2-Man	40.14 \pm 6.03	35.50 \pm 1.65	36.60 \pm 0.92	40.28 \pm 0.82	40.86 \pm 0.61
4-Man	0.50 \pm 0.13	0.50 \pm 0.05	7.97 \pm 1.10	0.42 \pm 0.01	0.46 \pm 0.03
6-Man	1.91 \pm 0.20	1.69 \pm 0.04	1.57 \pm 0.04	1.71 \pm 0.06	1.67 \pm 0.03
6-Glc	11.70 \pm 1.56	13.40 \pm 0.73	12.63 \pm 0.23	12.52 \pm 0.33	12.32 \pm 0.18
4-Glc	2.15 \pm 0.90	1.21 \pm 0.05	2.50 \pm 1.27	1.53 \pm 0.19	2.00 \pm 0.28
2,3-Man	1.35 \pm 0.41	1.12 \pm 0.12	0.71 \pm 0.52	0.71 \pm 0.55	1.03 \pm 0.003
2,3-Glc	4.64 \pm 0.30	6.03 \pm 0.159	4.59 \pm 0.32	4.22 \pm 0.42	3.95 \pm 0.21
4,6-Man	0.20 \pm 0.11	0.14 \pm 0.02	0.43 \pm 0.04	0.14 \pm 0.003	0.14 \pm 0.01
2,6-Glc	8.30 \pm 1.06	9.61 \pm 0.74	9.17 \pm 0.13	9.17 \pm 0.06	9.07 \pm 0.20
3,6-Man	2.80 \pm 0.66	2.27 \pm 0.06	2.19 \pm 0.06	2.36 \pm 0.06	2.36 \pm 0.04

Table S12 Structural comparison of CSLK/CSLA proteins.

Species	Abbrev.	Protein Accession	Location	Start	End	orientation	Exon	Intron	CDS length (bp); without stop	Size (aa)	Expasy		TOPCONS		Signalpeptide
											pI	MW	TMDs	N-terminal end	
<i>Emiliana huxleyi</i>	Eh CSLK	Emihu11452714	scaffold_67	179975	181406	reverse	4	3	1023	341	8.14	37694.38	1	outside	No
<i>Porphyra umbilicalis</i>	Pum CSLK1	Pum0231s0012.1.p	scaffold_231	82115	84993	reverse	2	1	2052	684	5.35	75098.77	6	outside	No
	Pum CSLK2	Pum0372s0002.1.p	scaffold_372	18692	21568	forward	2	1	2334	778	6.26	82917.13	8	outside	No
<i>Chlamydomonas reinhardtii</i>	Cr CSLK	Cre06.g268550	Chr 6	2429364	2437056	forward	1	0	1947	649	7.91	70185.15	5	outside	No
<i>Micromonas pusilla</i>	Micpu CSLK	MicpuC2.e_gw1.16.71.1	scaffold_16	265399	266954	reverse	3	2	1335	445	7.49	50952.92	4	inside	No
<i>Mesostigma viride</i>	Mv CSLK	Mesvi205S04101	scaffold_205	157421	164551	forward	9	8	2001	667	8.63	75689.76	8		No
<i>Ostreococcus lucimarinus</i>	OI CSLK	O.lucimarinus v2.0 18489	Chr 16	137169	138714	forward	1	0	1542	514	8.24	59118.56	5	outside	No
<i>Chara braunii</i>	Cb CSLK	GBG73110.1	BFEA01_177	1401383	1403721	reverse	2	1	1899	633	5.4	70667.25	2	outside	No
<i>Physcomitrella patens</i>	Pp CSLA1	Pp3c17_19630v3	Chr 17	13157861	13161711	reverse	6	5	1614	538	9.05	62249.84	5	outside	No
	Pp CSLA2	Pp3c14_13750v3	Chr 14	8794188	8798427	forward	6	5	1614	538	9.09	62320.91	5	outside	No
	Pp CSLA3	Pp3c14_12270v3	Chr 14	7816158	7820693	reverse	6	5	1614	538	8.99	62119.89	5	outside	No
<i>Marchantia polymorpha</i>	Mpol CSLA	Mapoly0065s0098	scaffold_65	1146684	1156304	forward	9	8	1623	541	9.06	62377.49	5	outside	No
<i>Selaginella moellendorffii</i>	Sm CSLA	S.moellendorffii v1.0 230176	scaffold_3	414909	417588	reverse	9	8	1584	528	9.14	60545.13	5	outside	No
<i>Arabidopsis thaliana</i>	At CSLA1	KJ138750.1 (partial cds)	Chr 4	-	9349258	forward	-	-	1203	401	9.27	46281.85	4	inside	No
	At CSLA1	AT4G16590.1	Chr 4	9345150	9349258	forward	9	8	1560	520	9.5	60584.5	5	outside	No
	At CSLA2	AT5G22740.1	Chr 5	7554689	7560190	reverse	9	8	1602	534	9.4	61557.8	5	outside	No
	At CSLA3	AT1G23480.1	Chr 1	8332862	8336385	forward	9	8	1668	556	8.5	64401	6	inside	No
	At CSLA7	AT2G35650.1	Chr 2	14985470	14988362	forward	8	7	1668	556	9.01	63794.5	5	outside	Yes
	At CSLA9	AT5G03760.1	Chr 5	985622	985622	reverse	9	8	1599	533	9.2	60918.9	5	outside	No
	At CSLA10	AT1G24070.1	Chr 1	8516166	8519984	reverse	10	9	1656	552	8.64	64154	5	outside	No
	At CSLA11	AT5G16190.1	Chr 5	5283550	5286293	reverse	9	8	1329	443	9.24	51469	4	inside	No
	At CSLA14	AT3G56000.1	Chr 3	20783269	20785885	reverse	8	7	1605	535	6.71	61482.4	6	inside	No
At CSLA15	AT4G13410.1	Chr 4	7792016	7796041	reverse	10	9	1611	537	8.67	62374	6	inside	No	
<i>Amborella trichopoda</i>	Atr CSLA	XP_006851749.1	scaffold_00040	4600640	4607539	forward	9	8	1599	533	9.01	61173.78	5	outside	No
<i>Amorphophallus konjac</i>	Ak CSLA3	ADW77641.1	na	na	na	na	1	0	1557	519	8.86	60058.67	5	outside	No
<i>Dendrobium officinale</i>	Do CSLA6	XP_020673043.1	scaffold_412	770742	779370	forward	9	8	1608	536	9.08	61122.61	5	outside	No

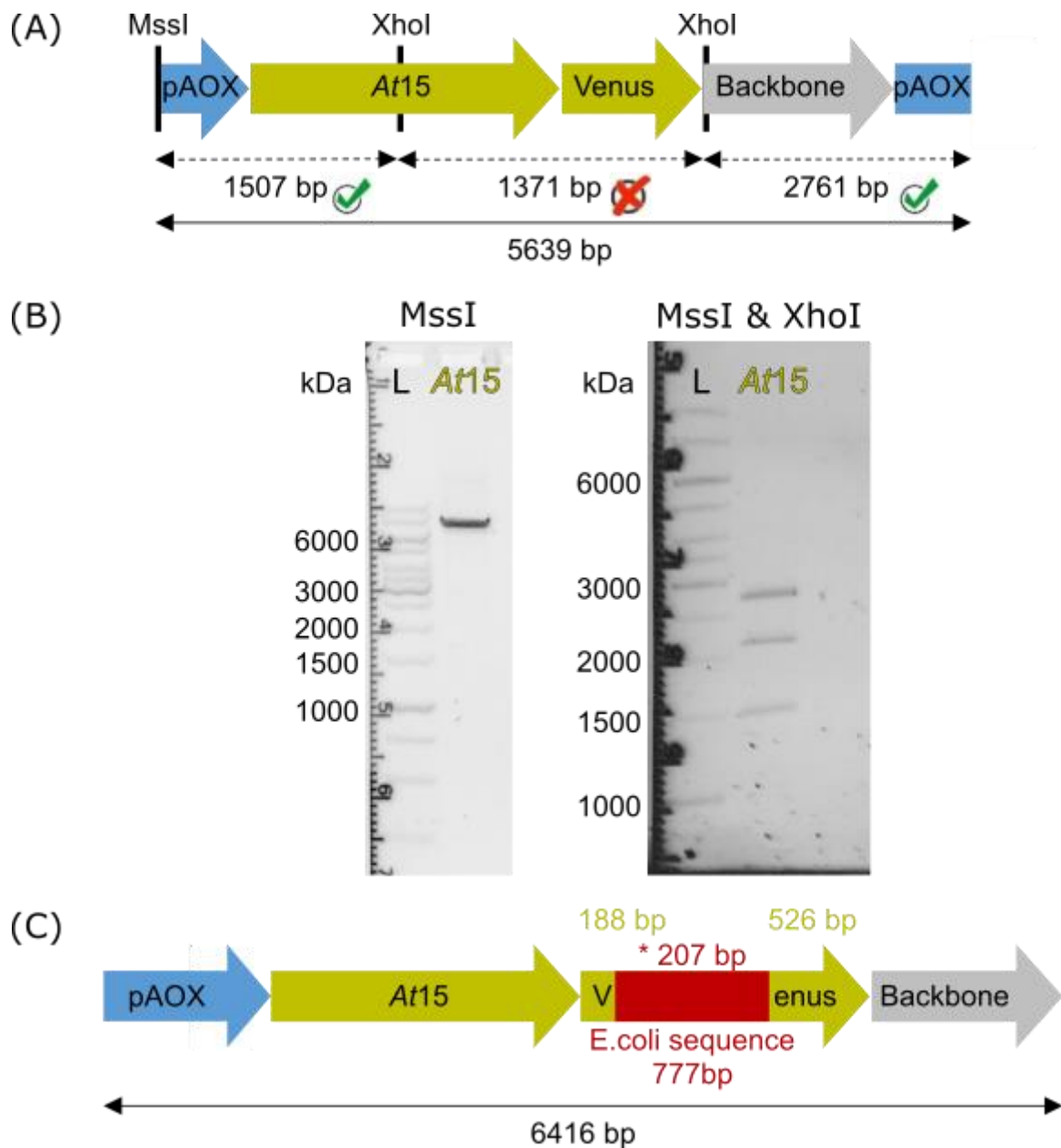


Figure S1: Verification of *AtCSLA15* (*At15*)-Venus DNA from *E. coli*. (A) Schematic overview of *AtCSLA15-Venus*, restriction sites for MssI and XhoI and expected size of the fragments. (B) Agarose gel image after *AtCSLA15-Venus* DNA was linearized with MssI (left) and digested with MssI and XhoI (right). (C) Schematic overview of the verified sequence of predicted *AtCSLA15-Venus* colony in *E. coli*. L: ladder; * denotes for a stop codon after 207 bp.

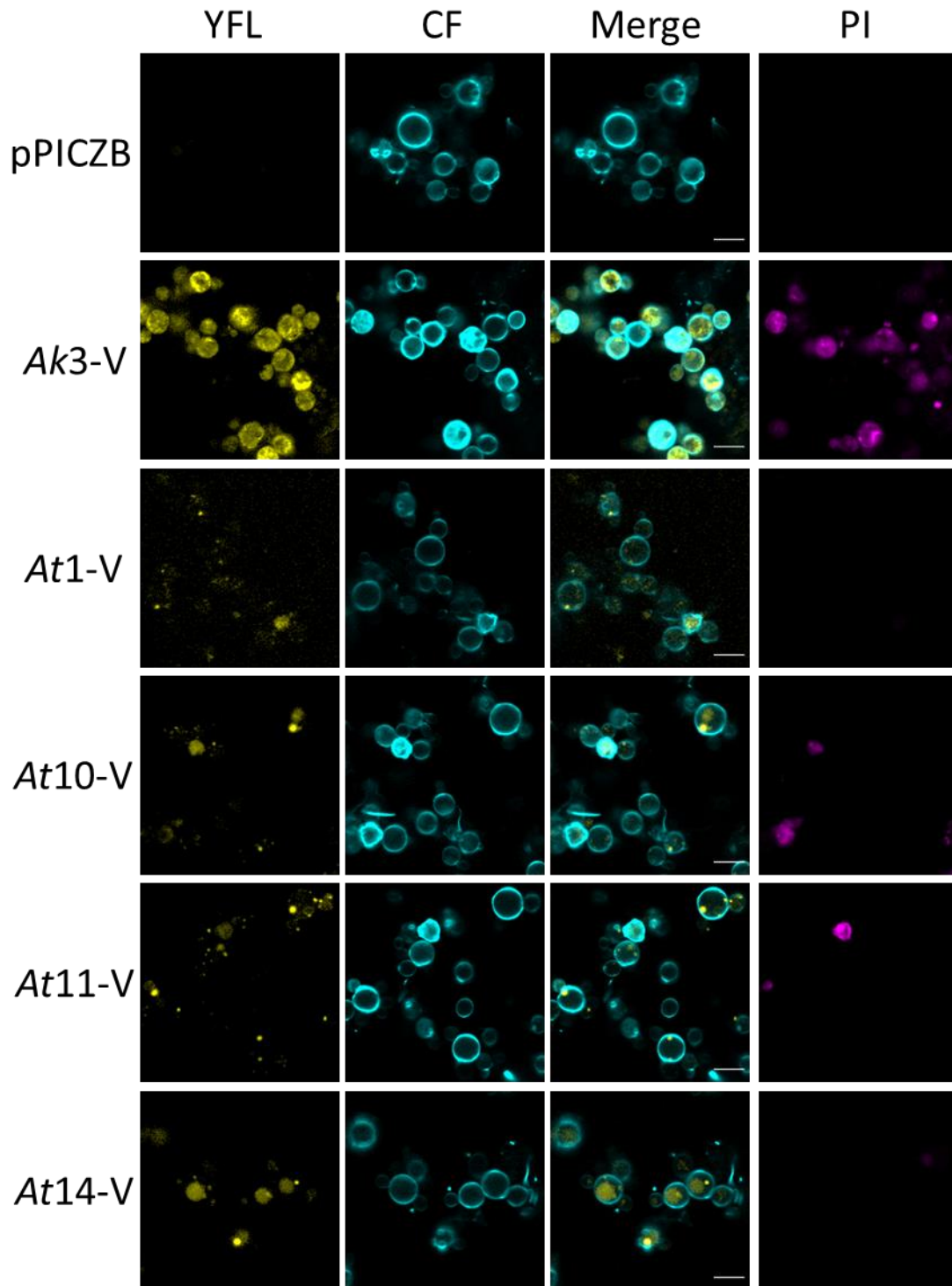


Figure S2: Localization of AkCSLA3-Venus and non-functional AtCSLA-Venus isoforms in *Pichia* cells. *Pichia* cultures were grown in 2 ml YPMG for 48h. Images were detected with a confocal microscope (LSM900). The plasma membrane of *Pichia* cells was stained with calcofluor white (CF) and propidium iodide (PI) was used to detect the viability of the cells. *Ak*: *Amorphophallus konjac*; *At*: *Arabidopsis thaliana*, V: Venus.



Figure S3: Relative monosaccharide composition of fluorescently tagged and untagged *AtCSLA10* and *AtCSLA15* expressing *Pichia*.

Coordinates glucan chains

<i>AtCSLA1</i> (AT4G16590.1):	V	Q	A	R	W	I	F	N	Y	H	F	K
<i>AtCSLA2</i> (AT5G22740.1):	V	Q	A	R	W	R	F	D	Y	H	F	T
<i>AtCSLA3</i> (AT1G23480.1):	V	Q	C	R	W	K	F	N	Y	H	F	V
<i>AtCSLA7</i> (AT2G35650.1):	V	Q	G	R	W	E	F	S	Y	H	F	T
<i>AtCSLA9</i> (Q9LZR3-1):	V	Q	A	R	W	K	F	D	Y	H	F	T
<i>AtCSLA10</i> (AT1G24070.1):	V	Q	A	R	W	R	F	N	Y	H	F	M
<i>AtCSLA11</i> (AT5G16190.1):	V	Q	A	R	W	K	F	N	Y	H	F	T
<i>AtCSLA14</i> (AT3G56000.1):	V	Q	A	G	W	K	Y	N	Y	H	F	A
<i>AtCSLA15</i> (AT4G13410.1):	V	Q	A	R	W	R	F	N	Y	H	F	M
<i>AkCSLA3</i> (ADW77641.1):	A	Q	A	R	W	N	F	D	Y	H	F	K

Figure S4: Comparison of conserved glucan chain coordinating motif between *AtCSLA* isoforms and glucomannan.

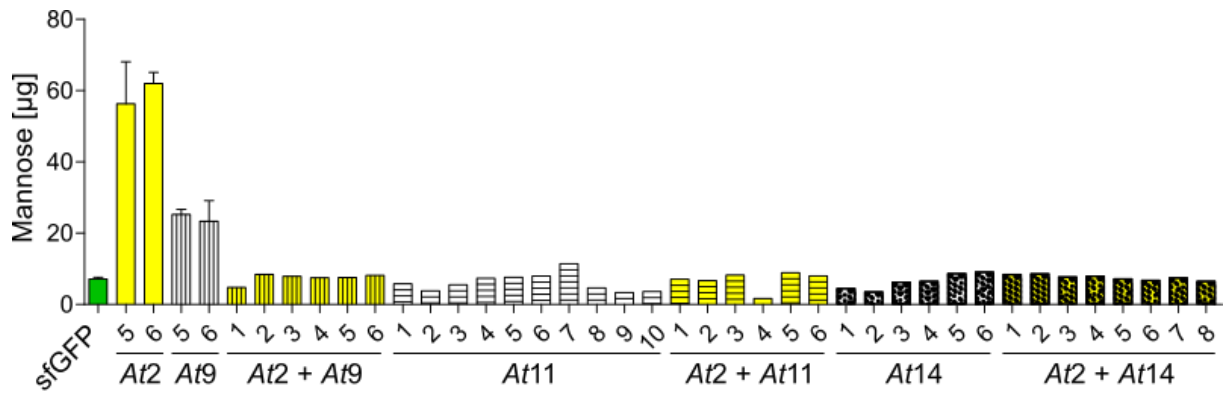


Figure S5: Multiple *Arabidopsis thaliana* CSLA expression in *Pichia pastoris*.

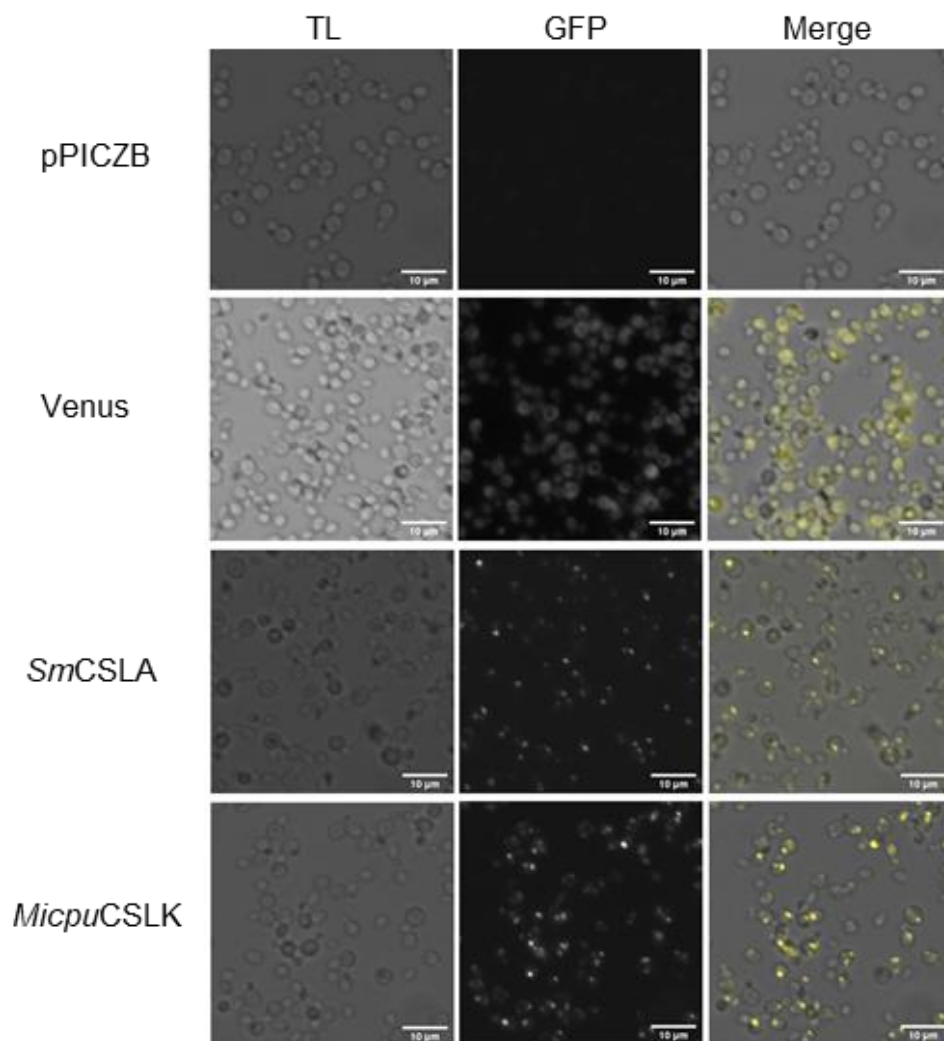


Figure S6: Localization of CSL-Venus constructs expressed in *Pichia pastoris*.

B Abbreviations

AA	Amino acid
Ac	Acetyl
AIR	Alcohol-insoluble residue
AKI	Alkaline-insoluble polymers
Ara	Arabinose
Araf	Arabinofuranose
BMGY	Buffered-glycerol-complex medium
CAZy	Carbohydrate active enzyme
CDS	coding sequence
CSC	cellulose synthase complexes
CF	Calcofluor white
CSL	Cellulose Synthase-Like
DP	Degree of polymerization
DS	Degree of substitution
ER	Endoplasmic reticulum
FL	Fluorescence
FP	Fluorescent Protein
Fuc	Fucose
FPP	Fluorescence protease protection
Fru6P	Fructose-6-phosphate
Gal	Galactose
Gal1P	Galactose-1-phosphate
GalT	Galactosyltransferase
GDP	Guanidine diphosphate
GFP	Green Fluorescent Protein
Glu1P	Glucose-1-phosphate
GGM	Galactoglucomannan
Glc	Glucose
GlcA	Glucuronic acid
Gly	Glycerol
GONST	Golgi-Localized Nucleotide Sugar Transporter

GT	Glycosyltransferase
GTP	Guanine triphosphatase
HG	Homogalacturonan
HM	Heteromannan
IsLB	Low salt Lennox B
LSM	Laser scanning confocal microscopes
MAGT/MUC10	Mannan α -Galactosyltransferase1/Mucilage-Related10
Man	Mannose
Man1P	Mannose-1-phosphate
MBGT1	Mannan β -Galactosyltransferase1
MeGlcA	Methylglucuronic acid
MOAT	Mannan O-acetyltransferase
MoClo	Modular cloning
MLG	Mixed-linkage glucans
MSR	Mannan Synthase-Related
NMR	Nuclear magnetic resonance
NST	Nucleotide sugar transporter
OD ₆₀₀	Optical density at 600 nm
P2A	Self-cleaving 2A peptide from porcine teschovirus
p35S	Cauliflower mosaic virus 35S promoter
PB	Phosphate buffer
PCW	Primary cell wall
PI	Propidium iodide
PMAA	Partially methylated alditol acetates
PMM	Partially methylated monosaccharide
RFP	Red fluorescent protein
RG	Rhamnogalacturonan
RGS4	Regulator of G protein signalling 4
Rha	Rhamnose
RT	room temperature
SCW	Secondary cell wall
TB	Trypan blue
TBL	Trichome Birefringence-like

TFA	Trifluoroacetic acid
TMD	Transmembrane domains
TMD1	First TMD
UDP	Uridine diphosphate
UGE	UDP-Glucose 4-Epimerase
UGP	UDP-Glucose Pyrophosphorylase
USP	UDP-Sugar Pyrophosphorylase
UTr	UDP-Gal Transporter
w/o	without
XGA	Xylogalacturonan
XyG	Xyloglucan
Xyl	Xylose
XXT	XyG xylosyltransferases
YP	Yeast peptone
YPM	YP supplemented with methanol
YPMG	YP supplemented with glycerol

C *Curriculum vitae*

Personal Details

Name: Annika Grieb-Osowski

Date of birth and address: 18.03.1989 in Duisburg

Education and Research Experience

Since 05/2022

Research Associate

Heinrich-Heine University Düsseldorf

Institute Biological Data Science

Department: Biological Data Science

02/2019 – 04/2022

Doctoral Student

Leibniz Institute of Plant Biochemistry

Independent Junior Research Groups

Department: Designer Glycans

Supervisor: Cătălin Voiniciuc

10/2015 – 07/2018

Master of Science (Biology)

Heinrich-Heine University Düsseldorf

10/2011 – 10/2015

Bachelor of Science Biology

Heinrich-Heine University Düsseldorf