



MARTIN-LUTHER-UNIVERSITÄT
HALLE-WITTENBERG

Naturwissenschaftliche Fakultät I
"Biowissenschaften"

Bioengineering of Hemicellulose Synthesis and Detection

Domain Swaps and Carbohydrate-Binding Modules as
Tools for Designer Hemicelluloses

Dissertation zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

durchgeführt am Leibniz Institut für Pflanzenbiochemie



vorgelegt am 31. Juli 2023 von

Madalen Katharina Aude Robert

verteidigt am 20. März 2024

Gutachter 1: Prof. Steffen Abel

Gutachter 2: Assoc. Prof. Cătălin Voiniciuc

Gutachter 3: Prof. Marcel Quint

This doctoral thesis, spanning from February 2020 to July 2023, was supervised by Dr. Cătălin Voiniciuc. Most of the experimental work was carried out at the Leibniz Institute of Plant Biochemistry in Halle (Saale), Germany. The initial phase of the research (February 2020 - February 2022) was conducted in the independent junior research group "Designer Glycans", led by Assoc. Prof. Cătălin Voiniciuc. From March 2022 to July 2023, I was kindly hosted by the Jasmonate Signalling Laboratory led by Dr. Debora Gasperini, with a three-month interruption from October 2022 to December 2022, where I worked in the "Designer Glycans" group of Assoc. Prof. Cătălin Voiniciuc in the Horticultural Sciences Department of the University of Florida.

CONTENTS

CONTENTS	I
PUBLICATION	V
DECLARATION OF AUTHORSHIP	V
ACKNOWLEDGEMENT	VI
FIGURES	VII
TABLES	VIII
APPENDICES	VIII
FREQUENT ABBREVIATIONS	IX
ZUSAMMENFASSUNG	1
ABSTRACT	2
1 INTRODUCTION	3
1.1 INDUSTRIAL SIGNIFICANCE OF CELL WALL POLYSACCHARIDES	3
1.2 HETEROMANNANS	3
1.2.1 HETEROMANNAN FUNCTIONS AND BIOSYNTHESIS	5
1.3 SWAPPING CSLA DOMAINS TO ENHANCE HEMICELLULOSE PRODUCTION	6
1.4 USING CARBOHYDRATE-BINDING MODULES TO DETECT HEMICELLULOSES	8
1.4.1 CARBOHYDRATE-BINDING MODULES	8
1.4.2 CBM SPECIFICITY AND BIOLOGICAL FUNCTIONS	9
1.4.3 BIOTECHNOLOGICAL APPLICATIONS OF CBMS	11
1.4.4 CURRENT MANNAN LABELLING TECHNIQUES	12
1.4.5 CBM RECOMBINATION FOR <i>IN VIVO</i> HM LABELLING	13
1.4.6 POTENTIAL LIMITATIONS OF RECOMBINANT CBM-PROBES	15
1.5 MODEL ORGANISMS	16
1.5.1 <i>ARABIDOPSIS THALIANA</i>	16
1.5.2 <i>NICOTIANA BENTHAMIANA</i>	16
1.5.3 <i>PICHIA PASTORIS</i>	17
1.6 OBJECTIVES	18
2 METHODS	19
2.1 GENE EXPRESSION AND YEAST CULTIVATION	19
2.1.1 MODULAR CONSTRUCTS IN <i>E. COLI</i> AND <i>PICHIA</i>	19

2.1.2	AKCSLA3-SFGFP LIBRARY CREATION	20
2.1.3	<i>PICHIA</i> CULTIVATION	21
2.1.4	GROWTH CURVE ANALYSIS	21
2.2	FLUORESCENCE ANALYSIS	22
2.2.1	ASSESSMENT OF OD AND FL IN <i>PICHIA</i> CULTURES AND SUPERNATANT	22
2.2.2	<i>PICHIA</i> CONFOCAL FLUORESCENCE IMAGING	22
2.2.3	SINGLE CELL FLUORESCENCE ANALYSIS VIA FLOW CYTOMETRY	23
2.3	CARBOHYDRATE ANALYSIS	24
2.3.1	EXTRACTION OF ALKALINE INSOLUBLE POLYMERS FROM <i>PICHIA</i> CELL WALLS	24
2.3.2	SULFURIC ACID AND TFA HYDROLYSIS	24
2.3.3	MONOSACCHARIDE ANALYSIS AND OLIGOSACCHARIDE PROFILING	25
2.3.4	LINKAGE ANALYSIS	26
2.4	PLANT METHODS	27
2.4.1	PLANT GROWTH CONDITIONS	27
2.4.2	MODULAR CONSTRUCT ASSEMBLY AND TRANSFORMATION IN <i>A. TUMEFACIENS</i>	27
2.4.3	TRANSIENT EXPRESSION IN <i>N. BENTHAMIANA</i> AND LEAF PLASMOLYSIS	29
2.4.4	STABLE GENE EXPRESSION IN <i>ARABIDOPSIS</i>	29
3	RESULTS	30
3.1	CSLA DOMAIN SWAPS	30
3.1.1	ENGINEERING OF CELLULOSE SYNTHASE-LIKE ENZYMES AND CELL WALLS	30
3.1.2	INFLUENCE OF HIGH YIELDING CHIMERAS ON (GLUCO)MANNAN COMPOSITION	34
3.1.3	IMPACT OF CSLA EXPRESSION ON YEAST GROWTH AND MORPHOLOGY	35
3.2	DESIGN CBM-PROBES TO DETECT AND LOCALIZE HM <i>IN VIVO</i>	40
3.2.1	OPTIMIZATION OF YEAST GROWTH AND CBM EXPRESSION	40
3.2.1.1	Constitutive CBM Expression is Toxic to Yeast Cells	40
3.2.1.2	Optimizing Culture Volume and Methanol Concentration	42
3.2.1.3	Minimal Growth Media Enables High Gene Expression	43
3.2.1.4	Carbon Co-Feeding for Optimal Growth and Gene Expression	45
3.2.2	RECOMBINATION OF CBMS WITH FLUORESCENT PROTEINS FOR HM DETECTION	47
3.2.2.1	CBM Secretion Decreases upon Heteromannan Presence	47
3.2.2.2	CBM-Probe Behaviour is Unaffected by Transformation Method	50
3.2.2.3	CBM-Probes Co-Localize with CSLAs in Yeast	51
3.2.2.4	Influence of CSLA Expression on Yeast Golgi Morphology	55
3.2.2.5	Immunolabelling of Heterologous Heteromannan in Yeast	62

3.2.3	INTRODUCING A RATIO-METRIC PROBE TO ASSESS CBM-EXPRESSION LEVELS	65
3.2.3.1	Influence of Signal Peptides on CBM Behaviour in Yeast	68
3.2.3.2	Influence of Glucomannan Abundance on PaCBM35 Secretion	70
3.3	APPLICATION OF THE NEW CBM PROBES IN YEAST AND IN PLANTS	73
3.3.1	APPLICATION IN YEAST: DISCOVER NEW FUNCTIONAL HM-PRODUCING PROTEINS	73
3.3.1.1	Characterization of Mutated CSLAs	74
3.3.1.2	CBM-Probes Detect New Functional CSLA Variants	76
3.3.2	APPLICATION IN PLANTS: CBMs HELP ELUCIDATE NATIVE HM DISTRIBUTION	83
3.3.2.1	Transient Expression of CBM-Probes in <i>N. benthamiana</i>	83
3.3.2.2	Non-Invasive Detection of Heteromannans in <i>Arabidopsis</i>	86
4	DISCUSSION	90
4.1	CSLA DOMAIN SWAPS FOR HM BIOSYNTHESIS TAILORING	90
4.2	DESIGN OF A CBM-PROBE FOR <i>IN VIVO</i>	92
4.2.1	CBM-PROBE SECRETION FROM YEAST DEPENDS ON HM QUANTITIES	93
4.2.2	IN YEAST, CBM-PROBES CO-LOCALIZE WITH CSLAs	93
4.2.3	OPTIMIZATION OF CBM-PROBES USING A RATIO-METRIC FLUORESCENT PROTEIN	95
4.2.4	CBM-PROBES HELP DETECT NEW CSLA VARIANTS	95
4.2.5	CBM-PROBES LABEL HETEROMANNANS <i>IN PLANTA</i>	97
4.3	CONCLUSION	100
5	REFERENCES	101
6	APPENDICES	116
	CURRICULUM VITAE	121

PUBLICATION

Robert M, Waldhauer J, Stritt F, Yang B, Pauly M, Voiniciuc C (2021) Modular biosynthesis of plant hemicellulose and its impact on yeast cells. *Biotechnol Biofuels* **14**: 140

Contributions: I carried out most of the experimental work using constructs that Julian Waldhauer and Fabian Stritt had already cloned. I drafted most Figures, Results, and Methods sections. Bo Yang carried out experiments related to *N. benthamiana*. The manuscript was finalized in close collaboration with Cătălin Voiniciuc and the collaborating authors.

Part of the introduction, results and discussion of the CSLA Swap experiments are presented in 1.3, 3.1 and 4.1.

DECLARATION OF AUTHORSHIP

I, Madalen Katharina Aude Robert, born 07. February 1992, hereby declare that I am the sole author of this dissertation entitled "Bioengineering of Hemicellulose Synthesis and Detection: Domain Swaps and Carbohydrate-Binding Modules as Tools for Designer Hemicelluloses". Furthermore, I affirm that all contents of this work either stem from me or have been acknowledged appropriately via citations in the text and in the list of references. I completed the experiments undertaken for this dissertation independently without unauthorized help. To the best of my knowledge, I declare that all information provided is true and accurate.

Date: 31. July 2023

Signature: _____

ACKNOWLEDGEMENT

This dissertation would never have been possible without the advice and assistance of a number of people that accompanied me through the past three and a half years.

Thank you, Cătălin for believing in me despite my lack of “wet lab” experience when I started back in 2020. You were a great example on how to work efficiently and your advice both experiment and editorial have been invaluable to me. Looking back, I am amazed at how much I have progressed professionally thanks to you.

To my Designer Glycans Group lab mates - Annika, Steffi M., and Bo - I cannot thank you enough for your guidance through the early stages of my PhD-student time. Bo, saying “It does not matter” gave me confidence in moments of uncertainty. Steffi M., thank you for stepping in, when I felt overwhelmed. Annika, your knowledge of the lab and on practical work greatly enhanced my own skills. Thank you, Steffi C., for your helping me during my move to the new lab. The labels you prepared in that time still make me smile. Also, Steffi M., Annika and Steffi C. I loved our teamwork and our chats during incubation times. Talia and Ulschan (and your family), you welcomed me so kind-heartedly in Florida and have made my time there an experience that I will forever look back on with happiness.

I extend my sincere gratitude to Debora Gasperini for welcoming me into your lab and for your valuable professional and personal. Working with Yunjing, Stefan, Verona, Mariem, and Andi was a pleasure and I deeply thank you for your help with plant work. A special note of thanks also goes to Andi, the best bench partner I could wish for.

I am forever grateful to my parents, Anne and Frank, who raised me to be confident enough to take on the journey of a PhD thesis. My parents and sisters, Aelez, Rozenn, and Nolwenn, supported and encouraged me in difficult times. I also deeply appreciate that my family and friends proof-reading my thesis, despite it being a distant subject for them.

Finally, a special thanks to my husband Darius. You were my rock, patiently listening during tough times, raising me back up when I wanted to quit and taking care of everything around me during work intensive times. Without you, I could not have done this.

And last but not least, as Snoop Dogg famously said:

“I want to thank me for believing in me, for doing all this hard work, for never taking a day off, for never quitting, for always giving more than I receive, for trying to do more right than wrong, for being me at all times.”

Thanks to all of you!

FIGURES

FIGURE 1 HETEROMANNAN: COMPOSITION AND STRUCTURE OF BACKBONE COMPONENTS.	4
FIGURE 2 BIOSYNTHESIS OF HETEROMANNAN.	5
FIGURE 3 CBM BINDING-TYPE TOPOLOGIES.	10
FIGURE 4 STRUCTURE OF ENGINEERED RECOMBINANT CBMs.	13
FIGURE 5 3-DIMENTIONAL REPRESENTATION OF CANDIDATE CBMs.	15
FIGURE 6 MODULAR ENGINEERING OF HETEROMANNAN BIOSYNTHESIS.....	31
FIGURE 7 ABUNDANCE AND COMPOSITION OF ENGINEERED <i>PICHIA</i> CELL WALL POLYMERS	33
FIGURE 8 FINE-STRUCTURE OF (GLUCO)MANNANS PRODUCED BY THE TOP CHIMERIC CSLAs.....	35
FIGURE 9 INFLUENCE OF CSLA EXPRESSION ON <i>PICHIA</i> GROWTH AND CELL WALL MORPHOLOGY.....	37
FIGURE 10 TOXIC EFFECTS OF AkCSLA3-DRIVEN GLUCOMANNAN SYNTHESIS ON <i>PICHIA</i> CELLS.....	39
FIGURE 11 CONSTITUTIVE AND METHANOL-INDUCIBLE EXPRESSION OF TmCBM27	41
FIGURE 12 OPTIMIZATION OF GROWTH FORMAT AND METHANOL (MeOH) CONCENTRATION.....	42
FIGURE 13 EFFECTS OF BMGY/BMMY AND YPD/YPM MEDIA ON CELL DENSITY AND GENE EXPRESSION	44
FIGURE 14 INFLUENCE OF DIFFERENT CARBON SOURCES ON GROWTH, EXPRESSION AND CBM SECRETION	46
FIGURE 15 WORKING PRINCIPLE OF THE NEW CBM-PROBES.	47
FIGURE 16 CORRELATION BETWEEN CBM SECRETION AND HM PRESENCE.	49
FIGURE 17 INFLUENCE OF TRANSFORMATION DIRECTION ON CBM SECRETION.	51
FIGURE 18 EXPRESSION AND SECRETION OF CBMs GROWN FOR MICROSCOPIC LOCALIZATION ANALYSIS.....	52
FIGURE 19 LOCALIZATION OF CBMs IN DIFFERENT <i>PICHIA</i> BACKGROUNDS.	54
FIGURE 20 EFFECTS OF DIFFERENT CHEMICAL TREATMENTS ON CELL DENSITY, GFP AND RFP SIGNAL.....	57
FIGURE 21 EFFECTS OF DIFFERENT CHEMICAL TREATMENTS ON MONOSACCHARIDE COMPOSITION	58
FIGURE 22 EFFECTS OF CHEMICAL TREATMENTS ON AkCSLA3-SFGFP LOCALIZATION	61
FIGURE 23 IMMUNOLABELLING OF HM IN ATCSLA2 AND AkCSLA3 EXPRESSING CELLS.....	64
FIGURE 24 EXPRESSION AND SECRETION OF CBMs WITH DIFFERENT SIGNAL PEPTIDES.	66
FIGURE 25 LOCALIZATION OF CBMs WITH DIFFERENT SIGNAL PEPTIDES IN YEAST.....	67
FIGURE 26 CBM EXPRESSION AND SECRETION WITH AND WITHOUT SIGNAL PEPTIDE.	69
FIGURE 27 INFLUENCE OF N- AND C-TERMINAL SFGFP TAG OF AkCSLA3 (Ak3) ON FLUORESCENCE AND MANNOSE CONTENT.	71
FIGURE 28 CBM EXPRESSION AND SECRETION IN AkCSLA3 (Ak3) BACKGROUND WITH C- OR N-TERMINAL SFGFP.....	72
FIGURE 29 VERIFIED STRAINS OF THE CSLA LIBRARY THAT WERE USED FOR PRELIMINARY EXPERIMENTS.	74
FIGURE 30 EXPRESSION AND MONOSACCHARIDE COMPOSITION OF VERIFIED CSLA LIBRARY STRAINS EXPRESSED IN WILD-TYPE	75
FIGURE 31 EXPRESSION AND MONOSACCHARIDE COMPOSITION OF VERIFIED CSLA LIBRARY STRAINS EXPRESSED IN A-PACBM35	77
FIGURE 32 LOCALIZATION OF MUTATED AkCSLA3 (GFP) VERSIONS COMPARED TO PARENTAL AkCSLA3.....	79
FIGURE 33 TOPCONS 2.0 AND ALPHA FOLD 2.0 PREDICTIONS OF CSLA LIBRARY VARIANT 1 ALIGNED WITH AkCSLA3 PARENT.	80
FIGURE 34 SINGLE CELL CYTOMETRY OF VERIFIED CSLA LIBRARY STRAINS.	82
FIGURE 35 MODEL OF HOW PLASMOLYSIS AFFECTS PLANT CELL ORGANIZATION.....	83

FIGURE 36 TRANSIENT EXPRESSION OF TPOLCBM16-1 AND PACBM35 WITH DIFFERENT SIGNAL PEPTIDES IN <i>N. BENTHAMIANA</i>	85
FIGURE 37 <i>A. TUMEFACIENS</i> -MEDIATED STABLE <i>ARABIDOPSIS</i> TRANSFORMATION AND PROPAGATION OVER THREE GENERATIONS. ...	86
FIGURE 38 STABLE EXPRESSION OF A-TPOLCBM16-1-SC-2A-TQ2, A-PMSCARLET-I AND PMSCARLET-I IN <i>ARABIDOPSIS</i>	88
FIGURE 39 EXPECTED VS. OBSERVED CBM-LOCALIZATION IN HM-PRODUCING YEAST.	94

TABLES

TABLE 1 SELECTED CBM CANDIDATES.....	14
TABLE 2 EXCITATION AND EMISSION DETECTION-.....	23
TABLE 3 EFFECTS AND CONCENTRATIONS OF CHEMICALS USED FOR MORPHOLOGY ANALYSIS OF GOLGI-APPARATUS.....	55
TABLE 4 POLARITY OF MUTATED CSLA LIBRARY AMINO ACIDS.....	78

APPENDICES

APPENDIX 1 PRIMER SEQUENCES USED FOR GOLDEN GATE AND MOcLO ASSEMBLY	116
APPENDIX 2 PRIMERS USED FOR GENOTYPING AND SEQUENCE VERIFICATION AFTER CLONING.	118
APPENDIX 3 EXAMPLE OF PIPELINE USED TO RECOMBINE TPOLCBM16-1.	118
APPENDIX 4 DILUTIONS AND TOTAL VOLUMES USED FOR PLATE READING OF OD600 AND FLUORESCENCE.....	120
APPENDIX 5 SEGREGATION RATIOS OF TRANSGENIC <i>ARABIDOPSIS</i> LINES SELECTED FOR MICROSCOPICAL ANALYSIS.	120

FREQUENT ABBREVIATIONS

Ak	<i>Amorphophallus konjac</i>
At	<i>Arabidopsis thaliana</i>
BFP	Blue Fluorescent Protein
CAZyme	Carbohydrate-Active Enzymes
CBM	Carbohydrate-Binding Module
CF	Calcofluor White
CjCBM35	Family 35 CBM from <i>Cellvibrio japonicus</i>
CSLA	Cellulose synthase like enzyme from family A
G /Gly	Glycerol
Gal	Galactose
GFP	Green Fluorescent Protein
GH	Glycosyl hydrolase
Glc	Glucose
GlcMan	Glucomannan
GT	Glycosyltransferase
HM	Heteromannan
HPAEC-PAD	High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection
IC	Ion Chromatography
INM	Ivory Nut Mannan
KGM	Konjac Glucomannan
LSM	Laser Scanning Microscope
Man	Mannose
OD600	Optical density (absorbance) at 600 nm
PaCBM35	Family 35 CBM from <i>Podospora anserina</i>
PI	Propidium Iodide
PMAAs	Partially Methylated Alditol Acetates
pmScarlet-I/Sc	monomeric Scarlet-I, codon optimized for expression in <i>Pichia</i>
pSmTq2 / Tq2	pSmTurquoise2 (monomeric superfolder Turquoise 2, optimized for expression in <i>Pichia</i>)
RFP	Red Fluorescent Protein
sfGFP	superfolder green fluorescent protein
TB	Trypan Blue
TL	transmitted light
TmCBM27	Family 27 CBM from <i>Thermotoga maritima</i>
TpolCBM16-1	Family 16-1 CBM from <i>Caldanaerobius polysaccharolyticus</i>

ZUSAMMENFASSUNG

Pflanzenzellwände sind für Funktion und Wachstum essenziell und hoch relevant für die Lebensmittel-, Nicht-Lebensmittel- und Textilindustrie. Hemizellulosen machen ein Drittel der Zellwandbiomasse aus und sind in Zellwand-, Wachstums-, Signalübertragungs- und Abwehrprozesse involviert. Lebensmittel- und Kosmetikindustrie nutzen ihre Gelier- und Verdickungseigenschaften. Im Golgi-Apparat synthetisieren Glykosyltransferasen (GTs) der Zellulose-Synthase-ähnlichen Familie A (CSLA), teils zusammen mit Mannan-Synthesis-Related Proteinen (MSR), Heteromannane (HMs), eine einfache Hemizellulosenklasse.

Trotz signifikanter Fortschritte in den letzten Jahren, ist noch wenig über das effiziente Maßschneiden von Hemizellulosen für biotechnologische Anwendungen bekannt. Die Entwicklung neuer Werkzeuge zur HM-Biosynthese und ihrer *in vivo* Markierung in dieser Arbeit zielte auf die Reduktion dieser Wissenslücke. Solche Fortschritte könnten das Design von Arzneimittelüberzügen und die Nährwertoptimierung von Nutzpflanzen revolutionieren.

Das Austauschen von Domänen aus Pflanzen-Mannan- und -Glucomannan-Synthasen erlaubte die Modifikation der HM-Biosynthese in der industriellen Hefe *Pichia pastoris*. Einige chimäre Enzyme produzierten mehr β -1,4-verknüpftes Mannan als ihre Eltern-Proteine, während andere die toxischen Effekte von verlängerter Glucomannan-Synthase-Expression reduzierten. Interessanterweise konnten die Zellgrößen von *Pichia pastoris* durch die Expression verschiedener CSLA-Proteinsequenzen beeinflusst werden.

Zur Überwindung der Einschränkungen aktueller Mannan-Markierungstechniken, wurden neuartige Bildgebungssonden zur nicht-invasiven, schnellen Pflanzenmannan-Markierung auf Basis von Kohlenhydrat-Bindenden Modulen (CBMs) entwickelt. Die Sekretion der CBM-Sonden – ausgestattet mit Sekretions-Signal-Peptiden und fluoreszierenden Proteinen – hing von der Menge an (Gluco)mannan ab, was vielversprechende Möglichkeiten für die Hochdurchsatzsuche nach neuen funktionalen CSLA-Varianten ermöglichen wird.

Um das Potenzial der neuen CBM-Sonden zu demonstrieren, wurde eine Bibliothek zufällig mutierter AkCSLA3-Varianten erstellt. Die CBM-Sonden sagten die HM-Menge präzise voraus, und Zytometrie erlaubte die Einzelzellanalyse der Variationen von Lebensfähigkeit und Fluoreszenz. Dieser Ansatz könnte künftig das Sortieren funktionaler und nicht-funktionaler CSLAs auf Grundlage der CBM-Retention begünstigen. Die Mannan-Spezifität der neuen Sonden wurde auch in wildtypischen und Mannan-defizienten *Arabidopsis thaliana* gezeigt.

Diese Arbeit bietet neue Methoden zur Erzeugung von maßgeschneiderten HMs und zur Aufklärung der *in planta* HM-Verteilung mit Hilfe von CBM-basierten Sonden. Diese Methoden könnten leicht auf jedes CSL-Enzym angewendet werden, was die Möglichkeiten für Feinabstimmung und *in vivo* Nachweis verschiedener Hemizellulosen erweitert.

ABSTRACT

Plant cell walls are crucial for plant function and growth and are highly relevant for the food, non-food and textile industry. Hemicelluloses make up one third of the cell wall biomass and are relevant for processes related to cell walls, growth, signalling, and defense. Their gelling and thickening abilities are utilized in food and cosmetics. Heteromannans (HM), a simple class of hemicelluloses, are synthesized in the Golgi apparatus, where their backbones are elongated by glycosyltransferases (GTs) from the cellulose synthase-like family A (CSLA), sometimes in collaboration with Mannan-Synthesis-Related proteins (MSR).

Despite significant advances in recent years, little is known about how to efficiently tailor hemicelluloses for biotechnological applications. This doctoral thesis aimed to address this knowledge gap by developing tools to tailor HM biosynthesis and label HM *in vivo*. Such advancements could revolutionize drug coating design and nutritional improvement of crops.

Swapping domains from plant mannan and glucomannan synthases enabled the modification of HM biosynthesis in the industrially relevant yeast *Pichia pastoris*. Some chimeric enzymes produced more β -1,4-linked mannan than their parental counterparts, and others reduced the toxicity of prolonged glucomannan synthase expression. Interestingly, *Pichia pastoris* cell sizes could be influenced by expressing different CSLA protein sequences.

To overcome the limitations of current mannan labelling techniques, novel imaging probes based on Carbohydrate-Binding Modules (CBMs) were developed to enable non-invasive, rapid labelling of plant mannans. Secretion of CBMs – recombined with secretion signal peptides and fluorescent proteins – was dependent on HM quantity, opening up promising opportunities for the high-throughput search of new functional CSLA variants.

The potential of the novel CBM-probes was demonstrated by creating a library of randomly mutated AkCSLA3 variants for analysis. The new CBM-probes accurately predicted HM levels, and cytometry enabled single-cell analysis of viability and fluorescence variations. This approach may enable the future sorting of functional and non-functional CSLAs based on CBM retention. Stable expression in wild-type and mannan-deficient *Arabidopsis thaliana* revealed the CBM-probe's specificity for mannans.

Overall, this research offers novel methods to create custom HM and to elucidate *in planta* HM distributions using CBM-based probes. These methods could be easily applied to any CSL enzyme, expanding the possibilities for fine-tuning and *in vivo* detection of various hemicelluloses.

1 INTRODUCTION

1.1 Industrial Significance of Cell Wall Polysaccharides

Plant cell walls are fundamental for the function and growth of plants and highly relevant to the industry. Therefore, the research on plant cell walls does not only focus on understanding their fundamental principles but also on their potential for industrial usage (Rajasundaram et al., 2014). For humans, plants and their seeds are fundamental food sources and serve as raw materials for various industrial uses (Bewley et al., 2013). In the food industry, galactomannans from carob (*Ceratonia siliqua*) and guar (*Cyamopsis tetragonoloba*) seeds are used to hydrate, thicken and stabilize processed foods (Ebringerová et al., 2005; Bewley et al., 2013). The non-food industry uses cell wall materials in toothpaste, dyes, and cosmetics. Their industrial significance is even more widely spread. They are also used in drug production as parts of hydrophilic matrices as well as for waterproofing stick dynamite (Bewley et al., 2013; Rajasundaram et al., 2014). Other industrial uses of glycan-rich cell walls include paper and timber manufacturing (Rajasundaram et al., 2014). Further, lignocellulosic plant biomass is used to produce second-generation biofuel (Abramson et al., 2010; Yamabhai et al., 2016).

1.2 Heteromannans

Heteromannans (HMs) are hemicelluloses widely prevalent in plants. Hemicelluloses make up around one third of the cell wall biomass and are typically stored in secondary cell walls (Scheller and Ulvskov, 2010; Bewley et al., 2013; Pauly et al., 2013; Voiniciuc et al., 2019). They form hydrogen bonds with the surface of cellulose microfibrils and are major structural components of the cell wall. Usually, hydrogen bonds contribute to the strengthening of structures. In the case of the hydrogen cross-links between hemicelluloses and cellulose microfibrils, there are indications of an opposite effect. Here, hemicellulose cross-links prevent the aggregation of cellulose, reducing the cell walls' mechanical strength. This phenomenon supports the expansion of the cell wall (Somerville et al., 2004; Obembe et al., 2007).

In some seeds, especially the ones from endospermic legumes, hemicelluloses are the primary form of stored carbohydrates. They can be present in such high amounts that starch can no longer be found in the respective tissue. As structural components that are laid down in thick layers, hemicelluloses contribute to the cell wall rigidity. This, for example, is the case for the endosperm of the ivory nut palm, date palm, and coffee (Bewley et al., 2013). Particularly in palm and coffee seeds, the water-insoluble heteromannans confer hardness to the tissues through strong intermolecular interactions (Buckeridge, 2010).

Depending on their backbones, hemicelluloses are categorized into xyloglucans, xylans, mannans, glucomannans, and β -1,3-1,4-linked glucans (Scheller and Ulvskov, 2010). Research indicates that xyloglucans and mannan-based polysaccharides both similarly interact with cellulose, suggesting their involvement in structural mediations of the cell wall (Obembe et al., 2007).

HMs are linear polysaccharides with backbones linked via β -1,4-glycosidic bonds. They can be composed of either only mannose or a combination of the hexose sugars mannose and glucose. Pure mannan is defined as containing more than 90% of mannose residues, which are C2 epimers of glucose (Schnupf et al., 2010) (Figure 1). In addition to mannose (Man) and glucose (Glc), galactose (Gal) residues can decorate (gluco)mannan backbones through an α -1,6-linkage to mannosyl residues, leading to four possible HMs: **mannan**, **galactomannan**, **glucomannan** and **galactoglucomannan** (Buckeridge, 2010; Scheller and Ulvskov, 2010; Bewley et al., 2013; Pauly et al., 2013; Voiniciuc et al., 2019). As components of thickened cell walls, glucomannans appear solely in endosperms of certain monocots, such as members of the Lilaceae and Iridaceae, while galactoglucomannans are the main constituents of secondary walls. The amount of galactose

(Gal) side chains and the number of glucose units incorporated into the backbone account for the high variability of the physicochemical properties of galactomannans, including water solubility, density and solution viscosity (Ebringerová et al., 2005; Bewley et al., 2013). With an increasing number of Gal side chains, galactomannans adapt a higher water solubility, leading to more and more mucilage-like consistencies (Wang et al., 2012; Bewley et al., 2013).

In addition to galactosylation, O-acetyltransferases are hypothesized to help branch the linear backbones to reduce self-aggregation in the Golgi apparatus. Through that, the O-acetyltransferases could support the extracellular secretion of the synthesized HM. In addition, acetylation was shown to alter the interaction of mannans with cellulose (Scheller and Ulvskov, 2010; Berglund et al., 2020).

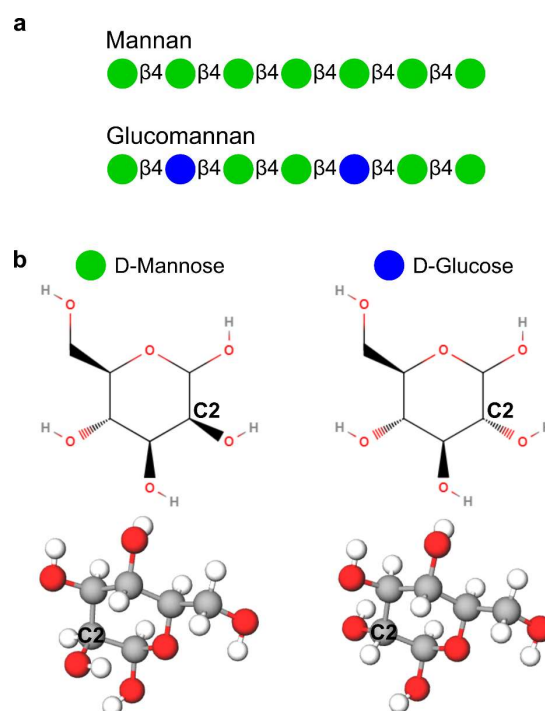


Figure 1 Heteromannan: composition and structure of backbone components. (a) Mannan consists of > 90% mannose residues, whereas glucomannan also incorporates glucose units. (b) Molecular structure of D-Mannose and D-Glucose, which are C2 epimers of each other, meaning that they have a different OH-group orientation at carbon 2. Molecular structures in (b) were retrieved from <https://molview.org/?cid=18950> (Man) and <https://molview.org/?cid=5793> (Glc).

1.2.1 Heteromannan Functions and Biosynthesis

Heteromannans are important hemicelluloses as their functions include the reinforcement of plant cell walls, cell expansion, storage, signalling, embryogenesis, tissue differentiation, and protection from pathogens (Buckeridge, 2010; Scheller and Ulvskov, 2010; Pauly et al., 2013).

In fact, missing the (gluco)mannan synthase AtCSLA7, a Cellulose Synthase-Like protein from family A, causes a lack of the respective HM. This can be lethal to *Arabidopsis* seeds as it is likely required for successful pollen tube growth (Goubet et al., 2003). Further, together with pectin and cellulose, hemicelluloses are present in the seed mucilage of *Arabidopsis thaliana* (Voiniciuc et al., 2015). The seed mucilage is a gel-like coat that forms around the seeds upon hydration. In seeds that contain a mutation of the cellulose synthase-like A protein from family 2 (CSLA2), glucomannan synthesis is impaired, which reduces the amount of HM present in the seed mucilage. Through that, the cellulose crystallinity seems to be reduced, and the spatial organization of the seed mucilage itself is altered (Yu et al., 2014).

Figure 2 briefly depicts the biosynthesis of heteromannans. It takes place in the Golgi apparatus and is mediated by several different proteins (Liepman et al., 2005). Depending on the HM that is synthesized, GDP-Man, GDP-Glc and/or UDP-Gal serve as substrates. Using these specialized precursors, glycosyltransferases (GTs) from the Golgi-located cellulose synthase-like family A (CSLA) synthesize the backbones of HM (Liepman et al., 2005; Davis et al., 2010; Scheller and Ulvskov, 2010; Voiniciuc et al., 2019). CSLA belong to the superfamily

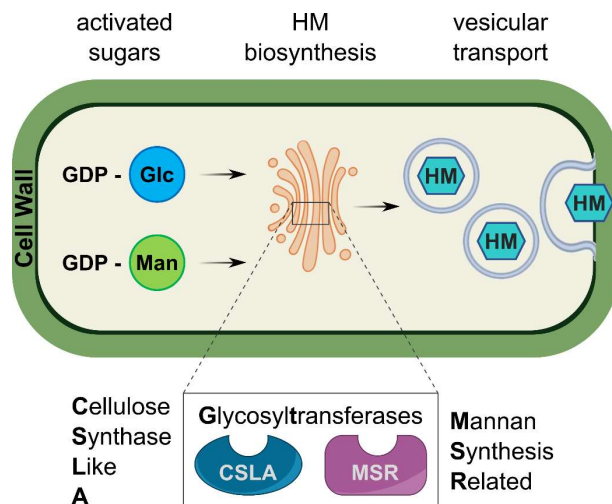


Figure 2 Biosynthesis of heteromannan. GTs from the CSLA family synthesize HM using GDP-Man and GDP-Glc. In some cases, MSR are involved in the process. Vesicles transport HM to the cell wall, where incorporation in the matrix takes place.

of cellulose synthases (CESA). Within that superfamily, they belong to the group of cellulose synthase-like proteins (CSL) and are involved in the synthesis of a variety of β -glycan polymers. From an evolutionary standpoint, the CSLAs likely share the same ancestral gene with the Cellulose Synthase-Like family C (CSLC), as both families have similar structural and physicochemical properties. Members of the CSLC family, however, play a role in the biosynthesis of β -1,4-glucan backbones of xyloglucans and are thought to possibly also include other types of polysaccharide synthases (Dwivany et al., 2009; Liepman and Cavalier, 2012).

The relative concentration of UDP-Gal and GDP-Man determines how strongly galactomannans are branched. Galactomannan synthesis involves not only GTs but also galactosyltransferases (GalT), which add α -1,6-galactose to the mannan backbone (Liepman et al., 2005; Scheller and Ulvskov, 2010). In fact, experiments showed that the putative α -1,6-GalT MUC10 decorates backbones of galactoglucomannans (GGM) that were synthesized by CSLA2 (Voiniciuc et al., 2015; Amos and Mohnen, 2019).

Other cofactors involved in HM biosynthesis are Mannan-Synthesis-Related (MSR) GTs, first described as such by Wang et al. (Wang et al., 2012; Wang et al., 2013). They identified three MSR genes, namely TfMSR from fenugreek as well as AtMSR1 and AtMSR2 from *Arabidopsis thaliana*, all located in the Golgi apparatus. The role of MSR in the biosynthesis of HM is not yet entirely determined; however, there are indications that they might be involved in primer synthesis for the initiation of HM synthesis or in the stabilization and/or activation of CSLA proteins (Wang et al., 2013). More recently, it was found that they may interact with CSLAs to modify (gluco)mannan backbone elongation or could be involved in their post-translational modification. Sometimes MSRs were crucial for HM biosynthesis. In some cases, CSLAs could only produce HMs when co-expressed with AtMSR1. Also, AtMSR1 was able to increase and alter HMs when co-expressed with CSLAs (Voiniciuc et al., 2019).

Synthesized HMs are embedded into vesicles which transport them from the Golgi apparatus to the plant cell wall, where they are integrated into its matrix (Pauly and Keegstra, 2008; Franková and Fry, 2013; Voiniciuc et al., 2019).

1.3 Swapping CSLA Domains to Enhance Hemicellulose Production

While the basics of HM biosynthesis are known, the specificity of different CSLAs, meaning which type of HM they produce, remains to be elucidated. One common approach to investigate CSLA specificity is their expression in heterologous hosts such as *Arabidopsis thaliana*, *Nicotiana benthamiana* (tobacco), or the methylotrophic yeast *Pichia pastoris*. While the expression of CSLAs in plants is more convenient to decipher possible phenotypes caused by CSLA overexpression or CSLA knockouts (e.g., *cs/a2,3,9* triple mutant with no detectable mannan (Goubet et al., 2009; Liepman and Cavalier, 2012)), the expression in yeast provides a clean background (native sugars are different from HM (α -mannoproteins & β -1,3-glucans)) and gives a clear and clean view on what type of HM a certain CSLA might produce. However, as described in section 1.2.1, in nature, CSLAs work together with co-factors to elongate the HM backbones. In addition, the availability of pre-cursors plays a crucial role in determining the final product. Like that, *in vitro* experiments with AtCSLA2, which produces glucomannan in *Arabidopsis*, revealed that on its own, it incorporates only small amounts of Glc, causing it

to produce mannan and not glucomannan when heterologously expressed in yeast (Liepman et al., 2005; Yu et al., 2014; Voiniciuc et al., 2019).

While invaluable insight into CSLA specificity has been gained using the above-mentioned approaches, no research focusing explicitly on how distinct motifs within the CSLA proteins may influence the biosynthesis of (gluco)mannan has been conducted so far.

One interesting approach to investigate the functions of specific enzyme motifs or domains is to swap domains between similar enzymes. For instance, domain swap experiments from two studies allowed to determine CSLF amino acid motifs responsible for determining the ratio between β -1,3 and β -1,4-linkages in mixed-linkage glucans. By transiently expressing CSLF6 chimeras – obtained by swapping CSLF domains from different species – in *N. benthamiana* leaves, it was possible to narrow down the amino acid motifs responsible for the mixed-linkage glucan fine structure (Jobling, 2015; Dimitroff et al., 2016).

In addition to gaining fundamental knowledge, domain swaps could also be employed to optimize hemicellulose production by improving the respective enzymes. Optimized enzymes could ultimately be used to produce crops with higher nutritional value, thus improving their effect on human health (Voiniciuc et al., 2015). Furthermore, engineering living materials with custom-made enzymes could be utilized to equip them with desirable characteristics (Gilbert and Ellis, 2018).

The present study aimed to alter heterologous plant (gluco)mannan production in the industrially relevant yeast *Pichia* (see 1.5.3). AtCSLA2 (mannan production in yeast) and AkCSLA3 (glucomannan production in yeast) were selected as candidates for single domain swaps. Eight chimeric enzyme versions that were based on AtCSLA2 or AkCSLA3, respectively were heterologously expressed in *Pichia*, followed by a thorough investigation of the resultant (gluco)mannan yields and compositions. It was found that domain swaps enabled toxicity reduction of prolonged AkCSLA3 overexpression and could result in higher HM yields (Robert et al., 2021). One significant advantage of the modular approach is that it can be applied to various CSL enzymes and, by that, allow the sustainable production of tailored hemicelluloses for future applications.

1.4 Using Carbohydrate-Binding Modules to Detect Hemicelluloses

To decipher the distributions and dynamics of the produced HM, current labelling techniques have several shortcomings. Like that, immunolabelling techniques are specific to HM but also invasive and very laborious, ultimately preventing the easy labelling of HM *in vivo*. An easier approach is to use stains, but they are not specific to HM, since they usually stain an array of different polysaccharides (e.g., Calcofluor White) (Voiniciuc et al., 2018b).

In this work, a new probe for non-invasive and specific *in vivo* HM labelling was developed. It is based on Carbohydrate-Binding Modules (CBM), which were recombined with fluorescent proteins. Co-expressing native or chimeric CSL enzymes with CBMs, will enable the detection of functional enzymes by fluorescence screening. Further, fluorescence imaging will allow to elucidate the distribution and dynamics of CBM-labelled HM in yeast and *in planta*.

1.4.1 Carbohydrate-Binding Modules

Carbohydrate-Binding Modules (CBM) are domains of carbohydrate-active enzymes. They fold autonomously, their primary function being carbohydrate-binding activity (Tomme et al., 1998; Boraston et al., 2004; Abbott and Boraston, 2012; Lombard et al., 2014). They can also bind starch and glycogen, two insoluble polysaccharides with storage functions (Boraston et al., 2004). In rare cases, they have cellulosomal scaffolding properties or are just independent putative CBMs (Lombard et al., 2014).

Typically CBMs are part of carbohydrate-active enzymes (CAZymes) such as water-insoluble polysaccharide degrading glycoside hydrolases (GH) like cellulases, mannanases, and xylanases, assisting them in concentrating in proximity to their target polysaccharides (Tomme et al., 1988; Henrissat et al., 2001; Boraston et al., 2004; Obembe et al., 2007; Armenta et al., 2017). Linker sequences with various flexibility grades connect CBMs to other domains of a respective protein (example: PaCBM35 in Figure 5). CBMs can be found both at the C- and the N-terminus of a CAZyme, as well as in between other modules of the enzyme. CAZymes tend to have a modular structure, meaning they may contain more than one CBM and/or GH (Henrissat et al., 2001; Abbott and Boraston, 2012). Combinations of at least two CBM units, either from the same or different families, within one protein, are called tandem CBMs (Obembe et al., 2007). They enable the protein they are situated in to broaden the range of carbohydrates it can recognize (Abbott and Boraston, 2012; Armenta et al., 2017).

Depending on the type, CBMs differ in their binding site topology (Boraston et al., 2004; Armenta et al., 2017). In general, binding platforms of CBMs can be specific to precisely one or to several ligands, in which case they are called promiscuous CBMs. Promiscuous

recognition allows a more efficient cell wall degradation by broadening the location where their enzymes can modify the cell wall (Obembe et al., 2007).

CBMs are primarily found in plant-degrading organisms and can play a role in the modification of interactions between hemicelluloses and microfibrils (Levy et al., 2002). This was shown for the elongation of pollen tubes of peach (*Prunus persica* L.), where the tip zone of pollen tubes lost crystallinity when treated with a recombinant CBM (Shpigel et al., 1998). Further, some CBMs seem to modify the activity of cell wall-targeting enzymes and could be involved in processes related to resistance against pathogens and cell wall metabolism and structure (Obembe et al., 2007; Nardi et al., 2013; Nardi et al., 2015).

Over the years, many further experiments regarding CBMs have been carried out, all indicating that impairing their function reduces the activity of their respective enzymes significantly, which accounts for their high importance regarding the functionality of the cell wall polysaccharide degrading process (Boraston et al., 2004).

1.4.2 CBM Specificity and Biological Functions

Decisive factors for CBM specificity are the topography of their hydrophobic binding platforms, hydrogen bonds, calcium-mediated coordination, and other interactions. Depending on their binding specificities, CBMs are classified as planar type A, which binds surfaces of crystalline polysaccharides, endo type B, which binds the middle of single polysaccharide chains and exo type C, which binds small oligosaccharides at the end of polysaccharide chains (Boraston et al., 2004; Shoseyov et al., 2006; Gilbert et al., 2013; Armenta et al., 2017).

Hydrophobic interactions and CH- π interactions (hydrogen bonds between a soft acid and a soft base (Nishio, 2011)) formed by aromatic amino acids, such as tryptophan and tyrosine, mediate substrate binding in all CBMs types. In the case of type B and type C CBMs, interactions between binding sites and ligands are further stabilized through hydrogen bonds formed by polar amino acids or tyrosines (Boraston et al., 2004; Armenta et al., 2017).

The binding sites of CBMs type A, B, and C show some differences regarding their respective binding sites (Figure 3). **Type A** CBMs have hydrophobic planar or platform-like binding sites which mirror substrate topologies and can be joined to various GH. They are not capable of binding oligosaccharides but rather bind crystalline cellulose and chitin (Boraston et al., 2004; Gilbert et al., 2013; Armenta et al., 2017).

Type B binding sites are typically cleft-like but can also accommodate a slit-like conformation (Boraston et al., 2004; Armenta et al., 2017). The binding sites may be located near the end of the CBM and/or on the CBMs' concave side (Armenta et al., 2017). Binding sites at the end of the CBM enable enzymes to bind larger, branched polysaccharides that

match their respective catalytic module (Boraston et al., 2004; Abbott and Boraston, 2012). Contrary to type A CBMs, type B CBMs are specialized in binding glycan chains, such as galactans and starch, as well as xylans and mannans (Boraston et al., 2004; Gilbert et al., 2013; Armenta et al., 2017). The size of type B CBM binding sites is very variable, enabling some of them to bind xylose side chains while simultaneously interacting with their backbone molecules or even to hold a complete pyranose ring (Gilbert et al., 2013; Armenta et al., 2017). The specificity of type B CBMs is strongly determined by aromatic amino acid residues, of which the orientation is the most substantial specificity determining factor. These residues mediate substrate binding through hydrophobic CH- π interactions. Another crucial specificity factor of type B CBMs are hydrogen bonds forming between binding sites and ligands. These hydrogen bonds also exist in type A CBMs, yet with a less crucial role.

Pocket-like binding sites are typical for the **type C** CBMs, which are highly affine to smaller sugars, such as mono-, di- and tri-saccharides, and have lectin-like properties. Like types A and B, type C CBMs also form hydrogen bonds with their ligands, but for them, these hydrogen bonds play an even more important role than for the other two CBM types (Boraston et al., 2004; Armenta et al., 2017).

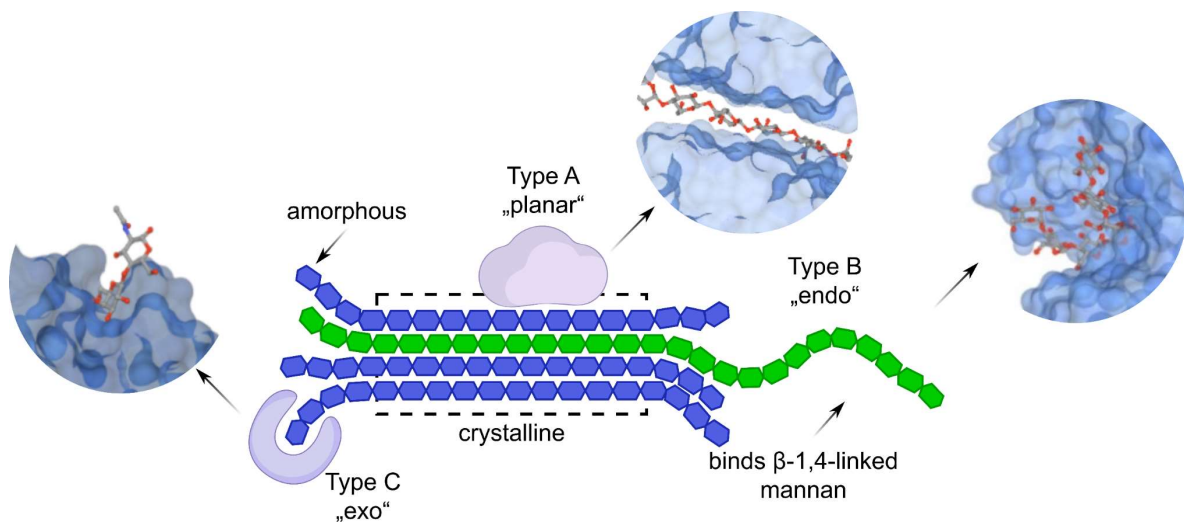


Figure 3 CBM binding-type topologies. **Type A** binds to crystalline cellulose in a planar manner (example: CBM63 from *Bacillus subtilis* expansin (EXLX1), with cellohexaose ligand (PDB: 4FER)). **Type B** binds the middle of mannan oligosaccharide chains (example: CBM58 from *Bacteroides thetaiotaomicron* α -amylase (susG-D498N) with maltopentaose ligand (PDB: 3K8L)). **Type C** binds the ends of (gluco)mannan backbones (example: CBM71 from *Streptococcus pneumoniae* β -galactosidase (BgaA) with β -D-galactopyranosol-1,4-N-acety-D-glucosamine (PDB: 4CUB)). The model was created based on a previous model that is shown on the following website: https://www.cazypedia.org/index.php/Carbohydrate-binding_modules. Examples for the three CBM-types were selected based on a previous review (Armenta et al., 2017).

Decisive factors for the specificity of CBMs are the topography of their hydrophobic binding platforms, hydrogen bonds, calcium-mediated coordination, and other interactions. Three types of CBM binding sites can be distinguished, namely planar, twisted, or sandwich-like. Examples of the different binding site topologies are depicted in Figure 3 (Armenta et al., 2017). Based on their specific configurations, CBMs are grouped into families. Over the years,

the number of families increased from 39 in 2004 to 97 in 2023 (<http://www.cazy.org/Carbohydrate-Binding-Modules.html>, 12.07.2023). Families with considerable resemblances are sometimes grouped into superfamilies or clans (Boraston et al., 2004; Obembe et al., 2007; Lombard et al., 2014).

In addition to being classified by types, CBMs are classified into fold families and further into CBM families. There are seven different fold families: β -sandwich, β -trefoil, cysteine knot, unique, OB fold, hevein fold and unique containing a hevein fold. The β -sandwich is the most common fold type. A CBM may belong to one of the 97 currently defined CBM families, and the ligand specificities substantially vary between the families (Boraston et al., 2004; Abbott and Boraston, 2012; Lombard et al., 2014; Armenta et al., 2017) (<http://www.cazy.org/Carbohydrate-Binding-Modules.html>, 12.07.2023).

CBMs comprise three different types of functions, namely proximity effects, specific targeting, and disruptive functions. Proximity effects increase the enzyme concentration on the polysaccharide surface. Thanks to their appended CBMs, CAZymes are concentrated efficiently and remain longer near their target substrates. This results in faster carbohydrate degradation (Bolam et al., 1998; Boraston et al., 2004; Shoseyov et al., 2006).

Targeting of hydrolytics to specific regions of the plant cell wall allows to produce modified plant fibers. Using these properties, molecular probes can be designed that are able to map the glycol-architecture of plant cell walls (Boraston et al., 2001; Notenboom et al., 2001; Boraston et al., 2004). In some cases, CBMs mediate the non-catalytic disruption of the crystalline cellulose structure (Din et al., 1994; Boraston et al., 2004).

Thus, CBMs can improve carbohydrate degrading activity, alter the physical and chemical characteristics of cellulosic fiber surfaces, and increase their tensile and burst indexes (Shoseyov et al., 2006). In addition to their recognition abilities of complex macromolecular structures, such as cell walls in plants, CBMs can recognize cellular glycans from bacterial pathogens during pathogenesis and colonialization (Armenta et al., 2017).

1.4.3 Biotechnological Applications of CBMs

CBMs are widely used for biotechnological applications (Armenta et al., 2017). In bioprocessing, CBMs that reversibly bind to cellulose may serve as affinity tags to produce and purify recombinant proteins and bioethanol and can also be used in phage display technology (Shoseyov et al., 2006). Using recombinant DNA technologies, CBMs can also be used to map plant cell walls, yielding valuable information about the composition and distribution of plant cell wall carbohydrates. Chimeric proteins that are fused to CBMs are used to increase the carbohydrate and biomass degrading activities. This can be applied to agricultural waste

products and can also be utilized to pre-process lignocellulosic material to fermentable sugars for bioethanol and biofuel production through saccharification (Armenta et al., 2017).

An interesting strategy to simultaneously decrease both the expression and activity of target enzymes involved in the cell wall catabolism is thought to be the overexpression of a promiscuous CBM. Furthermore, CBMs that are involved in the modulation of the cell wall metabolism of transgenic plants could be used to delay fruit ripening by slowing down cell wall degradation. This would keep fruits firm for a longer time and make them less susceptible to fungal attacks (Nardi et al., 2015).

The paper industry also profits from CBMs. The involvement of tandem CBMs in paper production yields considerably stronger paper with increased tensile strength, folding endurance, and burst indexes (Armenta et al., 2017). Cell immobilization for research purposes or fiber modifications utilizes CBMs that irreversibly bind to cellulose (Tomme et al., 1998; Boraston et al., 2004; Shoseyov et al., 2006).

Furthermore, CBMs are useful for the biomedical industry, where they may have various applications, e.g., regarding controlled drug release hydrogels or for producing biomaterials (Armenta et al., 2017). Other applications of CBMs involve targeting specific molecules, their engineering for applications such as detecting specific polysaccharides in plant cell walls and producing non-DNA microarrays (Shoseyov et al., 2006).

1.4.4 Current Mannan Labelling Techniques

For the research of cell wall polysaccharide distribution *in vivo*, various approaches have been taken over the past decades, as summarized in a recent review (Voiniciuc et al., 2018b). Among the different polysaccharide monitoring techniques are dyes, fluorescent protein tagging, immunolabelling, and click chemistry (Voiniciuc et al., 2018b). CBMs have also been used to label cell wall polysaccharides. Yet, the currently available probes are designed to label cellulose and xylan. Further, the methods using CBMs to label cell wall polysaccharides *in vivo* have a working pipeline that is very similar to immunolabelling, as they also require the use of antibodies (Blake et al., 2006; Khatri et al., 2016; Voiniciuc et al., 2018b).

While immunolabelling is well suited to label certain types of cell wall polysaccharides specifically, it is laborious and requires tissue disruption. It is, therefore, unsuitable to decipher the *in vivo* distribution of cell wall polysaccharides, let alone provide insight into their dynamics (Marcus et al., 2010; Voiniciuc et al., 2018b). An easier approach to label cell wall polysaccharides is the use of dyes such as Calcofluor White, which labels β -glucan structures (Anderson et al., 2010; Voiniciuc et al., 2018b), Ruthenium Red, which labels pectins (Voiniciuc et al., 2018b) or Congo Red which binds cellulose fibrils (Kerstens, 2002). While the dyes are

easy to use and do not require the destruction of the tissue to analyse, the binding range of polysaccharides is too broad to allow statements about the distribution of specific types of polysaccharides (Anderson and Wallace, 2012; Voiniciuc et al., 2018b).

Various specific labelling techniques are available for cellulose, pectin, xylan, and xyloglucan, yet, immunolabelling is the only available tool to specifically label HM (Voiniciuc et al., 2018b). Linear mannans can be labelled using the LM22 mAb, while the LM21 mAb can label both linear and branched mannans (Marcus et al., 2010).

To sum up, currently available labelling techniques for mannans are sparse. The ones available are laborious and invasive, making them unsuitable to label HM *in vivo*. From that, the need arises for a new, HM-specific labelling probe that is easy to use and non-invasive.

A promising approach to non-invasively label cell wall polysaccharides *in vivo* is to recombine CBMs with fluorescent proteins and signal peptides for secretion and express them in the target organism that should be investigated. Previously, recombinant CBMs enabled xylan labelling in *Arabidopsis thaliana* (Zhang, 2014). To use this approach for mannan labelling, mannan-binding CBM candidates were selected, recombined, and validated in *Pichia* (section 3.2). Then, as proof of concept for the two target applications, the CBMs were used to find new functional CSLA variants in yeast (section 3.3.1) and to elucidate the native mannan distribution in living plants (section 3.3.2).

1.4.5 CBM Recombination for *in vivo* HM Labelling

CBMs can be recombined with different types of proteins, depending on the target purpose. They can be recombined with adhesion which could strengthen the bond between bacterial cellulose and human microvascular endothelial cells (Andrade et al., 2010) or with his-tags to indirectly label cell wall polysaccharides (McCartney et al., 2004).

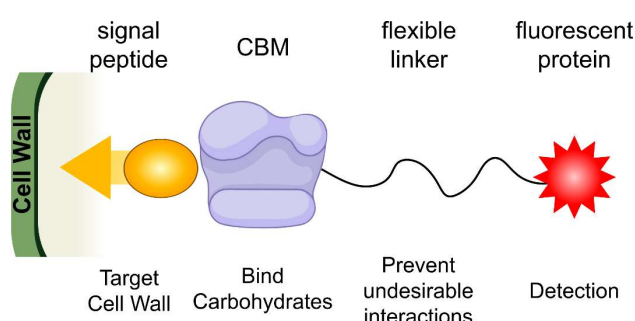


Figure 4 Structure of engineered recombinant CBMs. A signal peptide (yellow) targets the recombinant CBM protein to the cell wall. The CBM (purple) can then bind the respective carbohydrates. A fluorescent protein (red) is attached to the CBM, sometimes via a small flexible linker (black), to prevent undesirable interactions between the fluorescent protein and the CBM.

Recombinant CBMs can also be used to engineer their expression in the yeast *Pichia pastoris* and the model plant *Arabidopsis thaliana* to label HM. They consist of a signal peptide, a CBM, a linker, and a fluorescent protein (Figure 4). The signal peptide uses the secretory pathway to guide the fusion protein to the cell wall (Zhang, 2014). Taking advantage of the CBM binding specificities, their function within a recombinant protein is to detect cell

wall polysaccharides. Through that, they enable the documentation of the distribution and dynamics of cell wall polysaccharides. To enable optical detection of the CBMs, they are recombined with fluorescent proteins that can easily be imaged (Knox, 2008; Zhang, 2014). To prevent any possible undesirable effects potentially caused by directly linking the tested CBM to the fluorescent protein and to ensure the preservation of the CBMs biological activity, linker sequences like the flexible (15)-proline-threonine linker ((15)-PT-linker) may be introduced between the components (Poon et al., 2007; Han et al., 2013).

Successful recombination of CBMs with fluorescent proteins for polysaccharide labelling was demonstrated several times. However, none of the tested CBMs was specific for mannan. Further, the tested CBMs sometimes caused dwarf phenotypes and cell morphology alterations when heterologously expressed in *Arabidopsis* or *N. benthamiana* (Obembe et al., 2007; Zhang, 2014). As the CBM usage is only intended as a labelling tool, it should not affect the plant phenotypes, since this would affect the environment that should be monitored. Therefore, the newly designed recombinant CBMs will have to be investigated carefully for possible morphological effects on the heterologous hosts before using them as labelling tools.

To find suitable candidate CBMs for HM binding *in vivo*, the CAZy database (<http://www.cazy.org/>) was searched for CBMs known to bind (hetero)mannan (components). Natively, all selected CBM candidates are part of β -mannanases and primarily bind mannans. Except for PaCBM35, which comes from the fungus *Podospora anserina*, all candidate CBMs are found in different bacteria (Table 1).

Table 1 Selected CBM candidates for the design of new recombinant CBMs-probes for *in vivo* HM-detection.

CBM	PROTEIN	ORGANISM	UNIPROT/PDB	SUBSTRATES	REF.
TpoICBM16-1	β -mannanase 5A	<i>Caldanaerobius polysaccharolyticus</i>	Q9ZA17/ 3OEB	mannopentaose cellopentaose β -1,4-manno- oligosaccharides	(Bae et al., 2008; Su et al., 2010)
TmCBM27	β -mannanase 5	<i>Thermotoga maritima</i>	Q9RIK9/ 1OF4	carob galactomannan konjac glucomannan	(Boraston et al., 2003)
CjCBM35	endo- β -1,4- mannanase 5C	<i>Cellvibrio japonicus</i>	B2AEP0/ 2BGO	Internal regions of β -mannan	(Tunncliffe et al., 2005)
PaCBM35	endo- β -1,4- mannanase26A	<i>Podospora anserina</i> <i>Smat+</i>	E2GHW2/ 3ZM8	β -1,4-mannan (xylan)	(Couturier et al., 2011; Couturier et al., 2013)

The selected CBMs display similar conformations (Figure 5), accounting for the fact that they all are type B CBMs with a β -sandwich fold and highly similar ligands. The candidate CBMs were recombined with fluorescent proteins and signal peptides using the modular cloning systems GoldenPiCs (Prielhofer et al., 2017) for heterologous expression in the yeast *Pichia pastoris* and the MoClo system (Weber et al., 2011; Engler et al., 2014; Marillonnet and Grütznert, 2020) for expression in the model plants *Nicotiana benthamiana* (tobacco) and *Arabidopsis thaliana*.

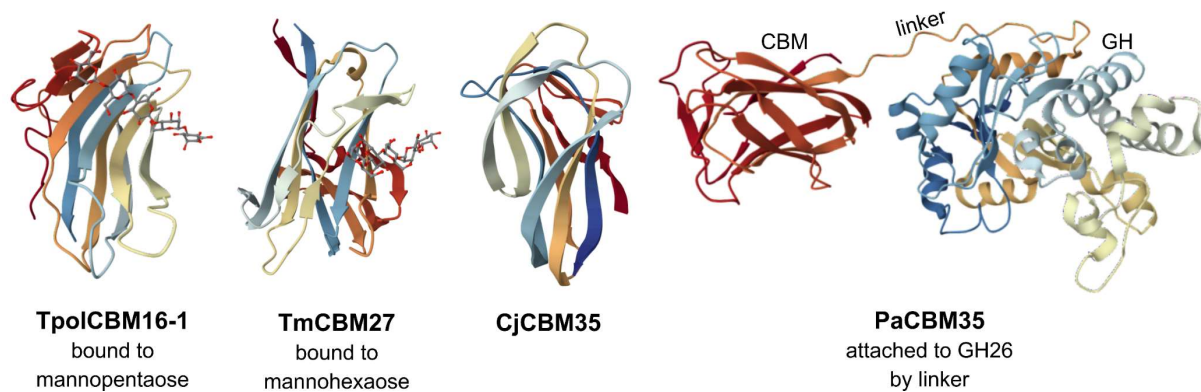


Figure 5 3-dimensional representation of candidate CBMs. TpolCBM16-1 and TmCBM27 are shown with mannopentaose and mannohexaose binding respectively. PaCBM35 is attached to its corresponding endo- β -1,4-mannanase 26 (GH26). 3D-structures were retrieved from the PDB database and adjusted for better visualization (<https://www.rcsb.org/>).

1.4.6 Potential Limitations of Recombinant CBM-Probes

In addition to the use as probes to label cell wall polysaccharides, recombinant CBMs can benefit a wide range of industries. Recombinant CBM technology can be used to enhance paper properties, improve the affinity of dye in the textile industry, help functionalize biomaterials in medicine and modulate plant growth, to name a few (Oliveira et al., 2015).

However, recombinant CBMs may face some challenges when developing a new CBM-based probe. The glycosylation process occurring in the heterologous host *Pichia pastoris* can sometimes reduce the substrate affinity of a recombinant CBM, ultimately reducing its potential as a labelling probe (Boraston et al., 2001; Oliveira et al., 2015). Further, it was reported that CBMs can sometimes irreversibly bind to cellulose, so re-using the probes would not be possible (Oliveira et al., 2015). CBMs can also be toxic to their hosts and lead to reduced biomass or phenotype alteration (Obembe et al., 2007; Zhang, 2014; Oliveira et al., 2015).

Another possible challenge is the accessibility of the pectin-embedded mannans *in vivo*. When immunolabelling was used to label heteromannans in cell walls, it was found that the pectin had to be removed first before the antibodies LM21 and LM22 could access and thus label the mannans in the plant cell wall (Marcus et al., 2010).

1.5 Model Organisms

1.5.1 *Arabidopsis thaliana*

Arabidopsis thaliana (*Arabidopsis*) is one of the most used model organisms in plant research. It is a small flowering angiosperm with a short generation time of approximately 12 weeks (Krämer, 2015; Provart et al., 2016). Major advantages of *Arabidopsis* include fast growth and short generation time, a high number of seeds, and easy transformation and crossing (The Arabidopsis Genome Initiative, 2000; Provart et al., 2016). The most significant advance in *Arabidopsis* research was the full sequencing of its relatively small genome in the year 2000, where “The Arabidopsis Genome Initiative” identified 25,498 genes, laying the base for more comprehensive and systematic research regarding plants and eukaryotes in general (The Arabidopsis Genome Initiative, 2000).

Different mutant lines, such as the mannan-deficient *cs1a2,3,9* mutant, were generated for HM research. They helped identify which roles CSLAs may play in synthesizing (gluco)mannans in *Arabidopsis* (Goubet et al., 2009; Liepman and Cavalier, 2012). Among other types of mutants, the triple mutant has since then been used to decipher the functions of further CSLAs and thus allowed a range of comprehensive studies regarding the biosynthesis of HM in *Arabidopsis* (Voiniciuc et al., 2015; Voiniciuc et al., 2018a; Terrett et al., 2019).

While *Arabidopsis* presents many advantages, it lacks industrial relevance. This limits the findings made using this model plant to being precursors to understanding and optimizing crops. Nevertheless, *Arabidopsis* is very useful for detecting general principles that could enable informed experimental design for crop research (Provart et al., 2016).

1.5.2 *Nicotiana benthamiana*

Nicotiana benthamiana (related to tobacco) is a model organism mostly used to investigate plant-pathogen interactions, as it is highly susceptible to virus infections. Since it is easy to silence its defense mechanisms and thus easy to transform transiently, the popularity of *N. benthamiana* as a model organism to study protein expression is constantly increasing (Goodin et al., 2008). In addition to studying protein expression, fluorescently tagged proteins can be used for localization studies. Even though *Arabidopsis* has a shorter generation time than *N. benthamiana* it still requires weeks before a result can be obtained (Krämer, 2015; Schultink et al., 2019). Compared to that, the method of transiently transforming *N. benthamiana* leaves yields results after only 2-5 days, making it one of the fastest ways to obtain information on protein expression in plants (Goodin et al., 2008; Krennek et al., 2015; Schultink et al., 2019).

However, studies suggest that there might be a genetic variability between the *N. benthamiana* accessions used in laboratories throughout the world, potentially leading to different results and conclusions on the same proteins in different laboratories (Goodin et al., 2008; Derevnina et al., 2019; Schultink et al., 2019). Therefore, research is ongoing in forward genetics to elucidate the genome of *N. benthamiana*. To quickly screen candidate genes and proteins prior to stable transformation in *Arabidopsis*, *N. benthamiana* is an excellent model organism (Schultink et al., 2019).

1.5.3 *Pichia pastoris*

Pichia pastoris (*Pichia*), also known as *Komagataella phaffii* (Barone et al., 2023) is a well-established heterologous host to study cell wall biosynthesis as it can produce heterologous hemicelluloses (Cocuron et al., 2007; Davis et al., 2010; Voiniciuc et al., 2019). The native cell wall components of *Pichia* primarily comprise β -1,3- and -1,6-glucans, mannoproteins, and trace amounts of chitin, which differ substantially from the heterologously produced β -1,4-linked HM, allowing to distinguish yeast native from heterologous polysaccharides easily (Aguilar-Uscanga, 2003; Voiniciuc et al., 2019).

Also, it is easy to manipulate genetically and grows very stably. As a eukaryotic organism, *Pichia* likely post-translationally modifies heterologously produced proteins, ensuring their correct folding and functionality. In addition to dextrose and glycerol, *Pichia* can also use methanol as a carbon source, which allows to induce the expression of certain genes at a specific time point using methanol-inducible promoters such as pAOX1. The latter and other strong promoters for foreign gene expression, as well as the ease in inserting foreign DNA into the genome of *Pichia*, enable easy and cost-efficient production of target proteins in higher amounts than in other eukaryotic and mammalian systems such as *E. coli*. The α -signal peptide from the yeast *Saccharomyces cerevisiae* enables heterologously expressed proteins to pass the secretory system of *Pichia* so that they can be targeted to its cell wall (Daly and Hearn, 2005; Graumann and Premstaller, 2006; Juturu and Wu, 2018; Zou et al., 2022).

Pichia-based protein production includes biopharmaceutical proteins, some of which are already approved by the US Food and Drug Administration (FDA), membrane proteins, and antimicrobial peptides. It is known that the glycosylation process in *Pichia* differs from the one in mammalian cells, which might affect the functionality of a heterologously produced drug or its availability to the human organism. Despite the mentioned limitations, *Pichia* is well suited to investigate hemicellulose biosynthesis (Gille et al., 2011; Juturu and Wu, 2018; Voiniciuc et al., 2019).

1.6 Objectives

This work comprises three main objectives that aim to enhance our understanding of heteromannan biosynthesis and their *in vivo* distribution and dynamics. First, it was investigated how heteromannans could be tailored for a desired use. Second a new *in vivo* HM-labelling tool was developed and verified in yeast. Third, proof of concept experiments were conducted. They aimed to verify the performance of the new HM-labelling tools for the high throughput search of new CSLA variants in yeast and to visualize HM *in planta*.

1. Use Domain Swaps to Modify (Gluco)mannan Biosynthesis in Yeast

Domains of the mannan synthase AtCSLA2 and the glucomannan synthase AkCSLA3 were swapped. Chimeric enzymes were expressed in *Pichia* for heterologous HM production and subsequent HM composition and fine structure analysis.

2. Develop CBM-Based Probes to Label Heteromannans

Mannan-binding CBMs TpoCBM16-1, TmCBM27, and PaCBM35 were recombined with the α secretion signal peptide and fluorescent proteins. CBM-probes were expressed in wild-type, mannan-producing, and glucomannan-producing yeast to assess if they would secrete in an HM-dependent manner and to visualize heterologous HM *in vivo*.

3. Applying CBM-Probes in Yeast and Plants

Verified strains from a library of randomly mutated AkCSLA3 were expressed alone and in conjunction with the new CBM-probes to demonstrate their ability to predict CSLA functionality. To demonstrate the CBM-probe abilities to label HM *in planta*, they were stably expressed in *Arabidopsis* lines with different mannan abundance.

2 METHODS

2.1 Gene Expression and Yeast Cultivation

2.1.1 Modular Constructs in *E. coli* and *Pichia*

Candidate mannan-binding CBMs were selected using the CAZy database, which aims to gather and classify an array of different carbohydrate-active enzymes such as glycosyl hydrolases (GH), glycosyltransferases (GT), Carbohydrate-Binding Modules (CBMs) (Lombard et al., 2014). Sequences were then retrieved from the UniProt database (The UniProt Consortium et al., 2023) (Figure 5). CSLA swap sequences were domesticated to remove undesirable type IIS cut sites (Robert et al., 2021). CBM sequences were optimized by removing restriction enzyme sites that are unwanted for GoldenPiCs cloning (ApaI, AscI, BamHI, BglII, BpiI, Bsp119I, BstXI, Eco31I, Eco72I, EcoRI, Esp3I, KpnI, MssI, NotI, SacI, SacII, SgsI, SpeI, XhoI). In addition, CBM sequences were codon optimized for optimal expression in *Pichia* using a codon usage frequency table (GenScript, 2023) before *de novo* synthesis by GeneArt from ThermoFisher Scientific.

CBM constructs and genes used for the CSLA Swap experiments for expression in *Pichia* were cloned using the GoldenPiCs Kit (gifted by Gasser/Mattanovich/Sauer group (Addgene kit #1000000133), which allows a modular assembly of genes of interest (GOI) with promoters, fluorescent proteins and terminators (Prielhofer et al., 2017). Except for pPICZ system-based constructs AtCSLA2, AkCSLA3-sfGFP, sfGFP-AkCSLA3, as well as ScGOS1-mRuby2 which were a gift by Annika Grieb-Osowski, the CSLA swap coding sequences, which were prepared by Julian Waldhauer and Fabian Stritt, and constructs containing CjCBM35 as well as pGAP:TmCBM27, the individual parts of each construct were first amplified via Phusion PCR using high fidelity Phusion DNA Polymerase by Thermo Fisher Scientific. The primers used for the CSLA swap project and the new CBM constructs can be found in Appendix 1. The individual parts were then assembled using the Golden Gate-based GoldenPiCs method (Prielhofer et al., 2017), with the exception that here, FastDigest restriction enzymes (*BsaI*, *BpiI*) were used in combination with a Thermo Fisher Scientific T4 DNA ligase.

For the CBM constructs, BB1_12 vectors were assembled with the methanol-inducible pFDH1 promoter only, or together with the α or the AtEXP10 signal peptide, except for TmCBM27 and CjCBM35 for which the α signal peptide was added to BB1_23. Otherwise, the CBMs were cloned into the BB1_23 vector on their own. The terminator module BB1_34 containing the ScCYC1tt terminator was customized to also contain either mRuby2 or the pmScarlet-I-2A-pSmTq2 fluorescent tag. The latter was assembled by fusing pmScarlet-I - to

which the first half of the P2A peptide and a custom fusion site was added with a Phusion primer – with pSmTq2 – to which the second half of the P2A peptide and the corresponding custom fusion site was N-terminally added in the same manner (example in Appendix 3). For assembly of the BB1 vector, 25 ng of each DNA part were added to a 10- μ L reaction mix and then incubated for a minimum of five 5-minute digestion cycles at 37 °C before ligation for 10 min at 22 °C. After final digestion at 75 °C for 10 min, the assembly mix was transformed into chemically competent *E.coli TOP10F'* using the heat-shock method.

To isolate the DNA, colonies resistant to the BB1-level selection marker kanamycin were selected for verification via colony PCR using genotyping primers (vector- and/or gene-specific) (Appendix 2) and Red Taq DNA polymerase master mix (VWR International). Positive colonies were grown in LsLB + antibiotic for approximately 24 h before DNA was extracted with the GeneJet Plasmid Miniprep Kit from Thermo Fisher Scientific. To rule out the presence of any potential undesired mutations, plasmids were verified via Sanger Sequencing over the complete GOI, using gene- or vector-specific primers (Appendix 2). Verified transcriptional CBM units were then assembled into the BB3rN_14 vector (Nourseothricin resistant) together with the pFDH1 promoter (unless stated otherwise) and - if no mRuby2 was present in the transcriptional unit - with the customized fluorescent terminators (example in Appendix 3). CSLA swap constructs were assembled into the BB3aZ_14 vector (Zeocin resistant) together with the methanol-inducible pAOX1 promoter and the RPP1Btt terminator. For stable expression, 150 ng of each BB3 plasmid were linearized prior to transformation in *Pichia X-33* or *Pichia* strains containing CSLA or CBM expressing backgrounds, using the condensed electroporation method (Lin-Cereghino et al., 2005). 3 days after growth on selection marker containing agar plates with YPDS medium, colonies were verified via colony PCR using the same primers used to verify the BB3 plasmids after DNA extraction from *E.coli*.

2.1.2 AkCSLA3-sfGFP Library Creation

The AkCSLA3-sfGFP Library was created by error-prone PCR of the AkCSLA3 portion of the gene carried out by AZENTA Life Sciences. Plasmids were created with at least 3 random mutations per 1000 bp across the 1557 bp AkCSLA3 sequence with a theoretical diversity of 10^7 . By verifying 48 clones, AZENTA confirmed the library accuracy with a confidence of $\geq 80\%$. After resuspending the dried 5 μ g of Plasmid DNA obtained, the heat shock method was used to transform 1 ng of DNA into *E.coli TOP10F'*. Then, 10 Zeocin-resistant colonies were selected for expression in *Pichia X-33* and the α -PaCBM35-pmScarlet-I-2A-pSmTq2 background and subsequent detailed analysis. To verify the accuracy of these 10 selected plasmids, DNA was sequenced via whole plasmid sequencing by Plasmidsaurus (<https://www.plasmidsaurus.com/>).

2.1.3 *Pichia* cultivation

For each experiment, 2-5 biological replicates of each *Pichia* strain to be analysed (except for Figure 12c and d, with one biological replicate) were grown in liquid culture for 24-72 h at 30 °C and 250 rpm in a Thermo Scientific MaxQTM 6000 shaker or 230 rpm in a 42 INNOVA® Incubator Shaker. Respective replicate numbers are specified in the figure legends. Unless stated otherwise, cells for the CSLA swap experiments grew for 48 h in 2 mL medium in reusable, autoclavable polypropylene square 24-deep well enzyscreen plates (EnzyScreen CR1424a) with metal covers (CR1224b), sealed with micropore tape. For experiments with the new CBM probes, 72 h and 300-600 µL culture volume in sterile round 24-well plates were the preferred growth conditions as they yielded the optimal results for CBM expression.

In this study, *Pichia* growth was optimized for high biomass and gene expression. Therefore, various growth conditions were tested. Preliminary experiments were mainly conducted using a combination of BMGY (buffered-glycerol-complex medium containing 100 mM potassium phosphate (pH 6.0), 1.25% glycerol (v/v), and the following ingredients in (w/v): 1% yeast extract, 2% peptone, 1.34% YNB, 0.0004% biotin) for biomass production and BMMY (replacement of glycerol with 0.5% methanol) to induce gene expression. When using BMGY/BMMY, cells were grown in 2 mL medium in sterile plastic culture tubes or sterile glass tubes with aluminium lids (14 mL) for 24 h in BMGY. Cells were then pelleted at 2000 g for 5 min to exchange BMGY to BMMY under sterile conditions. The same procedure was applied when cells were grown in YPD only or YPD/YPM (yeast-peptone base supplemented with 0.5% dextrose and 1.5% methanol). To determine the optimal carbon source, YPM was supplemented with the following carbon sources (v/v): 0.5% glycerol (YPM + G), 0.5% glucose (YPM + Glc), 0.5% mannose (YPM + Man).

After cultivation, an aliquot of each culture was saved for fluorescence measurements. For subsequent carbohydrate analyses, the remaining culture was transferred to 2 mL Eppendorf tubes and centrifuged at 2000 g for 5 min for cell harvest.

2.1.4 Growth Curve Analysis

The OD₆₀₀ of 2 pre-cultured biological replicates of each tested genotype were measured in a 1:10 dilution with water using a BioSpectrometer (Eppendorf) to analyse growth curves. Each biological replicate was split into three replicates grown in 300 µL YPM in sterile 48-well plates with a starting OD₆₀₀ of 0.1. Then, cells were further cultivated in a Tecan Spark 10 M fluorescent plate reader at 29 °C and 360 rpm with 1.5 mm orbital amplitude. The growth was monitored by measuring the OD₆₀₀ of each well 5 x 5 times in a filled circle pattern and a 700 µL border every 30 min until a total cultivation time of 72 h.

2.2 Fluorescence Analysis

2.2.1 Assessment of OD and FL in *Pichia* Cultures and Supernatant

After cultivation, OD600 (optical density = absorbance at 600 nm) and fluorescence measurements (Excitation/Emission nm: 434/474 (BFP), 485/511 (GFP), 569/593 (RFP)) were carried out with the Tecan Spark 10, Tecan M1000 or the Biotek Synergy H1. Cultures were measured undiluted in 48-well plates or diluted appropriately for the anticipated fluorescence and OD600 values in 96-well plates. For a detailed list of which dilutions were used for which experiments, see Appendix 4. When necessary, fluorescence was normalized to OD600 (FL/OD). Cultures were centrifuged for 2 min at 6000 g to obtain the supernatant fluorescence. Aliquots of the supernatant were measured the same way as the total culture. To calculate the CBM secretion, the fluorescence correction formula was adjusted to account for the high OD600 difference between the culture and the supernatant. The corrected red fluorescence was calculated as follows:

$$\text{corrected RFP} = (1 + \text{OD600}) * \text{uncorrected RFP}$$

2.2.2 *Pichia* Confocal Fluorescence Imaging

Microscopical imaging was done with several microscopes, namely laser scanning confocal Zeiss microscopes LSM700 (Figure 9, Figure 10) with an MBS405/488/555/639 beam splitter, LSM780 (Figure 25, Figure 26, Figure 28, Figure 38) with an MBS488/561 beam splitter, LSM900 (Figure 19, Figure 22, Figure 23, Figure 36), and a Leica SP5 (Figure 32). Cells were stained with Calcofluor White (CF), Propidium Iodide (PI), or Trypan Blue (TB), depending on the purpose of the experiment, by mixing cell dilutions with 0.01% (w/v) of dye.

Microscopy on the Zeiss LSM instruments was done using 20x, 40x, or 63x (water immersion) objectives using separate acquisition channels for transmitted light and different fluorescent proteins. Details on the settings for each fluorescent dye or protein imaged are listed in Table 2.

Table 2 Excitation and emission detection-wavelengths used for acquisition of fluorescent dye and protein signal.

FLUORESCENT DYE/ PROTEIN	MICROSCOPE	EXCITATION	DETECTION
Calcofluor White	LSM700/900	405 nm	420–550 nm
Calcofluor White	Leica SP5	405 nm	415-480 nm
Propidium Iodide	LSM900	639 nm	< 640 nm
Propidium Iodide	Leica SP5	543/633 nm	610-690 nm
Trypan Blue	LSM700	639 nm	< 640
pSmTq2	LSM780	405 nm	440-495 nm
pSmTq2	LSM900	405 nm	410-510 nm
Venus, sfGFP	LSM700/780	488 nm	490–555 nm
sfGFP, AlexaFluor488	LSM900	488 nm	510-570 nm
sfGFP	Leica SP5	458 nm	520-550 nm
pmScarlet-I	LSM780	561 nm	590-640 nm
mRuby2, pmScarlet-I	LSM900	561 nm	576-700 nm

2.2.3 Single Cell Fluorescence Analysis via Flow Cytometry

To analyse the fluorescence and size of single cells, CSLA library strains were first grown in YPM + G. For cytometry with the BD Accuri™ C6, cell densities were then adjusted to a concentration of 107 cells/mL using a Neubauer Chamber. The initial cell concentrations were then plotted against the measured OD600. This allowed to determine to which OD600 the cultures had to be adjusted in order to contain the desired cell concentration, which was 0.2 for the *Pichia* cells analysed in this work. Before cell dilution to the desired cell concentration, they were first stained with PI, which had previously been shown to be effective in labelling dead cells (Hohenblum et al., 2003).

For that, PI (0.01%) was added to the cultures in a 1:1 ratio and mixed for 5 min at room temperature. Then, residual PI was removed by 5 min centrifugation at 3000 g and decanting of the supernatant, followed by two washes with 1 mL water. After a final wash, the supernatant was removed, and the sample was re-filled with water to the starting volume.

For cytometry, 2 mL of solution for each strain were prepared, and 100,000 per strain were then analysed with the BD Accuri™ C6 (University of Florida, ICBR). For assessment of the sfGFP signal, laser excitation was 488 nm, and the fluorescence was acquired in the FL1 (FITC/GFP) channel, equipped with a 533/30 nm filter. Red fluorescence of PI was measured FL2 (PE/PI) channel using the same laser excitation but this time a 585/40 nm filter. In addition

to fluorescence, cell size was recorded via forward and side scattering (FSC, SSC). The data was analysed using the online analysis tool floreada (<https://floreada.io/analysis>). From this, fluorescence plots were exported and processed in Inkscape for better visualization. In addition, the relative number of cells per designated quadrant in the graph plotting green against red fluorescence was determined and plotted using GraphPad Prism 9.

2.3 Carbohydrate Analysis

2.3.1 Extraction of Alkaline Insoluble Polymers from *Pichia* Cell Walls

Extraction of alkaline insoluble (AKI) polymers from *Pichia* cell pellets was carried out essentially as described previously ((Voiniciuc et al., 2019), this time using a neoMix 7-921 thermomixer. After treatment with hot NaOH for 90 min, samples were neutralized with acetic acids before washing them with water to remove alkaline soluble polymers and residual NaOH and acetic acid. Pellets were resuspended in 600 μ L in water and homogenized with small glass beads(Sigma Aldrich, Cat# G8772-500G) and a ball mill (Retsch MM400 mill) in a ball mill (except for the experiment described in 3.2.2.4, where volumes of NaOH and acetic acid, as well as resuspension volumes, were reduced 12-fold for samples with very low cell density) and analysed directly, refrigerated at 4°C for short, or frozen at -20 °C for long term storage.

To further reduce the amount of native *Pichia* cell wall polymers and, by that, enrich their heterologously produced (gluco)mannan containing portion AKI samples were enzymatically treated with Zymolyase 20T[®] (from *Arthrobacter luteus*; USBiological) to remove unwanted β -1,3-glucans and obtain the enriched mannan fraction (EM). Three biological replicates of AKI samples were pooled, pelleted, and resuspended in 300 μ L potassium phosphate buffer (PBS, 0.2 M, pH 7.0). Then, 300 μ L water-based enzyme mixture with 10 μ g sodium azide and 125 μ g Zymolyase 20T[®] were added to the samples prior to mixing for 48 h at 37 °C and 250 rpm in a MaxQ 6000 shaker (Thermo Scientific). Pellets were obtained by centrifugation for 5 min at 16,000 g and washed twice with 1 mL of water. To further clean the samples, they were mixed gently with 300 μ L of acetone before and dried to ensure that no material would be lost. After resuspending the dry samples in 1000 μ L water, they were homogenized using a ball mill before direct analysis or storage as described above.

2.3.2 Sulfuric Acid and TFA Hydrolysis

For monosaccharide composition analysis, polysaccharides of AKI material were fractionated into monosaccharides by acid hydrolysis with either 4% sulfuric acid (Yeats et al., 2016) (Figure 7, Figure 16, Figure 28) or 2 M TFA (Figure 8, Figure 21, Figure 30, Figure 31).

In both cases, 50 μL AKI were used as starting material and 800 μL of 30 $\mu\text{g}/\text{mL}$ Ribose solution were added as internal standard, except when digested samples were analysed. In these cases, 150 μL of supernatant from the digested samples was used.

For the one-step hydrolysis with 4% sulfuric acid, samples were mixed with 72% sulfuric acid (final concentration of 4%) and heated for 60 min at 120 $^{\circ}\text{C}$ in a Techne™ Dri-Block™ in 2 mL screw cap tubes to ensure that the tubes would stay closed during that process. Standards consisting of 0, 1, 2, 5, 10, 25, 50, 75, 100, and 125 μg sugar standard (Glc, Man, Gal) were prepared the same way. After acid treatment, samples and standards were cooled on ice and centrifuged at 20,000 g for 15 min to collect all material at the bottom of the tubes.

When samples were partially digested or required a gentler acid treatment due to low amounts of starting material, 2 M TFA hydrolysis was employed. For digestion, 50 μL of AKI were mixed with 100 μL of 0.2 M potassium phosphate buffer (pH 7.0) with 1 U of endo-1,4- β -Mannanase (Megazyme, E-BMABC) and then digested at 40 $^{\circ}\text{C}$ for 30 min in a thermomixer. Then, samples were centrifuged for 2 min at 20,000 g, then dried in a Techne™ Dri-Block™ with gentle air or nitrogen flow. 100 μL of supernatant could then be hydrolysed with 2 M TFA.

For hydrolysis, dry samples were resuspended in 150 μL 2 M TFA, followed by a 90 min hydrolysis step at 120 $^{\circ}\text{C}$. Then, samples were centrifuged as described for the 4% sulfuric acid treatment and dried again before resuspension in 300 μL Ribose solution with a ball mill. Then, samples were briefly centrifuged again at 20,000 g to collect any insoluble material at the bottom of the tubes. For monosaccharide analysis, 150-200 μL of sulfuric acid or TFA hydrolysed samples and standard were transferred to IC vials.

For oligosaccharide profiling of EM samples, E-BMABC was additionally heat inactivated for 10 min at 90 $^{\circ}\text{C}$ and the samples were not hydrolysed but directly analysed.

2.3.3 Monosaccharide Analysis and Oligosaccharide Profiling

Monosaccharides of hydrolysed samples were separated via HPAEC-PAD using a Metrohm 940 Professional IC Vario system equipped with a Metrosep Carb 2–250/4.0 analytical and guard columns. Following the injection of 10 μL sample, carbohydrates were separated into Glc, Man, and Gal using a 30 min protocol that consisted of a 20 min isocratic 2 mM sodium hydroxide (NaOH) + 3.4 mM sodium acetate (NaAce) separation step before a rinse with 80 mM NaOH + 136 mM NaAce for 3 min, followed by 4 min of re-equilibration with 2 mM sodium hydroxide (NaOH) + 3.4 mM sodium acetate (NaAce). Using the MagIC Net 3.2 software by Metrohm, calibration and integration of detected monosaccharide peaks was carried out automatically unless manual correction was required.

To analyse oligosaccharide profiles, the eluent gradients were adjusted to 15.6 mM NaOH, followed by an increase to 78 mM NaOH for 5 min, before a 25 min linear increase to 78 mM NaOH + 50 mM NaAce. Following these adjustments, a 15-minute re-equilibration period was given under conditions of 15.6 mM NaOH.

2.3.4 Linkage Analysis

The linkages between the individual monosaccharides were analysed to elucidate the fine structure of heterologously produced (gluco)mannans. This involved sample methylation (Ciucanu and Kerek, 1984) followed by 2 M TFA hydrolysis and then reduction and acetylation to PMAAs using a pipeline similar to the one described earlier (Pettolino et al., 2012).

100 μ L of AKI samples and 1mg/mL polysaccharide standards samples were dried in glass tubes and mixed overnight with 200 μ L of DMSO (pre-dried with molecular sieves). While maintaining an N₂ atmosphere, polymers were methylated with 200 μ L of a NaOH/DMSO slurry and 100 μ L of methyl iodide for 2–3 hours, followed adding 2 mL water to quench the reactions. Then, samples were gently bubbled with N₂ bubbling until the solution turned clear before adding 2 mL of dichloromethane. Subsequently, around 1.5 mL of the organic phase of each sample was transferred into a new tube and dried.

After 2 M TFA hydrolysis of the methylated polymers into monomers, samples were dried again, and myo-inositol was added as an internal standard. Using 200 μ L of fresh sodium borodeuteride (10 mg/mL) in 1 M ammonium hydroxide, samples were reduced at room temperature for 60 min and then neutralized with acetic acid and a series of methanol washes. Following a drying step, samples were acetylated with 50 μ L acetic anhydride and 50 μ L pyridine for 20 minutes at 120 °C. Next, the obtained PMAAs were dried and washed twice with 200 μ L of toluene. After a final cleaning with 1.2 mL ethyl acetate and 5 mL water, the samples were dried in fresh tubes and resuspended in 300 μ L acetone.

GC-MS analysis of 2 μ L of organic PMAAs per injection was performed using an Agilent Technologies 6890 N GC system, coupled with a Supelco SP-2380 column (30 m \times 0.25 mm \times 0.2 μ m) and an Agilent 5975 quadrupole EI detector. The initial temperature setting was 80 °C, held for 3 minutes, then ramped up to 170 °C (at 30 °C/minute), and then to 240 °C (at 4 °C/minute) with a hold time of 15 minutes per run.

PMAAs underwent semi-automated quantification utilizing Agilent MSD Chemstation Classic Data Analysis (G1701FA). This process was grounded on the retention time of the glycosidic linkage peaks sourced from polysaccharide standards alongside their respective ion spectra or data found in the CCRC Spectral Database for PMAAs (<https://www.ccrcc.edu/specdb/ms/pmaa/pframe.html>).

2.4 Plant Methods

2.4.1 Plant Growth Conditions

N. benthamiana plants were cultivated from seeds of non-treated parental plants, sown directly in vermiculite soil in 10 cm pots, and grown in long-day conditions (16 h of light), with light supplied from both natural and artificial sources at room temperature (18-22 °C).

For stable transformation or growth to maturity *Arabidopsis* Col-0 and the *csla2,3,9* mutant (Goubet et al., 2009; Liepman and Cavalier, 2012) were sown on vermiculite soil in round 8 cm pots and stratified for 2-3 days in the dark at 4 °C. After that, they were grown in a phytochamber with constant light of 100–120 $\mu\text{mol m}^{-2}\text{s}^{-1}$ 22°C and 60% relative humidity. The stable transformation was done on inflorescences of plants that started to flower (2-4 weeks post sowing) and then grown to maturity (10-12 weeks total growth time).

For root imaging via confocal microscopy, *Arabidopsis* seeds were transferred to 1.5 mL Eppendorf tubes and sterilized. This was done by washing them with 500 μL 70% EtOH, 0.5% Triton X-100 solution, followed by a wash with 500 μL absolute EtOH. During each wash, seeds were gently shaken for 10 min. After EtOH removal, seeds were dried on a heat block at 40 °C under sterile conditions for 20 min or until all EtOH was evaporated. Then, seeds were sown on 0.85% square agar plates containing 0.5 x MS adjusted to pH 5.7 with 10 M KOH (2.155 g Murashige Skoog (MS - Duchefa M0221.0050) + 0.25 g 2-(N-Morpholino) ethanesulfonic acid (MES - Sigma (M8250-100g) per liter). Then, plates were stratified for 3 days prior to growth in plant growth chambers (PERCIVAL) at 21 °C, 100 $\mu\text{E m}^{-2}\text{s}^{-1}$ light intensity, and 14 h light per day. Confocal imaging was then performed on 5-7 days old seedling roots.

2.4.2 Modular Construct Assembly and Transformation in *A. tumefaciens*

For expression in *N. benthamiana* and *Arabidopsis*, constructs were first cloned using the modular cloning system MoClo which allows the modular assembly of multigene constructs (Weber et al., 2011). For that, α -TpolCBM16-1-pmScarlet-I-pSmTq2, α -PaCBM35-pmScarlet-I-pSmTq2, AtEXP10-TpolCBM16-1-pmScarlet-I-pSmTq2, AtEXP10-PaCBM35-pmScarlet-I-pSmTq2, α -pmScarlet-I, and pmScarlet-I coding sequences were first amplified via Phusion PCR to introduce the FsB and FsC fusion sites and assemble them into the MoClo level 0 vector pAGM4031 (example in Appendix 3). Phusion PCR was carried out with gene-specific primers that would introduce the FsB and FsC fusion sites (Appendix 1), a Thermo Scientific high fidelity Phusion DNA polymerase, and 20 fmol DNA in a total volume of 20 μL . Initial denaturation at 98 °C for 30 sec was followed by 6 cycles of annealing and extension for 10

sec at 98 °C and 1:50 min at 72 °C. Then, 30 more cycles of annealing and extension were performed this time for 30 sec and 1:30 min, respectively, before final extension at 72 °C for 5 minutes and cooling to 12 °C. Subsequently, the PCR products were run on an agarose gel to recover the Phusion amplicons (0.7% agarose concentration, 50 min in 1xTAE, 100V).

After cutting out the Phusion amplicons from the gel, they were cleaned using a Zymoclean Gel DNA Recovery Kit and eluted in 12 µL HyClone water. Then, amplicons were assembled into the MoClo level 0 vector pAGM4031 using 25 ng/µL DNA dilutions and following the procedure described in 2.1.1. After genotyping (for primers, see Appendix 2) and sequence verification via Sanger Sequencing, the pAGM4031 vectors containing the CBM or pmScarlet-I control genes were assembled with the pICH75055 backbone, the pICH51277 vector that carried the constitutive p35S promoter and the pICH41432 vector containing the tOCS terminator (example in Appendix 3).

The resulting Level 1 constructs were transiently expressed in *N. benthamiana* (see 2.4.3). For stable expression in *Arabidopsis* (see 2.4.4), the pICH75055 vectors harbouring the CBM or pmScarlet-I control were assembled into the MoClo Level M vector pAGM8031 together with an RFP selection marker cassette (pAGT56318) and the end linker vector pICH50881 (example in Appendix 3). Plasmids were transformed into *E. coli* to generate DNA before transformation in *Agrobacterium tumefaciens* (GV3101) for plant expression.

Plasmids were transformed into *A. tumefaciens* calcium competent cells, which were prepared based on a previously described protocol (Berestovoy et al., 2018). Instead of 95 mL LsLB medium for culture growth, 50 mL were used, and the solutions to induce calcium competency were simplified to only 20 mM calcium chloride. Calcium-competent cells were used directly or frozen at -70 °C in 50 µL aliquots for long-term storage.

For transformation, 200 ng of each plasmid DNA and 100 ng of the helper plasmid pSOUP (Hellens et al., 2000) were added to the competent cells, gently mixed, and put on ice for 15 min. Tubes were frozen in liquid N₂ for 5 min, followed by a 5 min heat shock at 37 °C in a water bath. After returning to ice for 5 min, 1 mL LsLB was added, and the mixture was incubated by shaking at 30 °C for 4 h. The transformed cell mixtures were centrifuged for 5 min at 2000 g, and the supernatant removed so that approximately 200 µL were left in the tubes. Cells were carefully resuspended, and 100 µL plated on LsLB plates with appropriate antibiotics. After growing for 3 days at 30 °C, selection marker-resistant colonies were selected for genotyping (vector- and/or gene-specific (Appendix 2)) and transformation into plants.

2.4.3 Transient Expression in *N. benthamiana* and Leaf Plasmolysis

A. tumefaciens with GOIs were used as chassis to transiently express genes in *N. benthamiana* leaves using a previously established protocol (Grefen et al., 2010). After pre-culture growth and mixing with the P19 viral suppressor, both adjusted to 0.7 OD600, *N. benthamiana* leaves of 5-6 weeks old plants were infiltrated with the *A. tumefaciens* mixture using 1 mL syringes. Each tested strain was infiltrated in at least three individual spots on different leaves from 3-4 different plants to ensure that observed fluorescence intensities and localizations were not caused by the specific leaf type or plant. After three days, leaf discs of 0.5 cm diameter were excised from transformed spots for confocal fluorescence imaging.

To discern potential localization differences of the expressed genes, notably between the fluorescent proteins alone versus with the α signal peptide and between α and AtEXP10 in TpolCBM16-1 and PaCBM35, leaf discs were plasmolysed with 1 M NaCl solution. For that, leaf discs were immersed in 1 M NaCl in a 2 mL Eppendorf tube. Immersing was done by placing a 50 mL syringe on top of the tube so that no air could penetrate the tube. Then, the syringe was drawn up to create a vacuum in the tube. Once the syringe was released, the leaf discs got infiltrated with the salt solution. After plasmolysis, leaf discs were imaged directly via confocal microscopy with an LSM900, as described in 2.2.2, using a 20x objective.

2.4.4 Stable Gene Expression in *Arabidopsis*

For stable gene expression, *Arabidopsis* Col-0 and *cs1a2,3,9* (Goubet et al., 2009; Liepman and Cavalier, 2012) were transformed using the floral dip transformation method, and *A. tumefaciens* strains carrying the genes of interest to generate generation T₀. For that, *A. tumefaciens* strains were grown overnight at 28 °C in 20 mL YEP medium (5 g bacto peptone + 5 g bacto yeast extract + 2.5 g NaCl in 50 mL water). Then, cultures were centrifuged for 15 min at 2500 g before decanting the supernatant and resuspending the cells in 20 mL of a 5% sucrose and 0.05% Silwett L-77 solution. Inflorescences of *Arabidopsis* plants were dipped carefully into the transformation solution, then laid down and covered overnight. Finally, plants were uncovered the next day and grown to maturity.

Once dry, seeds were harvested and selected based on the RFP signal emitted by successfully transformed seeds using a Leica M165 FC with a DsRED filter (excitation filter: 545/30 nm (510–560 nm); barrier filter: 620/60 nm (590–650 nm)). Per genotype, approximately 40 seeds with RFP signal were selected and transferred to 1.5 mL Eppendorf tubes. Then, seeds were sterilized and grown on vertical agar plates (see 2.4.1) for 5 days prior to screening via confocal imaging as described in 2.2.2. using a 40x objective. For imaging, seedlings were mounted onto a 60 mm coverslip with 0.5 x MS medium, and covered with a 32 mm coverslip.

3 RESULTS

3.1 CSLA Domain Swaps

3.1.1 Engineering of Cellulose Synthase-Like Enzymes and Cell Walls

Data from this chapter was published in Biotechnology for Biofuels (Robert et al., 2021)

The first step to designing new chimeric proteins with the potential to modify (gluco)mannan production and quantity involved comparative analyses of the parental AtCSLA2 and AkCSLA3 sequences. Topology predictions were performed using TOPCONS (Tsirigos et al., 2015), which employs a range of algorithms to predict protein topologies. AtCSLA2 and AkCSLA3 exhibited highly similar topologies. Not only do they both possess five transmembrane domains, but they also share a cytosol-facing catalytic site. Furthermore, they share three regions with identical amino acid motifs. These shared regions were used as borders to divide each protein into four parts (Figure 6). The similarity between the four regions ranged from 38–73% amino acid similarity, with the highest similarity (73%) predicted for the conserved Pfam PF13641 GT2 domain, which transfers glycosyl residues. The topology predictions indicated that swapping domains between AtCSLA2 and AkCSLA3 could potentially result in alterations of the heteromannans produced.

Domesticated AtCSLA2 and AkCSLA3 sequences were partitioned into the four designated parts to verify that hypothesis. Subsequently, the sections were seamlessly reassembled into single-domain swaps using the GoldenPiCs modular cloning toolkit (Prielhofer et al., 2017), creating eight new mutation-free combinations. Each of these new combinations was labelled based on the domains it consisted of. For instance, a chimeric AtCSLA2 protein in which the first domain had been replaced by the first domain of AkCSLA3 was labelled as 3222 (Figure 6). Each construct was equipped with the strong methanol-inducible promoter pAOX1 for optimal expression. Linearized constructs were then stably introduced into *Pichia* for stable expression prior to cell wall analysis.

In an attempt to simplify the comparison of their products, AtCSLA2, AkCSLA3, and their respective chimeras, were C-terminally fused to a superfolder green fluorescent protein (sfGFP). However, while AkCSLA3 tolerated the fusion to sfGFP, AtCSLA2-sfGFP failed to produce elevated amounts of mannan (Robert et al., 2021). Therefore, the subsequent experiments were carried out with untagged parental and domain-swapped CSLAs. Cells were then grown according to a streamlined protocol using a growth medium based on yeast extract and peptone (YP) supplemented with at least one carbon source. After 48 h of growth, the

alkaline-insoluble residue from the cell walls of the engineered yeast strains was isolated, fractionated into monosaccharides, and analysed (Figure 7a).

Both AtCSLA2 and AkCSLA3 were able to produce substantial amounts of alkaline-insoluble mannose even when only methanol (M) was added to the growth medium. However, adding glycerol to the medium (YPM + G) resulted in a substantial biomass increase of 60% on average after 48 h of growth (Robert et al., 2021). Therefore, YPM + G was used for cultivation in the following experiments. Given that independent transformants of the same constructs yielded comparable heteromannan quantities (Robert et al., 2021), the focus was set on the top-producing colonies of the parental CSLAs and the chimeric proteins.

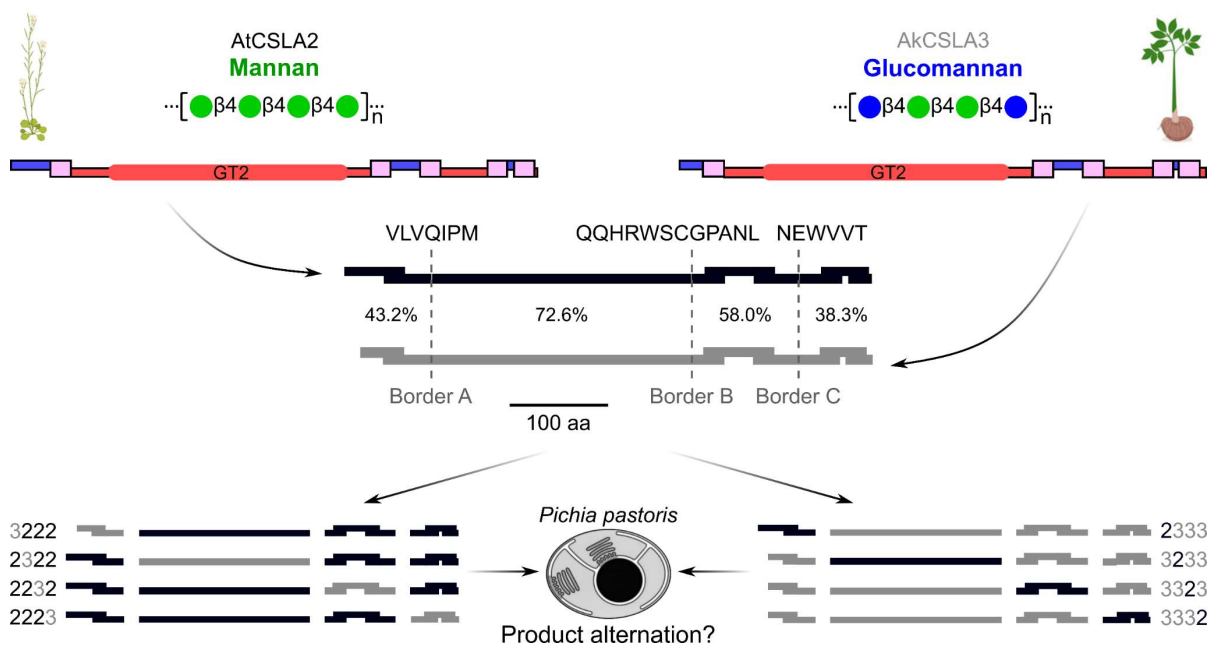


Figure 6 Modular engineering of heteromannan biosynthesis. The topologies of AtCSLA2 and AkCSLA3 enzymes were visualized with TOPCONS 2.0, which shows transmembrane domains (pink boxes), and regions inside (red) the membrane (Golgi) and outside (blue) the membrane (cytosol). GT2 is the conserved Pfam PF13641 GT2 domain. Dashed lines delineate the three regions with 100% amino acid identity that were selected as borders for single domain swapping of CSLAs. The DNA of the chimeric CSLAs was assembled using the GoldenPiCs cloning system. Then, chimeras were expressed in the methylotrophic yeast *Pichia pastoris* followed by cell wall analysis.

Different effects on the amount of mannose produced could be observed depending on the swapped domains. When the first or the last domains of AtCSLA2 (3222, 2223) or the first, N-terminal domain, of AkCSLA3 (2333) were swapped, the mannose production was significantly reduced compared to the parental CSLAs (Figure 7b). These findings align with the reduced mannan production observed upon C- or N-terminal tagging with fluorescent proteins, suggesting that these domains play a crucial role in CSLA functionality. The levels of mannose observed in the 3233 and the 3323 constructs that contained the second or third domain of AtCSLA2 were higher than the negative Venus control, yet they did not reach the mannose levels observed in the parental CSLAs. Interestingly, despite most domain swaps reducing the heteromannan yield, three chimeras, namely 2322, 2232, and 3332,

outperformed the parental CSLAs. Remarkably, after digestion with endo- β -mannanase, the 2232 chimera exhibited mannan quantities reaching 1 mg per well (2 mL) - a level that is 2.4 times greater than that of AtCSLA2 and 1.6 times more than that of AkCSLA3 (Robert et al., 2021). This is a strong indicator that domain swaps could be a promising approach to discovering new and more performant CSLA variants.

To elucidate how the fine structures of (gluco)mannans produced by the chimeric CSLAs would compare to the parental structures, alkaline-insoluble polymers were subjected to derivatization to partially methylated alditol acetates (PMAAs) for linkage analysis using gas chromatography-mass spectrometry (GC-MS) (Pettolino et al., 2012). In the Venus negative control, most of the PMAAs (> 80%) could be attributed to native yeast polymers consisting of 3- and 6-Glc (unbranched) and 2,3- and 3,6-Glc (singly branched). Also, only minor quantities of unbranched 4-Man were detected, underscoring the cleanliness of the *Pichia* background when it comes to the analysis of heterologously produced unbranched β -1,4-linked mannans. This is in line with previous findings where strains without plant (gluco)mannan production showed similar compositions (Voiniciuc et al., 2019). As expected, the parental CSLAs AtCSLA2 and AkCSLA3 showed elevated amounts of 4-Man, accounting for their ability to produce (gluco)mannan. In addition, 4,6-Man, an indicator of branched mannan, was found to be below 1%, suggesting that the produced (gluco)mannan was mostly unbranched (Robert et al., 2021).

However, it remained to be elucidated how the degree of polymerisation (DP) of the (gluco)mannan produced by the chimeric enzymes would compare to the AtCSLA2 and AkCSLA3 parents as well as insoluble mannan extracted from the ivory nut. Given that the molecular weight of the produced (gluco)mannans can only be measured when they are fragmented (Voiniciuc et al., 2019), the polymer length was estimated from the ratio between the linkages attributable to heteromannan, i.e., 4-Man and 4,6-Man and terminal mannan located and the chain ends (t-Man). Compared to the Venus negative control (DP < 1), all expressed enzymes caused an increase in DP. However, only the top-performing swaps 2322, 2232, and 3332 achieved DPs of 13-22, similar to the ones of the parental enzymes (Figure 7c), accounting for the reduced ability of the other chimeric enzymes to synthesize long (gluco)mannan chains. Considering that ivory nut mannan is expected to have a DP of 15-20 (Megazyme Knowledge Base, Product P-MANIV), the observed chain lengths for the top five (gluco)mannan were in a plausible range.

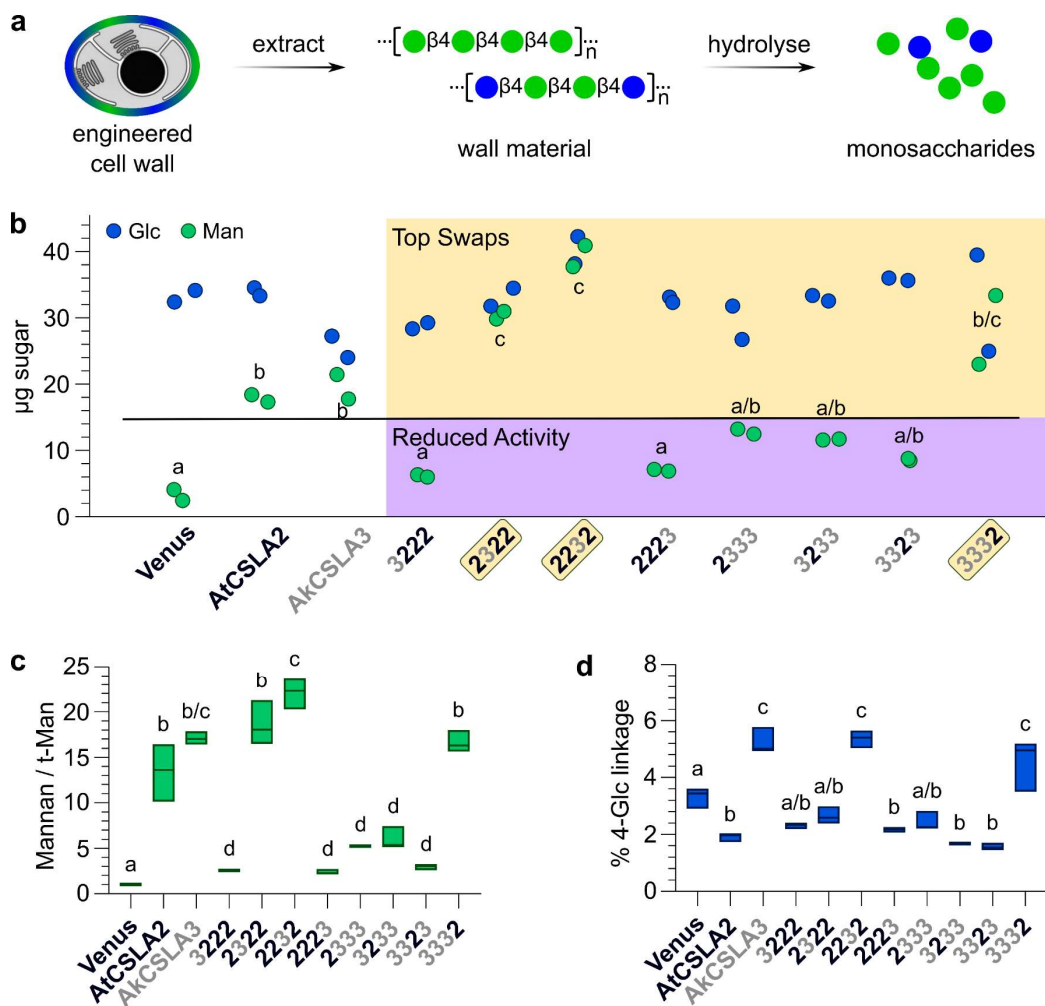


Figure 7 Abundance and composition of engineered *Pichia* cell wall polymers and influence of AtCSLA2 and AkCSLA3 domain swaps on (gluco)mannan linkages. (a) monosaccharide extraction and analysis from engineered yeast cell walls. Alkaline-insoluble residues (AKI) are hydrolysed into monosaccharides for analysis using sulfuric acid or TFA. (b) Glucose (Glc) and mannose (Man) abundance of parental AtCSLA2 and AkCSLA3 compared to chimeric enzymes and the negative Venus control (sulfuric acid hydrolysis of AKI). Cells grew in YPM + G (YP + 0.5% glycerol + 1.5% methanol) for optimal biomass production. Dots show values of 2 biological replicates and letters denote significant differences between samples obtained using one-way ANOVA with Tukey test, $P < 0.05$. Note that in (b) all Glc levels were similar. (c) Ratio between total mannan (4-Man + 4,6-Man) and terminal-Man (t-Man) in alkaline-insoluble polymers calculated following their derivatization to partially methylated alditol acetates. (d) relative amounts of 1,4-linked Glc compared to the total area of glycosidic linkages. Boxes show min to max range of 3 biological replicates with the inner horizontal line showing the median value. Letters denote between samples obtained using one-way ANOVA with Tukey test in with $P < 0.001$ in (a) and $P < 0.05$ in (b).

Elevated amounts of 4-Glc above 4.5%, indicative of glucomannan presence (Pauly et al., 2019), were only found in AkCSLA3 and the 2232 and 3332 swaps (Figure 7d). Digestion of alkaline-insoluble residues confirmed the glucomannan production by AkCSLA3 and the 3332 swap. In that context, no increased Glc was found for the 2232 swap indicating its incapability of β -1,4-glucomannan production in yeast (Figure 8).

3.1.2 Influence of High Yielding Chimeras on (Gluco)mannan Composition

A more detailed structure and characteristics analysis was performed on the swaps 2322, 2232, and 3332, as these yielded the highest amount of (gluco)mannan. Contrary to the observations made upon linkage analysis (Figure 7c and d), the digestion of alkaline-insoluble residues with endo- β -mannanase showed elevated Glc only for the AkCSLA3 parent and the 3332 swap. This was in line with the results from linkage analysis on pooled samples that had been enriched for (gluco)mannan by Zymolyase 20T[®] to further minimize the yeast Glc background (Robert et al., 2021). The AtCSLA2 parent and the swaps 2322 and 2232 did not show such an increase in Glc but exhibited elevated amounts of Man (Figure 8b). To gain deeper insight into the Glc:Man ratio in the top-performing constructs compared to the parental CSLAs, oligosaccharide profiling was carried out on enriched samples partially digested with endo- β -1,4-mannanase via HPAEC-PAD (Figure 8c and d). As controls, commercial konjac glucomannan (KGM) and ivory nut mannan (INM) underwent the same treatment. As predicted in the literature (Verherbruggen et al., 2021), KGM showed not only the small manno-oligosaccharides with a DP \leq 4, which was also observed in INM but also displayed additional peaks indicative of oligosaccharides with a DP \geq 6 (Figure 8a). The absence of additional peaks in the INM control was corroborated by the fact that they remained consistent even when a tenfold higher amount of INM was analysed (Robert et al., 2021). Quantifying the peak areas revealed that 48% of the total peak area in the KGM sample could be attributed to the glucomannan diagnostic peaks, whereas 99% of INM total peak area were attributable to the smaller mannan peaks (Figure 8d).

The oligosaccharides derived from the parental CSLAs (AtCSLA2, AkCSLA3) respectively resembled the profiles of the INM and KGM standards (Figure 8c). With 4–7% of the total peak area, the glucomannan peaks detected for 2322 and 2232 did not significantly differ from the AtCSLA2 profile, consistent with the previous findings that swaps did not cause significant product alteration (Figure 8c and d). With 22–26% of the total peak area of AkCSLA3 and 3332 being attributable to glucomannan oligosaccharides, these enzymes reached around 50% of the KGM control. Consequently, the GT domain bearing regions 2 and 3 do not seem sufficient to alter the mannan composition produced by AtCSLA2. However, co-expression of the CSLAs with the MSR co-factor was able to boost (gluco)mannan production and increase glucose incorporation during synthesis (Robert et al., 2021), demonstrating that, although the CSLA swaps themselves were not able to alter the (gluco)mannan fine structure, it is possible to adjust these structures through the introduction of additional co-factors.

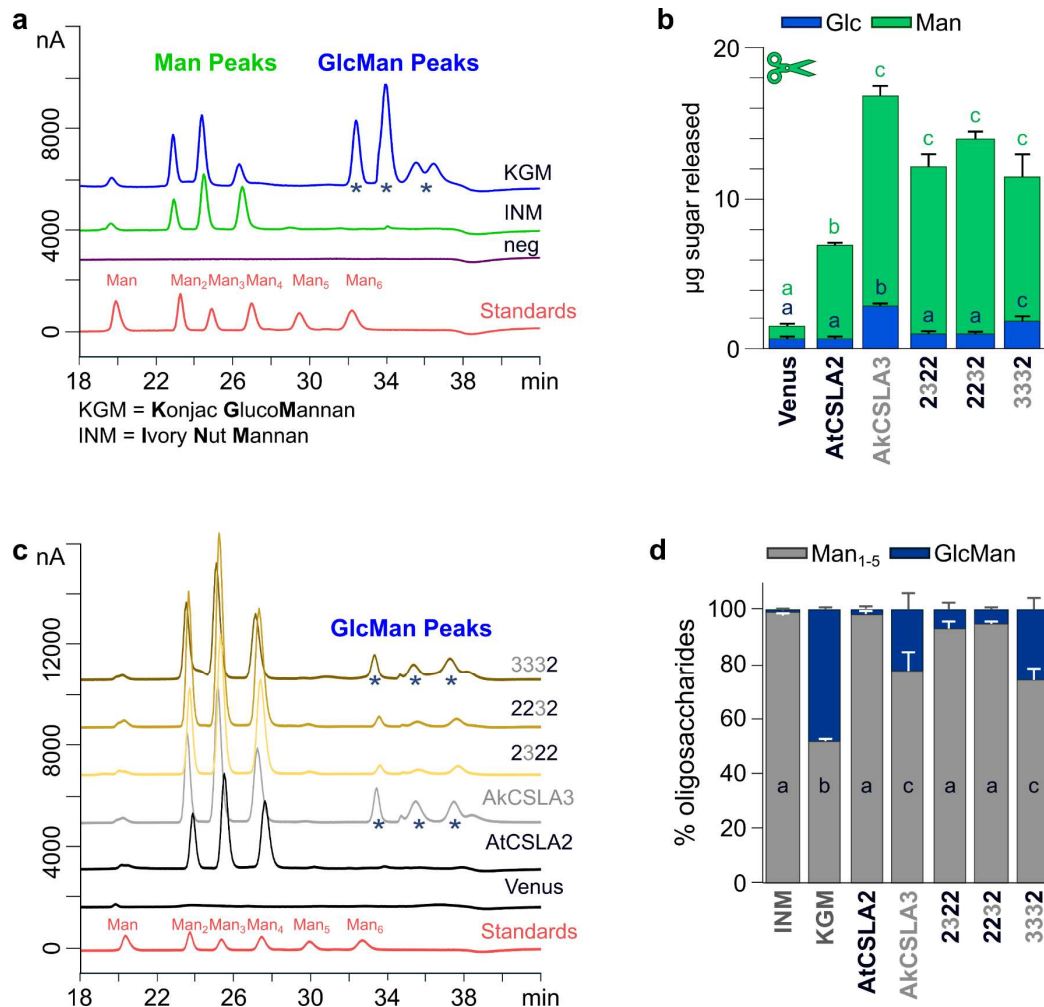


Figure 8 Fine-structure of (gluco)mannans produced by the top chimeric CSLAs. (a) Konjac glucomannan (KGM) and ivory nut mannan (INM) oligosaccharide profiling with glucomannan (GlcMan) diagnostic peaks marked by asterisks. (b) Carbohydrates released by digestion of alkaline-insoluble polymers with endo- β -1,4-mannanase. Solubilized carbohydrates were hydrolysed with TFA and analysed with HPAEC-PAD. Data show the mean + SD of 3 re-grown biological replicates of the top chimeric CSLAs based on results shown in Figure 7. (c) HPAEC-PAD oligosaccharide profiles of mannanase-treated samples quantified in (d). Peaks diagnostic for GlcMan are highlighted by asterisks. (c) Relative peak area of mannan (Man₁₋₅) and glucomannan (GlcMan) oligosaccharides released from *Pichia* EM compared to INM and KGM standards. Data show the mean + SD of 2 measurements. In (b) and (d), letters denote significant differences between samples and were determined using one-way ANOVA with Tukey test, $P < 0.05$.

3.1.3 Impact of CSLA Expression on Yeast Growth and Morphology

A crucial aspect of heterologous polymer production is the influence of gene expression and/or product on the host organism. For instance, any potential toxicity to the host could reduce the overall product yield or even be lethal, thus completely abolishing production. Therefore, the effects of producing linear plant (gluco)mannan in *Pichia* on biomass and morphology were investigated.

As expected, the biomass production of strains carrying the engineered or native CSLAs did not experience changes when the genes were not expressed. Prolonged overexpression of the Venus control did not affect biomass. Conversely, overexpression of

AkCSLA3 in an inducible medium containing methanol (YPM and YPM + G) resulted in a significant reduction of biomass regardless of the utilized cloning system (data not shown) (Voiniciuc et al., 2019; Robert et al., 2021). Coupled with the fact that previous attempts to constitutively express AkCSLA3 in yeast failed (Prielhofer et al., 2017), these results imply that high levels of linear glucomannan might exert a toxic effect on the cells.

To further investigate this phenomenon, the growth patterns of the top-performing chimeras 2322, 2232, and 3332 and the parental AtCSLA2 and AkCSLA3 were closely monitored over a 70 h period and compared to the Venus negative control (Figure 9a). The objective was to identify when the growth of AkCSLA3 would start to diverge from that of the Venus control and the other tested enzymes. The Venus fluorescence served as a control for gene expression (data not shown). The growth curves, obtained by measuring the optical density at 600 nm (OD₆₀₀), revealed that AkCSLA3 started deviating from the other CSLAs and the Venus control after around 24 h. This deviation became more evident when the optical density of AkCSLA3 plateaued at about 70% of that of the other strains. This pattern remained consistent, even after prolonged expression for 60 h, as evidenced by the persistent 30% difference in optical density between AkCSLA3 and the Venus control.

To investigate the cell morphology, they were stained Trypan Blue (TB), which labels β -glucans (Suzuki et al., 1997). While no obvious morphological differences were observed after 24 h of gene expression (Robert et al., 2021), notable differences emerged when cells were grown in YPM + G for 72 h. Despite TB not being able to penetrate the plasma membrane, it was taken up by most AkCSLA3-expressing cells in addition to cell wall labelling. On the other hand, only a few AtCSLA2 and Venus cells displayed internal TB labelling, while cells expressing the swapped enzyme versions 2322, 2232, and 3332 showed intermediate phenotypes (Figure 9b & (Robert et al., 2021)). Venus cells without fluorescence took up TB, which is associated with compromised cell wall integrity (Robert et al., 2021).

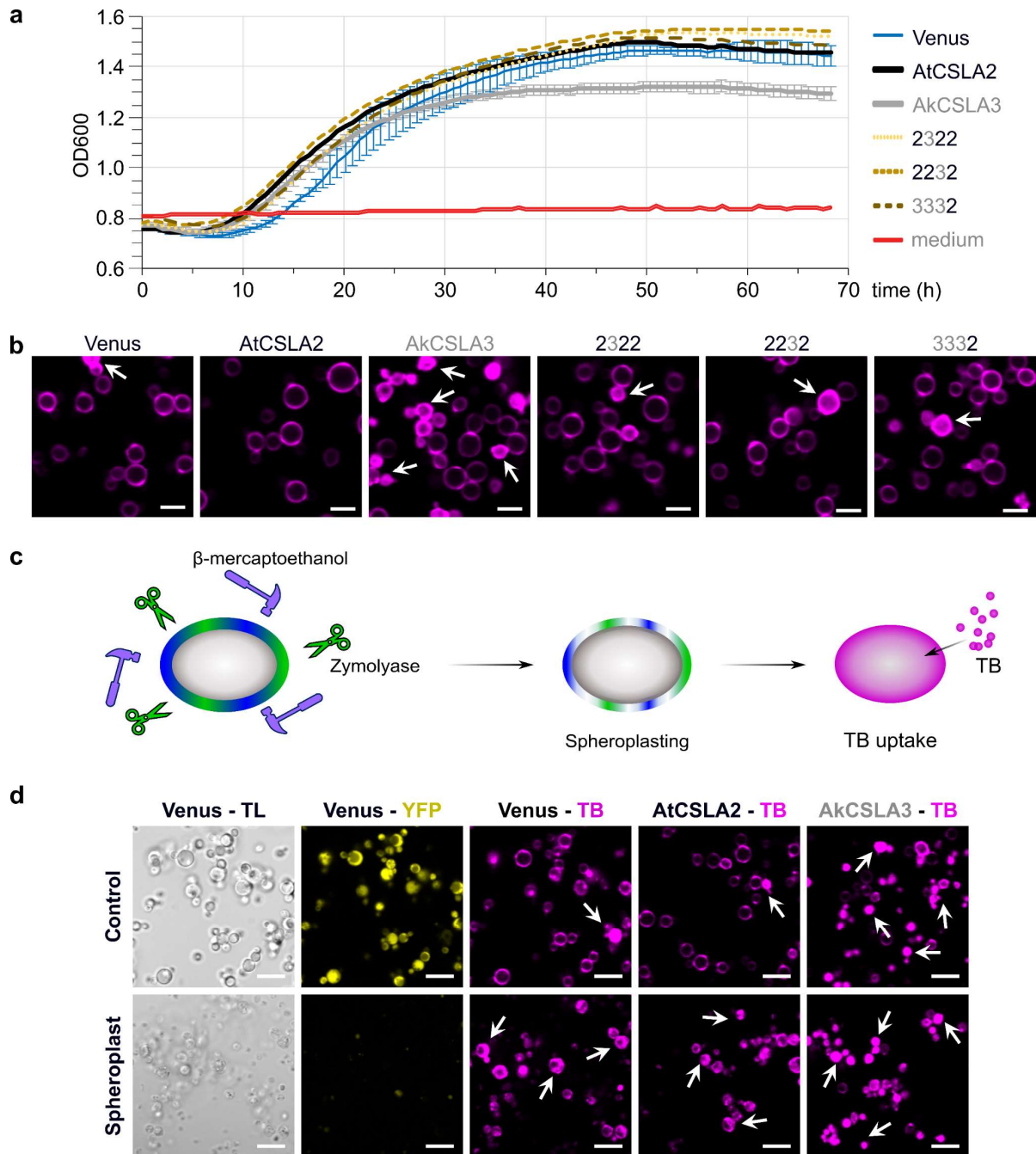


Figure 9 Influence of CSLA expression on *Pichia* growth and cell wall morphology. (a) Growth curve of CSLA parents and top swaps after transfer to YPM in a 48-well plate. Optical density (OD600) was monitored with a plate reader every 30 min for 70 h. Data show the mean \pm SD of at least three biological replicates. Due to space limitations, error bars are only shown for Venus and AkCSLA3. However, the other data points had a coefficient of variance below 5% after 32 h. (b) Morphology of cells grown for 72 h in YPM + G. β -glucans were stained with Trypan Blue (TB). (c) β -mercaptoethanol and Zymolyase 20T[®] were applied to induce intentional cell wall damage, which causes TB uptake into spheroplasted cells. (d) Zymolyase and β -mercaptoethanol treatment successfully spheroplasted Venus cells based on Transmitted Light (TL) and yellow protein fluorescence (YFP) imaging. Spheroplasted AtCSLA2 and AkCSLA3 cells show high TB uptake (arrows) Scale bars = 5 μ m in (b), 10 μ m in (d).

To verify if TB uptake in AkCSLA3-expressing cells might indeed be related to cell wall damage caused by glucomannan production, cell wall integrity was intentionally damaged through partial digestions with Zymolyase 20T[®] and β -mercaptoethanol. It was hypothesized that partially spheroplasted cells would absorb TB in a manner similar to untreated AkCSLA3 and Venus without yellow fluorescence (Figure 9c). Indeed, spheroplasting resulted in the loss of yellow fluorescence in the majority of Venus cells. In addition, it caused an increased TB uptake, a pattern also observed in AtCSLA2 (Figure 9d). The similarity between the observed TB uptake in untreated AkCSLA3 compared to AtCSLA2 and Venus suggests that prolonged AkCSLA3 overexpression may be damaging the cell wall integrity. Interestingly, the 3332 swap showed lower TB uptake than AkCSLA3 while demonstrating only minor changes in the Glc:Man ratio (Figure 8, Figure 9d, Figure 10a and c), indicating that swapping CSLA domains could mitigate cell wall damage-related toxicity on the host organism while still producing considerable amounts of (gluco)mannan.

Besides the cell wall integrity damage associated with extended AkCSLA3 overexpression, it was found that, compared to the Venus control, AkCSLA3 expression caused a 12.9% cell area reduction. Interestingly, despite rescuing the TB uptake phenotype, the 3332 swap showed a similar size decrease. On the other hand, AtCSLA2 expression led to a 4.5% cell area increase compared to the Venus control. The 2322 and 2232 swaps, containing elements of both parental enzymes, showed intermediate cell sizes (Figure 10b). This could indicate that a single AkCSLA3 domain may be sufficient to influence cell size.

The observed loss of cell wall integrity in AkCSLA3 could be a sign of cell death, potentially explaining why prolonged AkCSLA3 overexpression is toxic to the host cells. To explore this hypothesis, cell wall β -glucans were stained with Calcofluor White (CF). Propidium Iodide (PI) was used to stain dead cells, as this dye penetrates them and intercalates in the DNA, ultimately labelling their nuclei (Suzuki et al., 1997).

A significant number of nuclei of AkCSLA3-expressing cells were indeed labelled with PI (Figure 10a), which was suggestive of cell death. The ratio between the total number of cells - determined by CF labelling - and the cells with PI-stained nuclei allowed to estimate cell viability for each strain. Notably, AtCSLA2, 2232, and the Venus control strains showed 27-fold lower PI labelling frequency than AkCSLA3 (Figure 10c). Interestingly, 2322 and 3332 showed intermediate viability phenotypes, indicating that the second AkCSLA3 domain might contribute to the toxicity exerted on the host. Importantly, the results show that replacing the last domain of AkCSLA3 with that of AtCSLA2 can reduce toxicity without significantly altering the resulting product. Supported by further staining experiments with Congo Red that showed consistent results (Robert et al., 2021), it could be concluded that high glucomannan levels are toxic to the yeast cells.

In contrast, mannan production driven by AtCSLA2 seems to be tolerated well and even leads to cell size increases. Interestingly, no such effects were observed when AkCSLA3-sYFP was transiently expressed in leaves of the model plant *Nicotiana benthamiana* (Robert et al., 2021). This is plausible, given that *Pichia* cells contain many mannan precursors and thus allow much greater mannan accumulation than *N. benthamiana*, which is known to harbour only small mannan amounts (Le Mauff et al., 2017).

Together, the results demonstrate the suitability of a CSLA domain swap approach for customizing heteromannan biosynthesis and promoting the understanding of how (gluco)mannan production can affect total biomass production and cell morphology.

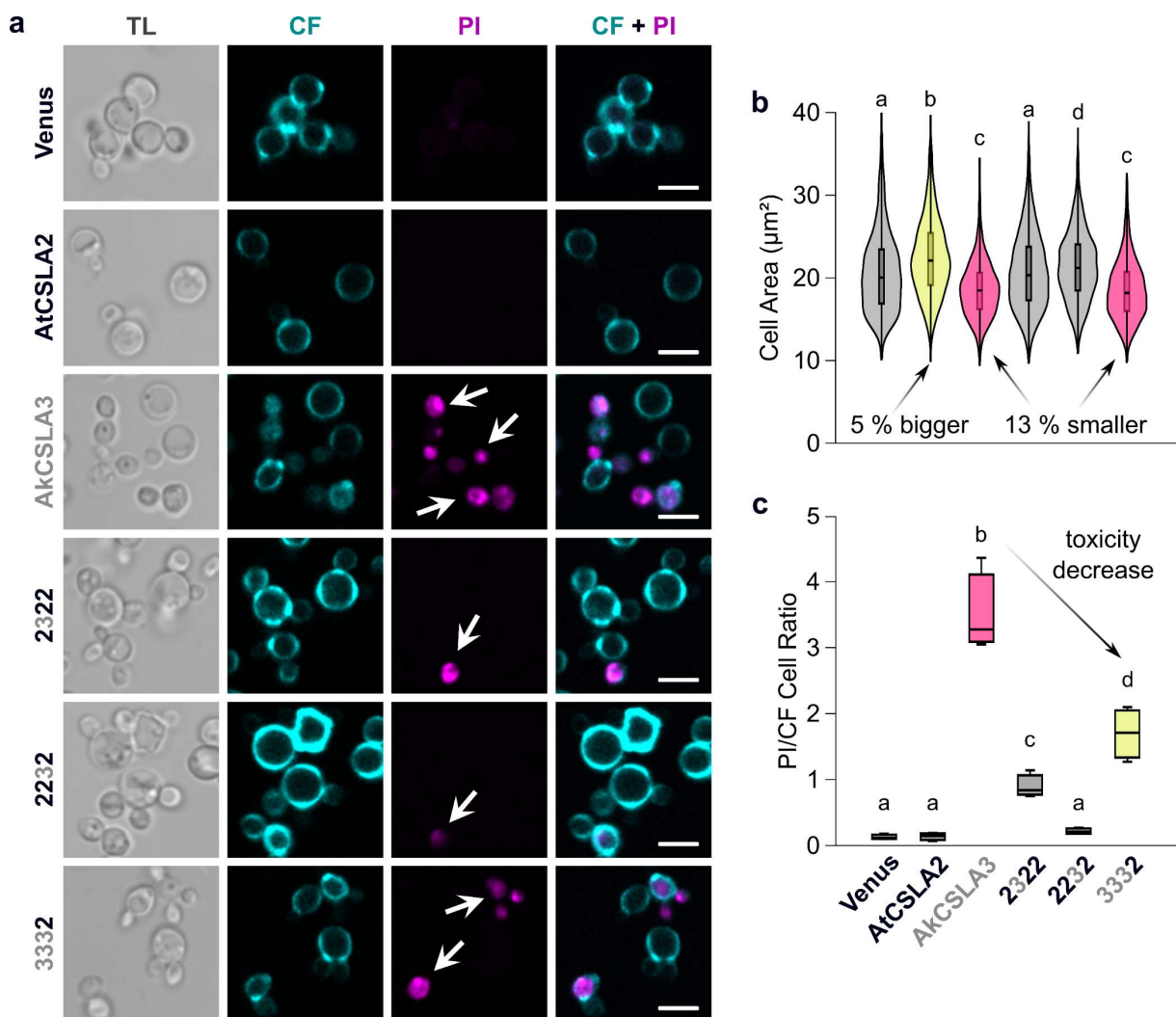


Figure 10 Toxic effects of AkCSLA3-driven glucomannan synthesis on *Pichia* cells. Cells were grown for 72 h in YPM + G. (a) Calcofluor White (CF) staining highlights the cells walls and Propidium Iodide (PI) marks dead cells. TL = transmitted light, Scale bar = 5 μm . (b) Cell areas (μm^2) of *Pichia* cells expressing CSLA parents and top chimeric CSLAs. Violin plot shows size distribution of at least 1200 cells per genotype (combining four biological replicates). Yellow demarcates increased AtCSLA2 cell size and pink delineates AkCSLA3 and 3332 with reduced cell sizes. (c) Ratio between PI- and CF-stained cells. Cells were segmented using Yeastspotter and counted with ImageJ. Boxes show 25–75% quartiles, median value (inner horizontal line). Whiskers delineate min/max values. Pink box shows high toxicity of AkCSLA3, and yellow box shows reduced toxicity of 3332. In (b) and (c) letters denote significant differences between samples (one-way ANOVA with Tukey test (b) $P < 0.001$ and (c) $P < 0.05$).

3.2 Design CBM-Probes to Detect and Localize HM *in vivo*

3.2.1 Optimization of Yeast Growth and CBM Expression

3.2.1.1 Constitutive CBM Expression is Toxic to Yeast Cells

Data from this chapter is unpublished and manuscript is currently under preparation.

The design of a new CBM-based probe for *in vivo* HM detection required optimal growth conditions to achieve high heterologous CBM expression in *Pichia* without compromising biomass. Due to its ability to bind β -1,4-mannooligosaccharides, carob galactomannan and konjac glucomannan, TmCBM27 from the *Thermotoga maritima* Mannanase 5 (Man5) (Boraston et al., 2003), was selected as a promising candidate for the binding of heteromannan *in vivo*. To quickly determine its expression levels in *Pichia* and to enable microscopical visualization, it was tagged with the fluorescent protein mRuby2. Constitutive gene expression is generally the most straightforward approach. Therefore, it was tested how constitutive TmCBM27 expression using the pGAP promoter affected growth compared to the constitutive expression of mRuby2 alone. After 24 hours of growth in YP media with dextrose (YPD) or methanol (YPM), YPD resulted in higher biomass than YPM (Figure 11a), likely because *Pichia* prefers dextrose as a carbon source (Weinhandl et al., 2014; Ben Azoun and Kallel, 2017). mRuby2 expression was highest in YPD and significantly reduced in YPM (Figure 11b). TmCBM27-mRuby2 expression was almost undetectable and adversely affected biomass production, suggesting that constitutive TmCBM27 overexpression is likely toxic to yeast cells.

To overcome the toxicity of TmCBM27 expression and reduce its negative effect on optical density, two strong, methanol-inducible promoters, namely pFDH1 and pAOX1 (Prielhofer et al., 2017), were tested as to their effects on CBM expression and yeast biomass. Since the CBMs should be designed to secrete out of the cells by default and only stay in the cells when HMs are present, the effects of the secretion signal peptide α on expression and biomass were also investigated. After 24 h of growth in YPM, the TmCBM27 expressing strains showed a slight reduction in biomass compared to the fluorescent controls and AkCSLA3-sfGFP (Figure 11c). Notably, the use of methanol-inducible promoters mimicked the behaviour of the constitutive promoter pGAP when grown in YPM. This indicates that it could be advantageous to combine the use of YPD and YPM. Using methanol-inducible promoters would allow to first grow cells in YPD for biomass production before adding methanol for gene expression, as shown earlier for other genotypes (section 3.1 and (Robert et al., 2021)).

Despite the reduction of optical density in colonies expressing TmCBM27 with methanol-inducible promoters, it was found that the highest expression of mRuby2 alone could

be achieved with pFDH1 (Figure 11d). This suggests that it might be a good candidate for high CBM expression, especially since the TmCBM27 constructs with pAOX1 and pFDH1 showed higher mRuby2 signals than constitutively expressed TmCBM27. Co-expression of pGAP:mRuby2 and pFDH1:mRuby2 with AkCSLA3-sfGFP did not cause a reduction of AkCSLA3-sfGFP expression (Figure 11e). However, co-expression of AkCSLA3-sfGFP with both CBM constructs reduced the green fluorescence to the level of the non-green fluorescence expressing mRuby2 tagged constructs in X-33, suggesting that AkCSLA3-sfGFP expression might be impaired upon co-expression with this CBM when grown for 24 h in YPM. Since it is vital to achieve successful co-expression of AkCSLA3-sfGFP and CBM in order to develop the new CBM-based probes, further growth condition optimization is needed.

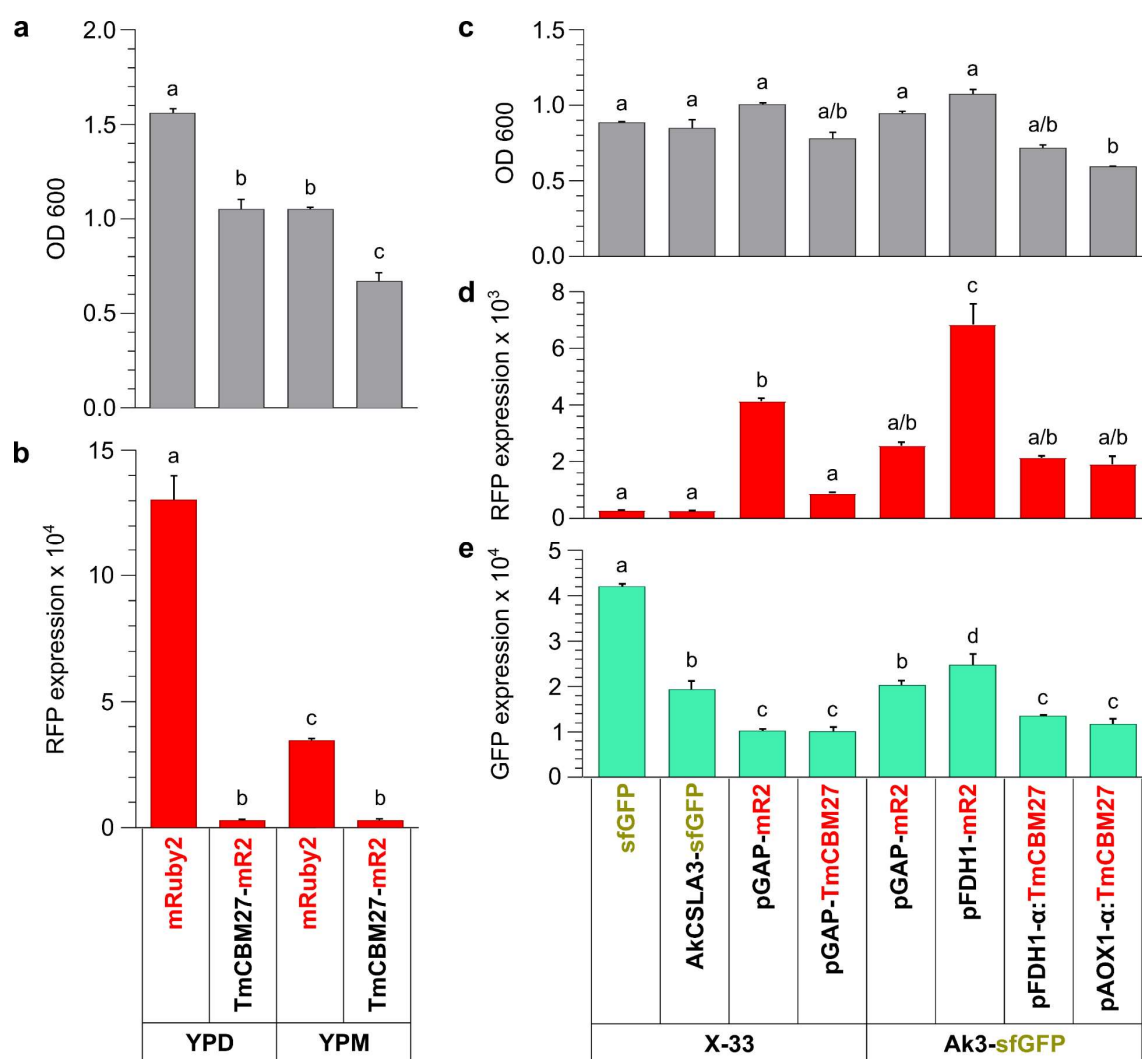


Figure 11 Constitutive and methanol-inducible expression of TmCBM27 with and without co-expression with AkCSLA3. (a) and (b) show OD600 and red fluorescence (RFP) of constructs that were grown for 48 h in 2 mL YPD or YPM. Bars represent mean + SD of two biological replicates that were measured twice each. Letters denote significant differences (one-way ANOVA with Tukey test, $P < 0.05$). (c), (d), and (e) show OD600, RFP and green fluorescence (GFP) of TmCBM27 in wild-type (X-33) with constitutive promoter (pGAP) and α :TmCBM27 with methanol-inducible promoters (pFDH1, pAOX1) in AkCSLA3-sfGFP background. Except for pGAP-TmCBM27-PT-mRuby2 (4 biological replicates), 2 biological replicates were grown for 24 h in 600 μ L YPM. Error bars show standard deviation, letters denote significant differences (one-way ANOVA with Tukey test, $P < 0.05$).

3.2.1.2 Optimizing Culture Volume and Methanol Concentration

The goal of using CBMs as probes for HM detection is to enable high throughput screening of its presence in multiple genotypes. To achieve this, a growth optimization test in 96-well plates was conducted, allowing the screening of a large number of genotypes in one single batch. In this context, it was assessed how two different plate covers – a silicon mat and an enzyscreen metal cover – would influence growth and gene expression. After 24 h of growth in 300 μ L YPM, some significant differences between OD600 values of different samples could be detected. However, in some cases, there was no significant difference between the optical density of the YPM medium control and the transgenic yeast strains tested (Figure 12a). This is consistent with the lack of significant GFP signal differences between sfGFP and no-GFP expressing strains and the YPM medium control (Figure 12b). Due to the suboptimal growth and gene expression in the 96-deep well enzyscreen plates, the following experiments were carried out in sterile 24- or 48-well plates. For proof-of-concept experiments conducted in this work, the 96-well plate format was not essential. However, optimizing growth in 96-well plates will be crucial to increase throughput once the CBM probe system is established.

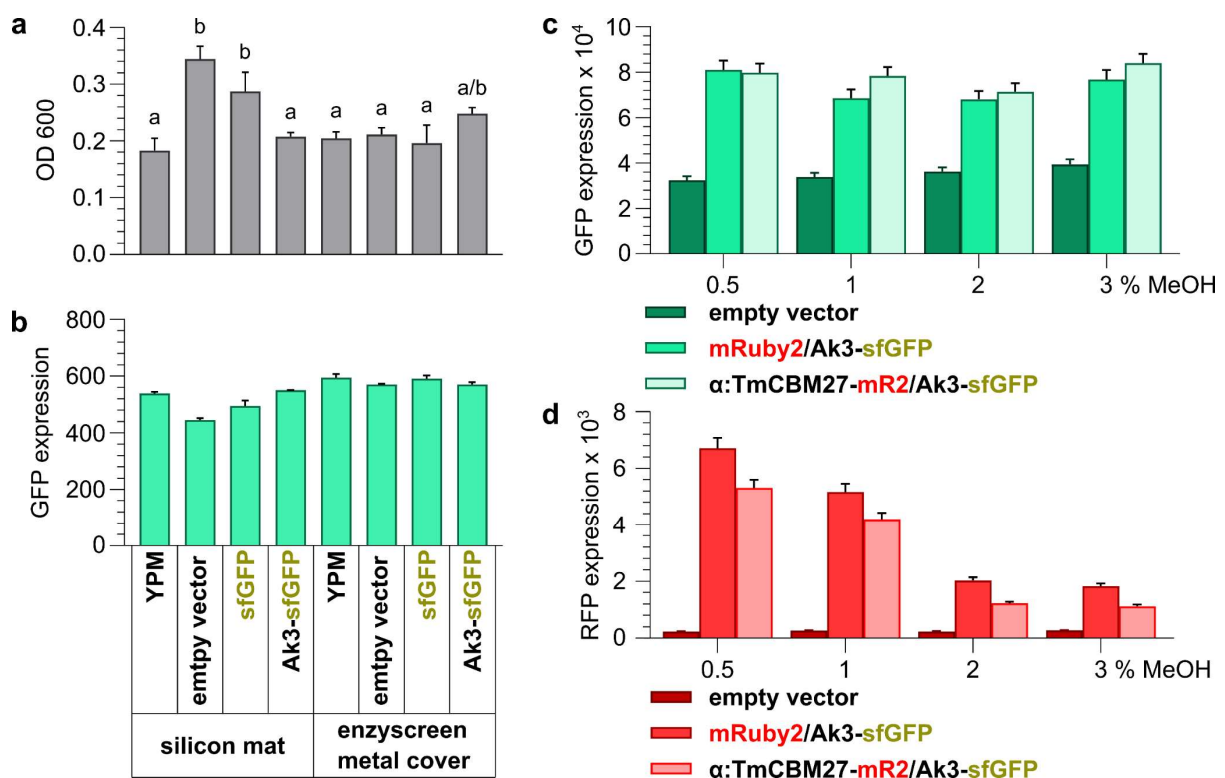


Figure 12 Optimization of growth format and methanol (MeOH) concentration in media for high CBM expression. OD600 and GFP expression (b) of colonies grown in a high throughput format in sterile 96-deepwell enzyscreen plates with two different plate covers (silicon mat and enzyscreen metal cover). Colonies grew in 300 μ L YPM for 24 h. Bars show mean + SD of 3 biological replicates. Letters denote significant differences determined via one-way ANOVA with Tukey test, $P < 0.05$. (c) and (d): expression of sfGFP (GFP) (c) and mRuby2 (RFP) (d) tagged constructs of colonies grown with different MeOH concentrations. Colonies grew in 24-well enzyscreen plates in BMGY for 60h before the medium was changed to BMMY with final methanol concentrations of 0.5, 1, 2 and 3%. Graphs show results for one biological replicate. Error bars show standard deviation of plate reading. Abbreviations: mR2 = mRuby2, Ak3 = AKCSLA3.

To optimize gene expression further, the influence of four different methanol (MeOH) concentrations in the media (0.5, 1, 2, and 3%) on biomass production and gene expression was investigated. Biomass production remained unaffected by different MeOH concentrations (data not shown). GFP levels also remained similar across the varying MeOH concentrations (Figure 12c). On the other hand, increasing methanol concentrations caused expression reductions for mRuby2 alone and pAOX1- α -TmCBM27-PT-mRuby2 in the AkCSLA3-sfGFP background (Figure 12d). Based on these results, 0.5% final MeOH concentration was selected for the cultivation media to achieve optimal gene expression.

3.2.1.3 Minimal Growth Media Enables High Gene Expression

The toxicity associated with constitutive CBM expression in yeast cells and the limitations of generating sufficient biomass and expression rates in 96-well plates made further growth condition optimization necessary. The optimization involved increasing the culture volumes and well sizes wells for sample growth. To overcome the toxicity issue, two different media combinations, namely BMGY/BMMY and YPD/YPM, were tested with a final MeOH concentration of 0.5%, as these combinations had previously demonstrated the highest expression levels for mRuby2 tagged constructs (Figure 12). The objective was to generate a substantial amount of biomass using methanol-free media prior to transferring the cells to methanol-containing media for gene expression induction.

The new growth conditions of 2 mL media in plastic culture tubes resulted in a considerable cell density increase, enabling the differentiation between genotypes. Constitutively expressing TmCBM27-mRuby2 in wild-type (X-33) reduced the optical cell density (OD600) significantly compared to the constitutive expression of mRuby2 alone. This observation aligns with the previous findings that constitutive CBM overexpression is toxic to yeast cells. This toxicity could be overcome by utilizing the methanol-inducible promoter pFDH1. When pFDH1 was employed, the cell densities of CBM-containing yeast cells were not significantly different from the ones achieved upon mRuby2 alone expression (Figure 13a). These results were consistent between YPD/YPM and BMGY/BMMY media.

Considering that the future application of the new CBM-based probes involves their co-expression with CSLAs to assess their functionality, the co-expression of AkCSLA3-sfGFP and α :TmCBM27-mRuby2 was also examined alongside the co-expression of mRuby2 and AkCSLA3-sfGFP. Except when AkCSLA3-sfGFP was co-expressed with α :TmCBM27-mRuby2, all AkCSLA3-sfGFP containing yeast strains showed significantly higher GFP signal in YPD/YPM compared to BMGY/BMMY (Figure 13b). Although it can be argued that YPD/YPM is more advantageous for the expression of AkCSLA3-sfGFP compared to BMGY/BMMY, the expression disparity of the tested mRuby2 tagged constructs in the different

media was less pronounced. Like that, a higher expression of mRuby2 alone in YPD/YPM compared to BMGY/BMMY was only observed when expressed under the constitutive pGAP promoter in X-33. pGAP and pFDH1-driven mRuby2 expression in the AkCSLA3-sfGFP background were similar in both tested media combinations. In the TmCBM27 constructs, the expression of α :TmCBM27-mRuby2 in the AkCSLA3-sfGFP background was significantly higher in YPD/YPM compared to BMGY/BMMY. However, the constitutive expression of TmCBM27-Ruby2 in X-33 did not differ significantly between the two tested media combinations (Figure 13c). In addition to the higher expression levels observed in YPD/YPM compared to BMGY/BMMY, the expression rates of the mRuby2 tagged constructs were more consistent when grown in YPD/YPM. Considering the lower complexity of YPD/YPM compared to BMGY/BMMY (Voiniciuc et al., 2019; Robert et al., 2021), YPD/YPM was selected as the foundation for further growth medium optimization.

While the expression of the CBM and CSLA constructs tested could be optimized using the described growth conditions, one issue remains to be addressed. The α signal peptide is widely used and known to work relatively well in *Pichia* most of the time (Lin-Cereghino et al., 2013; Zou et al., 2022). Therefore, complete secretion of CBMs with an N-terminal α signal peptide when no heteromannans are available for binding is expected. Interestingly, the secretion of α :TmCBM27-mRuby2 secretion from the glucomannan-producing AkCSLA3-sfGFP was

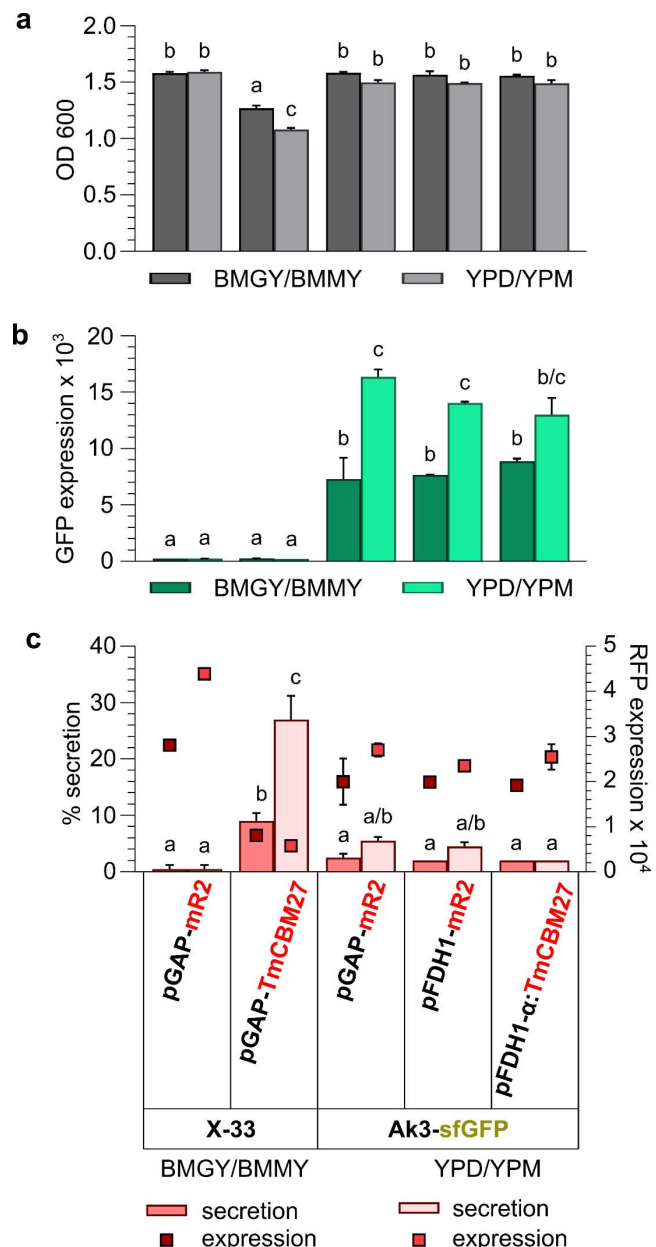


Figure 13 Effects of BMGY/BMMY and YPD/YPM media on cell density and gene expression of TmCBM27-mRuby2 (tagging highlighted by red colour) and AkCSLA3-sfGFP. Cultures grew in 2 mL YPD or BMGY for 24 h, before transfer to 2 mL YPM or BMMY for another 24 h. (a) OD600, (b) GFP expression, (c) squares denote RFP expression and bars show % secretion based on corrected RFP. Bars show mean + SD of 2 technical replicates and letters denote significant differences for OD600, GFP expression and RFP % secretion based on one-way ANOVA with Tukey test, $P < 0.05$.

minimal and significantly lower than the “secretion” of TmCBM27-mRuby2 without a signal peptide from the X-33. This provides an initial hint that α :TmCBM27-mRuby2 might be retained in the AkCSLA3-sfGFP- expressing cells due to heteromannan presence.

3.2.1.4 Carbon Co-Feeding for Optimal Growth and Gene Expression

To identify the ideal growth conditions for high biomass and efficient heterologous gene expression, various carbon sources were tested for biomass production prior to gene expression induction. Specifically, the effectiveness of co-feeding glycerol, glucose, or mannose with methanol in YP-based media was compared to feeding YPM only.

pDAS2: α -mRuby2 both in wild-type (X-33) and in the AkCSLA3-sfGFP background were tested alongside with pPOR1- α :TmCBM27:PT-mRuby2 in wild-type (X-33) and pPOR1- α :CjCBM35:L2-mRuby2 in the background of AtCSLA2. α :mRuby2 exhibited the highest growth rate when glycerol was used as a co-feeding carbon source, regardless of its expression background. Co-feeding of mannose and methanol significantly increased the optical density of both α -mRuby2 strains compared to growing the strains in YPM only. However, this increase was less pronounced than the biomass density increase in glycerol. For the CBM expressors, neither glycerol nor mannose caused significant changes in biomass.

In contrast, all genotypes showed a significant biomass decrease when D-glucose (dextrose) was used as carbon source compared to YPM alone. This is intriguing, as growing cells in YPD (YP + dextrose) generally yields higher biomass than growth in YPM. The lower biomass observed when co-feeding YPM and Glc compared to YPM alone suggests an unfavourable interplay between Glc and methanol during co-feeding. In light of this experiment, YPM + G would be the best medium for biomass production (Figure 14a). The choice to utilize YPM + G for subsequent experiments was fortified by the finding that this medium yielded the highest GFP signal for α -mRuby2 in the AkCSLA3-sfGFP background (Figure 14b).

Compared to OD600 and GFP, the RFP signal intensities exhibited more variation. While YPM only yielded the highest RFP signal for α : mRuby2 in wild-type (X-33), α :mRuby2 in AkCSLA3-sfGFP displayed its highest RFP signal in the YPM + G medium. On the other hand, the CBM expressors achieved their highest RFP signals when cultivated in YPM + Glc (Figure 14c). One of the paramount characteristics of the newly designed CBM probes will be their secretion behaviour. Ideally, the CBM secretion should be reduced when the cells contain higher levels of heteromannan. Hence, it was crucial to find a growth medium that would yield consistent, background-independent secretion of the α :mRuby2 control. While both YPM + G and YPM + Man fulfilled that criterion (Figure 14c), including mannose in the medium could potentially have adverse effects on future carbohydrate analyses as it is a major component of the mannan and glucomannan produced by AtCSLA2 and AkCSLA3-sfGFP.

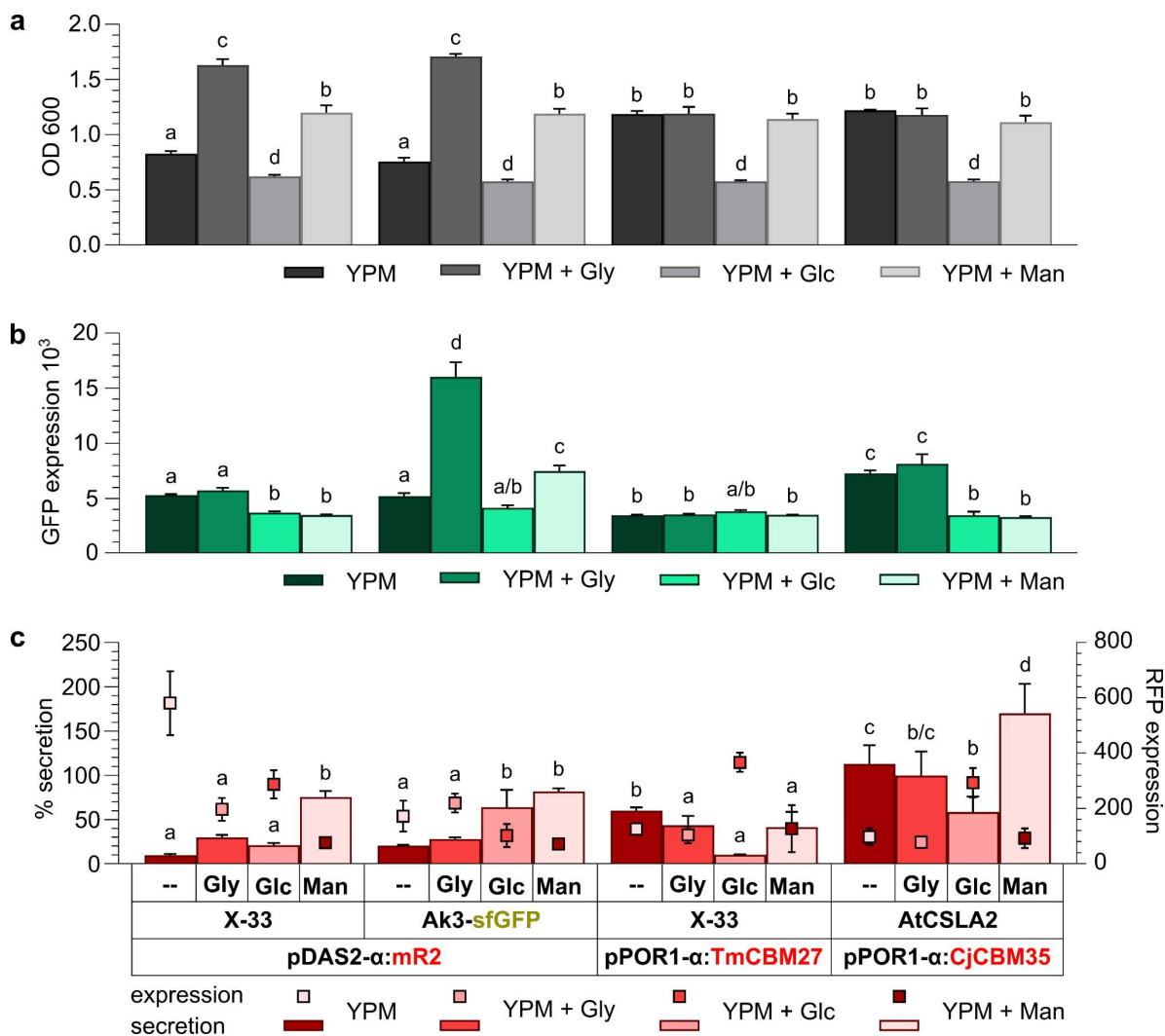


Figure 14 Influence of different carbon sources on growth, expression and CBM secretion. Colonies grew for 24 hours in 600 μ L YPM (YP + 1.5% methanol), YPM + G (YP + 0.5% glycerol (Gly) + 1.5% methanol), YPM + Glc (YP + 0.5% glucose + 1.5% methanol) and YPM + Man (YP + 0.5% mannose + 1.5% methanol). (a) OD600, (b) GFP expression, (c) RFP expression (squares) and secretion (bars). Bars and error bars depict mean + SD of 3 biological replicates. Letters denote significant differences between OD600, GFP expression and RFP secretion calculated using one-way ANOVA with Tukey test, $P < 0.05$. X-axis labelling from top to bottom: co-feeding carbon sources, expression background, construct. Abbreviations: Ak3 = AkCSLA3, mR2 = mRuby2, red colouring of CBMs symbolizes that they are C-terminally tagged with mRuby2.

Consequently, YPM + G represented the ideal balance between high biomass production and consistent gene expression and secretion of α :mRuby2. Furthermore, the CBM strains exhibited comparable biomass and expression levels, which would facilitate comparisons regarding the quantity of CBMs they secrete. Hence, the co-feeding of methanol and glycerol in the YP-based medium, termed YPM + G in the following, was chosen for subsequent experiments. In addition to being a simpler medium than BMGY/BMMY and the fact that YP-based media generally show higher expression levels than BMGY/BMMY (Figure 13 b and c), the co-feeding of glycerol and methanol eliminates the need for manual culture transfer from the biomass production medium to the gene expression induction medium, considerably simplifying cultivation workflow and increasing the possible throughput.

In addition to media optimization, different culture volumes were tested for their impact on CBM expression (results not shown). In summary, the optimized growth conditions were set to 300-600 μ L YPM + G in sterile 24- or 48-well plates at 30 °C and 250 rpm for 48-72 h.

3.2.2 Recombination of CBMs with Fluorescent Proteins for HM Detection

3.2.2.1 CBM Secretion Decreases upon Heteromannan Presence

For the CBM probe design, two new CBM candidates, in addition to TmCBM27 (Boraston et al., 2003), were selected due to their abilities to bind (hetero)mannan components. TpolCBM16-1 from the *Caldanaerobius polysaccharolyticus* β -mannanase 5A (Man5A) was shown to bind mannopentaose and cellopentaose (Bae et al., 2008; Su et al., 2010) and PaCBM35 from the *Podospora anserina* endo- β -1,4-mannanase26A (GH26) binds β -1,4-mannan and with a lower affinity also xylan (Couturier et al., 2011; Couturier et al., 2013).

The candidate CBMs were recombined with an N-terminal α signal peptide for secretion (Lin-Cereghino et al., 2013; Zou et al., 2022) to examine their responsiveness to HM presence. The hypothesis proposed that CBMs fused with the α secretion peptide would secrete from *Pichia* cells by default and remain within the cells upon heteromannan binding (Figure 15). This would enable the detection of heteromannan presence based on the presence or absence of CBMs within the *Pichia* cells. To facilitate the assessment of CBM presence within the cells, the candidate CBMs were C-terminally tagged with mRuby2 prior to expression. As proof of concept for the heteromannan presence-dependent secretion, TmCBM27, TpolCBM16-1, and PaCBM35 were individually expressed in three distinct *Pichia* backgrounds, namely X-33 wild-type and mannan (AtCSLA2) and glucomannan producing (AkCSLA3-sfGFP). For CBM expression, the methanol-inducible pFDH1 was selected, as it demonstrated similar expression rates across various colonies in preliminary experiments.

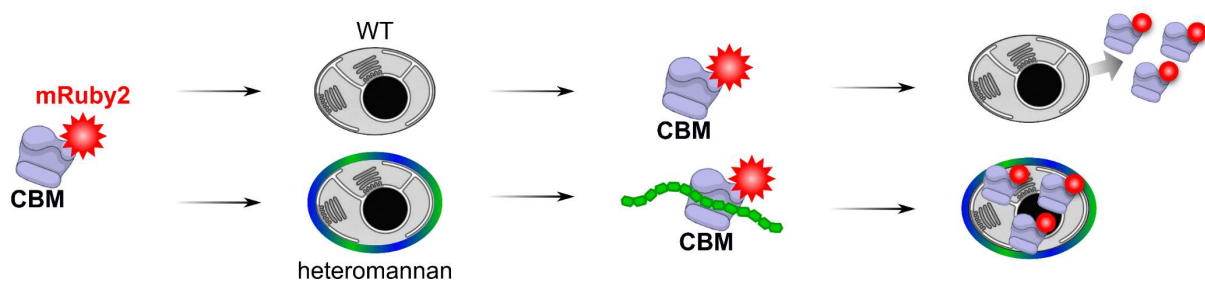


Figure 15 Working principle of the new CBM-probes. mRuby2-tagged CBMs equipped with a signal peptide for secretion are expressed in yeast. When expressed in wild-type (WT), CBMs secrete out of the cells. When expressed in a heteromannan producing background, CBMs bind to the middle of mannan oligosaccharide chains, which causes them to be retained in the cells. So, CBM fluorescence in the cells is an indicator for HM presence.

After elevated GFP levels for AkCSLA3-sfGFP confirmed CSLA expression and elevated RFP levels confirmed CBM expression, the correlation between CBM secretion and heteromannan presence was examined (Figure 16a and b). Although the α :mRuby2 expressed similarly in all three tested backgrounds, its secretion was lower than expected. Since α :mRuby2 lacks a CBM, its secretion should have reached 100% but remained below 50% (Figure 16b). One possible explanation could be an interference of α :mRuby2 with the yeast's secretory system due to very high expression. Conversely, the secretion was similar across all three backgrounds, indicating its independence from heteromannan presence.

In X-33, all three tested CBMs expressed similarly. However, only TpolCBM16-1 and PaCBM35 showed significant secretion, while TmCBM27 secreted poorly. The expression of the tested CBMs was generally lower than that of the α :mRuby2 control, except for TmCBM27 in the AtCSLA2 background, which was the highest expressor (Figure 16b). Regarding background-dependent secretion, TmCBM27 behaved similarly to α :mRuby2. Its secretion rates were low across all three tested *Pichia* backgrounds and showed no significant differences despite the detection of significantly elevated relative mannan amounts in the AtCSLA2 and AkCSLA3-sfGFP background (Figure 16d). Notably, the total amount of sugar detected in TmCBM27 expressing genotypes was similar to the levels detected in the other CBM/background combinations (Figure 16c). Together, the results obtained for TmCBM27 suggest that it might not secrete in dependence of the amount of heteromannan present in the cells and is therefore not suitable as a probe for heteromannan presence.

In contrast, TpolCBM16-1 and PaCBM35 exhibited highly promising outcomes. Both CBMs secreted approximately 80% from X-33, indicating nearly complete secretion in the absence of heteromannan availability for CBM binding. Interestingly, when expressed in the mannan-producing AtCSLA2 or the glucomannan-producing AkCSLA3-sfGFP background, TpolCBM16-1 and PaCBM35 displayed significantly reduced secretion. Excitingly, the secretion of both CBMs was significantly lower in the AkCSLA3-sfGFP background compared to the AtCSLA2 background. These findings are particularly interesting as they suggest that in addition to the sheer presence of heteromannan, the type or quantity of HM might influence the secretion rate of the CBMs (Figure 16b).

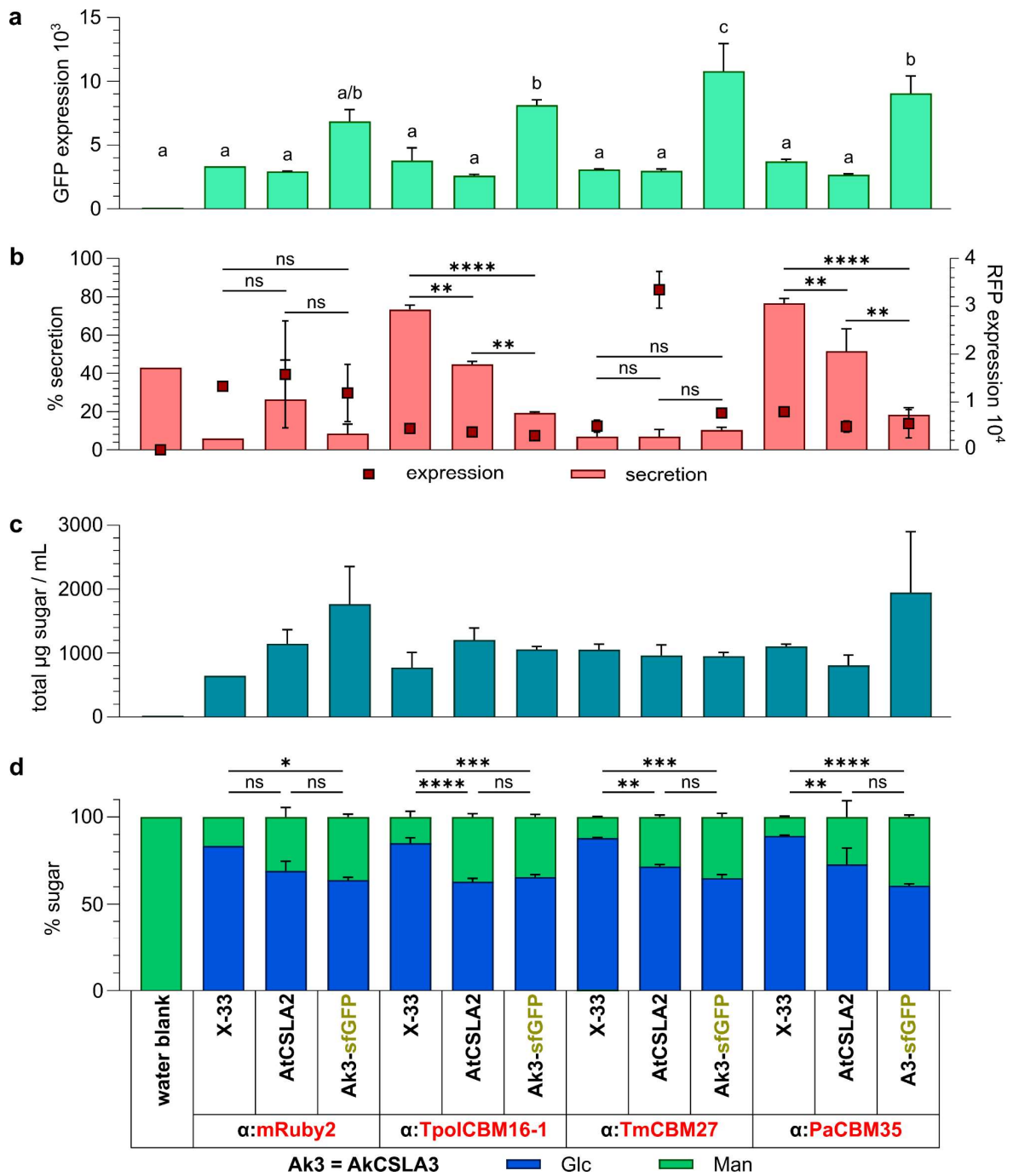


Figure 16 Correlation between CBM secretion and HM presence. Colonies were grown in sterile 48-well plates for 72 h, in 300 μL YPM + G and NTC (100 ng/ μL). mRuby2 control and CBM were expressed in X-33 wild-type, AtCSLA2 (mannan) and AkCSLA3-sfGFP (glucomannan) backgrounds under the control of the pDFH1 promoter. (a) GFP expression, (b) corrected RFP expression (squares) and secretion (bars) (c) total μg sugar/mL culture (no significant differences) (d) relative amounts of glucose (Glc) and mannose (Man). Bars and squares represent mean + SD of 3 biological replicates except for α :mRuby2 in AtCSLA2 and AkCSLA3-sfGFP, and α :TmCBM27 in X-33 (2 biol. Replicates) and α :mRuby2 in X-33 (1 replicate). mRuby2 tagging is highlighted through red colouring of the respective CBMs. Significant differences are denoted by letters (a) or asterisks (b) and (d) (significant differences between secretion rates) calculated with one-way ANOVA with Tukey test, $P < 0.05$.

To further investigate that matter and confirm the production of heteromannan in the strains where the CBMs secreted less, an AKI fraction was prepared from these cultures and subjected to 4% sulfuric acid hydrolysis for monosaccharide analysis using IC. While the total sugar content was comparable across all genotypes, cells expressing AtCSLA2 and AkCSLA3-sfGFP exhibited significantly elevated relative amounts of mannose compared to X-33. However, no significant difference between the relative amounts of mannose was detected between AtCSLA2 and AkCSLA3-sfGFP (Figure 16d and e). This suggests that the reduced secretion of TpolCBM16-1 and PaCBM35 from AkCSLA3-sfGFP compared to AtCSLA2 might not solely be rooted in the absolute amounts of mannose present but rather the specific type of heteromannan. Another possibility could be that the total amount of glucomannan produced by AkCSLA3-sfGFP is only partly reflected in the elevated mannose, as it also incorporates glucose. This was previously shown through linkage and oligosaccharide analyses of this genotype (Figure 7c and d). Therefore, both hypotheses are plausible.

Furthermore, the monosaccharide analysis revealed that co-expressing CBMs and CSLAs did not impact heteromannan production. This is crucial as a probe should ideally not interfere with the system it is probing. Based on their background-dependent secretion, both TpolCBM16-1 and PaCBM35 were selected as candidates for subsequent experiments.

3.2.2.2 CBM-Probe Behaviour is Unaffected by Transformation Method

Utilizing the new CBM-based probes serves two main objectives: the visualization of native heteromannan distribution in living organisms and the detection of new functional CSLA variants with potentially higher yields. By enabling the visualization of (gluco)mannan produced in yeast and, more importantly, in plants, the new probes will offer valuable insights into the *in vivo* distribution of native heteromannan. These insights could significantly contribute to our understanding of how changes in heteromannan type and quantity influence its localization and how these changes may affect the plant phenotype.

On the other hand, the new CBM probes offer a powerful tool for high-throughput detection of novel functional CSLA variants. This could, for instance, be achieved by introducing various mutated CSLAs into a CBM-expressing background. In this case, the presence or absence of CBM fluorescence would directly correlate with the presence or absence of heteromannan, enabling the rapid selection of functional variants. These selected functional variants could then be analysed to identify the specific mutations influencing the type and quantity of heteromannan made.

Hence, it is crucial to ensure consistent CBM secretion behaviour regardless of whether (a) they are expressed in a (gluco)mannan producer or (b) a (gluco)mannan-producing protein is expressed in a CBM background. To address this question, the expression and secretion of

PaCBM35 were examined for cases (a) and (b). The fluorescence measurements of the culture and supernatant revealed significantly reduced secretion rates for both tested genotypes compared to the α :PaCBM35:mRuby2 only control. In addition, the secretion rates of both

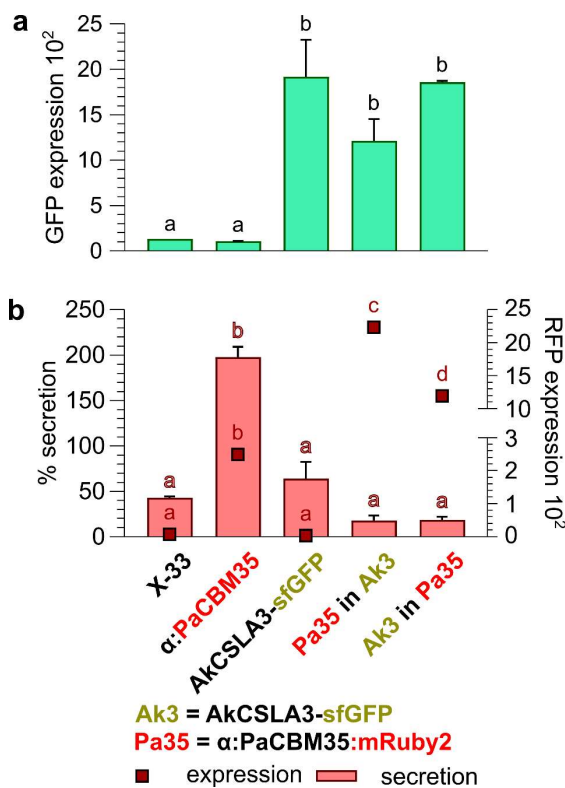


Figure 17 Influence of transformation direction on CBM secretion. Colonies grew in sterile 24 well plates in 600 μ L YPM+G for 72 h. Bars and squares represent mean + SD of 2 biological replicates. Letters denote significant differences (one way ANOVA with Tukey's pairwise ($p=0,05$)).

genotypes were similar, indicating that the transformation method transformation does not significantly affect the CBM-probe functionality (Figure 17b).

Further, AkCSLA3-sfGFP showed similar expression levels both with and without α :PaCBM35:mRuby2, supporting the previous finding that co-expressing AkCSLA3-sfGFP with a CBM does not affect its expression rate (Figure 17a). It is noteworthy that when co-expressed with AkCSLA3-sfGFP, α :PaCBM35:mRuby2 showed considerably higher fluorescence levels than when expressed alone. While that could be an indicator that co-expression with CSLAs might boost CBM expression, it is more likely, that in the tested CSLA/CBM co-expressing colonies, the expression was exceptionally high. In fact, α :PaCBM35:mRuby2 showed similar expression levels compared to mRuby2 alone

and TmCBM27 (data not shown), confirming that its expression levels were not particularly low. Despite the high difference in expression levels, the experiment demonstrates the suitability of the new CBM probes to detect native (gluco)mannan *in vivo* and identify new CSLA variants in yeast.

3.2.2.3 CBM-Probes Co-Localize with CSLAs in Yeast

Using fluorescence measurements of whole cultures and supernatant in combination with monosaccharide analysis enabled the identification of the top two CBM candidates, TpolCBM16-1 and PaCBM35, for convenient detection of heteromannan in *Pichia*. In addition to the detecting heteromannans, the new CBM-based probes should enable the visualization of heteromannans *in vivo*. As proof of concept for the CBM-mediated visualization of heteromannan *in vivo*, α :TpolCBM16-1:mRuby2 and α :PaCBM35:mRuby2 in X-33, AtCSLA2, and AkCSLA3-sfGFP were re-grown for microscopy analysis.

Although the growth conditions slightly differed from those utilized for the colonies shown in Figure 16, the results remained consistent. Notably, only α :TpoICBM16-1:mRuby2 exhibited a minor deviation, as no significant difference between the secretion from the AtCSLA2 and AkCSLA3-sfGFP backgrounds was observed in this experiment (Figure 18). However, the secretion from the (gluco)mannan-producing backgrounds was significantly lower than that from the X-33 wild-type, which aligns with previous results (Figure 16).

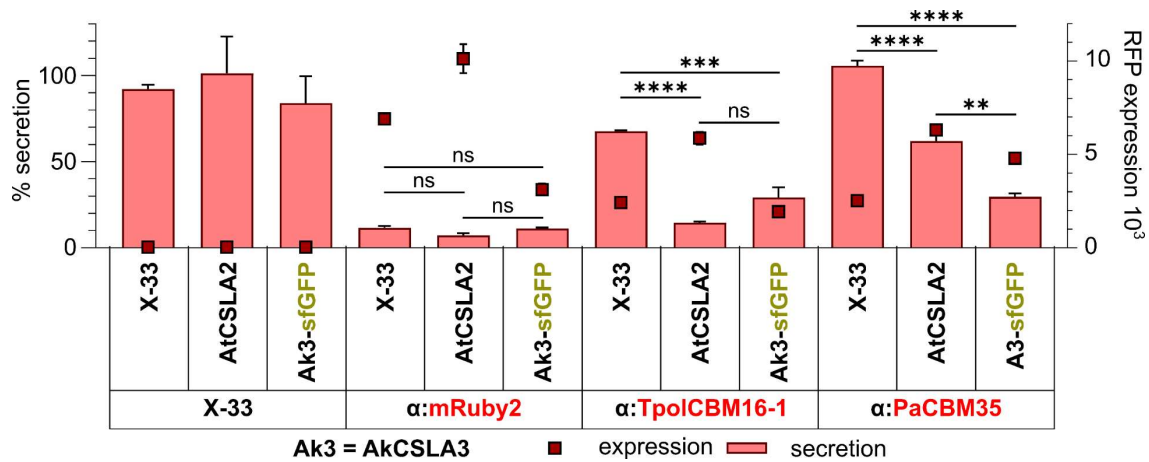


Figure 18 Expression and secretion of CBMs grown for microscopic localization analysis. Cells grew in 600 μ L YPM + G in sterile 24-well plates for 48 h. All tested colonies expressed the mRuby2 control and the CBMs under control of the pFDH1 promoter and were equipped with the α secretion peptide. The mRuby2 control and the CBMs were expressed in X-33 (wild-type), AtCSLA2 (mannan) and AkCSLA3-sfGFP (glucmannan) backgrounds. Bars (corrected secretion) and squares (corrected expression) represent mean + SD of 3 biological replicates. Note: most SD of expression are very low and therefore not visible. mRuby2 tagging is highlighted through red colouring of the respective CBMs. One-way ANOVA with Tukey test, $P < 0.05$ was used to calculate significant differences.

To investigate the subcellular localization of the CBMs, cells were counterstained with CF to image the cell walls and subsequently imaged via confocal microscopy using a Zeiss LSM900 (Figure 19). As anticipated, no red fluorescence signals were observed in X-33, and the AtCSLA2 and AkCSLA3-sfGFP expressors. In the case of AkCSLA3-sfGFP, its punctate localization throughout the cells was consistently observed, regardless of whether it was expressed alone or in combination with the α :mRuby2 control or the CBMs (Figure 19). This implies that the expression of additional constructs did not impede the intracellular distribution of AkCSLA3-sfGFP, which is essential to ensure that the new CBM-based probes do not perturb the studied system. The consistent α :mRuby2 signal observed in all three tested *Pichia* backgrounds agrees with the fluorescence measurements. It underscores the finding that even if the secretion of α :mRuby2 is much lower than expected, it is not influenced by the presence or absence of (gluco)mannan (Figure 19). However, the lack of α :mRuby2 secretion requires further optimization, such as exploring alternative fluorescent protein tags. The observed difficulties of α :mRuby2 to secrete could potentially be attributed to its high expression levels, which may negatively affect the *Pichia* secretory system. In fact, the expression levels of

α :mRuby2 are considerably higher than those of the tested CBMs, except in the AkCSLA3-sfGFP background. In the case of α :mRuby2 in AkCSLA3-sfGFP, the co-expression of both genes may reduce the expression of α :mRuby2. However, this particular type of co-expression might still be sufficient to impair the yeasts' secretory system, resulting in low α :mRuby2 secretion.

In contrast, the CBMs (α :TpolCBM16-1 and α :PaCBM35) efficiently secreted from the X-33 wild-type background, as evidenced by the absence of mRuby2 signal in these particular strains (Figure 19). When expressed in the AtCSLA2 or AkCSLA3-sfGFP backgrounds, α :TpolCBM16-1 and α :PaCBM35 were retained in the cells as expected (Figure 19). However, contrary to the expectation that the CBMs would localize to the cell walls where the (gluco)mannan was expected, they co-localized with the AkCSLA3-sfGFP in the cells and displayed punctate structures in the AtCSLA2 backgrounds. These observations suggest that the (gluco)mannan produced by AtCSLA2 and AkCSLA3-sfGFP might not primarily localize to the cell walls but rather remain in close proximity in the CSLAs. Alternatively, the CSLAs may interfere with the Golgi apparatus, potentially preventing CBM secretion. To verify the presence or absence of (gluco)mannan in the cell walls, heteromannans were immunolabelled on fixed and non-fixed cells in a later experiment (see: chapter 3.2.2.5).

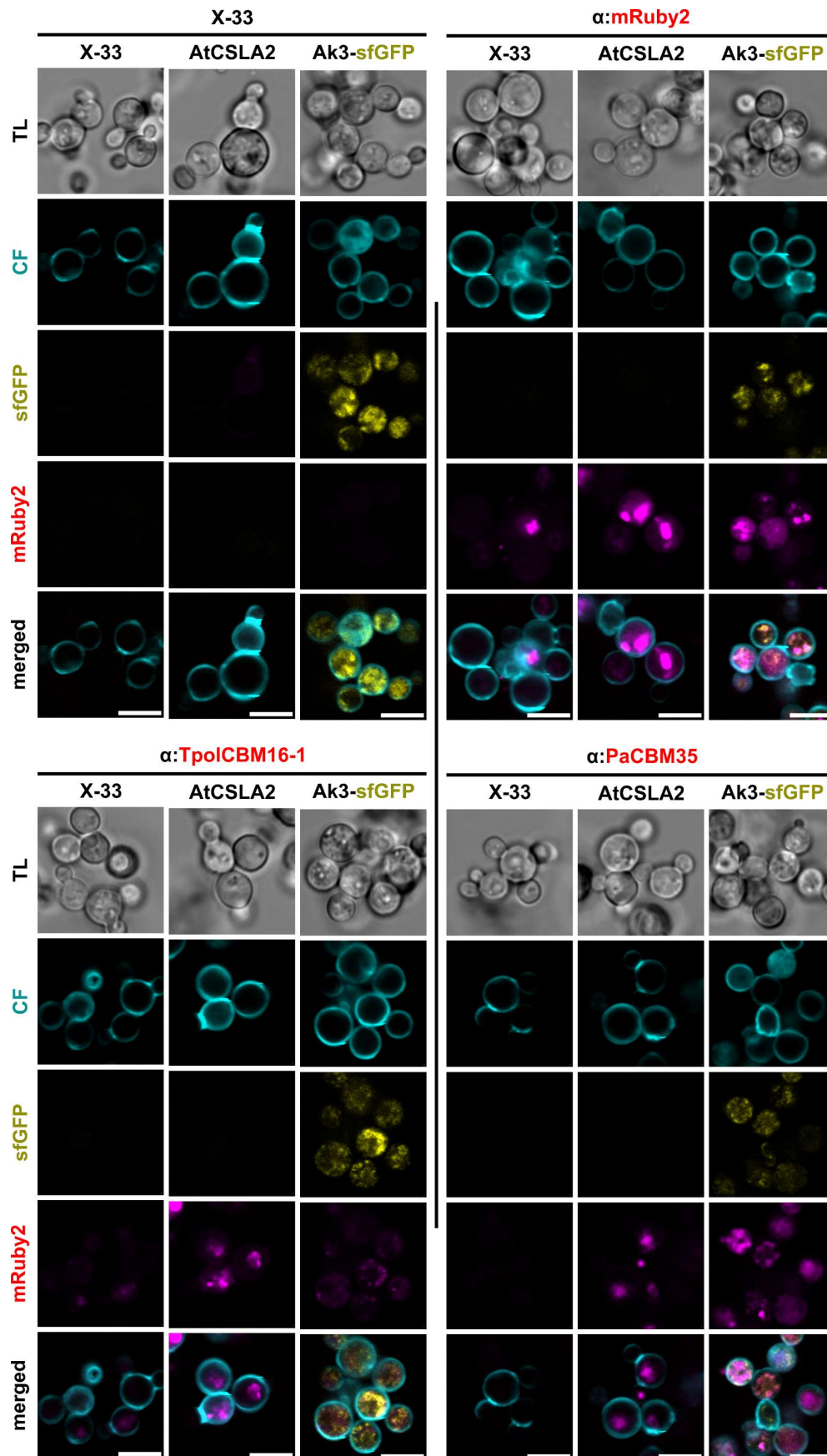


Figure 19 Localization of CBMs in different *Pichia* backgrounds. Colonies were grown and analysed as described in Figure 18. For microscopy, 10 μ L of undiluted culture were mixed with 10 μ L of 0.001% Calcofluor White (CF) and immediately with a Zeiss LSM900. Panels from top left to bottom right: X-33, AtCSLA2 and AkCSLA3-sfGFP alone, α :mRuby2, α :TpoICBM16-1 and PaCBM35 expressed in three different yeast backgrounds. TL = transmitted light; merged = merged images of all fluorescence channels.

3.2.2.4 Influence of CSLA Expression on Yeast Golgi Morphology

One possible reason for the co-localization of CBMs and CSLAs could stem from the potential damage that CSLA overexpression may inflict on the Golgi apparatus. As membrane-bound proteins, CSLAs anchor themselves to the membranes of the Golgi apparatus (Voiniciuc et al., 2019). Their overexpression could potentially impair the Golgi apparatus's secretory function, leading to CBM retention in the Golgi apparatus and their co-localization with the CSLAs. To scrutinize this hypothesis, the Golgi morphology of AkCSLA3-sfGFP-expressing yeast was examined by co-expressing it with the constitutively expressed Golgi-localized protein ScGOS1 (McNew et al., 2000) that was fused to mRuby2 for *in vivo* visualization. Following this, the colonies were subjected to various chemical treatments. After 24 h of cultivation in YPD for ScGOS1-mRuby2 expression, the cultures were transferred to YPM medium to additionally express AkCSLA3-sfGFP. The YPM medium was supplemented with one of the following chemicals: DMSO + water, Zymolyase 20 T[®], Anidulafungin, Brefeldin A. The individual effects and concentrations of each chemical in the growth medium are delineated in Table 3.

Table 3 Effects and concentrations of chemicals used for morphology analysis of Golgi-apparatus.

CHEMICAL	ACTIVITY	SOLVED IN	µg/mL	REFERENCE
WATER	Control for effect of DMSO alone	DMSO	8	--
ZYMOLYASE 20T[®]	fungal β -1,3-glucan hydrolysis	Water	2000	(Lacapere, 2017)
ANIDULAFUNGIN	β -1,3-glucan synthase inhibitor	DMSO	8	(Murdoch and Plosker, 2004; Kuti and Kuti, 2010)
BREFELDIN A	BFA, Golgi inhibitor	DMSO	8	(Fujiwara et al., 1988)

Previous research demonstrated that the overexpression of AkCSLA3-sfGFP elicits a toxic response in cells, evidenced by a reduced optical density compared to WT (Chapter 3.1 & (Robert et al., 2021)). The significantly reduced optical density of AkCSLA3-sfGFP with or without the co-expression with ScGOS1-mRuby2 compared to empty vector and ScGOS1-mRuby2 expression is consistent with these previous results. It is also noteworthy that the treatment with only DMSO + water caused a mild OD600 reduction of empty vector and ScGOS1-mRuby2 compared to YPD pre-culture. However, this minor optical density reduction might also be attributed to the media change, as direct growth in methanol poses a higher burden on yeast growth compared to dextrose (Jordà et al., 2013). The strongest impacts on growth were discerned when the cultures were treated with Zymolyase 20T[®] or Anidulafungin, where the OD600 did not significantly differ from the medium control (Figure 20a).

Given that both Zymolyase 20T[®] and Anidulafungin intrude upon the integrity of the cell wall, it could be speculated that the cells may either undergo cell death or encounter difficulties

in executing efficient division (Murdoch and Plosker, 2004; Kuti and Kuti, 2010; Lacapere, 2017). Conversely, the treatment with the Golgi inhibitor caused a much less pronounced effect on cell density. This suggests that the cell wall may assume a more critical role than the Golgi apparatus in the context of cell growth and division.

To account for the substantial optical density variations and allowing to draw reliable conclusions regarding the expression of ScGOS1-mRuby2 and AkCSLA3-sfGFP per cell, the measured GFP and RFP values were normalized to the optical densities of the cultures according to the formula below.

$$\frac{\text{Fluorescence Signal}}{\text{Optical Density (OD600)}} = \text{Normalized Fluorescence}$$

Except for Anidulafungin-treated cells, where AkCSLA3-sfGFP was almost completely abolished, which is evidenced by the lack of significant sfGFP signal compared to the empty vector control (Figure 20b), the expression of AkCSLA3-sfGFP was not affected. A possible hypothesis for that observation might be that Anidulafungin may not only impede the function of fungal β -1,3-glucan synthases but could potentially also interfere with the β -1,4-glucomannan synthase AkCSLA3, considering that both enzymes belong to the same clade of synthetic enzymes (Oehme et al., 2019).

Consistent with prior unpublished findings by Annika Grieb-Osowski, the expression of constitutively expressed ScGOS1-mRuby2 significantly dropped after cell transfer from YPD to YPM medium, indicating that YPM is not an optimal medium for the constitutive expression of this Golgi marker. While ScGOS1-mRuby2 alone showed no RFP signal significantly higher than the medium control, all treatments consistently showed a slightly but significantly higher RFP level compared to the YPM medium control when AkCSLA3-sfGFP was co-expressed with ScGOS1-mRuby2 (Figure 20c). This aligns with the ScGOS1-mRuby2 signal in the YPD pre-culture that was higher upon co-expression with AkCSLA3-sfGFP.

To elucidate the potential effects of the chemical treatments on the amounts of mannan present in the cells, the AKI of the cultures was prepared and subjected to 2 M TFA hydrolysis. As expected, the cultures with lower optical densities presented significantly diminished glucose and mannose levels compared to those with higher cell densities (Figure 21a). When treated with Brefeldin A, the levels of glucose and mannose in AkCSLA3-sfGFP and AkCSLA3-sfGFP/ScGOS1-mRuby2 resembled the ones detected in cells that had been treated with the DMSO + water control. This indicates that either the Brefeldin A-mediated Golgi inhibition might not affect AkCSLA3-sfGFP expression and glucomannan production or that the Brefeldin A treatment may have been unsuccessful. Conversely, the reduced optical densities of Brefeldin A-treated cultures (Figure 20a) suggest that this chemical exerts an effect regarding cell growth and/or proliferation. Of all treatments, only Anidulafungin resulted in a relative reduction of the

Glc:Man ratio, indicating a decrease in AkCSLA3-sfGFP activity (Figure 21b). This finding aligns with the hypothesis that due to structural similarities between AkCSLA3 and fungal β -1,3-glucan synthases, AkCSLA3 may be susceptible to the inhibitory action of Anidulafungin, akin to that exerted on β -1,3-glucan synthases. The absence of a significant AkCSLA3-sfGFP signal increase in Anidulafungin-treated cells reinforces that conclusion and points to a potential off-target effect of Anidulafungin.

The results suggest a pivotal role of cell walls in *Pichia* cell growth and division. In contrast, the disruption of the Golgi apparatus appeared to have a less drastic effect on cellular density and heteromannan content.

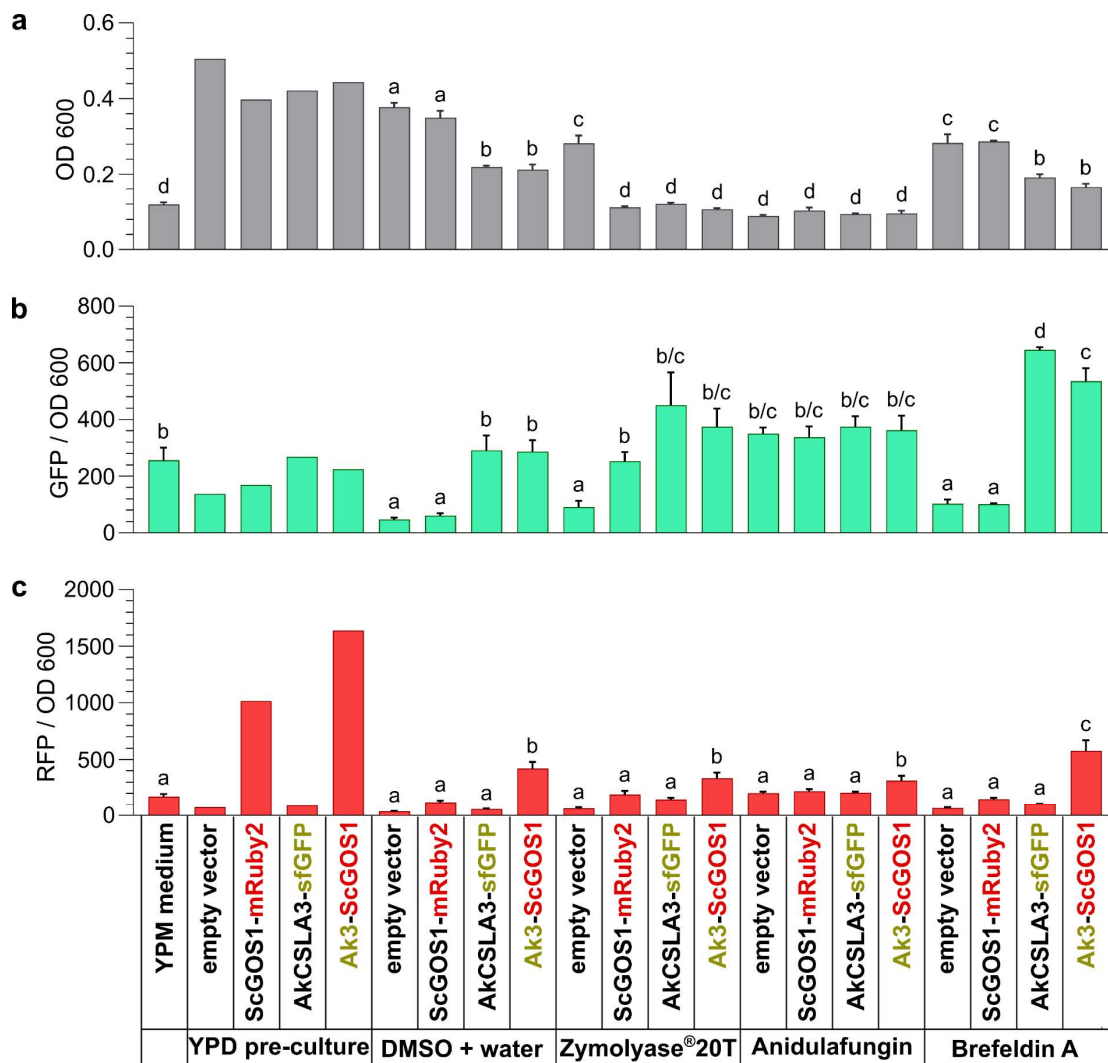


Figure 20 Effects of different chemical treatments on cell density, GFP and RFP signal of AkCSLA3 and ScGOS1. Pre-cultures (1 biological replicate) grew in 2 mL YPD in sterile glass tubes for 24 h and OD600 and fluorescence of small aliquots were measured using a Tecan M1000 and then inoculated in triplicates in 300 μ L YPM with appropriate concentrations of DMSO + water, Zymolyase 20 T[®], Anidulafungin and Brefeldin A (Table 3) for another 24 h in a sterile 48-well plate. Then, fluorescence and OD600 were measured with a Tecan M1000. (a): cell density (OD600). (b): GFP signal normalized to OD600. (c): RFP signal normalized to OD600 Error bars show mean + SD of 3 biological replicates. Letters denote significant differences (one-way Anova with Tukey's pairwise ($p=0.05$)).

To decipher how the chemical treatments would influence protein localization in the cells, treated cells that had grown in YPM with chemical were counterstained with 0.01% CF to visualize native yeast cell wall polysaccharides and then imaged via confocal microscopy with a Zeiss LSM900 (Figure 22).

After 24 h of growth in YPD, the Golgi marker ScGOS1-mRuby2 localized in small punctate throughout the cells, as expected. Following growth in YPM supplemented with DMSO + water for another 24 h, the ScGOS1-mRuby2 strongly diminished. The red signal observed in AkCSLA3-sfGFP and AkCSLA3-sfGFP/ScGOS1-samples can be attributed to an artefact, potentially caused by the intense CF in the respective cells. Nevertheless, in the AkCSLA3-sfGFP/ScGOS1-Ruby cells, a few ScGOS1-mRuby2 punctate co-localized with AkCSLA3-sfGFP. As a matter of fact, the punctate localization pattern of AkCSLA3-sfGFP, both when expressed alone or with ScGOS1-mRuby2, remained unaffected when the cells were grown in YPM with DMSO + water. This supports the assumption that the observations made with Anidulafungin and Brefeldin A treatment were not influenced by the fact that they were solved in DMSO.

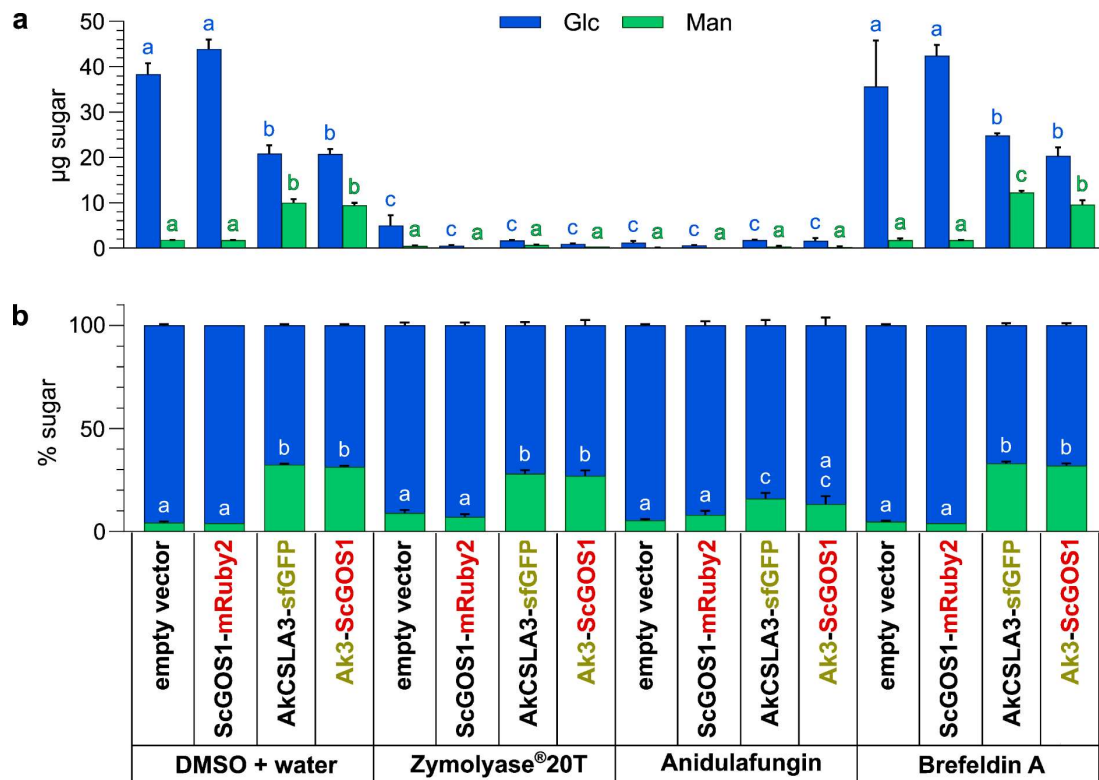


Figure 21 Effects of different chemical treatments on monosaccharide composition of AkCSLA3 and ScGOS1. AKI of colonies shown in Figure 20 was hydrolysed with 2 M TFA, to quantify mannose and glucose contents. (a) µg sugar / 50 µL AKI. (b): relative sugar composition (%). Error bars show mean + SD of 3 biological replicates. Letters denote significant differences obtained by one-way Anova with Tukey's pairwise ($p=0.05$).

While Zymolyase 20T[®] or Anidulafungin had comparable effects on the cell density, fluorescence, and absolute heteromannan content, their impact on cell morphology varied considerably. Like that, cells treated with Zymolyase 20T[®] that expressed the empty vector control closely resembled the wild-type morphology, which aligns with their ability to reach higher optical densities compared to the cells expressing ScGOS1-mRuby2 and/or AkCSLA3-sfGFP (Figure 20a). These latter cells all displayed substantial cell wall impairment, reflected by the lack of CF staining in cells with damaged walls. Only cells with intact cell walls exhibited CF cell wall labelling. In agreement with the plate reading measurements, ScGOS1-mRuby2 could not be detected via microscopy (Figure 20c), while AkCSLA3-sfGFP could still be identified. Interestingly, AkCSLA3-sfGFP maintained its punctate localization pattern, suggesting that the cell wall damage did not affect the Golgi morphology. This indicates that the degradation of β -1,3-glucans by Zymolyase 20T[®] does neither affect AkCSLA3-sfGFP localization nor functionality, which is supported by the relative monosaccharide compositions of AkCSLA3-sfGFP and AkCSLA3-sfGFP/ScGOS1-mRuby2 (Figure 21b).

Although both Zymolyase 20T[®] and Anidulafungin exert similar ultimate impacts on *Pichia* cell walls, namely a reduction of β -1,3-glucans, their different modes of action result in markedly different effects on cell walls. Zymolyase 20T[®] hydrolyses β -1,3-glucans after their production and cell wall incorporation, whereas Anidulafungin inhibits β -1,3-glucan synthesis, preventing polysaccharide formation and incorporation into the cell wall. Interestingly, the lack of β -1,3-glucans induced by Anidulafungin gave rise to cells that were 2-3 times larger than wild-type, which exhibited intense cell wall CF labelling that caused “bleeding” into the red imaging channel. This could indicate that β -1,3-glucans may be essential for proper cell division. While no AkCSLA3-sfGFP signal was observed in the cells expressing co-expressing AkCSLA3-sfGFP and ScGOS1-mRuby2, the signal from AkCSLA3-sfGFP expressed alone was noticeably mislocalized to the cytosol. This observation suggests that the reduced relative mannose content shown in Figure 21b might be attributed to impaired anchoring of AkCSLA3-sfGFP to the Golgi, which could, in turn, prevent it from functioning correctly.

Compared to Zymolyase 20T[®] and Anidulafungin, the impact of Brefeldin A on growth, fluorescence, and monosaccharide content was less pronounced, complicating the evaluation of its effectiveness (Figure 20, Figure 21). Imaging Brefeldin A treated cells unveiled that disruption of the Golgi apparatus prompted the mislocalization of AkCSLA3-sfGFP to cytosol, similar to the pattern observed under Anidulafungin treatment (Figure 22). Intriguingly, the mislocalization of AkCSLA3-sfGFP did not impact the mannan production significantly. Consequently, the loss of Golgi membrane attachment of AkCSLA3-sfGFP in Anidulafungin-treated cells may not be responsible for the observed lack of glucomannan production. A possible explanation could be the mode of action of Anidulafungin. Like that, it could be

possible that Anidulafungin specifically degrades the AkCSLA3 portion of AkCSLA3-sfGFP, and the fluorescence observed may originate from the remaining sfGFP component that we see could possibly be the remaining sfGFP by itself.

Together, the chemical treatments underscore the crucial roles of the cell wall and the Golgi apparatus for *Pichia* growth and cell division. Interestingly, damages to the Golgi apparatus did not affect the function of the Golgi-membrane-bound AkCSLA3-sfGFP as much as expected. The results agree with prior research suggesting that AkCSLA3-sfGFP overexpression toxicity on *Pichia* cells may stem from changes to the cell wall structure rather than from issues with the Golgi apparatus (Chapter 3.1 & (Robert et al., 2021)).

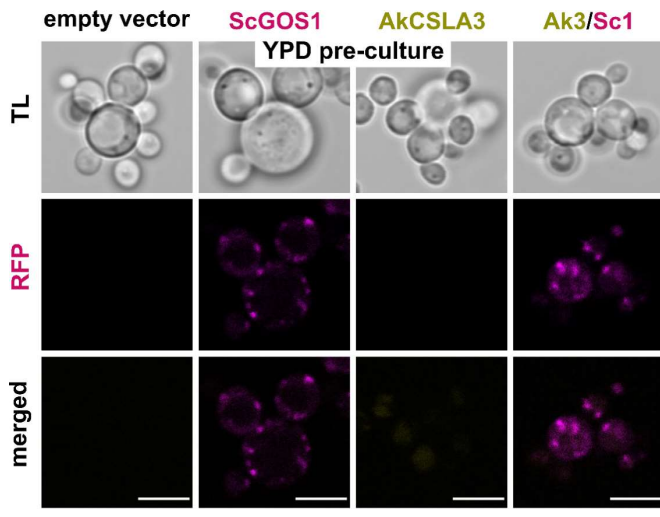
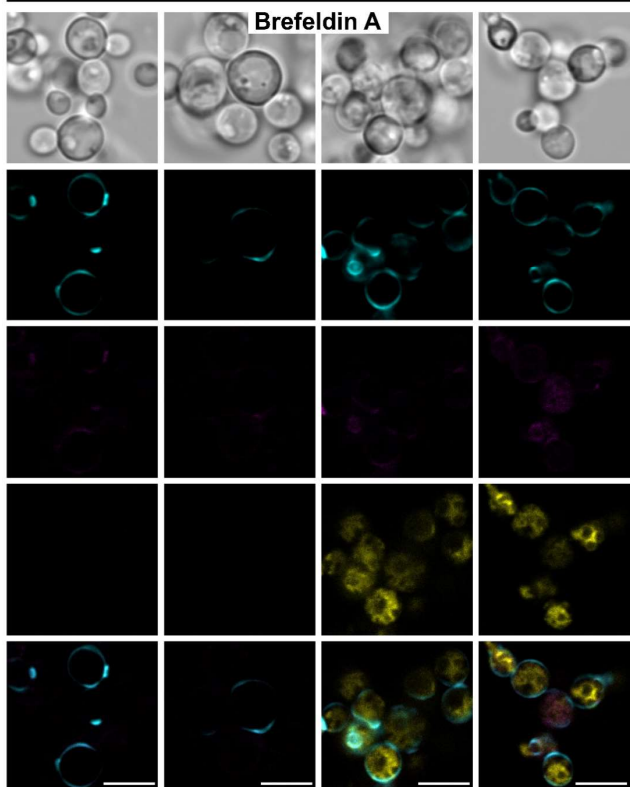
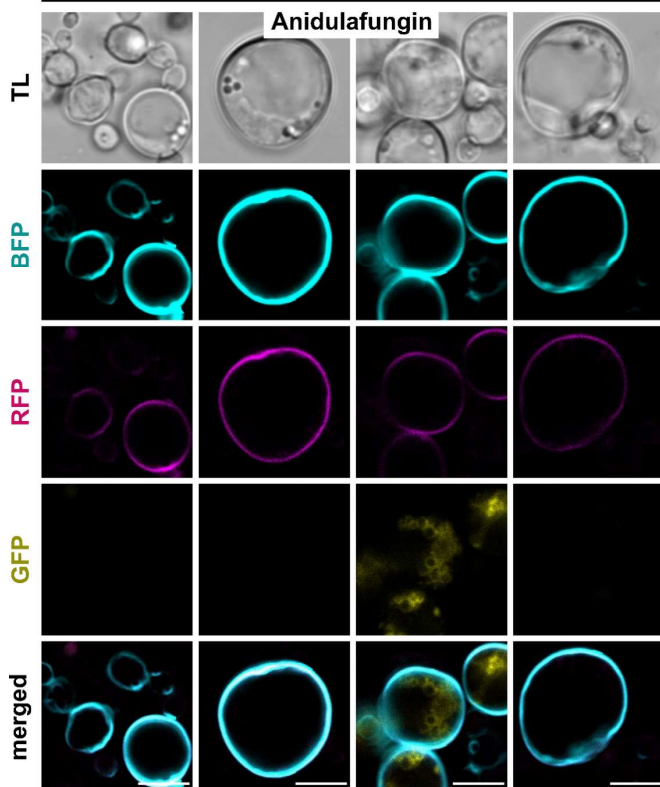
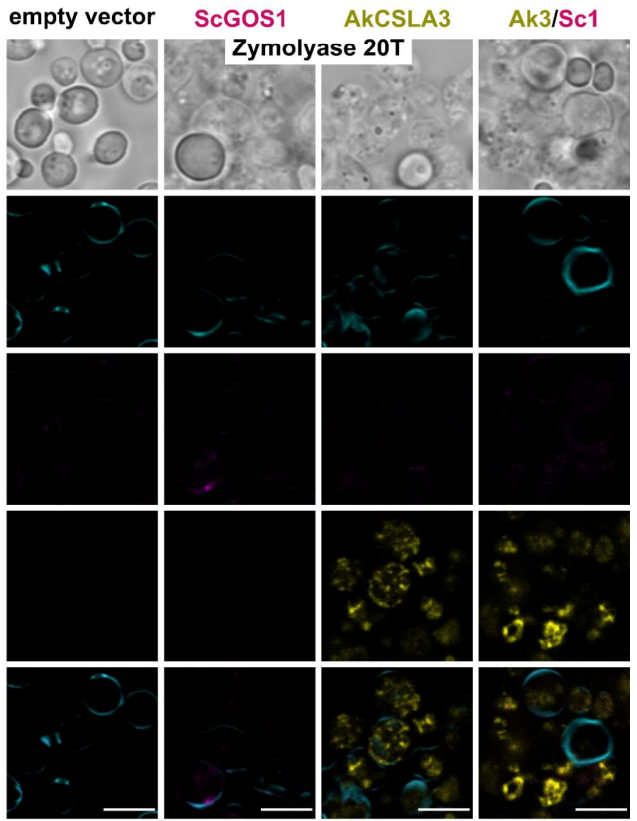
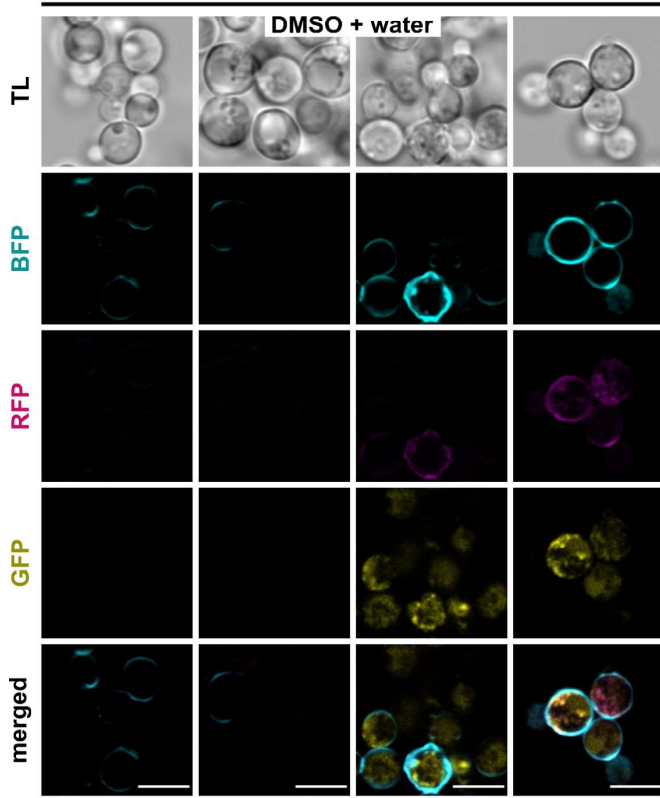


Figure 22 Effects of chemical treatments on AkCSLA3-sfGFP localization in *Pichia* cells. Imaged cells are from the colonies shown in Figure 20 and Figure 21. Prior to confocal imaging with a Zeiss LSM900, cells were counterstained with 0.01% CF (BFP) to label native yeast cell wall polysaccharides. Each panel shows a different treatment. From top to bottom left: YPD preculture, DMSO+ water, Anidulafungin; from top to bottom right: Zymolyase 20T[®], Brefeldin A. Scale bar = 5 μ m.



3.2.2.5 Immunolabelling of Heterologous Heteromannan in Yeast

The second hypothesis for the co-localization of CSLAs and CBMs (Figure 19, Figure 28c) proposed that contrary to what was expected, the majority of heteromannans produced by the CSLAs in yeast would not localize to the cell wall but instead remain in proximity to the CSLAs. It should be noted that previous experiments implied an influence of AkCSLA3 overexpression on yeast cell wall integrity, suggesting that at least part of the glucomannan synthesized by AkCSLA3 is likely incorporated into the cell wall (Chapter 3.1). However, this observation does not preclude the possibility that the majority of the glucomannan synthesized by AkCSLA3 may not be incorporated into the cell wall.

A co-cultivation approach was implemented to verify this hypothesis while bypassing the need to navigate CBMs through a potentially impaired Golgi apparatus due to CSLA expression. The aim of co-cultivating α :PaCBM35:mRuby2 and AkCSLA3-sfGFP expressing *Pichia* cells in a 1:1 ratio was to enable the labelling of glucomannan incorporated in the cell walls of AkCSLA3-sfGFP expressing cells.

In these co-cultures, the AkCSLA3-sfGFP signal and the α :PaCBM35:mRuby2 signal were reduced by 50% compared to monocultures. While this ratio could be confirmed via confocal imaging with a Zeiss LSM700, no α :PaCBM35:mRuby2 signal was detected at the edges of cells expressing AkCSLA3-sfGFP (results not shown). The lack of α :PaCBM35:mRuby2 signal at the cell walls of AkCSLA3-sfGFP expressors could be attributed to two possible reasons. Like that, the absence of α :PaCBM35:mRuby2 signal could result from a lack or minimal amount of glucomannan in the cell walls of AkCSLA3-sfGFP expressors. Alternatively, the concentration of secreted α :PaCBM35:mRuby2 in the medium may have been insufficient to efficiently bind to the glucomannan in the cell walls of AkCSLA3-sfGFP-expressing cells. To discern which of the aforementioned scenarios would be more likely, heteromannan immunolabelling of walls from CSLA-expressing cells was conducted.

For immunolabelling, previously investigated AtCSLA2 and AkCSLA3-expressing cells (paragraphs 3.2.2.1, 3.2.2.2, 3.2.2.3) were labelled with the anti-heteromannan monoclonal antibody LM21 (Marcus et al., 2010). Previous heteromannan immunolabelling in AkCSLA3-expressing yeast spheroplasts revealed that heteromannans produced by AkCSLA3 were deposited in the extracellular matrix (Voiniciuc et al., 2019). Since these results were inconsistent with the CBM localization observed in Figure 19 and Figure 28c, the experiment was repeated with AkCSLA3. Furthermore, the localization of glucomannans produced by AkCSLA3 was compared with that of mannan produced by AtCSLA2.

As LM21 cannot cross the cell membrane, yeast cells were spheroplasted with 2-mercaptoethanol and Zymolyase 20T[®] prior to immunolabelling. In addition to causing

spheroplasting, Zymolyase 20T[®] also served to partially remove native yeast polymers, increasing the accessibility of LM21 to heterologously produced heteromannans. To rule out that spheroplasting the cells might influence heteromannan presence or localization, a portion of the colonies was fixed with acetone immediately after spheroplasting as a control.

Following the primary antibody LM21 application, the secondary antibody Alexa Fluor 488 was applied to allow visualization of the LM21 localization. For confocal microscopy with a Zeiss LSM900, cells were counterstained with 0.01% CF (BFP) to detect native yeast cell wall polysaccharides and 0.01% PI to label dead cells.

As a control against potential unspecific cell labelling by Alexa Fluor 488, a subset of cells was treated with Alexa Fluor 488 only. As expected, neither the fixed nor the unfixed forms of any tested genotypes displayed Alexa Fluor 488 GFP signal. Intense PI labelling was observed in all cells except unfixed X-33 and fixed AtCSLA2, indicating cell death. CF staining yielded more varied results. Unfixed AkCSLA3 and fixed X-33 and AkCSLA3 all exhibited CF staining, whereas unfixed X-33 and AtCSLA2 and fixed AtCSLA2 were not labelled with CF (Figure 23). The weakness or absence of CF signal observed in most cells, regardless of whether they were stained with both antibodies or only with the secondary antibody, suggests that the Zymolyase 20T[®] treatment might have removed a significant portion of the native yeast polysaccharides. The GFP signals in cells labelled with LM21/Alexa Fluor 488 were notably stronger for the CSLA expressors than for the X-33 wild-type, demonstrating the absence of heteromannan in the X-33 wild-type.

In both fixed and unfixed cells, LM21 was observed in punctate structures at the yeast cell walls and, to some degree, within the cells. Notably, there were no obvious LM21 localization differences between AtCSLA2 and AkCSLA3 strains (Figure 23). This suggests that the post-biosynthesis cellular routing and cell wall incorporation of mannan from AtCSLA2 and glucomannan from AkCSLA3 may be highly similar or even identical.

The immunolabelling results suggest that heterologous HM might mostly be incorporated into the cell walls. On the other hand, LM21 might not be able to penetrate the cell walls due to its size (150 kDa (Charles A Janeway et al., 2001)), ultimately preventing the labelling of intracellular HM. If most HM were incorporated into the cell walls, it would not explain the co-localization of the CBMs and CSLAs. Since the CBMs and the CSLAs are expressed simultaneously, it could be speculated that the CBMs bind heterologous heteromannans directly after their biosynthesis and before initiation of their transport to the cell wall. These CBM/heteromannan complexes might then be hindered in their movement to the cell wall, which could lead to the observed CBM/CSLA co-localization.

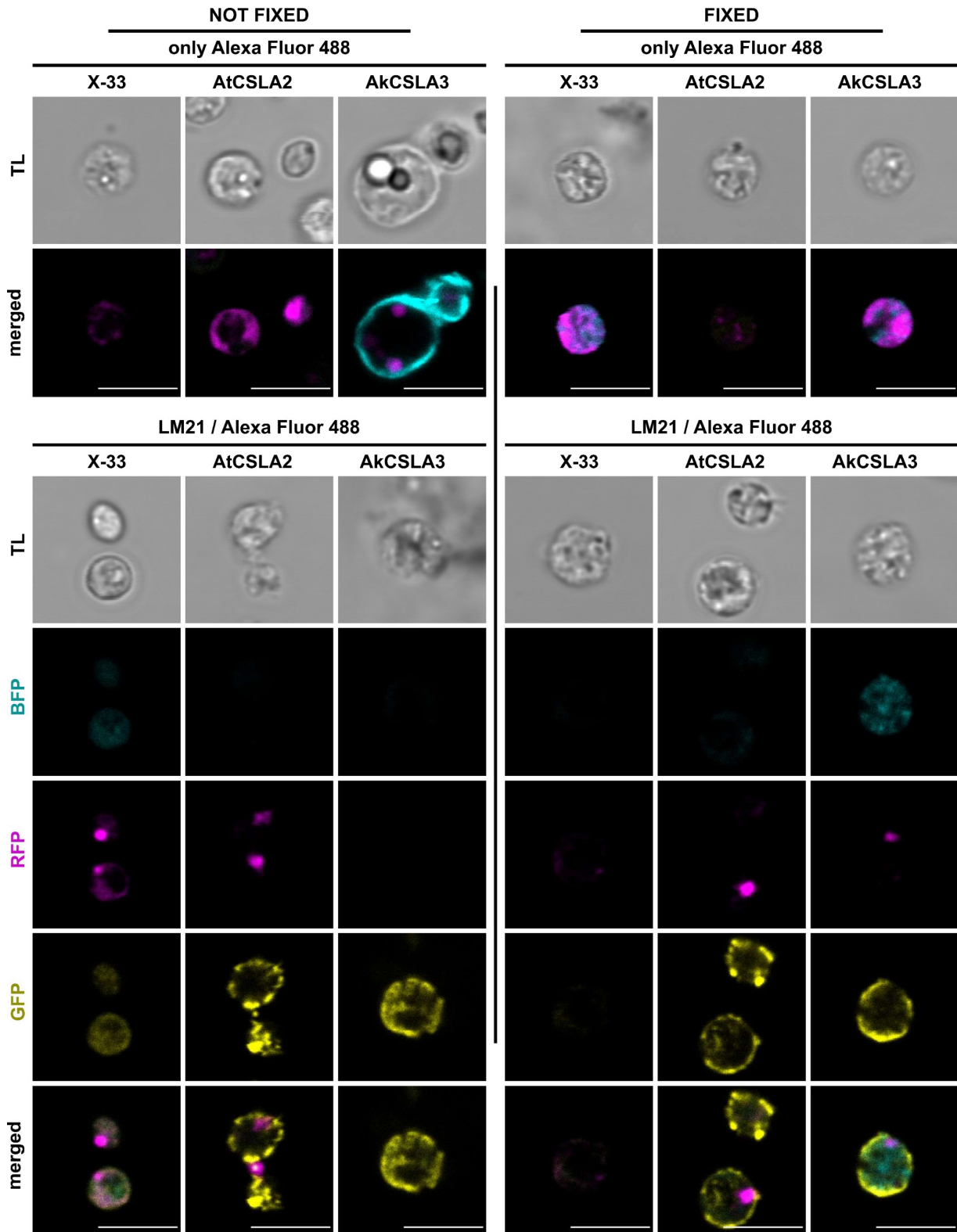


Figure 23 Immunolabelling of HM in AtCSLA2 and AkCSLA3 expressing cells using the primary anti-HM antibody LM21 combined with Alexa Fluor 488 as secondary antibody. Cells grew in sterile glass tubes in 3 mL YPM+G for 72 h. Prior to immunolabelling, cells were fixed using a 1:1 ratio of culture and acetone before to freezing at -20 °C for 20 min. Then, samples were washed with PBS buffer and water. For immunolabelling cells incubated with the primary antibody LM21 in the dark at roomtemperature for 90 min and were washed thoroughly with PBS. Alexa Fluor 488 (GFP) was then applied for 90 min before thorough washing with PBS. Cells were counterstained with 0.01% CF (BFP) to visualize native yeast cell wall polysaccharides and 0.01% PI (RFP) to label dead cells. Images were acquired with a confocal Zeiss LSM900. TOP: Controls with only Alexa Fluor 488. Only transmitted light (TL) and merged image of blue, green and red fluorescence are shown. BOTTOM: Both antibodies were applied to unfixed (left) and fixed (right) cells. Scale bar = 5 µm.

3.2.3 Introducing a Ratiometric Probe to Assess CBM-Expression Levels

A notable issue using mRuby2 for CBM tagging was the limited secretion of the α -mRuby2 control (Figure 16, Figure 18, Figure 19). Moreover, mRuby2 does not belong to the strongest red fluorescent proteins. To optimize the fluorescent tagging, mRuby2 was replaced with the stronger and more photostable fluorescent protein pmScarlet-I (Botman et al., 2019). To further enhance the optimization of the new fluorescent tag, the turquoise fluorescent protein pSmTurquoise2 (pSmTq2), one of the best performing blue/cyan fluorescent proteins (Botman et al., 2019), was fused to pmScarlet-I using a 2A peptide for cleaving (Ryan et al., 1991). The objective was for pSmTq2 to self-cleave from the CBM-pmScarlet-I complex and remain in the cytosol. By utilizing the pSmTq2 ratiometric probe, CBM expression could be assessed without relying on supernatant analysis. Instead, the presence of cytosolic pSmTq2 would serve as indicative markers of CBM expression (Figure 24 c). The ratiometric approach holds particular significance for future *in planta* experiments, where the analysis of secretion via assessment of fluorescence in the supernatant, as done in yeast, is not possible.

In addition to expressing CBMs in yeast to determine heteromannan levels, the new CBM probes should be tools to detect and evaluate the quantity and localization of (gluco)mannan in wild-type plants and heteromannan mutants. For later expression in plants, the performance of the effective *Saccharomyces cerevisiae* α -secretion peptide (Zou et al., 2022) was compared to the one of a new signal peptide derived from *Arabidopsis thaliana* Expansin 10 (AtEXP10). This signal peptide was previously used successfully to target recombinant proteins to the *Arabidopsis* cell wall (Cho and Cosgrove, 2000; Zhang, 2014). It was hypothesized that - when expressed *in planta* - a plant-native signal peptide may be more efficient than yeast native α -secretion peptide.

The performance of both signal peptides was assessed in *Pichia* X-33 using the new fluorescent tags. Quantitative fluorescence measurements revealed α -pSmTq2 alone secreted by approximately 60%, while the secretion of pSmTq2 without a signal peptide was negligible. Regardless of the signal peptide, the expression and secretion rates of TpolCBM16-1 and PaCBM35 were comparable. With approximately 30%, the secretion rates of pSmTq2 from CBM-expressing cells were intermediate between α -pSmTq2 and pSmTq2 alone (Figure 24a). This suggests that a significant portion of approximately 70% of pSmTq2 remained in the CBM-expressing cells. Moreover, the secretion rates of the CBM-pmScarlet-I complexes ranged from approximately 60–90% (Figure 24b). The high secretion of CBM-pmScarlet-I complexes compared to the low secretion of pSmTq2 in these strains indicated successful cleavage of pSmTq2 peptides from the CBM-pmScarlet-I complexes before their secretion out of the cells, aligning with the intended function of the self-cleaving pSmTq2 as ratiometric probes.

In addition to the added asset of a ratiometric probe for improved fluorescent tagging, the use of pmScarlet-I was able to eradicate the secretion problems encountered when using the mRuby2 tag. Like that, α -pmScarlet-I alone exhibited similar secretion levels as the CBMs in wild-type, unlike α -mRuby2 alone, which displayed significantly lower secretion rates than the mRuby2-tagged CBMs (Figure 16b, Figure 18, Figure 24b). Additionally, the comparability of the α -pmScarlet-I secretion and the secretion of the CBM equipped with the pmScarlet-2A-pSmTq2 tag demonstrates that the addition of the self-cleaving ratiometric probe does not influence the CBM secretion from the wild-type background. It is to be noted that the expression of pmScarlet-I, with or without signal peptide, was considerably higher than the expression of the CBMs (Figure 24b). This expression level difference was also observed between mRuby2-tagged CBMs and their fluorescent control and could be related to the greater complexity of the CBM constructs, resulting in a higher transcription and synthesis burden on the yeast cells (Figure 16b, Figure 18).

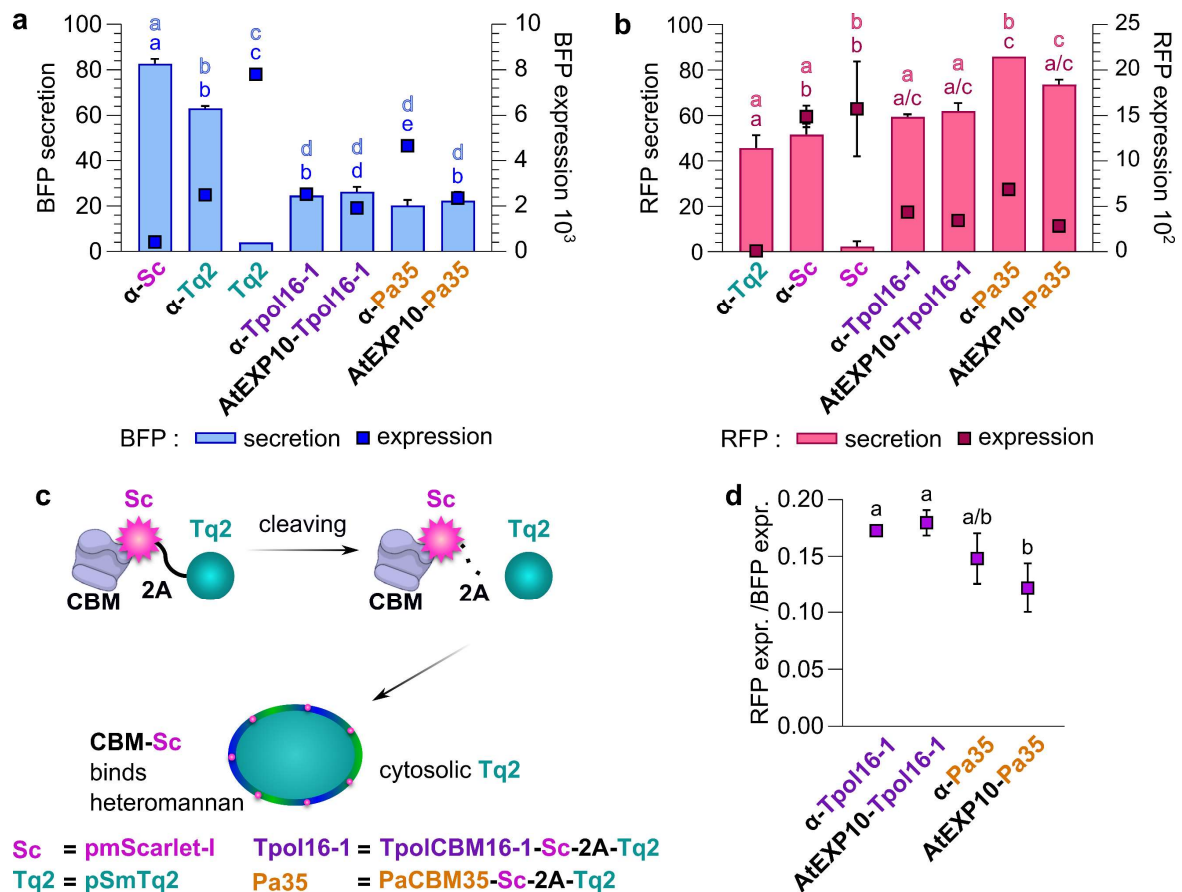


Figure 24 Expression and secretion of CBMs with different signal peptides. Colonies grew in sterile 48 well plates in 300 μ L YPM+G for 72 h. (a), (b) Expression (squares) and secretion (bars) of pSmTq2 and pmScarlet-I respectively, (c) mode of action of the self-cleaving pSmTq2 probe and anticipated localization of CBM-pmScarlet-I and pSmTq2 in yeast cells. 2A is a peptide for pSmTq2 self-cleaving. (d) ratio between pmScarlet-I (RFP) and pSmTq2 (BFP) expression in strains expressing TpolCBM16-1 or PaCBM35 with different signal peptides for secretion. Bars and squares represent mean + SD of 3 biological replicates. Letters denote significant differences obtained by one way ANOVA with Tukey's pairwise ($p=0,05$) for (a) and (b) top letters show significance for secretion and bottom letters show significant differences for expression.

Both tested signal peptides showed slightly different behaviours depending on whether they were used in conjunction with TpolCBM16-1 or PaCBM35. For TpolCBM16-1, which secreted less than PaCBM35 regardless of the signal peptide, no significant difference was observed in the performance of both signal peptides. In the case of PaCBM35, the α -signal exhibited a slight but significantly elevated secretion rate compared to the AtEXP10 signal peptide (Figure 24b). Given that the secretion of AtEXP10-PaCBM35 still remained significantly higher than the secretion of α -TpolCBM16-1 and AtEXP10-TpolCBM16-1, both signal peptides represented suitable candidates for further experiments *in planta*.

The calculated ratios between pmScarlet-I and pSmTq2 signals unveiled no significant differences between the signal peptide/CBM combinations, except for AtEXP10-PaCBM35 exhibiting a slightly lower ratio (Figure 24c). The consistent ratio between pmScarlet-I and pSmTq2 confirms that pSmTq2 is a suitable tool to measure total CBM expression.

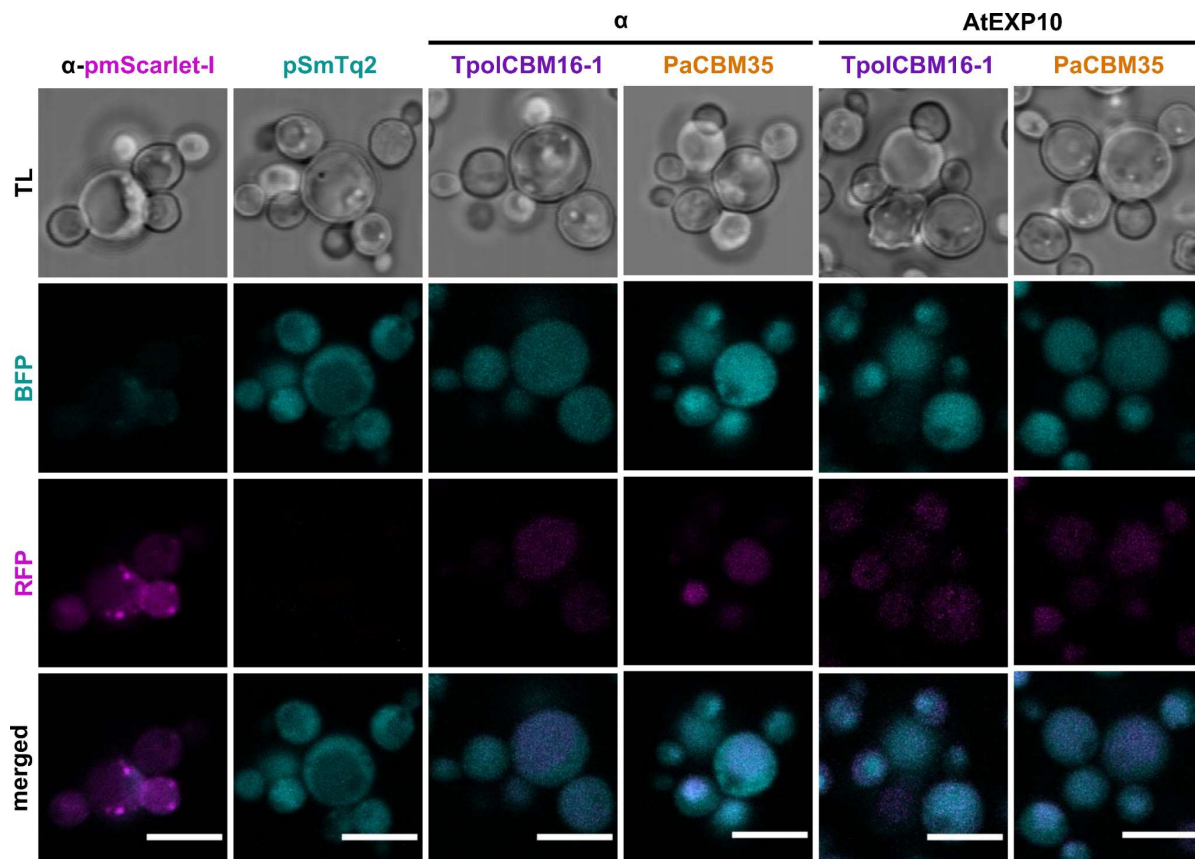


Figure 25 Localization of CBMs with different signal peptides in yeast. Selected colonies from Figure 24 were imaged using a Zeiss confocal microscope (LSM780). From left to right: α -pmScarlet-I, pSmTq2, TpolCBM16-1 and PaCBM35 with α -secretion peptide and TpolCBM16-1 and PaCBM35 with AtEXP10 secretion peptide. From top to bottom: transmitted light (TL), blue channel (BFP), red channel (RFP) and merge of blue and red channel. Scale bar = 5 μ m.

To confirm the cytosolic localization of pSmTq2, selected colonies from the ones shown in Figure 24 were imaged via confocal microscopy using a Zeiss LSM780. Compared to the non-pSmTq2 expressing α -pmScarlet-I, cells expressing pSmTq2 alone exhibited a clear

cytosolic pSmTq2 signal. In addition to being similar in intensity, the pSmTq2 signal in the CBM expressing strains resembled the localization observed in pSmTq2 alone. This observation underlines the robustness of the system set in place, as the expression of the TpolCBM16-1 and PaCBM35 – regardless of the employed secretion signal peptide - and the presence and localization of the pSmTq2 ratiometric probe do not interfere with each other. It is noteworthy that the strongly elevated α -pmScarlet-I expression compared to the expression of the CBM constructs resulted in a weak pmScarlet-I signal. However, since the signal was localized in the cytosol and not in punctate as observed when CBMs and CSLAs are co-expressed (Figure 19), it can be presumed that the presence of heteromannans does not cause the observed CBM retention but might rather be by the high α -pmScarlet-I expression levels.

3.2.3.1 Influence of Signal Peptides on CBM Behaviour in Yeast

The key feature of the new CBM-based probes lies in their capability to secrete out of non-heteromannan-producing cells like *Pichia X-33*. It was, therefore, crucial to verify that the employed signal peptides drove the CBM-probe secretion and if CBMs without a signal peptide would remain in *Pichia X-33* cells. Additionally, the expression of CBMs with no signal peptide enabled the evaluation of CBM localization in *Pichia X-33*. Another essential aspect of this experiment was to verify if the CBMs localized to the cytosol, distinct from their punctate localization in heteromannan-producing yeast strains.

Evaluation of the pSmTq2 signal and the total pmScarlet-I signal in the cultures showed comparable expression levels for TpolCBM16-1 and PaCBM35 with and without signal peptide, enabling a reliable comparison of their respective secretion rates (Figure 26a and b). The CBM secretion rate analysis confirmed the expectations that (a) α -TpolCBM16-1 and α -PaCBM35 were nearly completely secreted and that (b) absence of the α signal peptide resulted in minimal secretion of TpolCBM16-1 and PaCBM35, with levels below 10% (Figure 26a). These results confirm that the α signal peptide is the driving force of CBM secretion. In addition, the findings support the observed correlation between α -CBM retention and heteromannan presence in yeast cells as described in Chapters 3.2.2.1, 3.2.2.2, 3.2.2.3).

The localization of TpolCBM16-1 and PaCBM35 without signal peptide was examined via confocal microscopy (Zeiss LSM780). The presence of pSmTq2 in all analysed CBM-expressing strains confirmed the similar expression of α -CBMs and no-SP-CBMs (no signal peptide). In addition, the high signal intensity of pSmTq2 alone compared to the pSmTq2 signal in CBM-expressing cells was consistent with the fluorescence plate reading measurements (Figure 26b and c). The specificity of the pSmTq2 was confirmed by the lack of turquoise fluorescence in the pmScarlet-I control. Both pSmTq2 alone and the cleaved pSmTq2 from the CBMs localized to the cytosol, consistent with previous observations (Figure 25).

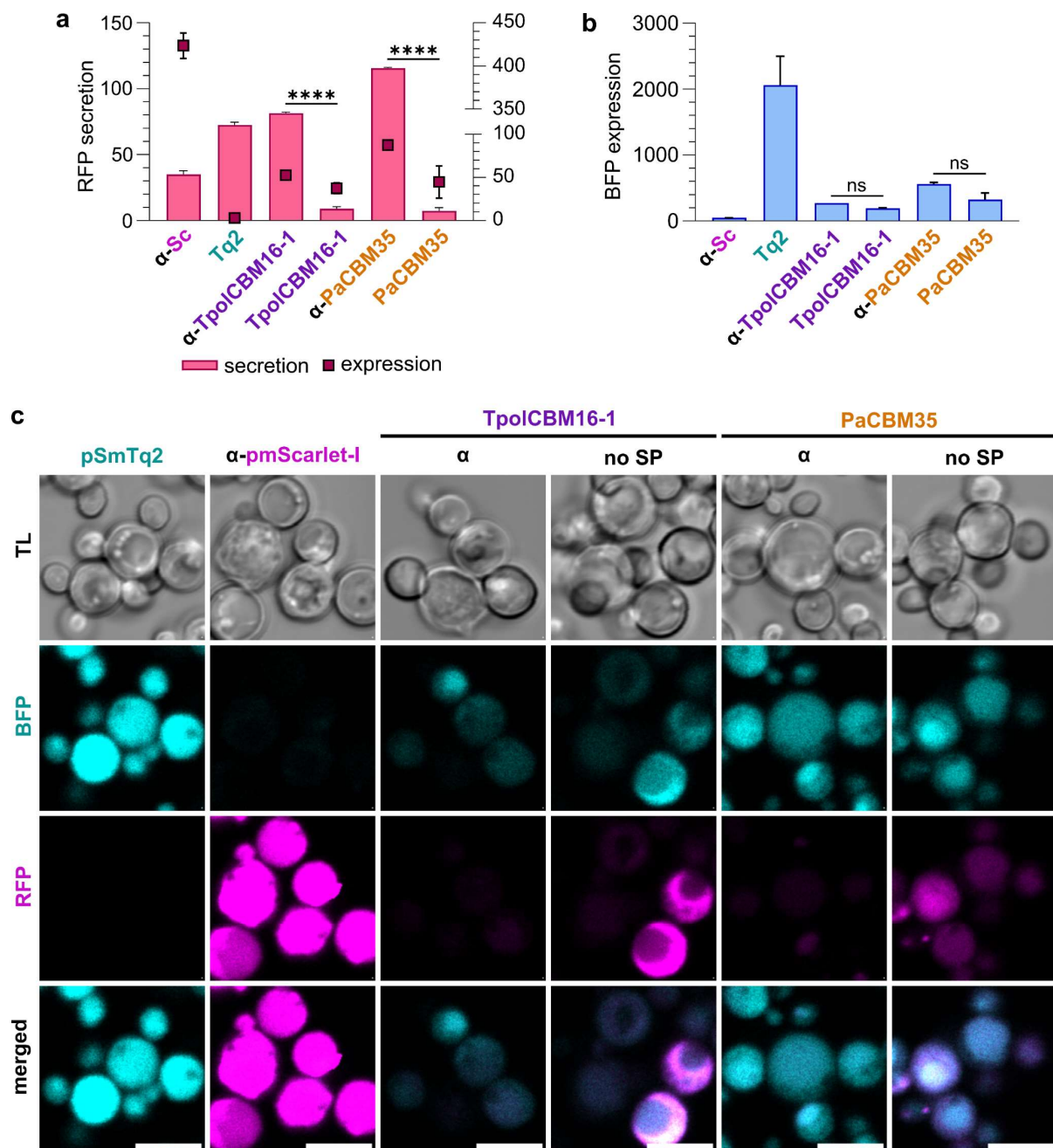


Figure 26 CBM expression and secretion with and without signal peptide. Cells were grown in sterile 24-well plates in 600 μ L YPM + G for 72 h. Fluorescence was measured with a Tecan M1000. (a) red fluorescence of pmScarlet-I representing expression (squares) and secretion (bars); mean + SD of 2-3 biological replicates. Significant differences between secretion of CBMs with and without signal peptide are shown as asterisks and were obtained by one-way Anova with Tukey's pairwise ($p=0.05$) (b) CBM expression assessed through levels pSmTq2 (Tq2). Bars show mean + SD of 2-3 biological replicates. No significant differences were found (ns) (c) TpolCBM16-1 and PaCBM35 with and without signal peptide tagged with pmScarlet-I-2A-Tq2. TL = transmitted light, BFP = pSmTq2 signal, RFP = pmScarlet-I signal. Images were acquired via confocal microscopy using a Zeiss LSM780.

The imaging results aligned with the CBM secretion rates determined via quantitative fluorescence analysis. TpolCBM16-1 and PaCBM35 were only observed when expressed without α , while no pmScarlet-I signal was detected in the cells expressing α -TpolCBM16-1 and α -PaCBM35. Just as the α -pmScarlet-I control, the no-SP-CBMs localized cytosolically. The high α -pmScarlet-I signal was attributed to its high expression and comparably low

secretion in this particular experiment. Nevertheless, the cytosolic localization of the no-SP-CBMs demonstrated that the tested CBMs only localize in punctate when expressed in heteromannan-producing yeast backgrounds. The results support the hypothesis that the presence of heteromannan or heteromannan-producing proteins is required for a punctate CBM localization, as shown in Figure 19 for mRuby2-tagged CBMs and in Figure 28 for CBMs tagged with pmScarlet-I-2A-pSmTq2.

In conclusion, this experiment highlights the effectiveness of the self-cleaving pSmTq2 as a tool to visually confirm CBM expression, even after their secretion from the cells.

3.2.3.2 Influence of Glucomannan Abundance on PaCBM35 Secretion

The secretion of TpolCBM16-1 and PaCBM35 correlated with the amount of mannose present in AtCSLA2 and AkCSLA3-sfGFP-expressing yeast strains. Notably, the CBMs secreted the least when expressed in the AkCSLA3-sfGFP background, which produced higher amounts of mannose than the AtCSLA2 background (Figure 16). These results indicate that the mannose amount might be determinative of the CBM secretion rate. This is supported by the fact that both TpolCBM16-1 and PaCBM35 are known to bind the middle of mannan oligosaccharide chains (Bae et al., 2008; Su et al., 2010; Couturier et al., 2011; Couturier et al., 2013), implying that they would bind the mannan parts of the glucomannan produced by AkCSLA3-sfGFP. To validate that hypothesis, it was investigated how the secretion rate of PaCBM35 would respond to different levels of the same product.

Therefore, PaCBM35 with the new pmScarlet-I-2A-pSmTq2 tag was expressed in two backgrounds that produce different amounts of glucomannan, namely AkCSLA3-sfGFP (high glucomannan) and sfGFP-AkCSLA3 (low glucomannan). Preliminary unpublished findings by Annika Griebß-Osowski had shown that C-terminal tagging of AkCSLA3 with sfGFP caused a 5.4-fold reduction in mannose compared to N-terminal tagging (Figure 27b). Importantly, the fold change of 5.9 between the sfGFP signals of AkCSLA3-sfGFP and sfGFP-AkCSLA3 was similar to the measured mannose amounts. This correlation between fold changes allowed a straightforward estimation of glucomannan abundance via sfGFP signal strength and, in turn, the assessment of the influence of glucomannan abundance on PaCBM35 secretion. To rule out a potential impact of the chosen signal peptide on the secretion differences of PaCBM35 between AkCSLA3-sfGFP and sfGFP-AkCSLA3, both the α and the AtEXP10 signal peptide were employed. This setup allowed to minimize unknown variables in the experimental setup and to enhance the reliability of the observed potential correlation of PaCBM35 secretion and glucomannan presence.

Fluorescent plate reading analysis confirmed the significantly lower expression of sfGFP-AkCSLA3 than AkCSLA3-sfGFP (Figure 28a). Consistent with the previously reported differences in glucomannan yield, α -PaCBM35 secretion was significantly increased in the sfGFP-AkCSLA3 background compared to the AkCSLA3-sfGFP background. As expected, the highest secretion of α -PaCBM35 was observed in the non-glucomannan-producing X-33 wild-type. For AtEXP10-PaCBM35, the secretion from the sGFP-AkCSLA3 was also higher than from the AkCSLA3-sfGFP background, although here, the difference in secretion rates was not statistically significant. Nevertheless, AtEXP10-PaCBM35 secretion was significantly reduced in the AkCSLA3 expressing strains compared to the secretion from X-33 (Figure 28b). The results demonstrate that both the α signal peptide and the AtEXP10 peptide promote high CBM secretion in the absence of heterologous heteromannan in *Pichia*. Additionally, the results suggest a tendency of the PaCBM35 secretion to correlate with the amount of glucomannan present. Hence, the CBMs can serve as a tool for rapid detection of glucomannan production in CSLA-expressing strains.

Qualitative assessment of the PaCBM35 signal intensity via confocal microscopy with a Zeiss LSM780 confirmed the correlation between PaCBM35 secretion and glucomannan amounts detected via fluorescent plate reading measurements (Figure 28a and b). While the α -pmScarlet-I control was efficiently secreted out of the AkCSLA3 expressing cells, regardless of the AkCSLA3 tagging strategy, PaCBM35 was retained in the same genotypes (Figure 28c). Consistent with previous results (Figure 19), α -PaCBM35 and AkCSLA3-sfGFP co-localized in the cells in punctate. The same localization pattern was also observed for AtEXP10-PaCBM35 demonstrating that AtEXP10 is not only a suitable signal peptide for secretion analysis but also that the PaCBM35 behaviour in AkCSLA3-sfGFP expressing strains is robust against the change of signal peptides. This is particularly important in the context of the potential application of CBM-probes *in planta* or other organisms where different signal peptides might be required for secretion.

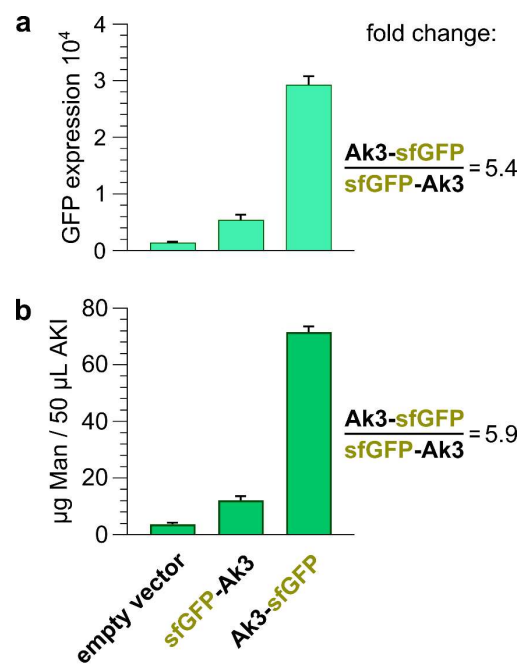


Figure 27 Influence of N- and C-terminal sfGFP tag of AkCSLA3 (Ak3) on fluorescence and mannose content. Experiment was carried out by Annika Grieß-Osowski, and figure was prepared by Madalen Robert. Cells were grown in 2 mL BMGY in plastic culture tubes for 24 h before transfer to BMMY for another 24 h. AKI was hydrolysed with sulfuric acid and submitted to HPAEC-PAD analysis. (a) GFP expression and (b) mannose (Man) content of 2 biological replicates. Bars represent mean + SD. Fold change was calculated by dividing the means of AkCSLA3-sfGFP by the means of sfGFP.

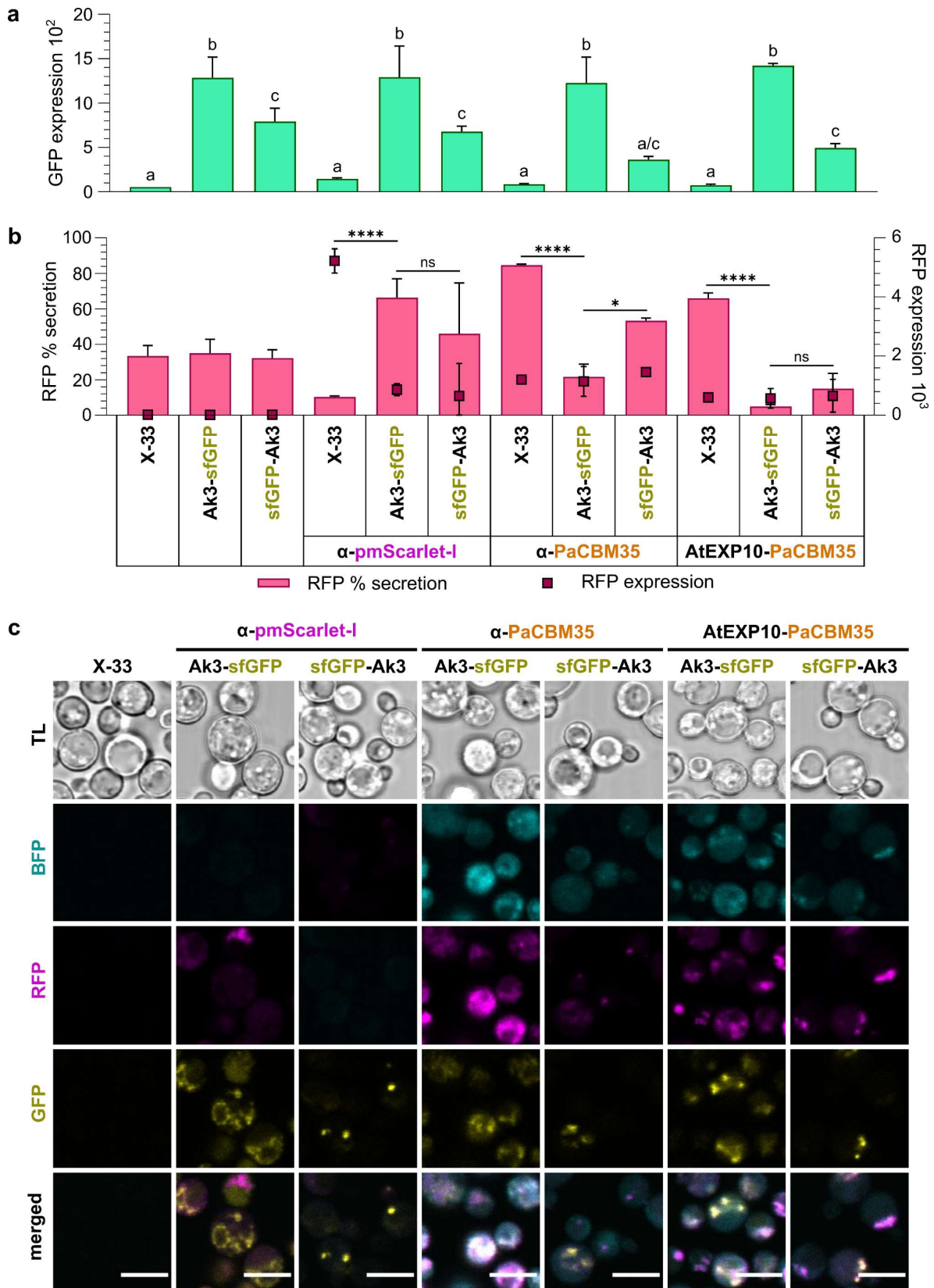


Figure 28 CBM expression and secretion in AkCSLA3 (Ak3) background with C- or N-terminal sfGFP. Cells grew in sterile 24-well plates in 600 μ L YPM+G for 72 h. Fluorescence was measured with a Tecan M1000. Bars represent mean + SD of 3 biological replicates of (a) GFP expression and (b) RFP % secretion (bars) and RFP expression (squares). Letters in (a) and asterisks in (b) denote significant differences based on one-way Anova with Tukey's pairwise ($p=0.05$). X-axis denotes CBM and pmScarlet-I control expression backgrounds. PaCBM35 was tagged with pmScarlet-I-2A-Tq2. (c) Localization of PaCBM35 in AkCSLA3 (Ak3) background with C- or N-terminal sfGFP. Confocal images were acquired with a Zeiss LSM780. Abbreviations: Transmitted light (TL), pSmTq2 (BFP), pmScarlet-I (RFP) and sfGFP signal (GFP). Merged image = overlay of fluorescence channels. Scale bar = 5 μ m.

Despite the lower signal of sfGFP-AkCSLA3 compared to AkCSLA3-sfGFP, the localization patterns of PaCBM35 with different signal peptides persisted when both CBM types were expressed in the sfGFP-AkCSLA3 background. The fact that neither AkCSLA3 nor PaCBM35 localization was affected by the different AkCSLA3 tagging strategies demonstrates the suitability of the newly developed CBM-based system to detect and localize heteromannan *in vivo* and its robustness against changes in signal peptide or CSLA-tagging strategy. It can be concluded that the CBM-based probes represent a robust and versatile tool that could potentially be applicable in various contexts where the assessment of heteromannan amounts and localization is required.

3.3 Application of the New CBM Probes in Yeast and in Plants

After successfully validating the novel CBM probes as indicators for HM presence in *Pichia*, their ability to perform in the context of two key applications was explored. Specifically, it was investigated if the new CBM probes could identify new functional CSLA variants and visualize the native HM distribution *in planta*. The discovery of new functional CSLA variants is crucial for understanding the relationship between amino acid changes and the type and quantity of HM. This could enable the design of higher yielding CSLAs and customized HM fine structures. Using CBMs as probes to visualize the native distribution of HM *in planta* will provide invaluable insights into the influence of HM distribution and dynamics on plant systems in terms of phenotype and nutritional value. This knowledge could be a first step towards engineering crops with higher yields and enhanced nutritional benefits.

3.3.1 Application in Yeast: Discover New Functional HM-Producing Proteins

To explore the potential of the novel CBM probes for high-throughput identification of novel and functional CSLA variants through single-cell analysis, a library of random AkCSLA3 mutants was generated. AkCSLA3 provides health benefits, especially regarding lifestyle diseases such as diabetes (Behera and Ray, 2016). Using error-prone PCR, random mutations, with an average rate of 5 mutations per 1000 base pairs, were introduced in the AkCSLA3 portion of AkCSLA3-sfGFP.

As proof of concept for recombinant CBMs as detectors for new functional CSLA variants and to confirm that the CBM probe secretion is related to heteromannan and not CSLA presence, 10 randomly selected mutated plasmids were used to (a) verify the mutation rates, (b) potentially identify new functional AkCSLA3 variants, (c) evaluate the CBM probe performance when co-expressed with non-functional AkCSLA3 variants.

3.3.1.1 Characterization of Mutated CSLAs

The sequences of the 10 randomly selected plasmids were determined via whole plasmid sequencing and aligned to the TOPCONS 2.0 prediction of the parental AkCSLA3 to assess where the mutations were localized. Additionally, the borders used to swap domains in Chapter 3.1 were also added to the alignment to evaluate in which of the domains selected for the domain swap experiments the mutations would localize (Figure 29). This examination is essential given the variable impacts of different AkCSLA3 domains on the protein's functionality. Exploring the potential correlation between AkCSLA3 functionality and the location of specific amino acid mutations could provide key insights into how mutations in different domains may influence AkCSLA3 functionality.

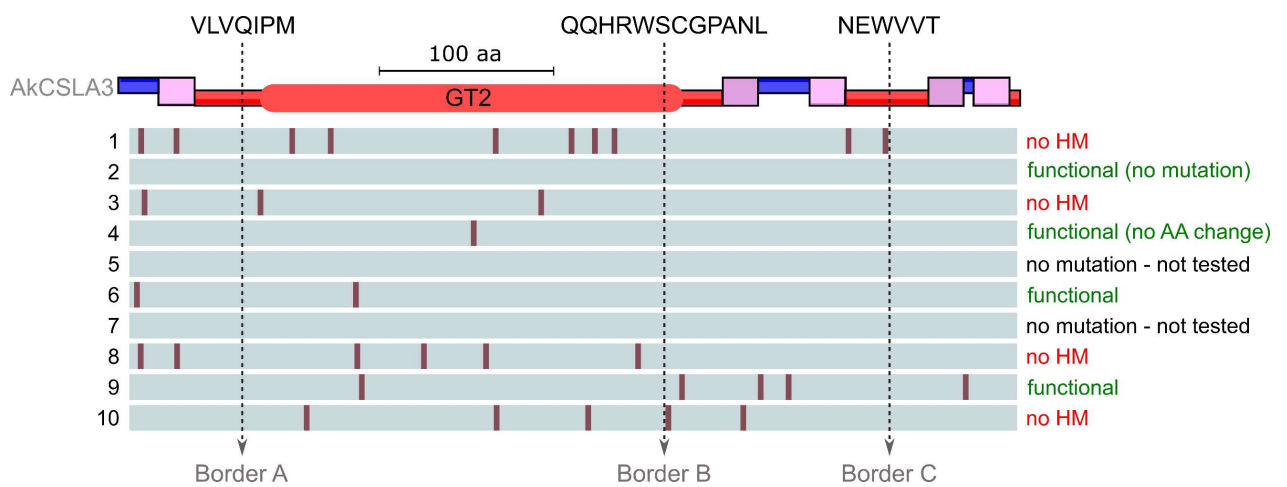


Figure 29 Verified strains of the CSLA library that were used for preliminary experiments. TOP: AkCSLA3 topology visualized using TOPCONS 2.0 with transmembrane domains (pink). Blue regions are outside the membrane, i.e., in the cytosol, and red regions are inside the membrane, i.e. within the Golgi apparatus. The GT2 domain shows the conserved Pfam PF13641 GT2 domain. The dashed lines represent the borders of 100% amino acid identity that were used as borders for the CSLA swap experiments (Chapter 3.1). BOTTOM: Mutations found through whole plasmid sequencing by Plasmidsaurus are denoted as red bars and aligned with the TOPCONS sequence. Notes on the right show the traits of the mutated versions that were later discovered.

Among the 10 sequenced library plasmids, three did not differ from the parental AkCSLA3 sequence. The number of bp changes ranged from 0-10, while the consecutive amino acid (AA) changes ranged from 0-8, as certain base pair mutations did not result in changed amino acids (Figure 30d). Two plasmids experienced frameshifts due to base pair insertions rather than mutations. Figure 29 shows the amino acid changes in relation to their localization in the gene, depicted as red vertical bars. It also specifies the numerical identifiers given to each plasmid and summarizes which plasmids were examined and which resulted in functional and non-functional mutated AkCSLA3 variants.

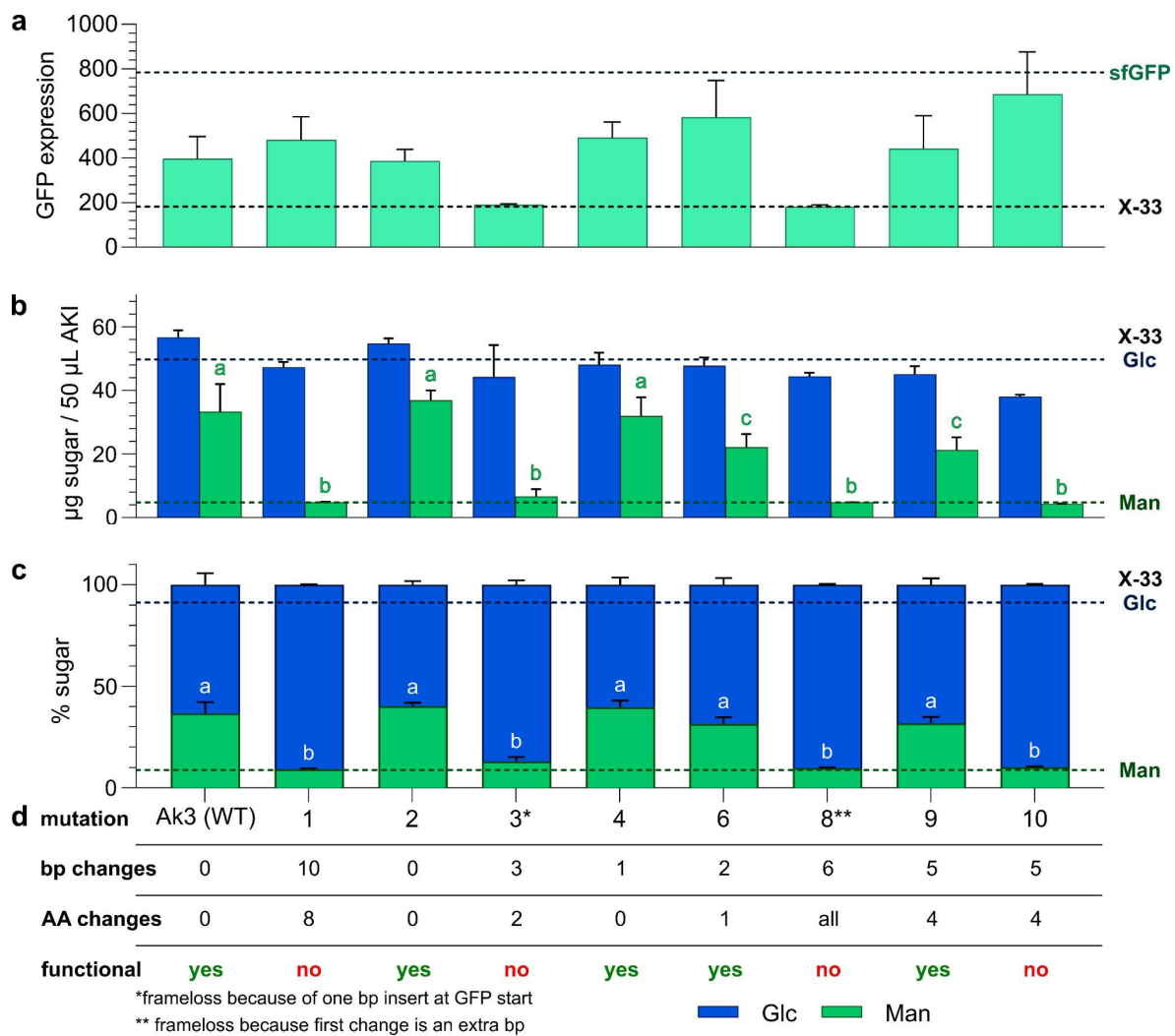


Figure 30 Expression and monosaccharide composition of verified CSLA library strains expressed in wild-type. Colonies grew in 600 µL YPM+ G in sterile 24-well plates for 72 h. Fluorescence (GFP) was measured using a Biotek Synergy H1 (a). GFP signal; levels wild-type X-33 and sfGFP control are shown as horizontal dashed lines. AKI of the samples hydrolysed with 2 M TFA to determine glucose (Glc) and mannose (Man) levels. (b) µg sugar per 50 µL AKI, (c) relative Glc/Man composition (%). Error bars show mean + SD of 4-5 biological replicates. Letters denote significant differences obtained by one-way Anova with Tukey's pairwise ($p=0.05$). X-33 Glc and Man levels are shown as horizontal dashed lines.

To determine the functionality and expression of the validated mutated AkCSLA3 variants, they were initially expressed and analysed in wild-type X-33. Predictably, all mutated variants without frameshift displayed sfGFP levels similar to the parental AkCSLA3, whereas variants with frameshift mutation showed no sfGFP signal (Figure 30a). Monosaccharide analysis following 2 M TFA hydrolysis of the alkaline insoluble residue of each variant revealed that neither the sfGFP signal nor the number of mutations were decisive factors for the amount of glucomannan produced.

Variants 9 and 10 behaved differently, although both had 5 bp changes and 4 AA changes. While variant 9 produced some glucomannan, variant 10 did not show a noticeable increase in mannose level compared to the X-33 wild-type (Figure 30b and c). This indicates that the position of the mutations might have a stronger influence on the CSLA functionality

than the sheer number of mutations. It is noteworthy that except for the functional variant 4, which had a single bp change and no AA change, the mannose content of functional variants 6 and 9 was slightly but significantly reduced compared to the parental AkCSLA3. However, the Glc:Man ratio did not deviate from the parental AkCSLA3, suggesting that while the mutations impacted the overall glucomannan yield, they may not have altered its composition and structure (Figure 30c).

As anticipated, the non-mutated CSLA (number 2) generated mannose amounts similar to the parent, and the frameshift variant 8 did not display an increased mannose level. Despite the frameshift only affecting the GFP portion of the gene, variant 3 also failed to produce significant amounts of mannose. This could imply that two essential amino acids were mutated in this case. Alternatively, the frameshift of sfGFP and the potential resulting misfolding might have caused misfolding of the AkCSLA3 portion, thereby abolishing its function.

3.3.1.2 CBM-Probes Detect New Functional CSLA Variants

After identifying functional and non-functional CSLA variants, they were expressed in the α -PaCBM35-Sc-2A-Tq2 background to verify if the CBM secretion rates would align with the results obtained through monosaccharide analysis (Figure 30).

Ultimately, a functioning pipeline like this could significantly increase the screening throughput to find new functional CSLA variants. By that, synthases with potentially higher efficiency could be detected easily. In addition, the non-functional variants could help identify the key amino acids for CSLA function. Given that different CBMs could be employed for any given HM type, this pipeline could be applied to determine the functionality of various CSL enzyme variants. Eventually, CBM-based probes could assist in discovering new CSL variants with desired functions, thereby fostering the design of new heteromannans, which could potentially find applications in industries such as healthcare and nutrition.

The sfGFP fluorescence of the selected library strains expressed in wild-type (Figure 30a) and the α -PaCBM35 background (Figure 31a) were similar, confirming that CSLA expression was not influenced when co-expressed with α -PaCBM35. Analysis of the pmScarlet-I fluorescence in whole cultures and the supernatant revealed that α -PaCBM35 exhibited the highest secretion rates when co-expressed with non-functional CSLA variants. Conversely, secretion decreased when the mannose content was higher. The lowest secretion rate was detected for the AkCSLA3 parent. The secretion from functional CSLA variants was slightly elevated compared to the AkCSLA2 parent but still considerably lower than the secretion from non-functional CSLA variants (Figure 31b). This is consistent with earlier findings that CBM secretion rates correlate with the mannose amounts in the cells (Figure 16).

Most importantly, the results confirm that the reduction in CBM secretion is indeed driven by the presence of HM and not by the mere presence of more or less CSLA proteins since the expression of non-functional CSLAs did not reduce the secretion of α -PaCBM35.

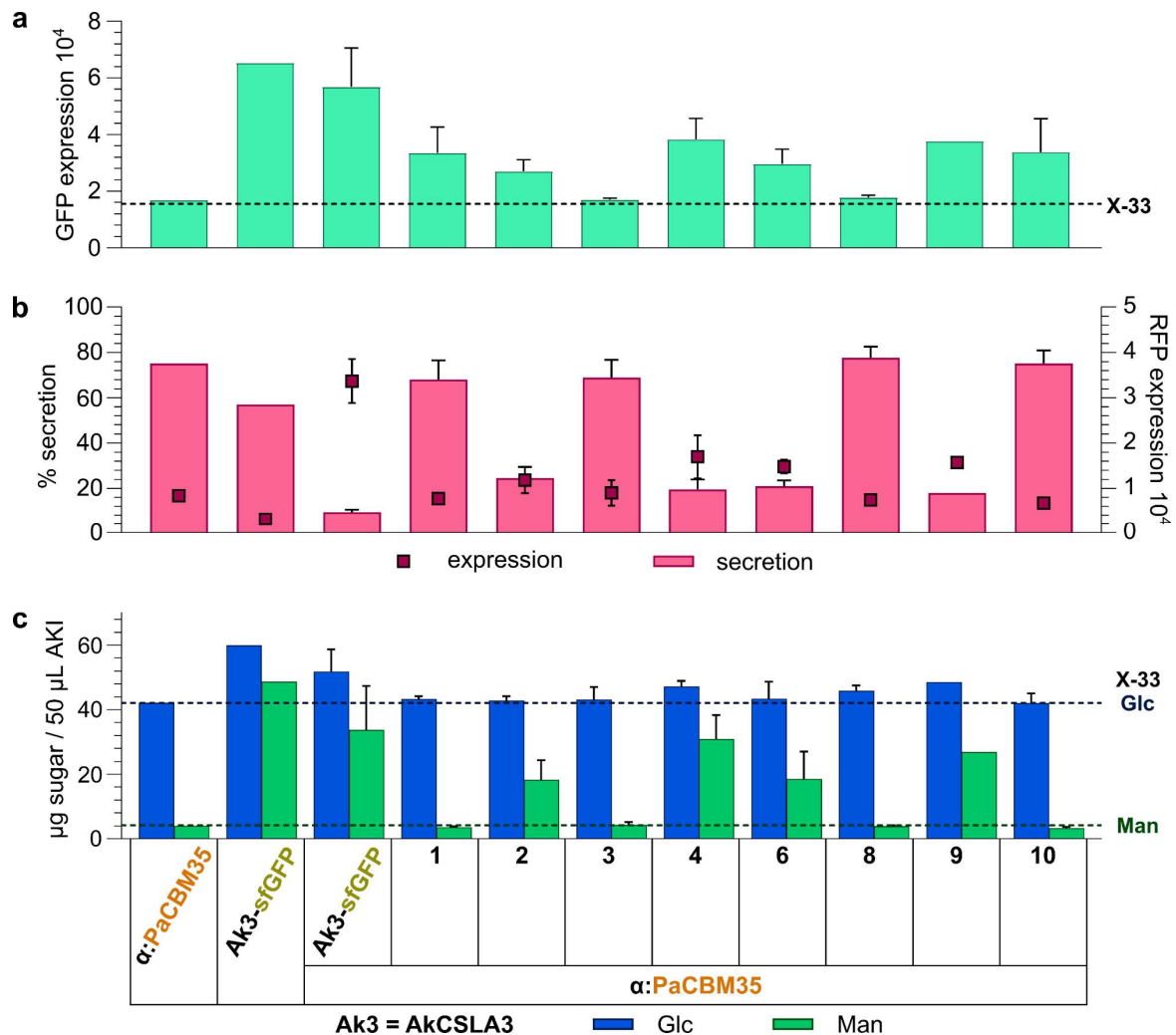


Figure 31 Expression and monosaccharide composition of verified CSLA library strains expressed in α -PaCBM35 tagged with-Sc-2A-Tq2. Colonies grew in 600 μ L YPM+ G (sterile 24-well plates) for 72 h. Fluorescence of whole cultures and supernatant was measured with a Biotek Synergy H1. (a) GFP signal shows AkCSLA3-sfGFP and variant expression. (b) RFP shows α -PaCBM35-Sc-2A-Tq2 expression (squares) and secretion (bars). Sample AKI was hydrolysed with 2 M TFA to determine glucose (Glc) and mannose (Man) levels. (c) μ g sugar per 50 μ L AKI. Error bars show mean + SD of 4-5 biological replicates, except for single construct colonies (1 biological replicate).

Considering that some mutations in the verified CSLA library strains were located near or within transmembrane regions of the proteins, it was hypothesized that these could influence the CSLA variant localization. To examine this, cells were imaged with a confocal Leica SP 5 microscope. Before imaging, cells were counterstained with PI to detect dead cells and CF to visualize native yeast cell wall polysaccharides. As anticipated, the variants without AA changes localized in small punctate structures akin to the parental AkCSLA3. Similarly, functional variants 6 and 9 also displayed that type of localization.

Conversely, the non-functional variants 1 and 10 displayed larger aggregates compared to the punctate structures observed in functional variants (Figure 32). This could point to a mislocalization of non-functional variants. Variant 1 harbours mutations in the first and third predicted transmembrane domain (Figure 29) and might therefore be compromised in its ability to anchor to the Golgi membrane, thus causing a mislocalization in the cell. However, variant 10 lacks mutations in the predicted transmembrane regions (Figure 29), suggesting that protein aggregation may also be due to improper protein folding.

The possibility of improper protein folding could arise from alterations in the charges of the mutated amino acids compared to those in the parental AkCSLA3. The focus to investigate that possibility was primarily directed towards variants 9 and 10, since variant 9 was functional while variant 10 was non-functional, despite both mutations having 4 amino acid changes. Upon examination, it was found that in variant 9, all changed amino acids exhibited a different polarity than the parent, while 3 out of 4 mutations in variant 10 did not change the polarity. This strongly suggests a lack of correlation between the polarity of the new amino acids in relation to parental counterparts in the context of glucomannan production (Table 4).

Table 4 Polarity of mutated CSLA library amino acids(n = non-polar; p = polar; a = acidic; b = basic)

VARIANT 9	functional	AA change	D139G	A324T	V369A	A488T
		polarity	a → n	n → p	n → n	n → p
VARIANT 10	no HM	AA change	P108Q	V270M	A317P	F359T
		polarity	n → p	n → n	n → n	n → n

Given that alterations in amino acid polarity did not appear to influence the functionality of the library CSLA, it was investigated if the mutations could cause protein misfolding, which would result in function loss. However, *de novo* protein folding calculations are typically time-consuming and require substantial computational resources. To gain a preliminary understanding of the structural differences between the CSLA variants and their parent AkCSLA3, AlphaFold 2.0 was used to predict protein folding. This tool utilizes neural network computing and complex algorithms to accurately predict protein structures by matching a given amino acid sequence to known 3D protein structures (Jumper et al., 2021; Akdel et al., 2022).

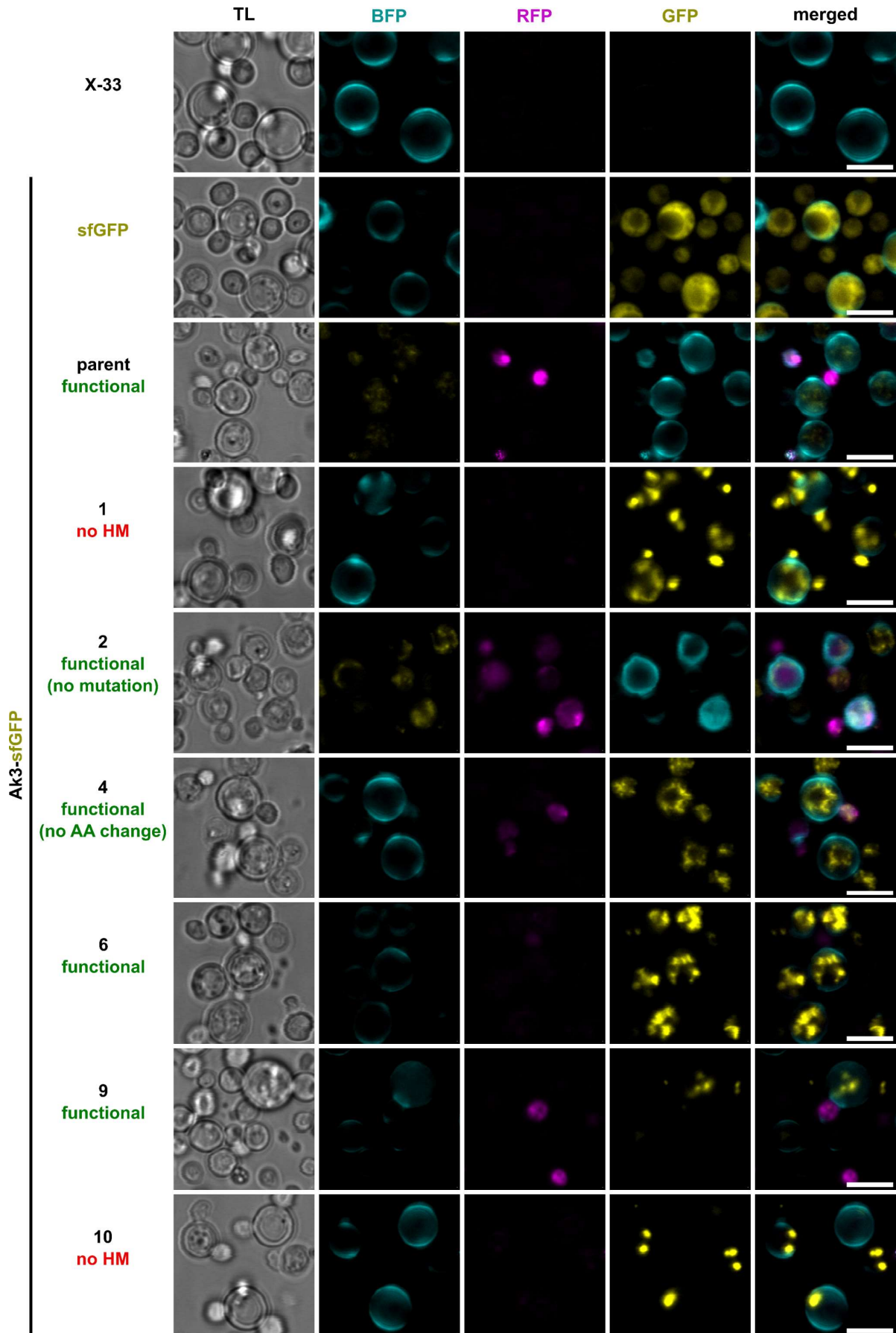


Figure 32 Localization of mutated AkCSLA3 (GFP) versions compared to parental AkCSLA3. Cells were grown in 600 μ L YPM+ G in sterile 24-well plates for 72 h and imaged with a Leica SP5 microscope. Cells were stained with 10 μ L of 0.01% PI (to visualize dead cells (RFP)) and 0.01% CF (to visualize native yeast polysaccharides (BFP)). Scale bar = 5 μ m.

Sequence superimposition revealed only minor structural differences at the sites of mutated amino acids. Even variant 1, which harbours 8 amino acid changes, was highly similar to the parent in terms of protein folding (Figure 33). However, the AlphaFold 2.0 mode of operation could lead to an underestimation of the effects of small amino acid changes. To verify if the lack of structural changes holds true for predictions of the transmembrane domains, the sequences were submitted to topology prediction with TOPCONS 2.0 (results not shown). There were no discernible differences between the mutated library proteins and the parental AkCSLA3, aligning with the AlphaFold 2.0 results (Figure 33). In summary, the computational analysis results did not reveal why some amino acid mutations seem tolerated, whereas others completely abolish the protein function. Therefore, future endeavours should focus on a more detailed exploration of the impacts of specific amino acid alterations on protein function and structure.

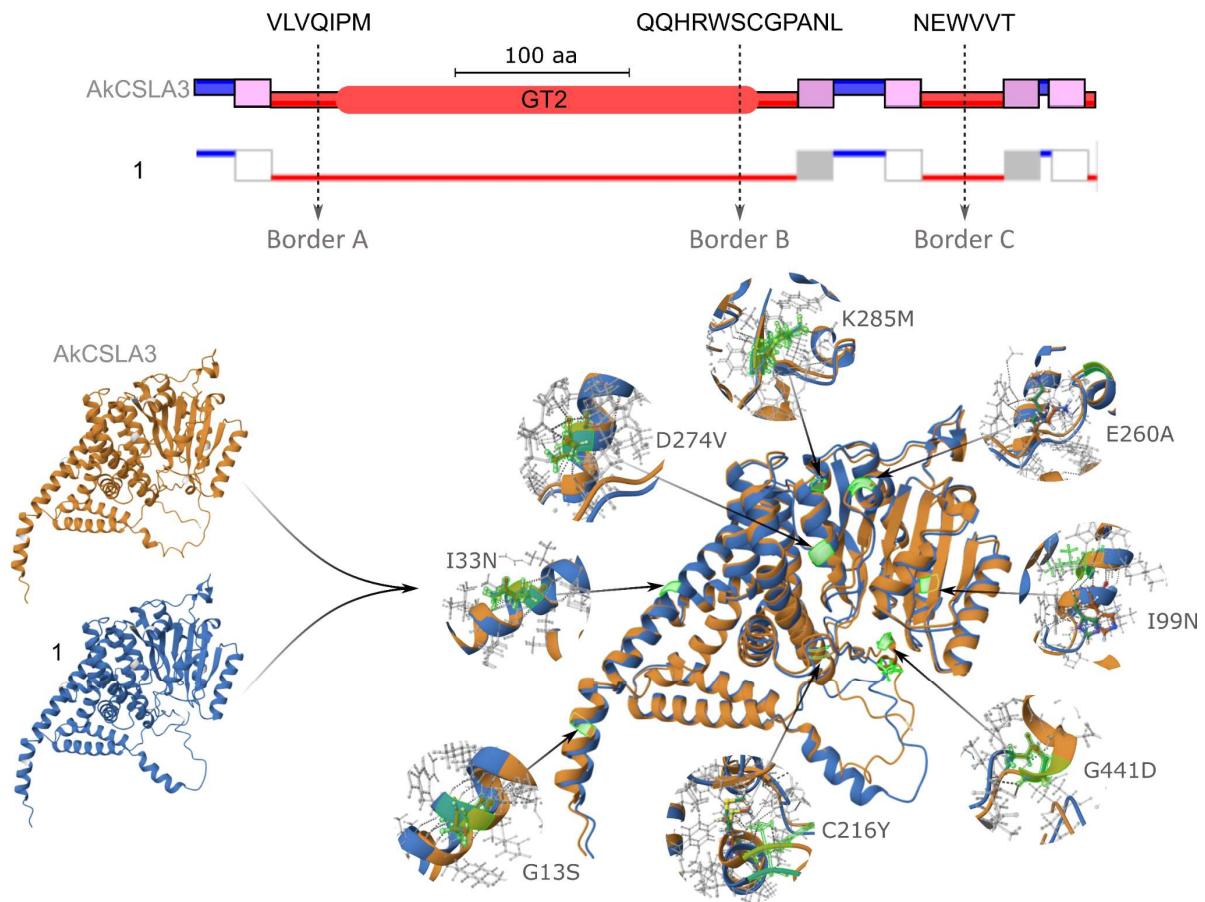


Figure 33 TOPCONS 2.0 and AlphaFold 2.0 predictions of CSLA library variant 1 aligned with AkCSLA3 parent. TOP: TOPCONS 2.0 topology prediction of transmembrane domains (pink / grey / white). Blue regions are outside the membrane, i.e., in the cytosol, and red regions are inside the membrane, i.e. within the Golgi apparatus. The GT2 domain shows the conserved Pfam PF13641 GT2 domain. The dashed lines represent the borders of 100% amino acid identity that were used as borders for the CSLA swap experiments (Chapter 3.1). BOTTOM: AlphaFold 2.0 structure prediction of AkCSLA3 (orange) and CSLA library variant 1 (blue). Green highlighted regions in overlay of both structures show mutated amino acids. Circles show effects of mutated amino acids on 3D – protein structure.

When analysing complete cultures consisting of one specific yeast strain, one essential piece of information that is overlooked is the level of uniformity within each population with respect to gene expression and viability which can be analysed in more detail using single-cell cytometry (Hohenblum et al., 2003). Therefore, proof of concept experiments for using cytometry and single-cell analysis to find new CSLA variants were carried out. These experiments also aimed to gather more information about the variability within one population expressing the same gene.

Cultures expressing the verified CSLA library variants in *Pichia* wild-type were grown in YPM + G according to the usual growth protocol, stained with PI to detect dead cells, and subjected to cytometry using the Accuri 6. Considering both the sfGFP and PI signal, it could be distinguished between the following three categories of AkCSLA3-sfGFP and CSLA library expression: (a) sfGFP signal + HM production, (b) sfGFP signal + no HM production, (c) no sfGFP signal + no HM production (Figure 34c). Like that, there were clear distribution differences between unstained wild-type, sfGFP, and AkCSLA3-sfGFP populations.

For X-33, almost 100% of cells were viable, and had no sfGFP signal. The viability of cells expressing sfGFP was also nearly 100%, as was the number of cells with sfGFP signal, underlining the uniformity of that population. In contrast, the toxicity of AkCSLA3-sfGFP was evidenced by the fact that over 30% of cells exhibiting GFP signal were also PI stained (Figure 34a and b). More importantly, non-functional AkCSLA3-sfGFP variants mirrored the sfGFP alone signal distribution and cell viability (Figure 34d), while cell viability and sfGFP signal distribution of functional AkCSLA3-sfGFP variants resembled the parental AkCSLA3-sfGFP (Figure 34e). Non-functional and non-fluorescent variants mimicked the wild-type, except for the frameshift variant 3, which displayed a high number of dead cells (Figure 34b and f).

The conducted experiment shows the potential of cytometry, which should, later on, be paired with cell sorting as a valuable tool for gene-library screening to detect new functional CSLA variants once the variants will be co-expressed with the CBM probes. In addition, cytometry + cell sorting could be very useful to quickly assess cell viability, thereby providing valuable insights into the potential toxicity of certain CSLAs.

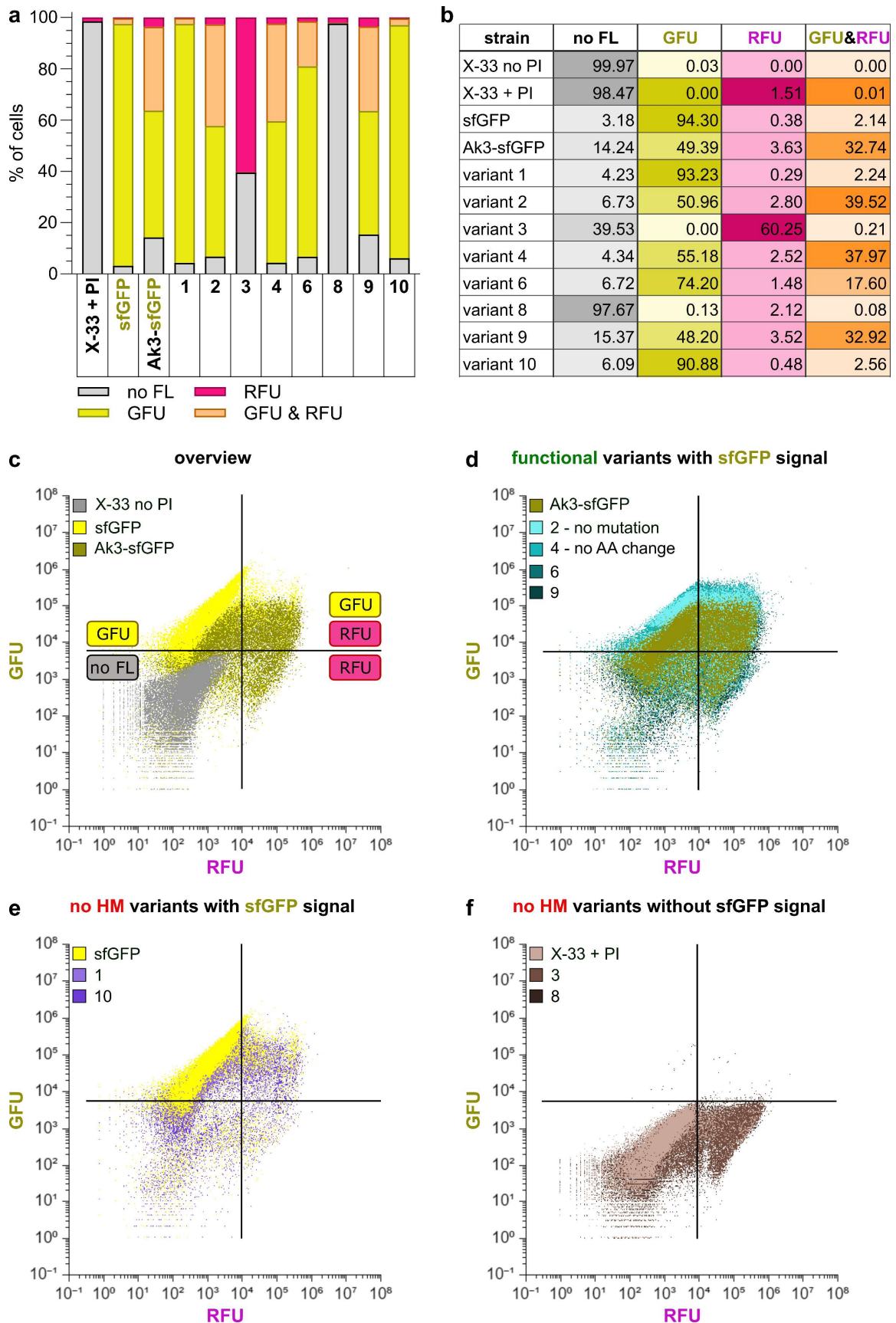


Figure 34 Single cell cytometry of verified CSLA library strains. (a) relation between cells with no (no FL), only GFP, only RFP or RFP and GFP fluorescence in%. (b) Table of percentages shown in (a). (c) Green fluorescence (GFU) and PI signal (RFU) of unstained X-33 non-fluorescent control (X-33 no PI), sfGFP and AkCSLA3-sfGFP. In the graph cell types are clustered dependent on emission of sfGFP and/or PI signal as follows: No FL = bottom left; sfGFP = top left; sfGFP + PI = top right; PI = bottom right. (d) sfGFP emitting functional variants compared to AkCSLA3-sfGFP parent. (e) sfGFP emitting variants that do not produce HM, compared to sfGFP. (f) variants neither sfGFP nor HM production compared to PI-stained wild-type (X-33 + PI). For each genotype, 100,000 cells were analysed and (c)-(f) were then plotted using the floreada analysis tool (<https://floreada.io/analysis>).

3.3.2 Application in Plants: CBMs Help Elucidate Native HM Distribution

To assess the ability of the newly developed CBM probes to uncover native heteromannan distribution and dynamics in living plants, the probes were expressed in the model plants *Nicotiana benthamiana* and *Arabidopsis thaliana*. Subsequently, the CBM expression patterns were imaged and analysed via confocal microscopy.

N. benthamiana is typically transformed transiently and therefore allows for rapid assessment of the CBM probe functionality. In contrast, the extensive collection of well-characterized *Arabidopsis thaliana* mutants would allow a more profound investigation of the CBM probe functionality. By expressing the CBM probes in heteromannan-deficient mutants and the Col-0 wild-type, it could be evaluated how well the new CBM probes can detect heteromannan quantity and distribution differences.

3.3.2.1 Transient Expression of CBM-Probes in *N. benthamiana*

The performance of the novel CBM-based heteromannan detection tool was first evaluated by transiently expressing TpoICBM16-1 and PaCBM35, both tagged with pmScarlet-I-2A-pSmTq2 and equipped either with the α or the AtEXP10 signal peptide *Nicotiana benthamiana* under the use of the constitutive p35S promoter. Leaf disc plasmolysis allows to contract cellular vacuoles, thus enlarging the remaining cellular structures (Figure 35) and enabling more precise investigation of potential localization differences between genotypes.

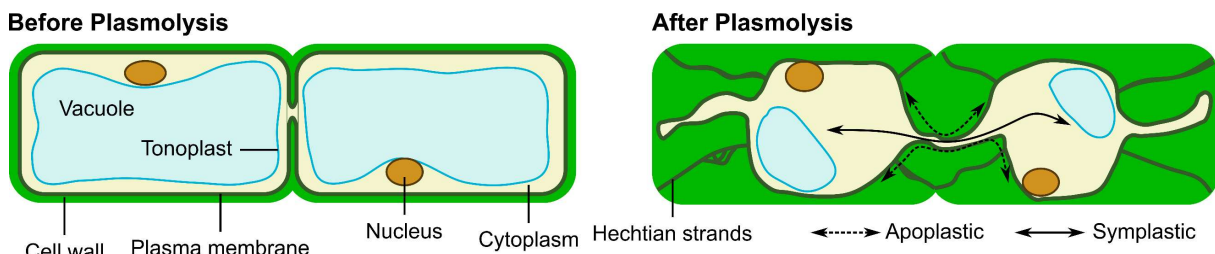


Figure 35 Model of how plasmolysis affects plant cell organization. The model was created based on a publication by (Giraldo and Valent, 2013) and shows how plasmolysis reduces the size of the vacuole unveiling Hechtian strands that connect the plasma membrane to the cell wall.

Figure 36 shows the images obtained for each tested genotype before and after plasmolysis. Note that the acquisition gain for the pmScarlet-I channel was reduced for plasmolysed α -PaCBM35 to ensure proper signal localization, given that the initial signal intensity would have saturated the image.

The fluorescent proteins pmScarlet-I and pSmTq2, components of the pmScarlet-I-2A-pSmTq2 tag, were expressed independently or in conjunction with the α signal peptide. In non-plasmolysed cells, it was challenging to discern localization differences influenced by the presence or absence of the α signal peptide. Following plasmolysis, it became apparent that the standalone fluorescent proteins likely localized predominantly at the plasma membrane or

within the cytosol. On the other hand, in addition to localization at the plasma membrane, the fluorescent proteins equipped with the α signal peptide appeared to localize in the vacuole or tonoplast and possibly the apoplast. The localization of α -equipped fluorescent proteins indicates that the α signal peptide may facilitate their secretion out of the cells.

The observed similarities or discrepancies in localization among the tested CBMs did not correlate with the CBMs themselves but seemed to coincide with the presence of different signal peptides. Like that, both α -TpolCBM16-1 and α -PaCBM35 exhibited pmScarlet-I signals at the cell peripheries in a punctate pattern, with the whole cells showing a uniform signal distribution with the exception of some small darker spots, possibly corresponding to mitochondria or peroxisome. The pSmTq2 signal mirrored that of its standalone expression, indicating successful cleaving from the CBM constructs. After plasmolysis with 1 M NaOH, both α -TpolCBM16-1 and α -PaCBM35 localized in punctate at the plasma membrane and showed uniform labelling at the vacuole periphery – possibly the tonoplast - and within the vacuole. In both cases, no obvious mannan labelling was observed in the cell wall.

On the other hand, when equipped with the ATEXP10 signal peptide, both TpolCBM16-1 and PaCBM35 exhibited a uniform localization pattern at the cell edges and in certain Hechtian strands or regions of the cytoskeleton. Full cell labelling was not observed. The pSmTq2 signal, once again, mirrored the location of pSmTq2 expressed independently, indicating successful cleavage. Upon plasmolysis with 1 M NaOH, AtEXP10-TpolCBM16-1 presented punctate localizations akin to those in α -pmScarlet-I (Figure 36). However, the higher number of punctate in the CBM sample suggested potential plasma membrane mannan labelling. Similar patterning was also observed for AtEXP10-PaCBM35, both regarding the pmScarlet-I and the pSmTq2 signal.

In conclusion, the results demonstrate that TpolCBM16-1 and PaCBM35 tagged with pmScarlet-I-2A-pSmTq2 can be successfully expressed transiently in *N. benthamiana* leaves and are promising candidates for stable expression in *Arabidopsis*. The results also highlight the influence of signal peptides on CBM localization. While it could not unequivocally be stated that the CBMs label the mannan in the plasma membrane or cell walls of *N. benthamiana* epidermal leaf cells, the punctate structures at the plasma membranes could indicate mannan labelling. This underscores the importance of stable CBM expression in *Arabidopsis*, as unlike *N. benthamiana*, *Arabidopsis* offers cell wall mutants that can facilitate the study of CBM interaction with, for instance, mannan deficient cell walls.

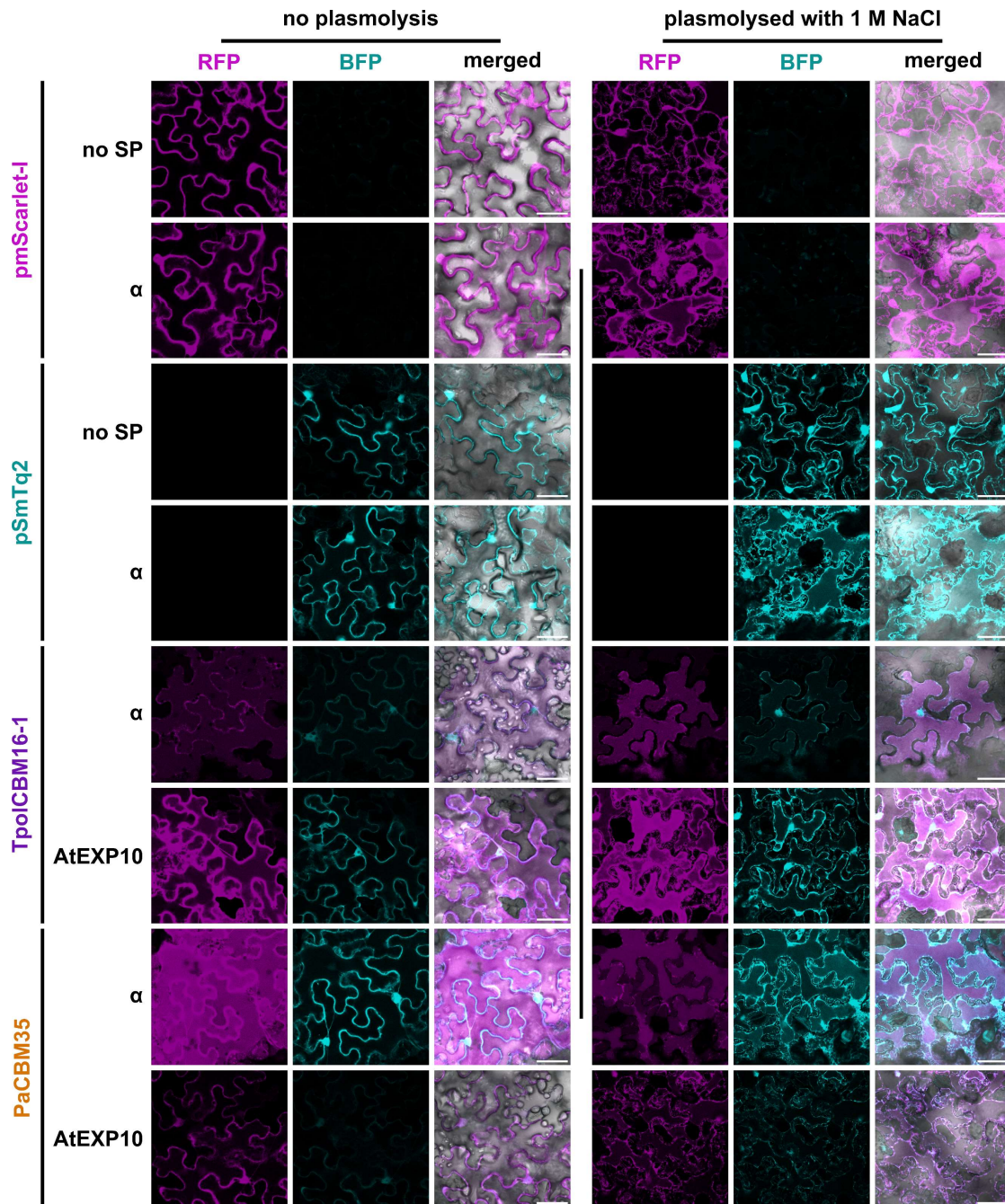


Figure 36 Transient expression of TpolCBM16-1 and PaCBM35 with different signal peptides in *N. benthamiana*. Leaves of 6 weeks old plants were infiltrated with *A. tumefaciens* containing CBM constructs tagged with Sc-2A-Tq2 and equipped with either α or AtEXP10 as a signal peptide. Fluorescent controls were expressed with and without α . Leaf discs were imaged 3 days post infiltration via confocal microscopy with a Zeiss LSM900. For plasmolysis, leaf discs were immersed in 1 M NaCl salt solution and imaged immediately. The negative control (infiltration solution only) was treated the same way as all samples and no fluorescence was detected (image not shown). Note that for plasmolysed α -PaCBM35 the pmScarlet-I signal was very high, which is why for this image only, the red channel gain was reduced. Scale bar = 50 μ m.

3.3.2.2 Non-Invasive Detection of Heteromannans in *Arabidopsis*

As proof of concept, for the ability of the novel CBM probes to detect the native distribution of heteromannans *in planta*, α -TpoICBM16-1-pmScarlet-I-2A-pSmTq2 was selected as a candidate, as it yielded promising results in both yeast and *N. benthamiana*. Alongside the expression and localization controls pmScarlet-I and α -pmScarlet-I, it was stably transformed into *Arabidopsis* using the constitutive p35S promoter. The constructs were first introduced into a MoClo vector, which contained an RFP selection marker gene that would allow the detection of positive T1 transformants based on the red fluorescence of the seeds. After stable agrobacterium-mediated transformation in *Arabidopsis* Col-0 and a *cs/a2,3,9* mutant background – known for its significantly reduced mannan production (Goubet et al., 2009; Liepman and Cavalier, 2012) - seeds were selected based on red fluorescence and grown on vertical agar plates for initial screening (Figure 37).

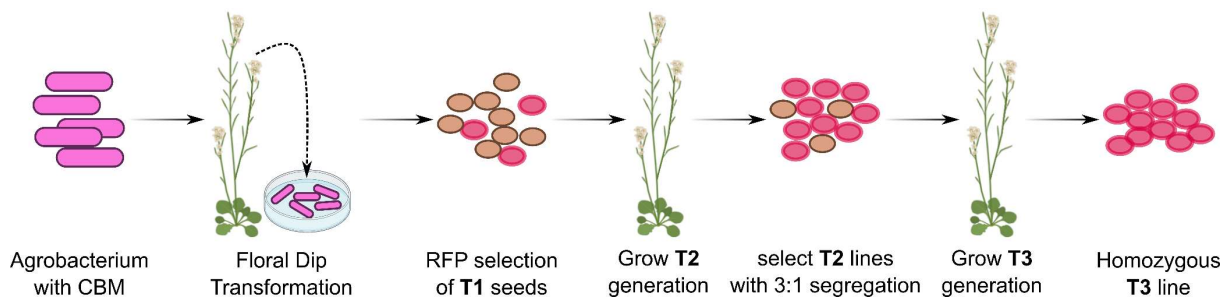


Figure 37 A. *tumefaciens*-mediated stable *Arabidopsis* transformation and propagation over three generations. From left to right: Transformation of *A. tumefaciens* with a gene of interest (for example a CBM), dipping of *Arabidopsis* inflorescences in a transformation solution containing *A. tumefaciens* with the gene of interest. After harvesting, seeds of the T1 generation are selected based on their RFP signal. From these seeds, the T2 lines are generated. Only T2 lines that show Mendelian segregation ratios of 3:1 transformed versus untransformed seeds are selected for growth of the T3 generation that is able to yield homozygous plants.

This first generation after transformation, termed T1, already yielded promising results in terms of fluorescence (results not shown). Using confocal microscopy, seedlings exhibiting fluorescence were selected and cultivated to maturity on soil. The resulting T2 lines were examined for Mendelian segregation; only the lines with an approximate 3:1 ratio (Appendix 5) of red and non-red fluorescent seeds were selected to generate the third generation, i.e., T3 lines. Of the selected lines, 30 fluorescent seeds were grown on vertical agar plates for 5 days under long-day conditions prior to root imaging. Between 8 to 12 seedlings were imaged using a confocal Zeiss LSM780 microscope. The images presented in Figure 38 reflect the typical observations among the screened seedlings.

The initial T1 line screening revealed that the pSmTq2 signal from α -TpoICBM16-1-pmScarlet-I-2A-pSmTq2 was barely discernible, despite a strong pmScarlet-I signal being observed. To ensure that an imaging issue did not cause this pSmTq2 signal absence, an additional control Cyan Fluorescent Protein (CFP) expressing *Arabidopsis* control was imaged

(Nelson et al., 2007). Moreover, *Pichia* strains expressing the same constructs as the T3 *Arabidopsis* seedlings were cultivated and imaged with particular attention to the pSmTq2 signal. The yeast control of α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 and the CFP control in *Arabidopsis* exhibited robust pSmTq2 signal (Figure 38a). Nonetheless, despite a strong pmScarlet-I signal, little to no pSmTq2 signal could be detected in the α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 expressing *Arabidopsis* lines, both in the Col-0 and in the *cs/a2,3,9* backgrounds (merged images in Figure 38b and c). This suggests an issue regarding pSmTq2 expression or stability under the conditions employed in this experiment. The blue channel was, therefore, not displayed separately in Figure 38b and c.

In the root elongation zone of 5-day-old Col-0 background seedlings, all genotypes showed a high pmScarlet-I signal compared to the negative control Col-0. When expressed in Col-0, pmScarlet-I was found ubiquitously in the epidermal cells, while no signal was detected in the stele of the root. On the other hand, α -pmScarlet-I in Col-0 presented a coarser labelling pattern that in some cells appeared as a smear towards the edges of certain epidermal cells, even though many cells were labelled entirely. In addition to these patterns, some nuclei were labelled, and faint labelling of the vasculature could also be observed (Figure 38c).

The localization of α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 differed notably from the localization of the pmScarlet-I controls. This genotype exhibited a filamentous punctate structure adjacent to epidermal cell walls. In addition, pericycle cells in the vasculature were strongly and completely labelled, except for the nuclei that remained unlabelled. When expressed in the *cs/a2,3,9* background, the pmScarlet-I signal was considerably weaker than in Col-0. However, the labelling patterns mirrored the patterns observed in the Col-0 background. Contrary to the localization observed in Col-0, α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 was not found in the vasculature.

These results are consistent with what might be expected due to the lack of heteromannan since many heteromannans are localized in xylem cells (Goubet et al., 2009), and a heteromannan deficient mutant such as *cs/a2,3,9* would likely exhibit a reduction or absence of heteromannan in the xylem. Regarding signal intensity, its reduction in *cs/a2,3,9* compared to Col-0 occurs both for the CBM and the fluorescent controls, suggesting that it may not be directly associated with the lack of heteromannan. To determine whether the reduced signal intensity is due to lower expression levels in the *cs/a2,3,9* mutant, future experiments should include a quantitative analysis of expression levels using techniques such as quantitative PCR (qPCR). This would provide a more detailed understanding of the potential correlation between gene expression and signal intensity in these plant lines.

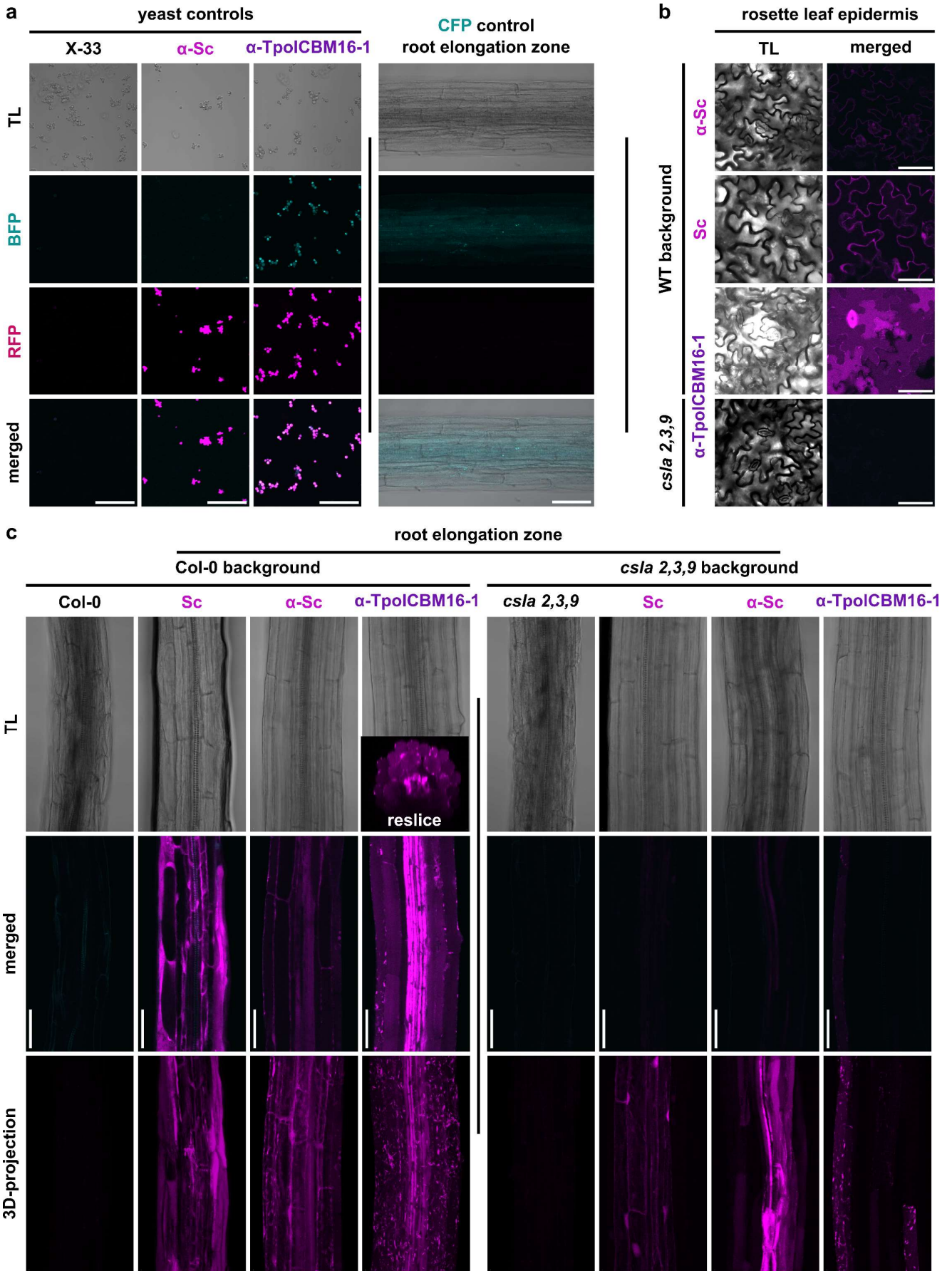
In leaves of 15-days-old seedlings, the observations made in the roots could be confirmed, underscoring the efficacy and reliability of the experimental design as well as the

α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 construct (Figure 38b). In fact, except for the lack of pSmTq2 signal, the pattern observed was highly similar to the pattern observed when the same construct was transiently expressed in tobacco leaves (Figure 36). Again, the absence of α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 signal in the *cs/a2,3,9* mutant could be due to the heteromannan deficiency or low expression. This stresses the importance of quantifying gene transcripts in the future.

Despite the absence of pSmTq2 signal, the results demonstrate the potential of recombinant CBMs for *in vivo* visualization of complex carbohydrates like heteromannans *in planta*. The selected construct α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 could successfully be expressed and localized differently depending on the background it was expressed in, suggesting a heteromannan-dependent localization pattern. In addition, the punctate localization of TpolCBM16-1 in the epidermal cell walls of Col-0 indicates that it might be able to access the mannans despite their embedding in the pectin matrix. This presents a significant advantage compared to more traditional techniques such as immunolabelling, where the antibody size restricts access to cell wall mannans. CBMs could therefore be a powerful tool for non-destructive *in vivo* labelling of plant carbohydrates, with potential applications in scientific research and industrial settings.

The following legend refers to the figure on the next page:

Figure 38 Stable expression of α -TpolCBM16-1-Sc-2A-Tq2, α -pmScarlet-I and pmScarlet-I in *Arabidopsis* Col-0 and *cs/a2,3,9*. Selected T3 lines showed Mendelian segregation ratios of ~ 3:1. All images were acquired with the same settings using a confocal Zeiss LSM780. 3D-projection signal intensities were adjusted in post-processing for best visualization. Negative controls were adjusted the same way as the darkest fluorescent signal. (a) Yeast fluorescent controls expressing α -TpolCBM16-1-Sc-2A-Tq2 and α -pmScarlet-I and *Arabidopsis* cyan fluorescence (CFP) control. (b) Epidermal cells of rosette leaves from 15 days old T3 seedlings, transferred to soil after 7 days of growth on vertical agar plates. (c): 5 days old seedling root elongation zones. Note: TL panel of α -TpolCBM16-1-Sc-2A-Tq2 also shows resliced image of the red channel. Abbreviations: TL = transmitted light, BFP = blue fluorescent protein, RFP = red fluorescent protein, merged = merging of fluorescent channels (except for CFP control where all channels were merged); 3D-projection = layering of all planes to form a 3-dimensional image in ImageJ. Scale bar = 50 μ m.



4 DISCUSSION

Synthetic biology greatly profits from employing microorganisms such as bacteria and yeast. Living materials can be engineered to exert novel functions, and polysaccharides like cellulose can be functionalized by engineering, such as cellulose-enzyme-networks (Gilbert and Ellis, 2018; Gilbert et al., 2021). They allow for rapid design-build-test-learn cycles due to their short generation times, which extend only over days, compared to plants' much longer generation times. This time efficiency makes microorganisms an invaluable resource for investigating and innovating polysaccharide production and functionalization.

The objective of the present study was threefold. First, single-domain swaps were employed in an attempt to tailor HM production in yeast. Second, new probes were developed to label HM *in vivo* non-invasively. These new probes were then used to demonstrate how they could help elucidate native HM distribution *in planta* and to discover new CSLA variants by using the new CBM-based probes as indicators of the functionality of different CSLA variants.

4.1 CSLA Domain Swaps for HM Biosynthesis Tailoring

To manipulate HM production, domains from the mannan-producing AtCSLA2 from the dicot *Arabidopsis* were swapped with AkCSLA3 domains from the monocot *A. konjac*. AtCSLA2 plays a role in *Arabidopsis* seed mucilage, with potential implications on seed mucilage characteristics, which may influence the salt tolerance of *Arabidopsis* seeds (Yu et al., 2014; Voiniciuc et al., 2015; Yang et al., 2021). Conversely, AkCSLA3 piques interest for its function in glucomannan biosynthesis during *A. konjac* development as glucomannan is currently used to manage lifestyle diseases such as diabetes (Gille et al., 2011; Behera and Ray, 2016). Modulating HM biosynthesis may serve dual purposes, potentially enhancing crop resilience to stress and augmenting the health benefits of hemicelluloses.

A new streamlined growth protocol (Robert et al., 2021) improved yeast growth times and increased HM yields. The Glc components of glucomannan produced by AkCSLA3 could be separated from the native yeast Glc by linkage analysis (Figure 7) and partial AKI digestion with β -mannanase. These techniques removed background yeast β -1,3-linked glucans from the β -1,4-linked heteromannans (Figure 8) (Aguilar-Uscanga, 2003; Voiniciuc et al., 2019).

Glycosidic linkage analysis in the domain-swapped enzymes and their parents AtCSLA2 and AkCSLA3 revealed an increase in 4-linked Man production in all chimeras relative to the wild-type, albeit marginal for five constructs (Figure 7b and c, Figure 8b). Swapping the N- or the C-terminal CSLA domain often led to a significant drop in HM production, except for 3332, which sustained HM quantities akin to its parent AkCSLA3.

Previous research highlighted the crucial role of transmembrane pores for the translocation and structure of xyloglucan (Davis et al., 2010), mixed linkage glucan (Jobling, 2015; Dimitroff et al., 2016), and cellulose synthases (Purushotham et al., 2020). Considering the high divergence at the N- and C-termini of AtCSLA2 and AkCSLA3 (Figure 6), these terminal regions may ensure proper protein structure and Golgi membrane anchoring. Altering these domains may compromise protein architecture, reducing or abolishing their function.

Surprisingly, chimeras 2322 and 2232 produced more mannan than their parent AtCSLA2 without incorporating significantly more Glc (Figure 7). This suggests that in a heterologous host, AtCSLA2 requires AtMSR1 to produce glucomannan (Voiniciuc et al., 2019), which was further supported by increased Glc incorporation and HM yield when 2322 and 2232 were co-expressed with AtMSR1 (Robert et al., 2021). These observations led to the hypothesis that the putative protein O-fucosyltransferase MSR1 might glycosylate and/or interact directly with the GT2 domain of CSLA enzymes.

Using cost-efficient materials and a single heterologous CSLA-based enzyme, it was possible to produce (gluco)mannan at mg / 2 mL levels. However, due to their insolubility (Berglund et al., 2020), assessing the molecular weight of intact heterologous HM was not feasible. Future research could consider co-expressing CSLAs with enzymes responsible for HM galactosylation and/or acetylation, which is hypothesized to help branch mannans (Scheller and Ulvskov, 2010) and, by that, increase HM solubility.

Prior studies on CSL chimeras disclosed that swapped-domain cereal CSLF6 proteins expressed transiently in *N. benthamiana* leaves generated variable β -1,3-1,4-linked-glucan amounts but consistent DP3:DP4 oligosaccharide ratios, generally matching the ratios observed in both parents or intermediate ratios (Jobling, 2015; Dimitroff et al., 2016). By using a similar approach and comparing the Glc:Man ratios of chimeric and parental enzymes, it was possible to distinguish between glucomannan, produced by AkCSLA3 and 3332, and pure mannan, produced by AtCSLA2, 2322, and 2232.

Contrary to the present study, where five CSLA swaps produced significantly less HM than their parents (Figure 7), CSLF6 chimeras from prior studies were all functional. This might stem from the tighter evolutionary kinship among cereal CSLF6s compared to the selected CSLAs. Notably, AtCSLA3 and not AtCSLA2 is the ortholog of AkCSLA3 (Gille et al., 2011). Yet, no specific amino acids related to glucomannan synthase activity were identified.

Extended overexpression of AkCSLA3-driven glucomannan production reduced growth compared to the wild-type, aligning with the impossibility of constitutive AkCSLA3 expression, despite available constitutive promoters (Prielhofer et al., 2017). AtCSLA2 and mannan-producing chimeras did not display such growth reduction (Figure 9). 24 h of AkCSLA3 overexpression sufficed to lower cell density but not to induce morphological

changes. However, after 72 h, most AkCSLA3-expressing cells lost viability, indicated by increased TB uptake (Figure 9b). Additional tests with PI (Figure 10) and Congo red (Robert et al., 2021) corroborated the viability loss in AkCSLA3 colonies.

Replacing the C-terminal domain of AkCSLA3 with the one of AtCSLA2 (3332) partially restored cell viability (Figure 9, Figure 10), with minimal impact on Glc:Man ratio (Figure 8). Also, TmCSLC4 expression, a β -1,4-glucan synthase putatively producing xyloglucan in yeast (XyG) (Cocuron et al., 2007; Schultink et al., 2014), did not affect cell viability (Robert et al., 2021). Thus, the viability drop upon prolonged AkCSLA3 overexpression could be glucomannan specific, possibly due to GDP-Glc precursor competition and/or intracellular glucomannan accumulation. In the future, this could be tested by expressing high-yielding CSLA variants to elucidate if this would intensify the defects. As a first step, an AkCSLA3 variant library was created. Combined with new CBM-based probes, new, higher-yielding CSLA variants could be discovered.

4.2 Design of a CBM-Probe for *in vivo*

Investigating CSLAs in yeast presents a notable challenge: only a select few maintain functionality when tagged with a fluorescent protein, thus preventing *in vivo* localization analysis of most CSLAs (Voiniciuc et al., 2019; Robert et al., 2021). In addition, the position of the fluorescent tag may strongly influence the productivity of a given CSLA (Figure 27).

Heteromannan labelling requires a specialized tool that can simultaneously fluoresce and anchor itself to HM to allow the visualization of its precise localization. Existing techniques employing these dual capabilities are unspecific stains or antibodies which require tissue disruption (Voiniciuc et al., 2018b). In this work, the shortcomings of the current labelling techniques were addressed by developing a new, user-friendly, non-invasive, specific probe to label HM. They are based on Carbohydrate-Binding Modules (CBMs) recombined with fluorescent proteins similar those previously developed for xylan labelling in *Arabidopsis* (Zhang, 2014). The new probes will be efficient instruments for detecting functional enzymes when co-expressed with native or chimeric CSL enzymes in heterologous systems and to elucidate the native distributions of HM *in planta*. Expressing CSLA libraries in CBM-probe-expressing yeast could help rapidly identify new functional CSLA variants. This approach could uncover key amino acids essential for CSLA function and potentially lead to minimally altered CSLAs with higher yields in a high throughput manner.

4.2.1 CBM-Probe Secretion from Yeast Depends on HM Quantities

In this work, probes for *in vivo* heteromannan binding were developed using mannan-binding CBMs that were recombined with C-terminal fluorescent proteins and N-terminal signal peptides for extracellular secretion. This ensured that the recombinant CBMs secreted from the yeast cells unless they were retained by HM presence (Figure 16a).

It was found that the amount of heterologous (gluco)mannan was decisive of the relative TpolCBM16-1 and PaCBM35 secretion (Figure 27, Figure 28), demonstrating the ability of CBMs not only to detect HM but also to give insight into the produced HM quantities. In contrast, TmCBM27 consistently showed low secretion rates across all tested backgrounds.

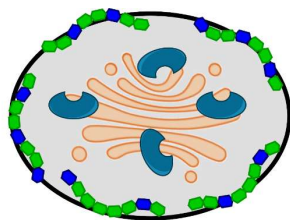
While the α -signal peptide enabled successful secretion of TpolCBM16-1 and PaCBM35 from the wild-type, this signal peptide may not be suitable for TmCBM27 secretion. Recent studies uncovered potential shortcomings of the α -signal peptide and proposed alternative, potentially better performing, signal peptides tailored for *Pichia* (Barrero et al., 2018; Shen et al., 2022). The new signal peptides could improve TmCBM27 secretion. As, from the selected CBM candidates, TmCBM27 seems to exhibit the highest specificity towards β -1,4-mannooligosaccharides, the use of alternative signal peptides for its further exploration could enhance the optimization of mannan-binding CBM-probes (Table 1).

Previous investigations regarding TmCBM27 binding to β -1,4-mannooligosaccharides, carob galactomannan, and konjac glucomannan (Boraston et al., 2003; Boraston et al., 2004) were recently confirmed computationally, where the formation of TmCBM27/mannopentaose complex was demonstrated (Zou et al., 2023). Yet, no studies were found in which the binding of TmCBM27 to differently linked poly- or oligosaccharides was tested. So, the limited TmCBM27 secretion could either result from its specific configuration, or it could potentially bind to native yeast polymers such as β -1,3- and 1,6-glucans or mannoproteins (Aguilar-Uscanga, 2003). Considering the promiscuous carbohydrate recognition of several CBMs (Obembe et al., 2007), future research could explore the potential recognition of non- β -1,4-linked carbohydrates by TmCBM27. A computational approach, such as a recently introduced deep learning-based system (Zou et al., 2023), could be the starting point for testing diverse ligands, followed by empirical ligand binding experiments.

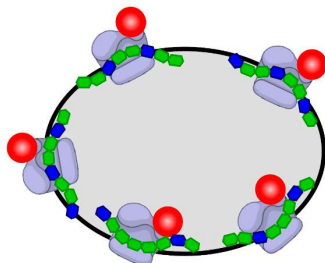
4.2.2 In Yeast, CBM-Probes Co-Localize with CSLAs

Surprisingly, when co-expressed with CSLAs, the CBMs did not localize to the cell walls - where heterologous (gluco)mannan was expected - but rather co-localized with the CSLA enzymes (Figure 19, Figure 28). This raised questions about the extent of heterologous HM incorporation into the yeast cell walls (Figure 39).

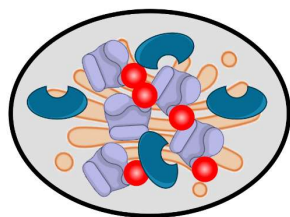
Expected localization of (Gluco)mannan in yeast after biosynthesis by CSLA



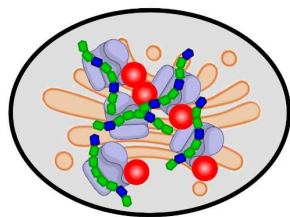
Expected CBM localization in (Gluco)mannan-producing yeast



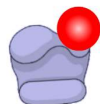
Observed CBM localization in (Gluco)mannan-producing yeast



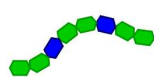
Hypothesized (Gluco)mannan localization in CSLA/CBM co-expressing yeast



CSLA



Fluorescently tagged CBM



Linear (Gluco)Mannan

Figure 39 Expected vs. observed CBM-localization in HM-producing yeast. From top to bottom: (Gluco)mannan produced by CSLAs is expected to localize to the yeast cell walls. Therefore, CBMs were expected to localize to the cell walls too. However, they were found to co-localize with the CSLAs in the Golgi-Apparatus. It is therefore hypothesized that a significant portion of heterologous (Gluco)mannan may remain in proximity to the CSLAs.

Immunolabelling untagged AtCSLA2 and AkCSLA3 expressing strains revealed (gluco)mannan presence at the cell walls, suggesting that a significant portion of (gluco)mannan may be incorporated into the cell walls. On the other hand, its considerable size (150 kDa (Charles A Janeway et al., 2001)) may prevent LM21 from penetrating the cell wall, preventing intracellular HM labelling.

If most HM were incorporated into the cell walls, the co-localization of the CBMs with the CSLAs would remain unexplained. Co-expression of CBMs and CSLAs could cause CBM/HM binding immediately after biosynthesis and before initiation of vesicular HM transport to the cell wall as CBM/HM may struggle to translocate through the Golgi and to the cell wall. This could be caused by the CBMs themselves. CBM/HM complexes would display different characteristics than HM alone, potentially preventing the HM from being recognized as components that should be incorporated into vesicles for transport to the cell wall (Franková and Fry, 2013). Considering the DPs of 13-22 for heterologous (gluco)mannan (Figure 7, (Robert et al., 2021)), the binding of TpolCBM16-1 to mannopentaose in a 1:1

ratio (Bae et al., 2008; Su et al., 2010), and the fact that mannanase derived family 35 CBMs bind mannopentaose (Tunncliffe et al., 2005), it could be speculated, that 3-4 recombinant CBMs would bind the same (gluco)mannan chain. The overall complex would exceed 200 kDa (α -CBM-FP \approx 50 kDa), possibly preventing it from passing the Golgi membrane.

4.2.3 Optimization of CBM-Probes Using a Ratiometric Fluorescent Protein

The top CBM-probes harbouring TpolCBM16-1 and PaCBM35 were further optimized by introducing a new tag comprising pmScarlet-I (Botman et al., 2019) and a self-cleaving pSmTq2 as a ratiometric probe (Figure 24). The ratiometric probe was designed for plant expression, considering that unlike in yeast cultures, no supernatant is available to measure fluorescence in plants. pSmTq2 was intended to confirm CBM expression, even if they were to secrete out of the plant due to lack of mannan, as expected in the mannan-deficient *Arabidopsis csla2,3,9* triple mutant (Goubet et al., 2009; Liepman and Cavalier, 2012).

Cleaved pSmTq2 remained in the cytosol while the CBMs with the pmScarlet-I tag still secreted out of the cells. The ratios of total pSmTq2 signal to total pmScarlet-I signal were consistent across the tested genotypes (Figure 24), suggesting that pSmTq2 could be a reliable indicator for CBM expression in plants.

In yeast, the secretion efficiency of a native *Arabidopsis* secretion signal peptide (AtEXP10) (Cho and Cosgrove, 2000; Zhang, 2014) was lower than α -mediated CBM secretion. This accounts for the fact that AtEXP10 is a plant signal peptide. It might have difficulties passing the yeast secretory pathway as efficiently as the one of *Arabidopsis*. Still, its secretion in yeast was relatively high in the wild-type, especially when considering CBM-probes without signal peptide, which showed no secretion at all (Figure 24, Figure 26).

In wild-type, recombinant PaCBM35 localized in the cytosol regardless of the secretion signal peptide (Figure 25, Figure 26), confirming that CBM-probes only localized in punctate when co-expressed with CSLAs. Similar to the behaviour of the α -signal peptide, AtEXP10 also resulted in a (gluco)mannan quantity-dependent secretion (Figure 27, Figure 28), underscoring its potential for successful use in plants. Also, the results show that plant secretion peptides are able to function when heterologously expressed in yeast.

4.2.4 CBM-Probes Help Detect New CSLA Variants

After validating the new CBM-probes in yeast, they were tested as to their ability to enable the high throughput discovery of new CSLA enzymes. Four of the 10 selected AkCSLA3-sfGFP library strains to analyse had no amino acid changes, while the others varied from 1-8 mutations. Two strains underwent frameshift mutations, impairing their function and sfGFP fluorescence. Variants 9 and 10 both had the same number of bp and AA changes. Yet, only variant 9 showed elevated amounts of mannose (Figure 30). This raised questions about the impact of mutation locations on the ability of non-functional variants to anchor to the Golgi membrane and the potential influence of altered AA polarity on enzyme functionality.

Indeed, non-functional variants formed larger aggregates than the parental AkCSLA3 and the functional variants, which localized in punctate (Figure 32). It was hypothesized that function loss and mislocalization could be caused by mutations in the transmembrane regions, which could affect protein anchoring to the Golgi membrane. Another hypothesis was that protein aggregation could be due to improper protein folding (Figure 29) caused by polarity changes of the mutated AA. In fact, AA polarity and polar interactions play a crucial role in protein-folding (Shanahan and Thornton, 2005; Pace and Gao, 2013), especially in transmembrane regions (Partridge et al., 2002).

However, there was no correlation between AA polarity and functionality (Table 4). In addition, AlphaFold2.0 prediction showed that despite eight AA changes, the enzyme structure of non-functional variant 1 was similar to the wild-type AkCSLA3 (Figure 33). Yet, AlphaFold 2.0 may overpredict α -helices and β -strands (Stevens and He, 2022), which might cause variant 1 to seemingly match the AkCSLA3-sfGFP parent, despite its potential architectural impairment. Future experiments with more variants are required to provide greater insights into AA or AA combinations essential for CSLA function. In addition, experimentally resolving structures of select CSLA variants would facilitate comparisons with AlphaFold2.0 predictions, establishing a benchmark for its precision in predicting CSLA structures.

PaCBM35-probe secretion rates correlated with the amounts of mannan made by the verified library strains (Figure 31), underscoring the ability of the new CBM probes to predict the CSLA (variant) performance as already demonstrated in section 3.2.3.2..

The independence of CBM secretion from sheer CSLA expression was a significant discovery. Notably, when co-expressed with non-functional CSLA variants, PaCBM35 secretion was similar to wild-type (Figure 30), affirming that CBM secretion depends on (gluco)mannan presence and not CSLA expression.

In whole yeast cultures, the variability within populations often goes unexamined. One key objective of CBM-probe application is to detect new CSLA variants at the single-cell level. Therefore, it is crucial to understand the distribution of cell size, gene expression, and viability of a yeast population. Experiments with AkCSLA3, AtCSLA2, and the CSLA swaps provided valuable insight into individual CSLA domain functions. However, the limited sample size (~1 μ L) and the narrow acquisition plane in microscopy only allowed the analysis of around 1000 cells per genotype (Chapter 3.1 & (Robert et al., 2021)). Moreover, microscopy does not allow the recovery of single cells, making it unsuitable for CBM-mediated CSLA variant searching.

Single-cell analysis via cytometry coupled with subsequent cell sorting could overcome microscopical limitations (Hohenblum et al., 2003). In such a setup, only cells with functional CSLAs would show fluorescence from retained CBMs, enabling high throughput cell sorting based on CSLA functionality. This approach could also help pinpoint influential amino acids

responsible for the fine structure of produced (gluco)mannans and overall CSLA functionality. As proof of concept, single-cell cytometry was tested as to its ability to help identify functional AkCSLA3-sfGFP library variants in wild-type cultures stained with PI to identify dead cells.

The sfGFP signal allowed to discern which populations expressed AkCSLA3-sfGFP or a variant. By examining the PI signal, it could also be identified which populations suffered from the (gluco)mannan accumulation toxicity, enabling to differentiate between functional and non-functional AkCSLA3 variants (Figure 34). The clear sfGFP and PI distribution patterns between unstained wild-type, sfGFP, and AkCSLA3-sfGFP populations were mirrored in the CSLA variants with the respective traits (sfGFP signal, PI staining). This demonstrated that even without the CBM-probes, which will allow more accurate CSLA function predictions in the future, it was possible to distinguish functional from non-functional CSLA variants.

Hence, combining cytometry with cell sorting offers a promising tool for high-throughput gene-library screening, particularly once the genes to screen will be co-expressed with the CBM-probes. Furthermore, this approach presents a rapid method for cell viability assessment, granting insights into the toxicity of specific (gluco)mannan structures for *Pichia*.

4.2.5 CBM-Probes Label Heteromannans *in planta*

The novel CBM-probes are also envisioned to provide unprecedented insight into the *in vivo* distribution of plant HM in wild-type plants and those with mutations affecting HM production. This will improve the understanding of how they affect HM distribution.

In *N. benthamiana* leaves, the pSmTq2 expression control localized differently than the CBM-probes, indicating successful cleaving, as previously observed in yeast (Figure 25, Figure 26). Interestingly, TpolCBM16-1 and PaCBM35 localized similarly in plasmolysed cells. Localization differences were rather tied to the different signal peptides (Figure 36).

The fact that the signal peptides had a stronger influence on localization than the CBMs underscores the similar substrate specificities of TpolCBM16-1 and PaCBM35 (Table 1). However, in this context, it remains uncertain if the CBMs labelled HM. Yet, when plasmolysed, AtEXP10-TpolCBM16-1 and AtEXP10-PaCBM35 showed more punctate structures than the α -pmScarlet-I control, suggesting at least some level of mannan-binding.

Mannan labelling could be hindered by pectin that embeds cell wall mannans (Marcus et al., 2010). While the recombinant CBMs (\approx 50 kDa) are smaller than typical antibodies (\approx 150 kDa) used to immunolabel mannans (Charles A Janeway et al., 2001), theoretically improving *in vivo* mannan access, the CBMs may still be too large compared to typical cell wall dyes. On the other hand, mannan acetylation sometimes protects them against CBM-binding and thereby reduces or prevents their degradation by glycosyl hydrolases (Montanier et al.,

2009; Marcus et al., 2010; Berglund et al., 2020). Consequently, even if the CBMs are sufficiently small to bypass the pectin, acetylation might still prevent them from labelling mannan effectively. Additionally, the relatively low mannan content in *N. benthamiana* leaves (Nguema-Ona et al., 2012) could contribute to a lack of CBM movement to the cell walls. In fact, in the roots of *Arabidopsis* Col-0 seedlings, which contain higher amounts of mannan, punctate structures at the cell walls suggested successful mannan labelling (Figure 38).

The top CBM candidate, α -TpolCBM16-1, was stably expressed in the *Arabidopsis* Col-0 wild-type and the *cs/a2,3,9* triple mutant. While the signal strength in Col-0 was significantly higher than in *cs/a2,3,9*, TpolCBM16-1 localization differences between the backgrounds were clearly discernible (Figure 38). Both pmScarlet-I controls (with and without α -signal peptide) primarily localized in smooth patterns at the epidermal cell edges in both *Arabidopsis* backgrounds. TpolCBM16-1, on the other hand, localized in punctate at the edges of the epidermal cells and showed very strong xylem labelling in Col-0 (Figure 38), aligning with the known abundance of mannans in the xylem and vasculature (Handford et al., 2003; Goubet et al., 2009; Marcus et al., 2010; Kim and Daniel, 2012).

In contrast, TpolCBM16-1 labelled only a few epidermal cells of the *cs/a2,3,9* mutant. The TpolCBM16-1 labelling patterns in these cells were similar to the ones in Col-0, suggesting that, at least in these cells, (gluco)mannan distribution may be similar in Col-0 wild-type and *cs/a2,3,9*. From these observations, it could be proposed that the mannan deficiency in the *cs/a2,3,9* mutant may not necessarily stem from a universal (gluco)mannan content reduction in each cell but rather from a complete absence in most epidermal cells and wild-type like (gluco)mannan production in a few cells.

These observations contrast with an earlier immunolabelling study on mature inflorescence *Arabidopsis* stems of the *cs/a2,3,9* mutant, which suggested a complete absence of mannan (Goubet et al., 2009). However, this study had a limited scope of cell analysis due to the laborious preparation of each stem section for immunolabelling. In contrast, CBM-probe expression enabled rapid microscopical screening of entire seedlings without preliminary sample preparation. Hence, it is not unlikely that previous studies may have overlooked the few cells in *Arabidopsis* stems that still produced mannan or that the observations in the present study are seedling-root-specific. In addition, the previous study showed that the presence of mannan in a few cells could be overlooked in carbohydrate analysis.

On the other hand, TpolCBM16-1 labelled neither xylem cells nor the vasculature in *cs/a2,3,9*, which aligns with the lack of mannan in this mutant (Goubet et al., 2009; Liepman and Cavalier, 2012) and underscores the specificity of CBM-based probes for heteromannans *in planta*. Consequently, this study highlights the value of non-invasive mannan labelling probes to increase the precision of *in planta* mannan distribution and content analysis.

Despite promising results in terms of HM labelling, certain challenges remain to be addressed. No pSmTq2 labelling was detected in the seedling roots, which could not be attributed to an imaging issue, as the signals from the yeast control pSmTq2 and a peroxisomal CFP *Arabidopsis* control (Nelson et al., 2007) were clearly visible. The lack of pSmTq2 signal may have been pH-related since fluorescent proteins like GFP respond to pH changes (Bizzarri et al., 2009). Here, the pSmTq2 signal was stable in *N. benthamiana* leaves even in elevated pH after plasmolysis (Figure 36). In addition, previous research showed that in *N. benthamiana* mesophyll cells, mTurquoise2 was stable in a pH range from 5-8 and only showed a slight signal reduction at pH 4 (Stoddard and Rolland, 2019). Given that the pH variations in *Arabidopsis* fall in this pH range (apoplast \approx pH (5-6); cytosol \approx pH 7) (Wegner and Shabala, 2020; Zhou et al., 2021), a pH-related lack of pSmTq2 signal seems unlikely, especially since mTq2 has already successfully been expressed in *Arabidopsis* seedlings previously (Schürholz et al., 2018; Gehrke et al., 2023). Since the codon optimization for expression in *Pichia* had no detrimental effect on the expression in *N. benthamiana* either, the cause for the absence of the pSmTq2 signal in the present study should be investigated further in the future.

Another concern was the high difference in CBM signal strength between Col-0 and *cs/a2,3,9*, which could possibly be an expression issue. Previous findings unveiled a resistance of a *rat4 Arabidopsis* mutant against *A. tumefaciens*-driven transformation. The cause of this resistance was identified as being a T-DNA insertion in an untranslated region of the CSLA9 gene, which also caused CSLA9 transcript reduction (Zhu et al., 2003). Unpublished findings from the “Designer Glycans” group also uncovered a difficulty of *A. tumefaciens* transformation of *cs/a2,3,9* mutant lines, suggesting that the CSLA9 gene impairment may have the same effect in the *cs/a2,3,9* mutant and the *rat4* mutant.

In the future, it will be vital to confirm that the absence of TpolCBM16-1 from the xylem in *cs/a2,3,9* is indeed due to the lack of HM and not low expression levels. In addition, it is important to consider that screened T3 generation has yet to be analysed as to the insertion loci of TpolCBM16-1 and that at this stage, the signal intensity differences could also be due to different CBM-probe gene integration loci and the resulting differing expression strengths. Using RT-qPCR, the transcripts of the CBM-probes, the fluorescent controls, and the (Gluco)mannan-producing CSLAs should be analysed, and expression levels correlated to the mannan content and CBM fluorescence signal intensity. This will allow to verify the observations made in this work. Additionally, the performance of new CBM-probes in other *Arabidopsis* mutants, like the cellulose-reduced mutant *kor1-4* (Mielke et al., 2021) or CSLA overexpressing mutants (Goubet et al., 2009), could provide intriguing insights, potentially unveiling the influence of cellulose absence on HM localization or how cell walls might reorganize upon increased heteromannan presence respectively.

Prior investigations with mannan-specific antibodies focused predominantly on stems, where mannan was found to be distributed evenly in a significant portion of primary and secondary cell walls. This patterning differs from the punctate localizations observed in this study (Handford et al., 2003; Marcus et al., 2010), making additional investigations with respect to the observed CBM-localizations *in planta* crucial, especially in stem sections.

4.3 Conclusion

In sum, the results demonstrate that CSLA alteration via domain swaps or the creation of CSLA libraries can facilitate the discovery of new, favourable heteromannans with tailored fine structures and positive effects on yeast growth. A new *in vivo* heteromannan-labelling tool based on CBMs was also successfully developed. The novel CBM-probes were validated in yeast, and their potential to assist high throughput discovery of new CSLA variants and elucidate HM distribution *in planta* was demonstrated through proof-of-concept experiments.

The key advantage of these findings is that the approaches used in this work could be replicated for other CSLA enzymes, thereby facilitating the fine-tuning and *in vivo* detection of a wide array of hemicelluloses.

In the future, CBM-probes could be used for various purposes related to mannan probing, heteromannan biosynthesis tailoring, and biotechnological applications. They could serve to decipher distributions and dynamics of heteromannans not only in model plants but also in industrially relevant crops like maize or cereal, aiding the discovery of variants with higher nutritional value due to increased heteromannan presence.

Furthermore, coupling cytometry with cell sorting could enable high throughput CBM-probe-based screening for new CSLA variants and uncover the key amino acids that influence heteromannan fine structure. This could enable precise heteromannan biosynthesis fine-tuning and subsequent production of heteromannans with desirable traits for improved crop nutrition, biocapsules with tailored drug release times (Homayun et al., 2019).

Beyond this, CBMs could impart new functions to engineered living materials. CBMs could be recombined with antimicrobial peptides like LL-37 (Dürr et al., 2006) and serve as junctions between scaffold carbohydrates like mannans (Kang et al., 2021) or living materials, leading to (engineered living) materials with antimicrobial traits.

With the continuous discovery of new CBMs, the number of potential future applications increases daily. Beyond the mentioned potential applications, exploring these versatile carbohydrate binders will certainly unveil possible applications across various fields.

5 REFERENCES

- Abbott DW, Boraston AB** (2012) Quantitative approaches to the analysis of carbohydrate-binding module function. *Methods in enzymology* **510**: 211–231
- Abramson M, Shoseyov O, Shani Z** (2010) Plant cell wall reconstruction toward improved lignocellulosic production and processability. *Plant Science* **178**: 61–72
- Aguilar-Uscanga B** (2003) A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. *Lett Appl Microbiol* **37**: 268–274
- Akdel M, Pires DEV, Pardo EP, Jänes J, Zalevsky AO, Mészáros B, Bryant P, Good LL, Laskowski RA, Pozzati G, et al** (2022) A structural biology community assessment of AlphaFold2 applications. *Nat Struct Mol Biol* **29**: 1056–1067
- Amos RA, Mohnen D** (2019) Critical Review of Plant Cell Wall Matrix Polysaccharide Glycosyltransferase Activities Verified by Heterologous Protein Expression. *Frontiers in plant science* **10**: 915
- Anderson CT, Carroll A, Akhmetova L, Somerville C** (2010) Real-Time Imaging of Cellulose Reorientation during Cell Wall Expansion in Arabidopsis Roots. *Plant Physiology* **152**: 787–796
- Anderson CT, Wallace IS** (2012) Illuminating the wall: Using click chemistry to image pectins in Arabidopsis cell walls. *Plant Signaling & Behavior* **7**: 661–663
- Andrade FK, Costa R, Domingues L, Soares R, Gama M** (2010) Improving bacterial cellulose for blood vessel replacement: Functionalization with a chimeric protein containing a cellulose-binding module and an adhesion peptide. *Acta biomaterialia* **6**: 4034–4041
- Armenta S, Moreno-Mendieta S, Sánchez-Cuapio Z, Sánchez S, Rodríguez-Sanoja R** (2017) Advances in molecular engineering of carbohydrate-binding modules. *Proteins* **85**: 1602–1617
- Bae B, Ohene-Adjei S, Kocherginskaya S, Mackie RI, Spies MA, Cann IKO, Nair SK** (2008) Molecular Basis for the Selectivity and Specificity of Ligand Recognition by the Family 16 Carbohydrate-binding Modules from *Thermoanaerobacterium polysaccharolyticum* ManA. *Journal of Biological Chemistry* **283**: 12415–12425

- Barone GD, Emmerstorfer-Augustin A, Biundo A, Pisano I, Coccetti P, Mapelli V, Camattari A** (2023) Industrial Production of Proteins with *Pichia pastoris*—*Komagataella phaffii*. *Biomolecules* **13**: 441
- Barrero JJ, Casler JC, Valero F, Ferrer P, Glick BS** (2018) An improved secretion signal enhances the secretion of model proteins from *Pichia pastoris*. *Microb Cell Fact* **17**: 161
- Behera SS, Ray RC** (2016) Konjac glucomannan, a promising polysaccharide of *Amorphophallus konjac* K. Koch in health care. *International Journal of Biological Macromolecules* **92**: 942–956
- Ben Azoun S, Kallel H** (2017) Investigating the effect of carbon source on rabies virus glycoprotein production in *Pichia pastoris* by a transcriptomic approach. *MicrobiologyOpen* **6**: e00489
- Berestovoy M, Tyurin A, Kabardaeva K, Sidorchuk Y, Fomenkov A, Nosov A, Goldenkova-Pavlova I** (2018) Transient Gene Expression for the Characteristic Signal Sequences and the Estimation of the Localization of Target Protein in Plant Cell. *BIO-PROTOCOL*. doi: 10.21769/BioProtoc.2738
- Berglund J, Kishani S, Morais De Carvalho D, Lawoko M, Wohler J, Henriksson G, Lindström ME, Wågberg L, Vilaplana F** (2020) Acetylation and Sugar Composition Influence the (In)Solubility of Plant β -Mannans and Their Interaction with Cellulose Surfaces. *ACS Sustainable Chem Eng* **8**: 10027–10040
- Bewley JD, Bradford KJ, Hilhorst HWM, Nonogaki H** (2013) Structure and Composition. *In* JD Bewley, KJ Bradford, HWM Hilhorst, H Nonogaki, eds, *Seeds*. Springer New York, New York, NY, pp 1–25
- Bizzarri R, Serresi M, Luin S, Beltram F** (2009) Green fluorescent protein based pH indicators for in vivo use: a review. *Anal Bioanal Chem* **393**: 1107–1122
- Blake AW, McCartney L, Flint JE, Bolam DN, Boraston AB, Gilbert HJ, Knox JP** (2006) Understanding the Biological Rationale for the Diversity of Cellulose-directed Carbohydrate-binding Modules in Prokaryotic Enzymes. *Journal of Biological Chemistry* **281**: 29321–29329
- Bolam DN, Ciruela A, McQueen-Mason S, Simpson P, Williamson MP, Rixon JE, Boraston A, Hazlewood GP, Gilbert HJ** (1998) *Pseudomonas* cellulose-binding

domains mediate their effects by increasing enzyme substrate proximity. The Biochemical journal **331** (Pt 3: 775–781

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. The Biochemical journal **382**: 769–781

Boraston AB, Creagh AL, Alam MM, Kormos JM, Tomme P, Haynes CA, Warren RA, Kilburn DG (2001) Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A. Biochemistry **40**: 6240–6247

Boraston AB, Revett TJ, Boraston CM, Nurizzo D, Davies GJ (2003) Structural and Thermodynamic Dissection of Specific Mannan Recognition by a Carbohydrate Binding Module, TmCBM27. Structure **11**: 665–675

Botman D, de Groot DH, Schmidt P, Goedhart J, Teusink B (2019) In vivo characterisation of fluorescent proteins in budding yeast. Sci Rep **9**: 2234

Buckeridge MS (2010) Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. Plant physiology **154**: 1017–1023

Charles A Janeway J, Travers P, Walport M, Shlomchik MJ (2001) The structure of a typical antibody molecule. Immunobiology: The Immune System in Health and Disease. 5th edition

Cho H-T, Cosgrove DJ (2000) Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **97**: 9783–9788

Ciucanu I, Kerek F (1984) A simple and rapid method for the permethylation of carbohydrates. Carbohydrate Research **131**: 209–217

Cocuron J-C, Lerouxel O, Drakakaki G, Alonso AP, Liepman AH, Keegstra K, Raikhel N, Wilkerson CG (2007) A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase. Proceedings of the National Academy of Sciences of the United States of America **104**: 8550–8555

Couturier M, Haon M, Coutinho PM, Henrissat B, Lesage-Meessen L, Berrin J-G (2011) *Podospora anserina* Hemicellulases Potentiate the *Trichoderma reesei* Secretome for Saccharification of Lignocellulosic Biomass. Appl Environ Microbiol **77**: 237–246

Couturier M, Roussel A, Rosengren A, Leone P, Stålbrand H, Berrin J-G (2013) Structural and Biochemical Analyses of Glycoside Hydrolase Families 5 and 26 β -(1,4)-

Mannanases from *Podospora anserina* Reveal Differences upon Manno-oligosaccharide Catalysis. *Journal of Biological Chemistry* **288**: 14624–14635

Daly R, Hearn MTW (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *Journal of molecular recognition* : *JMR* **18**: 119–138

Davis J, Brandizzi F, Liepman AH, Keegstra K (2010) Arabidopsis mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane. *The Plant journal : for cell and molecular biology* **64**: 1028–1037

Derevnina L, Kamoun S, Wu C (2019) Dude, where is my mutant? *Nicotiana benthamiana* meets forward genetics. *New Phytol* **221**: 607–610

Dimitroff G, Little A, Lahnstein J, Schwerdt JG, Srivastava V, Bulone V, Burton RA, Fincher GB (2016) (1,3;1,4)- β -Glucan Biosynthesis by the CSLF6 Enzyme: Position and Flexibility of Catalytic Residues Influence Product Fine Structure. *Biochemistry* **55**: 2054–2061

Din N, Damude HG, Gilkes NR, Miller RC, Warren RA, Kilburn DG (1994) C1-Cx revisited: intramolecular synergism in a cellulase. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 11383–11387

Dürr UHN, Sudheendra US, Ramamoorthy A (2006) LL-37, the only human member of the cathelicidin family of antimicrobial peptides. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1758**: 1408–1425

Dwivany FM, Yulia D, Burton RA, Shirley NJ, Wilson SM, Fincher GB, Bacic A, Newbiggin E, Doblin MS (2009) The CELLULOSE-SYNTHASE LIKE C (CSLC) family of barley includes members that are integral membrane proteins targeted to the plasma membrane. *Molecular plant* **2**: 1025–1039

Ebringerová A, Hromádková Z, Heinze T (2005) Hemicellulose. *In* T Heinze, ed, *Polysaccharides I*. Springer-Verlag, Berlin/Heidelberg, pp 1–67

Engler C, Youles M, Gruetzner R, Ehnert T-M, Werner S, Jones JDG, Patron NJ, Marillonnet S (2014) A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synth Biol* **3**: 839–843

- Franková L, Fry SC** (2013) Biochemistry and physiological roles of enzymes that “cut and paste” plant cell-wall polysaccharides. *Journal of experimental botany* **64**: 3519–3550
- Fujiwara T, Oda K, Yokota S, Takatsuki A, Ikehara Y** (1988) Brefeldin A causes disassembly of the Golgi complex and accumulation of secretory proteins in the endoplasmic reticulum. *Journal of Biological Chemistry* **263**: 18545–18552
- Gehrke F, Ruiz-Duarte P, Schindele A, Wolf S, Puchta H** (2023) An inducible CRISPR-KILL system for temporally controlled cell type-specific cell ablation in *Arabidopsis thaliana*. *New Phytologist* **nph.19102**
- GenScript** (2023) Codon Usage Frequency Table. GenScript Codon Usage Frequency Table(chart) Tool, <https://www.genscript.com/tools/codon-frequency-table>
- Gilbert C, Ellis T** (2018) Biological Engineered Living Materials: Growing Functional Materials with Genetically Programmable Properties. *ACS Synthetic Biology*. doi: 10.1021/acssynbio.8b00423
- Gilbert C, Tang T-C, Ott W, Dorr BA, Shaw WM, Sun GL, Lu TK, Ellis T** (2021) Living materials with programmable functionalities grown from engineered microbial co-cultures. *Nat Mater* **20**: 691–700
- Gilbert HJ, Knox JP, Boraston AB** (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Current opinion in structural biology* **23**: 669–677
- Gille S, Cheng K, Skinner ME, Liepman AH, Wilkerson CG, Pauly M** (2011) Deep sequencing of voodoo lily (*Amorphophallus konjac*): an approach to identify relevant genes involved in the synthesis of the hemicellulose glucomannan. *Planta* **234**: 515–526
- Giraldo MC, Valent B** (2013) Filamentous plant pathogen effectors in action. *Nat Rev Microbiol* **11**: 800–814
- Goodin MM, Zaitlin D, Naidu RA, Lommel SA** (2008) *Nicotiana benthamiana* : Its History and Future as a Model for Plant–Pathogen Interactions. *MPMI* **21**: 1015–1026
- Goubet F, Barton CJ, Mortimer JC, Yu X, Zhang Z, Miles GP, Richens J, Liepman AH, Seffen K, Dupree P** (2009) Cell wall glucomannan in *Arabidopsis* is synthesised by

- CSLA glycosyltransferases, and influences the progression of embryogenesis. *The Plant Journal* **60**: 527–538
- Goubet F, Misrahi A, Park SK, Zhang Z, Twell D, Dupree P** (2003) AtCSLA7, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in *Arabidopsis*. *Plant physiology* **131**: 547–557
- Graumann K, Premstaller A** (2006) Manufacturing of recombinant therapeutic proteins in microbial systems. *Biotechnology journal* **1**: 164–186
- Grefen C, Donald N, Hashimoto K, Kudla J, Schumacher K, Blatt MR** (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *The Plant Journal* **64**: 355–365
- Han R, Li J, Shin HD, Chen RR, Du G, Liu L, Chen J** (2013) Carbohydrate-binding module-cyclodextrin glycosyltransferase fusion enables efficient synthesis of 2-O-d-glucopyranosyl-l-ascorbic acid with soluble starch as the glycosyl donor. *Applied and Environmental Microbiology* **79**: 3234–3240
- Handford MG, Baldwin TC, Goubet F, Prime TA, Miles J, Yu X, Dupree P** (2003) Localisation and characterisation of cell wall mannan polysaccharides in *Arabidopsis thaliana*. *Planta* **218**: 27–36
- Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM** (2000) pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Molecular Biology* **42**: 819–832
- Henrissat B, Coutinho PM, Davies GJ** (2001) A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. *Plant molecular biology* **47**: 55–72
- Hohenblum H, Borth N, Mattanovich D** (2003) Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *Journal of Biotechnology* **102**: 281–290
- Homayun B, Lin X, Choi H-J** (2019) Challenges and Recent Progress in Oral Drug Delivery Systems for Biopharmaceuticals. *Pharmaceutics* **11**: 129
- Jobling SA** (2015) Membrane pore architecture of the CslF6 protein controls (1-3,1-4)- β -glucan structure. *Sci Adv* **1**: e1500069

- Jordà J, Suarez C, Carnicer M, Ten Pierick A, Heijnen JJ, Van Gulik W, Ferrer P, Albiol J, Wahl A** (2013) Glucose-methanol co-utilization in *Pichia pastoris* studied by metabolomics and instationary ¹³C flux analysis. *BMC Syst Biol* **7**: 17
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al** (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583–589
- Juturu V, Wu JC** (2018) Heterologous Protein Expression in *Pichia pastoris*: Latest Research Progress and Applications. *Chembiochem : a European journal of chemical biology* **19**: 7–21
- Kang S-Y, Pokhrel A, Bratsch S, Benson JJ, Seo S-O, Quin MB, Aksan A, Schmidt-Dannert C** (2021) Engineering *Bacillus subtilis* for the formation of a durable living biocomposite material. *Nat Commun* **12**: 7133
- Kerstens S** (2002) Cellulose Orientation in the Outer Epidermal Wall of Angiosperm Roots: Implications for Biosystematics. *Annals of Botany* **90**: 669–676
- Khatri V, Hébert-Ouellet Y, Meddeb-Mouelhi F, Beauregard M** (2016) Specific tracking of xylan using fluorescent-tagged carbohydrate-binding module 15 as molecular probe. *Biotechnol Biofuels* **9**: 74
- Kim JS, Daniel G** (2012) Immunolocalization of hemicelluloses in *Arabidopsis thaliana* stem. Part II: Mannan deposition is regulated by phase of development and its patterns of temporal and spatial distribution differ between cell types. *Planta* **236**: 1367–1379
- Knox JP** (2008) Revealing the structural and functional diversity of plant cell walls. *Current opinion in plant biology* **11**: 308–313
- Krämer U** (2015) Planting molecular functions in an ecological context with *Arabidopsis thaliana*. *eLife* **4**: e06100
- Krenek P, Samajova O, Luptovciak I, Doskocilova A, Komis G, Samaj J** (2015) Transient plant transformation mediated by *Agrobacterium tumefaciens*: Principles, methods and applications. *Biotechnology Advances* **33**: 1024–1042
- Kuti EL, Kuti JL** (2010) Pharmacokinetics, antifungal activity and clinical efficacy of anidulafungin in the treatment of fungal infections. *Expert Opinion on Drug Metabolism & Toxicology* **6**: 1287–1300

- Lacapere J-J, ed** (2017) Membrane Protein Structure and Function Characterization: Methods and Protocols. doi: 10.1007/978-1-4939-7151-0
- Le Mauff F, Loutelier-Bourhis C, Bardor M, Berard C, Doucet A, D'Aoust M-A, Vezina L-P, Driouich A, Couture MM-J, Lerouge P** (2017) Cell wall biochemical alterations during *Agrobacterium* -mediated expression of haemagglutinin-based influenza virus-like vaccine particles in tobacco. *Plant Biotechnol J* **15**: 285–296
- Levy I, Shani Z, Shoseyov O** (2002) Modification of polysaccharides and plant cell wall by endo-1,4- β -glucanase and cellulose-binding domains. *Biomolecular Engineering* **19**: 17–30
- Liepman AH, Cavalier DM** (2012) The CELLULOSE SYNTHASE-LIKE A and CELLULOSE SYNTHASE-LIKE C families: recent advances and future perspectives. *Frontiers in plant science* **3**: 109
- Liepman AH, Wilkerson CG, Keegstra K** (2005) Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 2221–2226
- Lin-Cereghino GP, Stark CM, Kim D, Chang J, Shaheen N, Poerwanto H, Agari K, Moua P, Low LK, Tran N, et al** (2013) The effect of α -mating factor secretion signal mutations on recombinant protein expression in *Pichia pastoris*. *Gene* **519**: 311–317
- Lin-Cereghino J, Wong WW, Xiong S, Giang W, Luong LT, Vu J, Johnson SD, Lin-Cereghino GP** (2005) Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *BioTechniques* **38**: 44–48
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B** (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research* **42**: D490-5
- Marcus SE, Blake AW, Benians TAS, Lee KJD, Poyser C, Donaldson L, Leroux O, Rogowski A, Petersen HL, Boraston A, et al** (2010) Restricted access of proteins to mannan polysaccharides in intact plant cell walls: Masking of mannan polysaccharides. *The Plant Journal* **64**: 191–203

- Marillonnet S, Grützner R** (2020) Synthetic DNA Assembly Using Golden Gate Cloning and the Hierarchical Modular Cloning Pipeline. *Current Protocols in Molecular Biology*. doi: 10.1002/cpmb.115
- McCartney L, Gilbert HJ, Bolam DN, Boraston AB, Knox JP** (2004) Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. *Analytical Biochemistry* **326**: 49–54
- McNew JA, Parlati F, Fukuda R, Johnston RJ, Paz K, Paumet F, Rothman JE** (2000) Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Cell* **101**: 407:
- Mielke S, Zimmer M, Meena MK, Dreos R, Stellmach H, Hause B, Voiniciuc C, Gasperini D** (2021) Jasmonate biosynthesis arising from altered cell walls is prompted by turgor-driven mechanical compression. *Sci Adv* **7**: eabf0356
- Montanier C, Money VA, Pires VMR, Flint JE, Pinheiro BA, Goyal A, Prates JAM, Izumi A, Stålbrand H, Morland C, et al** (2009) The Active Site of a Carbohydrate Esterase Displays Divergent Catalytic and Noncatalytic Binding Functions. *PLOS Biology* **7**: e1000071
- Murdoch D, Plosker GL** (2004) Anidulafungin: *Drugs* **64**: 2249–2258
- Nardi C, Escudero C, Villarreal N, Martínez G, Civello PM** (2013) The carbohydrate-binding module of *Fragaria × ananassa* expansin 2 (CBM-FaExp2) binds to cell wall polysaccharides and decreases cell wall enzyme activities *in vitro*. *Journal of plant research* **126**: 151–159
- Nardi CF, Villarreal NM, Rossi FR, Martínez S, Martínez GA, Civello PM** (2015) Overexpression of the carbohydrate binding module of strawberry expansin2 in *Arabidopsis thaliana* modifies plant growth and cell wall metabolism. *Plant molecular biology* **88**: 101–117
- Nelson BK, Cai X, Nebenführ A** (2007) A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. *The Plant Journal* **51**: 1126–1136
- Nguema-Ona E, Moore JP, Fagerström A, Fangel JU, Willats WGT, Hugo A, Vivier MA** (2012) Profiling the main cell wall polysaccharides of tobacco leaves using high-throughput and fractionation techniques. *Carbohydrate Polymers* **88**: 939–949

- Nishio M** (2011) The C-H \cdots O hydrogen bond in chemistry. Conformation, supramolecules, optical resolution and interactions involving carbohydrates. *Physical chemistry chemical physics* : PCCP **13**: 13873–13900
- Notenboom V, Boraston AB, Kilburn DG, Rose DR** (2001) Crystal structures of the family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A in native and ligand-bound forms. *Biochemistry* **40**: 6248–6256
- Obembe OO, Jacobsen E, Timmers J, Gilbert H, Blake AW, Knox JP, Visser RGF, Vincken J-P** (2007) Promiscuous, non-catalytic, tandem carbohydrate-binding modules modulate the cell-wall structure and development of transgenic tobacco (*Nicotiana tabacum*) plants. *Journal of plant research* **120**: 605–617
- Oehme DP, Shafee T, Downton MT, Bacic A, Doblin MS** (2019) Differences in protein structural regions that impact functional specificity in GT2 family β -glucan synthases. *PLoS ONE* **14**: e0224442
- Oliveira C, Carvalho V, Domingues L, Gama FM** (2015) Recombinant CBM-fusion technology — Applications overview. *Biotechnology Advances* **33**: 358–369
- Pace CJ, Gao J** (2013) Exploring and Exploiting Polar- π Interactions with Fluorinated Aromatic Amino Acids. *Acc Chem Res* **46**: 907–915
- Partridge AW, Melnyk RA, Deber CM** (2002) Polar Residues in Membrane Domains of Proteins: Molecular Basis for Helix-Helix Association in a Mutant CFTR Transmembrane Segment. *Biochemistry* **41**: 3647–3653
- Pauly M, Gawenda N, Wagner C, Fischbach P, Ramírez V, Axmann IM, Voiniciuc C** (2019) The Suitability of Orthogonal Hosts to Study Plant Cell Wall Biosynthesis. *Plants* **8**: 516
- Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G** (2013) Hemicellulose biosynthesis. *Planta* **238**: 627–642
- Pauly M, Keegstra K** (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant journal : for cell and molecular biology* **54**: 559–568
- Pettolino FA, Walsh C, Fincher GB, Bacic A** (2012) Determining the polysaccharide composition of plant cell walls. *Nat Protoc* **7**: 1590–1607

- Poon DKY, Withers SG, McIntosh LP** (2007) Direct demonstration of the flexibility of the glycosylated proline-threonine linker in the *Cellulomonas fimi* Xylanase Cex through NMR spectroscopic analysis. *The Journal of biological chemistry* **282**: 2091–2100
- Prielhofer R, Barrero JJ, Steuer S, Gassler T, Zahrl R, Baumann K, Sauer M, Mattanovich D, Gasser B, Marx H** (2017) GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. *BMC Syst Biol* **11**: 123
- Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S, et al** (2016) 50 years of *Arabidopsis* research: highlights and future directions. *New Phytol* **209**: 921–944
- Purushotham P, Ho R, Zimmer J** (2020) Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science* **369**: 1089–1094
- Rajasundaram D, Runavot J-L, Guo X, Willats WGT, Meulewaeter F, Selbig J** (2014) Understanding the relationship between cotton fiber properties and non-cellulosic cell wall polysaccharides. *PLoS one* **9**: e112168
- Robert M, Waldhauer J, Stritt F, Yang B, Pauly M, Voiniciuc C** (2021) Modular biosynthesis of plant hemicellulose and its impact on yeast cells. *Biotechnol Biofuels* **14**: 140
- Ryan MD, King AM, Thomas GP** (1991) Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *J Gen Virol* **72 (Pt 11)**: 2727–2732
- Scheller HV, Ulvskov P** (2010) Hemicelluloses. *Annual review of plant biology* **61**: 263–289
- Schnupf U, Willett JL, Momany F** (2010) DFTMD studies of glucose and epimers: anomeric ratios, rotamer populations, and hydration energies. *Carbohydrate Research* **345**: 503–511
- Schultink A, Liu L, Zhu L, Pauly M** (2014) Structural Diversity and Function of Xyloglucan Sidechain Substituents. *Plants (Basel)* **3**: 526–542
- Schultink A, Qi T, Bally J, Staskawicz B** (2019) Using forward genetics in *Nicotiana benthamiana* to uncover the immune signaling pathway mediating recognition of the *Xanthomonas perforans* effector XopJ4. *New Phytol* **221**: 1001–1009

- Schürholz A-K, López-Salmerón V, Li Z, Forner J, Wenzl C, Gaillochet C, Augustin S, Barro AV, Fuchs M, Gebert M, et al** (2018) A Comprehensive Toolkit for Inducible, Cell Type-Specific Gene Expression in Arabidopsis. *Plant Physiology* **178**: 40–53
- Shanahan HP, Thornton JM** (2005) Amino acid architecture and the distribution of polar atoms on the surfaces of proteins. *Biopolymers* **78**: 318–328
- Shen Q, Zhou X-T, Guo Q, Xue Y-Z, Xue Y-P, Zheng Y-G** (2022) Potential of the Signal Peptide Derived from the *PAS_chr3_0030* Gene Product for Secretory Expression of Valuable Enzymes in *Pichia pastoris*. *Appl Environ Microbiol* **88**: e00296-22
- Shoseyov O, Shani Z, Levy I** (2006) Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and molecular biology reviews : MMBR* **70**: 283–295
- Shpigel, Roiz, Goren, Shoseyov** (1998) Bacterial cellulose-binding domain modulates in vitro elongation of different plant cells. *Plant physiology* **117**: 1185–1194
- Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredes A, Persson S, Raab T, et al** (2004) Toward a systems approach to understanding plant cell walls. *Science (New York, NY)* **306**: 2206–2211
- Stevens AO, He Y** (2022) Benchmarking the Accuracy of AlphaFold 2 in Loop Structure Prediction. *Biomolecules* **12**: 985
- Stoddard A, Rolland V** (2019) I see the light! Fluorescent proteins suitable for cell wall/apoplast targeting in *Nicotiana benthamiana* leaves. *Plant Direct* **3**: e00112
- Su X, Agarwal V, Dodd D, Bae B, Mackie RI, Nair SK, Cann IKO** (2010) Mutational Insights into the Roles of Amino Acid Residues in Ligand Binding for Two Closely Related Family 16 Carbohydrate Binding Modules. *Journal of Biological Chemistry* **285**: 34665–34676
- Suzuki T, Fujikura K, Higashiyama T, Takata K** (1997) DNA Staining for Fluorescence and Laser Confocal Microscopy. *J Histochem Cytochem* **45**: 49–53
- Terrett OM, Lyczakowski JJ, Yu L, Iuga D, Franks WT, Brown SP, Dupree R, Dupree P** (2019) Molecular architecture of softwood revealed by solid-state NMR. *Nat Commun* **10**: 4978

- The Arabidopsis Genome Initiative** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796–815
- The UniProt Consortium, Bateman A, Martin M-J, Orchard S, Magrane M, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bye-A-Jee H, et al** (2023) UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**: D523–D531
- Tomme P, Boraston A, McLean B, Kormos J, Creagh AL, Sturch K, Gilkes NR, Haynes CA, Warren RAJ, Kilburn DG** (1998) Characterization and affinity applications of cellulose-binding domains. *Journal of Chromatography B: Biomedical Sciences and Applications* **715**: 283–296
- Tomme P, van Tilbeurgh H, Pettersson G, van Damme J, Vandekerckhove J, Knowles J, Teeri T, Claeysens M** (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *European journal of biochemistry* **170**: 575–581
- Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A** (2015) The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res* **43**: W401–W407
- Tunncliffe RB, Bolam DN, Pell G, Gilbert HJ, Williamson MP** (2005) Structure of a Mannan-specific Family 35 Carbohydrate-Binding Module: Evidence for Significant Conformational Changes upon Ligand Binding. *Journal of Molecular Biology* **347**: 287–296
- Verherbruggen Y, Boudier A, Vigouroux J, Alvarado C, Geairon A, Guillon F, Wilkinson MD, Stritt F, Pauly M, Lee MY, et al** (2021) The TaCslA12 gene expressed in the wheat grain endosperm synthesizes wheat-like mannan when expressed in yeast and *Arabidopsis*. *Plant Science* **302**: 110693
- Voiniciuc C, Dama M, Gawenda N, Stritt F, Pauly M** (2019) Mechanistic insights from plant heteromannan synthesis in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 522–527
- Voiniciuc C, Engle KA, Günl M, Dieluweit S, Schmidt MH-W, Yang J-Y, Moremen KW, Mohnen D, Usadel B** (2018a) Identification of Key Enzymes for Pectin Synthesis in Seed Mucilage. *Plant Physiol* **178**: 1045–1064

- Voiniciuc C, Pauly M, Usadel B** (2018b) Monitoring Polysaccharide Dynamics in the Plant Cell Wall. *Plant Physiol* **176**: 2590–2600
- Voiniciuc C, Schmidt MH-W, Berger A, Yang B, Ebert B, Scheller H V, North HM, Usadel B, Günl M** (2015) MUCILAGE-RELATED10 Produces Galactoglucomannan That Maintains Pectin and Cellulose Architecture in Arabidopsis Seed Mucilage. *Plant physiology* **169**: 403–420
- Wang Y, Alonso AP, Wilkerson CG, Keegstra K** (2012) Deep EST profiling of developing fenugreek endosperm to investigate galactomannan biosynthesis and its regulation. *Plant molecular biology* **79**: 243–258
- Wang Y, Mortimer JC, Davis J, Dupree P, Keegstra K** (2013) Identification of an additional protein involved in mannan biosynthesis. *The Plant journal: for cell and molecular biology* **73**: 105–117
- Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S** (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS ONE* **6**: 11
- Wegner LH, Shabala S** (2020) Biochemical pH clamp: the forgotten resource in membrane bioenergetics. *New Phytologist* **225**: 37–47
- Weinhandl K, Winkler M, Glieder A, Camattari A** (2014) Carbon source dependent promoters in yeasts. *Microb Cell Fact* **13**: 5
- Yamabhai M, Sak-Ubol S, Srila W, Haltrich D** (2016) Mannan biotechnology: from biofuels to health. *Critical reviews in biotechnology* **36**: 32–42
- Yang B, Hofmann F, Usadel B, Voiniciuc C** (2021) Seed hemicelluloses tailor mucilage properties and salt tolerance. *New Phytologist* **229**: 1946–1954
- Yeats T, Vellosillo T, Sorek N, Ibáñez A, Bauer S** (2016) Rapid Determination of Cellulose, Neutral Sugars, and Uronic Acids from Plant Cell Walls by One-step Two-step Hydrolysis and HPAEC-PAD. *BIO-PROTOCOL*. doi: 10.21769/BioProtoc.1978
- Yu L, Shi D, Li J, Kong Y, Yu Y, Chai G, Hu R, Wang J, Hahn MG, Zhou G** (2014) CELLULOSE SYNTHASE-LIKE A2, a Glucomannan Synthase, Is Involved in Maintaining Adherent Mucilage Structure in Arabidopsis Seed. *Plant Physiol* **164**: 1842–1856

- Zhang T** (2014) Using Carbohydrate-Binding Modules as Tools for In Vivo Visualization/Modulation of Plant Cell Wall Polysaccharides. Graduate Faculty of The University of Georgia, Georgia
- Zhou J-Y, Hao D-L, Yang G-Z** (2021) Regulation of Cytosolic pH: The Contributions of Plant Plasma Membrane H⁺-ATPases and Multiple Transporters. *International Journal of Molecular Sciences*. doi: 10.3390/ijms222312998
- Zhu Y, Nam J, Carpita NC, Matthyse AG, Gelvin SB** (2003) Agrobacterium-Mediated Root Transformation Is Inhibited by Mutation of an Arabidopsis Cellulose Synthase-Like Gene. *Plant Physiol* **133**: 1000–1010
- Zou C, Lu L, Wang S, Zhang C, Chen X, Lin Y, Huang Y** (2022) The α -mating factor secretion signals and endogenous signal peptides for recombinant protein secretion in *Komagataella phaffii*. *Biotechnol Biofuels* **15**: 140
- Zou Y, Wang R, Du M, Wang X, Xu D** (2023) Identifying Protein–Ligand Interactions via a Novel Distance Self-Feedback Biomolecular Interaction Network. *J Phys Chem B* **127**: 899–911

6 APPENDICES

Appendix 1 Primer sequences used for Golden Gate and MoClo assembly of constructs cloned for this thesis. Primers are clustered by subprojects they were used for. Primer names for the domain swap project relate to the four domains established in Figure 6. Recognition sites of the enzymes are highlighted in bold, fusion sites are coloured in yellow. Bold lower case letters in the fusion sites of primers denote the domestications (dom) employed to remove unwanted enzyme cut sites. To create the small cleaving peptide P2A, designated to cause self-cleaving of pSmTq2, the first and the second P2A halves were integrated in the respective assembly primers (blue). F and R specify forward or reverse primer orientation.

GOLDEN GATE AND MOCLO ASSEMBLY	
PRIMERS	5' to 3' sequence (fusion sites)
DOMAIN SWAP PRIMERS	
AtCSLA2 dom1 R	atat GGTCTC t GAg aTGCTCGGACGGCGAGAT
AtCSLA2 dom2 F	ataa GGTCTC a tC TTCGCGGCTGGAAT
AkCSLA3 dom1 R	cgac GGTCTC t At ACCTTCACGTACAGGCTCA
AkCSLA3 dom2 F	ataa GGTCTC a GTa TTCGCGGCGGCCAA
AkCSLA3 dom2 R	cggc GGTCTC t Tt CATATTTTATGTTACCCCATCC
AkCSLA3 dom3 F	acat GGTCTC a GAa ACCAGGGCCAACAGAAATG
AkCSLA3 dom3 R	atca GGTCTC t Gt CCGATCTCCGGGTTGTGGAT
AkCSLA3 dom4 F	atat GGTCTC a GGa CTCGCTCAGGCTCGCTGGAA
AtCSLA2-1 F	acat GGTCTC a CATG GACGGTGTATCACCAAAGT
AtCSLA2-1 R	cttc GGTCTC t GAGG ACGACGGGAAATT
AtCSLA2-2 F	taca GGTCTC a CCTC CCTCGTACAAATCCCAT
AtCSLA2-2 R	cgtc GGTCTC t ACCA TCTATGTTGCTGA
AtCSLA2-3 F	acgc GGTCTC t TGGT CTTGTGGACCT
AtCSLA2-3 R	atca GGTCTC t CGTT CGTTAGCCCTCCTGCC
AtCSLA2-4 F	agtc GGTCTC a AACG AGTGGGTAGTGACT
AtCSLA2-4 R	acgc GGTCTC t AAGC CTAACTCGGGACATAAGTCC
AkCSLA3 DOMAIN SWAPS PRIMERS	
AkCSLA3-1 F	ataa GGTCTC a CATG GCCATCGACTGGGC
AkCSLA3-1 R	cttc GGTCTC t GAGG ACCATTGGGTAGGC
AkCSLA3-2 F	taat GGTCTC a CCTC GTCCAAATACCCAT
AkCSLA3-2 R	catc GGTCTC t ACCA TTCGATGCTGTT
AkCSLA3-3 F	taca GGTCTC t TGGT CATGCGGGCC
AkCSLA3-3 R	atac GGTCTC t CGTT GACCCTCCCGATCT
AkCSLA3-4 F	atat GGTCTC a AACG AGTGGGTCGTCACA
AkCSLA3-4 R	cgcg GGTCTC t AAGC CTACTTTTCACTAGGAACAAAGGT
TmCSLC4 DOMAIN SWAP PRIMERS	
TmCSLC4-2 F	taat GGTCTC g CCTC GTCCAAATTCCTATGTGC
TmCSLC4-2 R	catc GGTCTC g ACCA CCATCTATGTTGTTGCTT
PRIMERS FOR α SIGNAL PEPTIDE INTRODUCTION IN BB1_12	
pFDH1 F	agcc GGTCTC a GGAG AAATGGCAGAAGGATCAGCC

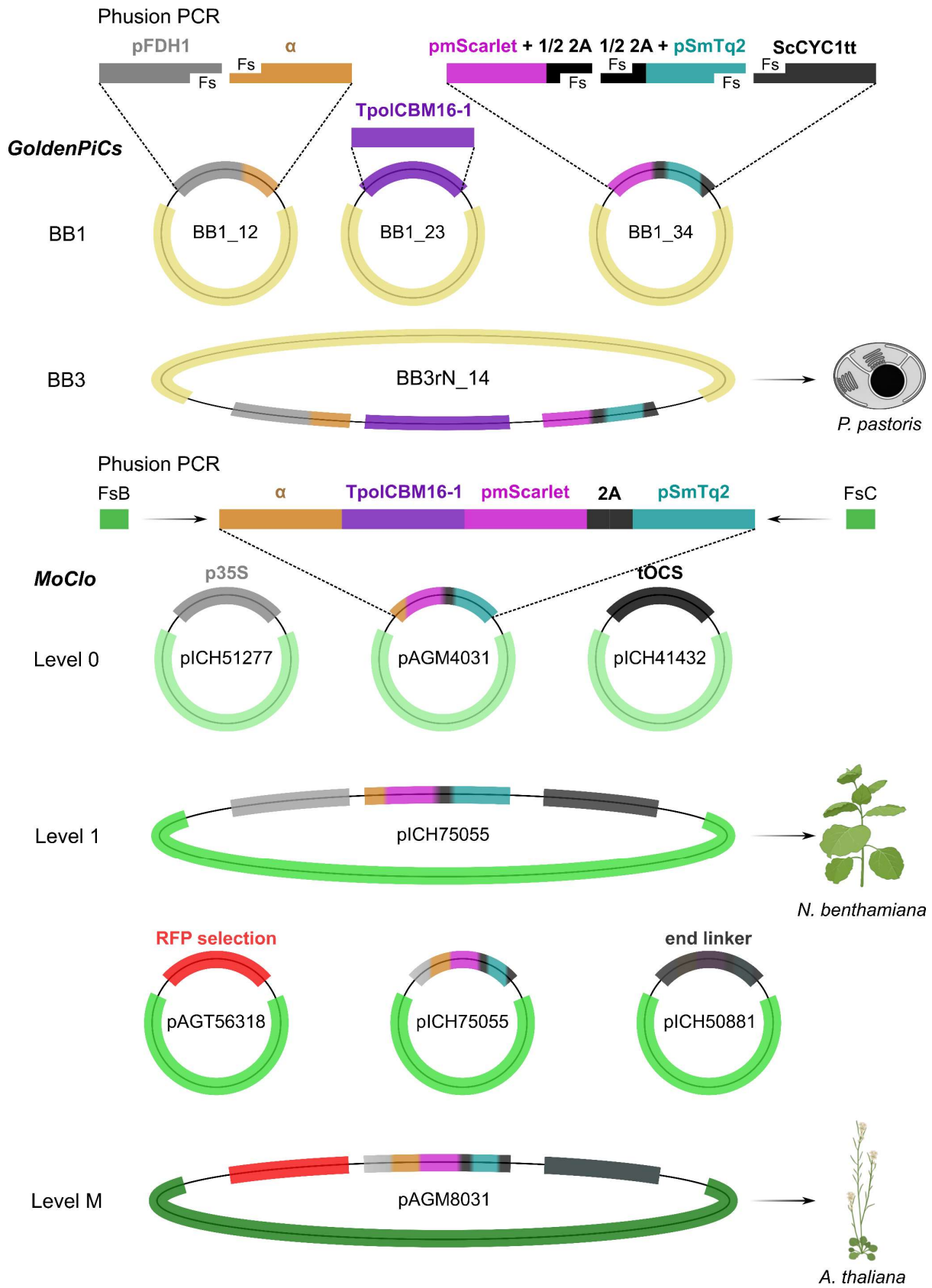
pFDH1 F	gcgcGGTCTCgCAACTGTTTAAGTGGGTGATGTTG
BB1_12 SPalpha F	gtctGGTCTCgGTTGATGAGATTTCTTCAATTTT
BB1_12 SPalpha R	acgaGGTCTCgCATGGCTTCAGCCTCTCTTTTCTC
PRIMERS FOR RECOMBINATION OF CBMS WITH A AND mRuby2 IN BB1_23	
BB1_23 SPalpha F	agccGGTCTCaCATGAGATTTCTTCAATTTTTAC
BB1_23 SPalpha R	aataGGTCTCaAGAACCCAGCTTCAGCCTCTCTTTTCT
CjCBM35 + L2 F	aggaGGTCTCaTTCTGCTGTTCCAGAAGGTAACCTCCTG
CjCBM35 + L2 R	agaaGGTCTCaCTCCTTCTTACCCTTAGAAACCATTGG
mRuby2 F	atatGGTCTCaGGAGTGTCCAAAGGAGAGGAGTTA
mRuby2 R	acgaGGTCTCgAAGCTTACTTATACAATTCATCCATACCACCG
TpoICBM16-1 F	aggaGGTCTCaTTCTAAATATGGTTTCCAATCCAGG
TpoICBM16-1 R	agaaGGTCTCaCTCCCACTTCAACTAAATAAACGTC
PRIMERS FOR CREATING RATIO-METRIC PROBE IN BB1_34	
pmScarlet-I + 1st half P2A F	ataGGTCTCgCAGCttgcttcaacagagaaaaattagtggcTTTGTACAATTCATCCATAC
2nd half P2A + pSmTq2 F	attaGGTCTCaGCTGgagatgtgaagagaaaccctgggtccaATGGTATCGAAAGGTGAGGAG
pSmTq2 R	cgcgGGTCTCgCTATTTTGTACAATTCGTCCATCC
(stop)ScCYC1tt R	acGGTCTCgATAGgctagTCATGTAATTAGTTATGTCAC
PRIMERS FOR BRINGING CBM CODING SEQUENCES INTO THE MOCLO SYSTEM	
AtEXP10SP F	cgctGAAGACacAATGATGGGCATCTTGGGTTCTTAG
SPalpha F	cgctGAAGACacAATGATGAGATTTCTTCAATTTTTAC
pSmTq2 (stop) R	acgaGAAGACcgAAGCttaTTTGTACAATTCGTCCATC

Appendix 2 Primers used for genotyping and sequence verification after cloning. Primers are clustered by subprojects they were used for. Primer names for the domain swap project relate to the four domains established in Figure 6. F and R specify forward or reverse primer orientation.

GENOTYPING AND SEQUENCING	
PRIMERS	5' to 3' sequence
CSLA SWAP PROJECT	
AtCSLA2-1 F	GAGGATGGAGATCACAGGCCAA
AtCSLA2-2 R	TGAAGGTCACCGAGGTAGAGAA
AtCSLA2-3 F	TGGACCTGCAAATCTCTTTAGGA
AtCSLA2-4 R	CCAGCCCACTGATGAAGAAA
AkCSLA3-1 F	CCTCTACATTTGGGGTCAGG
AkCSLA3-2 R	TCCCACGTAGACAAACTTCCAACC
AkCSLA3-3 F	CGTAGCTCATTTTCGTCACCT
AkCSLA3-4 R	AGGTGCCAACATAACCAAGTCC
TmCSLC4 R	GCCTTGATTCTCCAAACACC
TmCSLC4 F	ACTTTCTGCTCAGGGTGTCC
TmXXT2 R	CATTGTGACCTCTGCTCGTC
M13 F	TGTAAAACGACGGCCAGT
M13 R	CAGGAAACAGCTATGAC
CBM PROJECT	
TpolCBM16-1	CACGAATCCAAACCATCCTC
TmCBM27	CAACAACCTCCGATGTGCAAC
CjCBM35	AAGCAGTACCGTCCATGACC
PaCBM35	GTTTTGACCTGCATTGAGCA
pmScarlet-I	TCCCACCTTGAAACCTTCTGG
pSmTq2	TACACCCCAGGAAAGAGTGG
mRuby2	CAAATGACCTCCACCATCAA
pFDH1	TTCAGTGTGCTGACCTACACG
ScCYC1tt	CTTTTCGGTTAGAGCGGATG
pAGM8031 F	TAGGCACCCCAGGCTTTAC
pAGM8031 R	GGCACATACAAATGGACGAAC

NEXT PAGE:

Appendix 3 Example of pipeline used to recombine TpolCBM16-1. From top to bottom: Coding sequences (CDS) are prepared for introduction into GoldenPiCs BB1 level vectors with standard fusion sites at the outer edges (not shown) and customized fusion sites (Fs) between elements when more than one CDS is introduced into a vector. Note that the 2A signal peptide for self-cleaving of pSmTq2 is introduced via primers with custom fusion sites that have the first or the second half of it respectively. Elements from the BB1 Level are recombined into the GoldenPiCs Level 3 vector BB3rN_14, which carries a Nourseothricin resistance. BB3 vectors are linearized and transformed into *Pichia*. For plant expression, the α -TpolCBM16-1-pmScarlet-I-2A-pSmTq2 sequence is extracted from the BB3 vector via Phusion PCR using Phusion primers that introduce the appropriate fusion sites (FsB, FsC) for entry into the MoClo Level 0 vector pAGM4031. CDS is assembled into MoClo Level 1 together with the p35S promoter and the tOCS terminator. These constructs are then transiently expressed in *N. benthamiana*. For stable expression in *Arabidopsis*, the MoClo Level 1 construct is reassembled into a the MoClo level M vector pAGM8031 together with an RFP selection CDS, which induces red fluorescence in transformed seeds, and an end linker. Level M constructs were then expressed in *Arabidopsis*.



Appendix 4 Dilutions and total volumes used for plate reading of OD600 and fluorescence as described in 2.2.1. Table specifies the dilutions and total volumes for each figure. When supernatant was measured with a different dilution than the total culture, the supernatant dilutions are shown in parenthesis.

FIGURE	μL CULTURE	μL MILLIQ WATER
FIGURE 11, FIGURE 12 C/D, FIGURE 16, FIGURE 17, FIGURE 20, FIGURE 24	50	150
FIGURE 12 A/B	200	0
FIGURE 13	100	100
FIGURE 14	50 (100)	150 (100)
FIGURE 18, FIGURE 30, FIGURE 31	10	240
FIGURE 26, FIGURE 28	10	290

Appendix 5 Segregation ratios of transgenic *Arabidopsis* lines selected for microscopical analysis. For each genotype, seeds of 7-13 T2 lines were screened with respect to their segregation ratios. The three lines (except only two for α-pmScarlet-I) with the segregation ratios closest to 3:1 were selected for T3 growth and seedling root imaging via confocal microscopy.

construct	background	line	red seeds	no FL seeds	segregation ratio
<i>pmScarlet-I</i>	Col-0	1	45	15	3.0
		2	45	14	3.2
		3	45	15	3.0
<i>α-pmScarlet-I</i>		1	45	18	2.5
		2	45	15	3.0
		<i>α-TpoICBM16-1-Sc-2A-Tq2</i>	1	45	15
2			45	13	3.5
3			43	17	2.5
<i>pmScarlet-I</i>		csla2,3,9	1	45	16
	2		45	13	3.5
	3		45	13	3.5
<i>α-pmScarlet-I</i>	1		45	14	3.2
	2		45	16	2.8
	3		45	14	3.2
<i>α-TpoICBM16-1-Sc-2A-Tq2</i>	1		45	13	3.5
	2		45	14	3.2
	3		45	17	2.6

CURRICULUM VITAE

Personal Details

Name: Madalen Katharina Aude Robert

Education and Research

02/2020 - 05/2023 **Doctoral Student**

Leibniz Institute of Plant Biochemistry, Halle

Independent Junior Research Groups

Department: Designer Glycans

Supervisor: Cătălin Voiniciuc

01/10-23/12/2022 **DAAD fellowship**

for a research-exchange to the University of Florida, Gainesville (USA)

04/2017-09/2019 **M.Sc. Molecular Biosystems**

Otto von Guericke Universität, Magdeburg

Master Thesis: *Packing and Intermolecular Interactions in Small Organic Molecular Crystals and Natural Compounds*

(Max-Planck-Institute for Dynamics of Complex Technical Systems, Magdeburg)

09/2011-08/2016 **B.Sc. Bionik**

Westfälische Hochschule, Bocholt

Bachelor Thesis: *Analysis of the Deformation Behaviour of Planet Carriers to Create Analytic Formulas for their Torsional Stiffness*

(Siemens AG Division Process Industries and Drives, Bocholt)