

1 *Engineering and characterization of carbohydrate-binding modules to enable*
2 *real-time imaging of cellulose fibrils biosynthesis in plant protoplasts*

3 Dharanidaran Jayachandran¹, Peter Smith², Mohammad Irfan¹, Junhong Sun⁴, John M. Yarborough²,
4 Yannick J. Bomble², Eric Lam³, Shishir P.S. Chundawat^{1*}

5 ¹Department of Chemical and Biochemical Engineering, Rutgers-The State University of New Jersey,
6 Piscataway, NJ 08854, USA

7 ²Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado 80401, USA

8 ³Department of Plant Biology, Rutgers-The State University of New Jersey, New Brunswick, NJ 08901,
9 USA

10 ***Corresponding Author:** Shishir P. S. Chundawat (shishir.chundawat@rutgers.edu)

11 **ORCID:** Dharanidaran Jayachandran (<https://orcid.org/0000-0002-8672-2631>)

12 Peter Smith (<https://orcid.org/0000-0002-9276-5099>)

13 Mohammad Irfan (<https://orcid.org/0000-0002-7877-5530>)

14 Yannick J. Bomble (<https://orcid.org/0000-0001-7624-8000>)

15 Eric Lam (<https://orcid.org/0000-0001-8462-9794>)

16 Shishir P.S. Chundawat (<https://orcid.org/0000-0003-3677-6735>)

17 Summary

- 18 • Carbohydrate binding modules (CBMs) are non-catalytic domains associated with cell wall
19 degrading carbohydrate-active enzymes (CAZymes) that are often present in nature tethered to
20 distinct catalytic domains (CD). Fluorescently labeled CBMs have been also used to visualize the
21 presence of specific polysaccharides present in the cell wall of plant cells and tissues.
- 22 • Previous studies have provided a qualitative analysis of CBM-polysaccharide interactions, with
23 limited characterization of optimal CBM designs for recognizing specific plant cell wall glycans.
24 Furthermore, CBMs also have not been used to study cell wall regeneration in plant protoplasts.
- 25 • Here, we examine the dynamic interactions of engineered type-A CBMs (from families 3a and 64)
26 with crystalline cellulose-I and phosphoric acid swollen cellulose (PASC). We generated tandem
27 CBM designs to determine their binding parameters and reversibility towards cellulose-I using
28 equilibrium binding assays. Kinetic parameters - adsorption (k_{on}) and desorption (k_{off}) rate
29 constants- for CBMs towards nanocrystalline cellulose were determined using quartz crystal
30 microbalance with dissipation (QCM-D). Our results indicate that tandem CBM3a exhibits a five-
31 fold increased adsorption rate to cellulose compared to single CBM3a, making tandem CBM3a
32 suitable for live-cell imaging applications. We next used engineered CBMs to visualize *Arabidopsis*
33 *thaliana* protoplasts with regenerated cell walls using wide-field fluorescence and confocal laser
34 scanning microscopy (CLSM).
- 35 • In summary, tandem CBMs offer a novel polysaccharide labeling probe for real-time visualization
36 of growing cellulose chains in living *Arabidopsis* protoplasts.

37 **Keywords:** Arabidopsis plant protoplasts, Carbohydrate-binding module, Cellulose microfibrils, Cell wall
38 biosynthesis, Confocal laser scanning microscopy, Live-cell imaging, Quartz crystal microbalance with
39 dissipation

40 Introduction

41 Plant cell walls are structurally complex, metabolically dynamic, and extremely rich polysaccharide
42 repositories. These plant cell wall polysaccharides mainly comprise cellulose, hemicelluloses, and pectin,
43 forming diverse networks and extensively interacting with each other (Knox, 2008). Cellulose microfibrils
44 form the major component of the plant cell wall and are intertwined with xyloglucans and other pectic
45 polysaccharides. To access the different polysaccharides, plant cell wall hydrolases contain highly specific,
46 non-catalytic carbohydrate-binding modules (CBMs) along with catalytic domains (CD). These CBMs
47 increase the proximity between CDs and polysaccharides, resulting in efficient enzymatic hydrolysis
48 (Talamantes *et al.*, 2016). Different CBMs recognize different polysaccharides based on their amino acid
49 sequences and the topology of the binding site. According to the Carbohydrate-Active enZYmes (CAZy)
50 database (<http://www.cazy.org>), there are 94 sequence-based families of CBMs, many of which bind to
51 cell wall polymers (Cantarel *et al.*, 2009). For instance, type-A CBM recognizes crystalline polysaccharides

52 such as cellulose, chitin, and mannan. Likewise, type-B CBMs target individual glucan chains, and type-C
53 CBMs bind specifically to small sugars (mono- or disaccharides) (Boraston *et al.*, 2004). These CBMs are
54 highly specific, and their specificity depends on the target substrate of its accompanying CD. However,
55 some cellulose-binding CBMs have been reported to be components of enzymes that hydrolyze xylans,
56 mannans, and pectins, other than cellulases (Kellett *et al.*, 1990; McKie *et al.*, 2001).

57 In nature, non-catalytic CBMs tend to coexist with other non-catalytic CBMs besides CDs, forming tandem
58 repeats of CBMs. Such tandem CBMs increase the cell wall hydrolases' overall efficiency by enhancing
59 affinity through prolonged contact with the target substrate (Hashimoto, 2006; Guillén *et al.*, 2010; Møller
60 *et al.*, 2021). For example, a type-A tandem CBM with three CBM10 domains associated with mannanase
61 showed improved binding and spatial flexibility (Møller *et al.*, 2021). Similarly, type-B CBM tandems
62 (CBM17 and CBM28) from *Bacillus* sp. 1139 Cel5 and two family 4 CBMs (CBM4–1 and CBM4–2) from
63 *Cellulomonas* sp. exhibited very tight non-crystalline cellulose binding. Taken individually, these three
64 CBMs recognized different regions of non-crystalline cellulose. (Boraston *et al.*, 2003; Kognole & Payne,
65 2018). Aside from hydrolysis of celluloses, only a few of these tandem CBMs have been engineered to
66 visualize multiple polysaccharides associated with plant cell walls. For example, Herve *et al.* found that the
67 CBM3a-CBM2b-1-2 tandem constructs bound tightly to cell wall of stem sections where cellulose and xylan
68 were cross-linked closely (Hervé *et al.*, 2010). No reports indicate the use of tandem CBMs in imaging
69 single plant protoplasts. Living plant protoplasts provide a distinct advantage as a cell-based system that
70 can be readily used for performing genomics, transcriptomics, proteomics, metabolomics, and epigenetic
71 analyses (Xu *et al.*, 2022). This versatile system could be used to characterize cell-wall regeneration and
72 the polysaccharides associated with it. Additionally, regenerating plant cell walls in *Arabidopsis* contain
73 crystalline and amorphous regions at the surface (Ruel *et al.*, 2012). It is imperative that the CBMs' ability
74 to bind to various forms of cellulose is also characterized critically.

75 Extensive research has been conducted on CBMs attached to glucanases using bulk biochemical assays
76 (Chundawat *et al.*, 2021; Nemmaru *et al.*, 2021), single-molecule cellulase motility assays (Brady *et al.*,
77 2015), kinetic modeling (Levine *et al.*, 2010; Shang *et al.*, 2013), and molecular simulations (Beckham *et al.*,
78 2014; Vermaas *et al.*, 2019). However, most of these assays have been conducted on single CBMs
79 attached to fluorescent proteins or glucanases. Consequently, the biochemical behavior of most tandem
80 CBMs under the assays mentioned above has not been studied despite the prevalence of tandem CBMs
81 in nature (Boraston *et al.*, 2004; McCartney *et al.*, 2006; Hervé *et al.*, 2010). Ultimately, the potential of
82 tandem CBMs in the visualization and characterization of cell wall polymers remains mostly unexplored.

83 Here, using tandem modular constructs, we computed different binding parameters for tandem versus
84 single CBMs. We picked two model type-A CBMs (CBM3a and CBM64) and systematically engineered
85 tandem versions with and without green fluorescent protein (GFP). Using conventional pull-down assays,
86 we calculated the number of binding sites, binding constants, and partition coefficient towards cellulose-I
87 and PASC. Using quartz crystal microbalance with dissipation (QCM - D), we also measured the adsorption

88 and desorption rate constants (k_{on} and k_{off}) for CBM binding towards nanocrystalline cellulose-I. Finally, we
89 used the CBMs generated to visualize the regenerated cell walls in *Arabidopsis thaliana* plant protoplasts
90 to fully understand their potential in real-time visualization of plant cell wall growth using confocal laser
91 scanning microscopy (CLSM) and wide-field fluorescence microscopy.

92 **Materials and Methods**

93 ***Bacterial strains, CBM plasmids, and reagents***

94 Avicel cellulose-I was obtained from Sigma Aldrich under the label Avicel PH-101. Phusion Master Mix was
95 obtained from Thermo Fisher Scientific. Dpn1 and T4 DNA polymerase was obtained from New England
96 Biolabs. Nickel-charged magnetic beads and magnetic racks were purchased from GenScript (NJ).
97 Chemically competent cells were procured from various vendors: *E. coli* 10g cells from Lucigen (Madison,
98 WI) and *E. coli* BL21-CodonPlus-RIPL [λ DE3] from Stratagene (Santa Clara, CA). The pEC-GFP-CBM3a
99 vector was kindly provided by Dr. Brian Fox (University of Wisconsin, Madison, USA) and was used as
100 plasmid backbone to engineer all CBM constructs (Whitehead *et al.*, 2017; Chundawat *et al.*, 2021). The
101 primers were obtained from Integrated DNA Technologies, and DNA sequencing was performed by Azenta
102 Life Sciences (NJ). All other reagents were purchased from VWR, Thermo Fisher Scientific, and Sigma
103 Aldrich unless mentioned otherwise.

104 ***Synthesis and cloning of wild-type CBM genes into pEC-GFP vector***

105 *E. coli* expression vector pEC-GFP-CBM3a was kindly provided by the Fox lab (UW Madison) (Whitehead
106 *et al.*, 2017), which was used as the plasmid backbone to insert the CBM64 gene from *Spirochaeta*
107 *thermophila*. Cloning and insertion of CBM64 are detailed in our previous work (Nemmaru *et al.*, 2021).
108 Plasmid maps for pEC-GFP-CBM3a and pEC-GFP-CBM64 are outlined in Figure S1. Tandem CBMs (pEC-
109 GFP-CBM3a-CBM3a, pEC-CBM3a-CBM3a, pEC-GFP-CBM64-CBM64, and pEC-CBM64-CBM64) were
110 prepared from the original plasmids (pEC-GFP-CBM3a and pEC-GFP-CBM64) using Sequence and
111 Ligation-Independent Cloning (SLIC) protocol (Stevenson *et al.*, 2013). The schematic representation of all
112 the CBM constructs is outlined in Figure S2. The SLIC primers used in this study are tabulated in Table S1.
113 The sequences for all CBM constructs generated are summarized in Supplementary Text S1. Briefly, the
114 gene fragment containing the linker and CBM was amplified using polymerase chain reaction (PCR) from
115 the original plasmid available at the Chundawat lab. It was inserted after the 'GFP-CBM' region in the
116 original 'pEC-GFP-CBM' plasmids to create the 'tandem CBM' (also referred to as 'CBM-CBM') carrying
117 plasmids. The insert and vector PCR products were purified to remove unreacted nucleotides
118 (deoxynucleoside triphosphates or dNTPs), mixed at optimal ratios, and then transformed into chemically
119 competent *E. coli* 10G cells to get colonies for screening. The tandem plasmids developed were later used
120 to create non-GFP versions by removing the GFP from those plasmids. All the pEC-GFP-CBM, pEC-GFP-
121 CBM-CBM, and pEC-CBM-CBM plasmids were verified using Sanger sequencing, and the sequence-

122 verified plasmids were stored at -80°C for long-term storage. *E.coloni* 10G cells carrying the plasmid of
123 interest were stored as 15% glycerol stocks and maintained at -80°C.

124 ***Production and purification of recombinant His-tagged CBMs***

125 After sequence verification, all CBM constructs were transformed into *E.coli* BL21-CodonPlus-RIPL [λ DE3].
126 Glycerol stocks were prepared for the transformed strains and were stored at -80°C. These glycerol stocks
127 were later used to inoculate 10 mL Luria Bertani (LB) media in culture tubes containing kanamycin at a 50
128 μ g/ml concentration. This overnight starter culture was used to inoculate 300 mL LB media in 1 L shake
129 flasks containing kanamycin at a 50 μ g/mL concentration. The flasks were incubated at 37°C and 200 rpm
130 until the growth reached the exponential phase (OD_{600} ~0.6-0.8). Protein expression was then induced
131 using 0.5 mM IPTG at 25°C for 16 hours. Cells were harvested at 7000x g for 15 mins. 3g of cell pellet was
132 resuspended in 15 mL cell lysis buffer (20 mM phosphate buffer, 500 mM NaCl, 20% (v/v) glycerol, pH 7.4).
133 For every 3g of wet cell pellet, 200 μ L protease inhibitor cocktail (1 μ M E-64 (Sigma Aldrich E3132)) and
134 15 μ L lysozyme (Sigma Aldrich, USA) was added. Cells were lysed on ice using a Misonix™ sonicator 3000
135 for 5 mins of total process time at 4.5 output level and pulse settings (pulse-on time: 10 secs and pulse-off
136 time: 30 secs) to avoid sample overheating. The cell lysate carrying the protein of interest was separated
137 from the cell debris by centrifuging at 48,400x g for 45 mins at 4°C. The cell lysate was clarified using a
138 0.22 μ m non-sterile syringe filter after centrifugation. Since all the expressed CBMs contained an N-terminal
139 8X-HIS tag, the proteins were purified using IMAC (immobilized metal affinity chromatography). The
140 clarified cell lysate was incubated with 2 ml pre-equilibrated Ni²⁺ charged magnetic resin purchased from
141 GenScript. Briefly, the resin was equilibrated with five column volumes of buffer A (100 mM MOPS, 500
142 mM NaCl, 10 mM Imidazole, pH 7.4). The equilibrated resin was incubated with the clarified lysate at 4°C
143 for 120 mins with gentle agitation. The lysate was removed from the magnetic resin using a magnetic rack.
144 The resin was later washed with five-column volumes of buffer A twice, followed by five-column volumes of
145 (buffer A: buffer B = 95:5). Finally, the proteins were eluted out using five-column volumes of buffer B (100
146 mM MOPS, 500 mM NaCl, 500 mM Imidazole, pH 7.4). The purified proteins were later desalted into 10
147 mM MES buffer, pH 6.5. Proteins were also characterized for molecular weights and purity using SDS-
148 PAGE (Figure S3). The concentration of CBM proteins was estimated by measuring the absorbance at 280
149 nm. The molecular weight and extinction coefficient of all the CBM proteins used in this study are
150 summarized in Table S2. Proteins were later aliquoted, flash-frozen under liquid nitrogen, and stored at -
151 80°C for further assays.

152 ***Generation of PASC from Avicel Cellulose-I***

153 Phosphoric acid swollen cellulose (PASC) was prepared according to the protocol detailed by Zhang et al.
154 (Y.-H. Percival Zhang & Lee R. Lynd, 2006). Briefly, 0.6 ml of distilled water was added to 0.2 g of Avicel
155 cellulose-I (PH-101) to form a wet-cellulose slurry. To this slurry, 10 mL of 86.2% ice-cold phosphoric acid
156 was added with vigorous stirring. The cellulose mixture turned transparent, after which 40 ml of ice-cold

157 distilled water was added at a rate of 10 mL per min. The resulting white cloudy precipitate was removed
158 by centrifugation at 5000x g for 20 mins at 4°C. This step was repeated four more times to remove the
159 phosphoric acid. In addition, 0.5 mL of 2M sodium carbonate was added to neutralize the mixture. Finally,
160 the mixture was washed in ice-cold DI water until the pH was between 5 and 7. The regenerated PASC
161 was stored at 4°C for further binding assays.

162 ***GFP-CBM pull-down binding assays with crystalline cellulose-I and PASC***

163 Binding assays were performed as discussed in previous papers from our group (Chundawat *et al.*, 2021;
164 Nemmaru *et al.*, 2021). All binding assays were performed with at least six replicates in 300 μ L 96-well
165 round-bottomed polypropylene plates (USA Scientific). Each of these replicates was a 200 μ L reaction
166 mixture comprising appropriate volumes of CBM dilution to reach effective concentrations of 0–600 μ g/mL
167 for obtaining the full binding isotherm. In addition to the CBMs, the mixture comprised 2.5 mg of Avicel
168 cellulose-I, an effective BSA concentration of 2.5 mg/mL to prevent non-specific protein binding, and an
169 effective buffer concentration of 10 mM MES (pH 6.5). To account for protein loss due to denaturation or
170 non-specific binding to the microwells, control reactions without cellulose were also included to obtain total
171 protein concentration. The microplate was then sealed with a 96-well plate mat and incubated inside a USA
172 Scientific hybridization oven at 5 rpm for 60 mins at room temperature with end-over-end mixing. Never-
173 shaken control reactions were also prepared, similar to previously shaken control reactions. This control
174 was used to obtain the calibration curve relating GFP fluorescence and known protein concentration. After
175 incubating the plate for an hour, the microplates were centrifuged at 2000 rpm for 2 mins using an
176 Eppendorf™ 5810R centrifuge to separate cellulose from soluble supernatant. Finally, 100 μ L of soluble
177 supernatant was picked up from each microwell using a multi-channel micropipette and transferred to black
178 opaque microplates for measuring fluorescence at 480 nm excitation, 512 nm emission with 495 nm cut-off
179 using Molecular Devices™ UV spectrophotometer. Schematic representation of the biochemical assay
180 workflow is shown in Figure 1a.

181 ***GFP-CBM pull-down binding assay analysis***

182 Preliminary data analysis was performed using Microsoft Excel™ to obtain free protein (μ M) and bound
183 protein concentrations (μ mol/g cellulose). The data was fit to Langmuir one-site model using the non-linear
184 curve fitting tool in Origin for full-scale binding assays. Curve fitting was done using the Levenberg-
185 Marquardt algorithm with a tolerance of 1e-9.

186 ***Reversibility testing for single and tandem CBM binding to cellulose substrates***

187 A binding reversibility study was performed for all single and tandem CBMs after the completion of the
188 binding assays. After collecting the supernatant for measuring fluorescence in binding assays, 100 μ L of
189 reconstitution mixture (2.5 mg/ml BSA + 10 mM MES (pH 6.5)) was added to the remaining original reaction
190 mixtures and shaken controls. The microplate was sealed with a 96-well plate mat and incubated at 5 rpm

191 for 60 mins at room temperature. After incubation, the supernatant was collected, and fluorescence was
192 measured using the spectrophotometer, as mentioned above under the 'pull-down binding assays' section.

193 ***Preparation of nanocrystalline cellulose through acid hydrolysis for QCM-D analysis***

194 Nanocrystalline cellulose was prepared from Avicel cellulose-I using the procedure mentioned previously
195 (Nemmaru *et al.*, 2021). Briefly, 2 g of Avicel cellulose-I was added to 70 mL 4N hydrochloric acid (HCl) in
196 a glass beaker placed over a water bath maintained at 80°C. The slurry was stirred every 30 mins using a
197 spatula to ensure uniform suspension. After 4 hours, 50 mL of DI water was added to dilute the acid
198 hydrolysis mixture. The slurry was split across 50 mL tubes and centrifuged at 1600x g for 10 mins. The
199 supernatant was discarded, and the pellet was resuspended in 10 mL DI water. This wash step was
200 repeated multiple times until the solution turned hazy and the pH rose to around pH 3.3. The haziness of
201 the supernatant indicated the development of cellulose nanocrystals, and these supernatants were
202 collected for future usage.

203 ***Preparation of cellulose thin films for QCM-D***

204 QCM-D sensors (4.95 MHz quartz crystals, Biolin Scientific QSX-301) were prepared in a manner as
205 described previously (Nemmaru *et al.*, 2021). Briefly, quartz crystals were washed thoroughly with
206 deionized water, rinsed with ethanol, and dried with nitrogen gas. The sensors were then placed in a rack
207 and submerged in a solution of 0.02% poly(diallyl dimethyl ammonium chloride), from Sigma Aldrich, for at
208 least 60 mins with orbital mixing (150 rpm) at room temperature. The sensors were then blown-dry once
209 again, followed by spin-coating with 225 μ L of cellulose nanocrystal slurry using a Chemat Technology KW-
210 4A spin coater with a pre-cycle spin for 3 secs at 1500 rpm, followed by a spin cycle for 60 secs at 3000
211 rpm. This spin coating step is repeated 4-8 times to obtain a uniform cellulose film thickness of ~20-40 nm
212 (as measured using the QSoft software using the Sauer brey model and an assumed density of 1191
213 kg/m^3). Note that the number of necessary cellulose spin coating steps is dependent on the concentration
214 of cellulose nanocrystals in the prepared slurry and may need to be optimized on a case-to-case basis.

215 ***Quartz crystal microbalance with dissipation (QCM-D) based CBM-cellulose binding assay***

216 Binding and unbinding assays were performed on a QSense E4 instrument (Nanoscience Instruments).
217 The quartz sensors with a deposited cellulose film were loaded into the instrument and allowed to
218 equilibrate overnight in 10 mM MES buffer pH 6.5. Following the overnight incubation, the frequency and
219 resonance changes associated with harmonics 1, 3, 5, 7, and 9 were tracked and monitored for stability for
220 at least 5 mins in the MES buffer prior to loading CBMs. All CBM proteins were diluted to a concentration
221 of 5 μ M and passed over the sensors at a flow rate of 100 μ L/min for at least 10 mins or until saturation
222 was observed. Unbinding of proteins was tracked by flowing 10 mM MES buffer pH 6.5 at 100 μ L/min for
223 at least 20 mins. The sensors and Qsense chambers were then rinsed with 5% Contrad solution followed
224 by deionized water to remove any traces of residual protein. The frequency and dissipation traces were

225 analyzed using an in-house data analysis routine based on binding and unbinding equations derived
226 previously (Nemmaru *et al.*, 2021).

227 **Analysis of QCM-D-based binding assay data to obtain kinetic parameters**

228 Binding curves of QCM-D data were analyzed using RStudio based on equations derived previously
229 (Nemmaru *et al.*, 2021). Data corresponding to the frequency changes of the third harmonic of each
230 experiment was transformed with the Sauerbrey equation to produce a change in mass ($\text{ng}\cdot\text{cm}^{-1}$) associated
231 with CBM binding and un-binding. The area of the sensor gaining mass was found to have a radius of 0.5
232 cm, and this value was used to determine the total ng of protein binding/unbinding to the cellulose surface.

233 It was noted that not all the mass that is added to the cellulose sensor is removed during the unbinding
234 experiments, and that some CBM appears to remain bound even after extensive washing. We postulate
235 that this may be due either to a difference in the nature and interaction of CBMs with the nanocellulose
236 coating the surface compared to Avicel-based cellulosic materials in other pull-down experiments, or
237 perhaps due to incomplete coating of the sensor surface with nanocellulose, allowing some CBM to bind
238 irreversibly to the underlying poly(diallyl dimethylammonium chloride) layer. To accommodate this finding
239 and provide a more suitable fit for the unbinding equation, the equation used to fit the unbinding data
240 included a term ($\text{CBM}_{\text{remain}}$) which accounted for the irreversible CBM binding and represented the number
241 of CBM molecules still bound to the sensor surface after washing. The inclusion of this term produced the
242 following equation, which was used to fit the data.

$$243 \quad [EC] = A * (e^{-(k_{off} * t)} + \frac{\text{CBM}_{\text{remain}}}{A})$$

244 [EC] - Number of binding sites occupied by protein (μmol)

245 A - Number of available binding sites on nanocellulose film

246 k_{off} - Dissociation rate constant (min^{-1})

247 t - time (min)

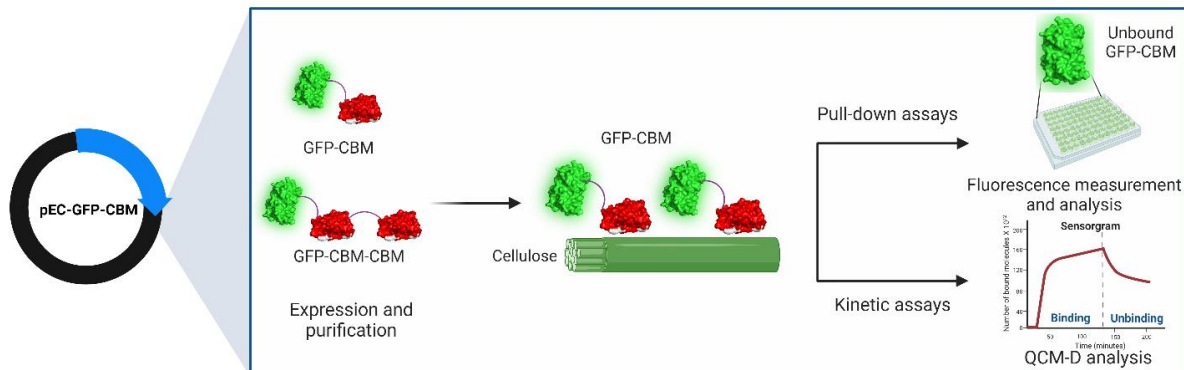
248 $\text{CBM}_{\text{remain}}$ - Number of CBM molecules that remain bound after washing

249 **Preparation of plant samples for imaging plant protoplasts**

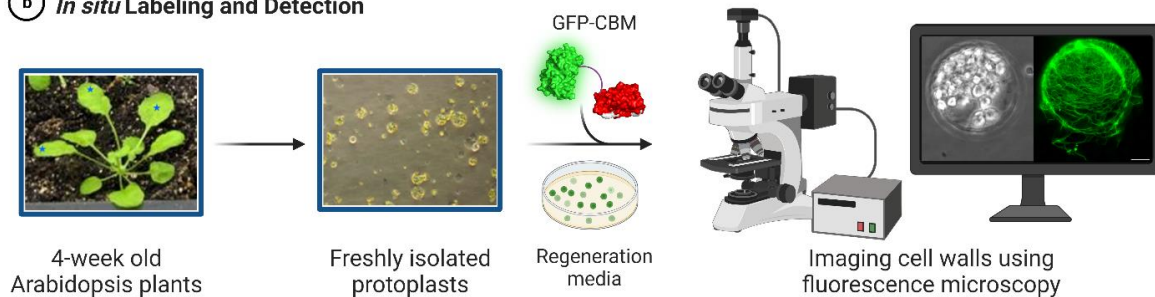
250 *Arabidopsis thaliana* Col-0 (Columbia) was used as plant material for all experiments. *Arabidopsis thaliana*
251 mesophyll protoplasts were isolated as described in detail by Yoo *et al.* (Yoo *et al.*, 2007). Briefly, 3–4-week-
252 old, uniformly sized leaves were selected to obtain consistently sized protoplasts (Figure 1b). Thin leaf
253 strips of 0.5-1 mm thickness were cut off from the middle portion of the leaves. These leaf strips were
254 immediately transferred to an enzyme solution (0.4 M Mannitol, 20 mM KCl, 20 mM MES, 1.5% Cellulase

255 R10 (Yakult, Japan), 0.4% Macerozyme R10 (Yakult, Japan), 10 mM CaCl₂, 5 mM 2-mercaptoethanol,
256 0.1% BSA). The leaf strips were vacuum infiltrated to perfuse it with the enzyme solution and then incubated
257 in the dark for 3-5 hours to digest the cell wall completely. The protoplasts were then diluted with an equal
258 volume of sterile W5 solution (154 mM NaCl, 125 mM CaCl₂, 5 mM KCl, and 2 mM MES in ultrapure water).
259 The undigested leaf material was filtered out using a pre-washed 75 μm nylon mesh. Protoplasts were then
260 collected by centrifugation at 1000 rpm for 3 mins at 4°C. The residual enzymatic solution was removed by
261 washing the protoplast cells with 10 mL of W5 solution twice. After washing, the protoplasts were
262 resuspended in 0.5 mL of W5 solution. The quality of the protoplasts was observed under a simple light
263 microscope (Figure 1b). The concentration of protoplasts was measured using a hemocytometer and then
264 adjusted to 2 × 10⁵ protoplasts mL⁻¹ of W5 solution. 200 μL of protoplasts at 2 × 10⁵ concentration was
265 incubated in cell wall regeneration media containing an equal volume of WI solution (0.5 M Mannitol, 4 mM
266 MES, 20 mM KCl) and 2M2 media (Gamborg's B-5 basal medium with minimal organics 6.4g/L (Sigma),
267 0.8M Trehalose, 0.1 M Glucose, 2 μM 3-naphthalene acetic acid (NAA) pH 5.7). The protoplasts were
268 incubated for 17 hours at room temperature under a Philips hue lamp. After incubation, the regenerated
269 protoplasts were isolated by spinning at 300x g for 3 mins at 4°C. The isolated protoplasts were then fixed
270 by incubation on ice for 10 mins in 200 μL of 1% glutaraldehyde. The protoplasts were then washed with
271 250 μL of 12% sorbitol twice. Samples were resuspended in 50 μL of 12% sorbitol and stored on ice until
272 use.

a Biochemical Assays



b In situ Labeling and Detection



273

274 **Figure 1.** Schematic representation of the workflow in analyzing and utilizing GFP-CBMs to visualize
275 regenerated plant cell walls. (a) pEC-GFP-CBM plasmid used for expressing GFP-CBM3a (*Clostridium*
276 *thermocellum*) and GFP-CBM3a-CBM3a proteins. Purified GFP-CBMs were used for estimating
277 biochemical and kinetic parameters using pull-down assays and QCM-D analysis. (b) Medium-sized leaves
278 from 4-week-old Arabidopsis plants (marked with blue stars) were used to isolate similarly sized protoplasts.
279 Isolated protoplasts were incubated in the regeneration media along with GFP-CBMs. On the right,
280 regenerated protoplasts with cell walls labeled with GFP-CBM3a as observed under fluorescence
281 microscopy. CBM, carbohydrate-binding module; GFP, green fluorescent protein; QCM-D, Quartz crystal
282 microbalance with dissipation.

283 **Fluorescence and Confocal Laser Scanning Microscopy**

284 Calcofluor white solution (100 μ L of 0.001%) was added to 200 μ L of regenerated protoplasts and incubated
285 at room temperature for 5 mins. Excess calcofluor solution was removed after spinning down the cells at
286 1000 rpm for 2 mins. The stained cells were washed with 12% sorbitol solution twice. 10 μ L of the
287 regenerated protoplast samples were placed on a glass slide that was layered with a coverslip. Calcofluor-
288 stained cell fibers were visualized under the DAPI channel in Olympus FSX 100 microscope. Staining using
289 CBMs was done in a similar manner, where the final concentration of CBMs in the solution was 100 nM.
290 Excess CBMs were removed, and the cells were observed under the GFP channel for GFP-CBMs and the
291 red channel for DSR-CBMs.

292 For CLSM, the samples were prepared as mentioned above, and the glass slides were then placed onto
293 the inverted platform of a Zeiss LSM 710 confocal microscope, and the cells were imaged under the 488-
294 nm channel for GFP CBMs and 561 nm channel for DSR-CBMs. All images were processed with Zeiss
295 imaging software, ImageJ, or Adobe Photoshop.

296 **Results**

297 **Type-A tandem CBMs show reduced binding affinity compared to single CBM counterparts**

298 We prepared two type-A tandem CBMs, CBM3a (*Clostridium thermocellum*) (Lehtiö *et al.*, 2003) and
299 CBM64 (*Spirochaeta thermophila*) (Schiefner *et al.*, 2016; Pires *et al.*, 2017) by fusing the linker-CBM region
300 to the N-terminal region of GFP-CBM as shown in Figure S2. Additionally, we studied the binding and
301 activity of GFP-CBM and GFP-CBM-CBM tandem constructs toward Avicel cellulose-I and PASC,
302 respectively (Figure 2a; Figure S4-5).

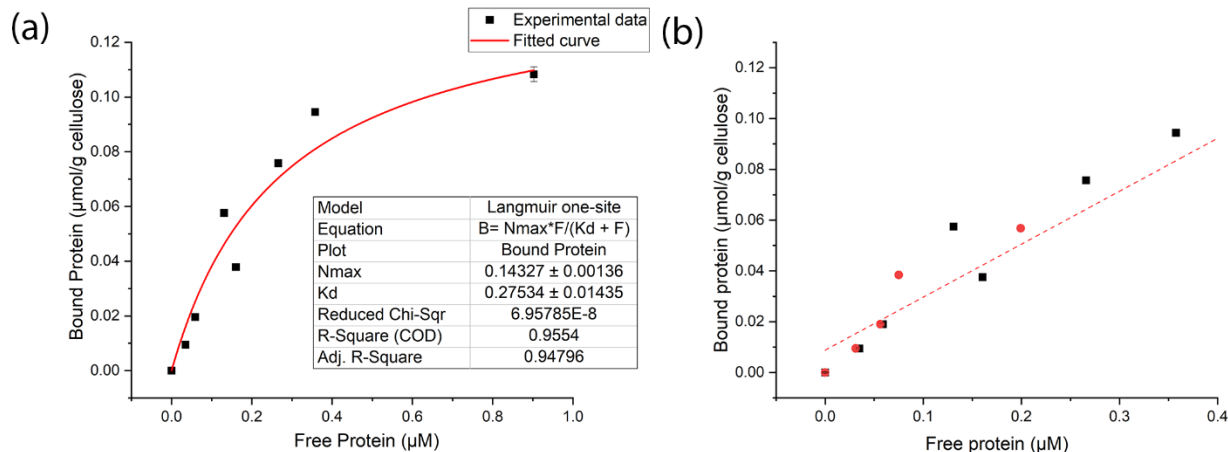
303 First, we performed fluorescence-based pull-down binding assays for GFP-CBM3a and GFP-CBM64 –
304 single and tandem versions against Avicel cellulose-I at varying protein concentrations (ranging between 0
305 and 600 μ g/ml). The resulting data were fit to Langmuir isotherm one-site models to obtain the maximum
306 number of binding sites (N_{max}) on the substrate, binding dissociation constant (K_d), and partition coefficient

307 (see Table 1). The number of available binding sites was the maximum for GFP-CBM3a compared to other
308 CBM versions. Regardless of the substrate (cellulose-I or PASC), GFP-CBM3a always showed a significant
309 similarity in binding affinity (i.e., the inverse of K_d). The binding affinity was also observed to be higher for
310 GFP-CBM3a compared to GFP-CBM3a-CBM3a against cellulose-I. Nevertheless, the binding affinity
311 parameters are comparable between the single and tandem versions of CBM64. The order of affinity
312 towards cellulose-I is GFP-CBM3a~GFP-CBM64~GFP-CBM64-CBM64>GFP-CBM3a-CBM3a>GFP-
313 CBM17.

314 Similarly, we subjected all the CBM constructs to binding assays against PASC. We observed a 7-fold and
315 a 1.25-fold increase in the number of available binding sites (N_{max}) for CBM3a compared to CBM64 towards
316 cellulose-I and PASC, respectively. Similar to cellulose-I, there was no significant difference in the number
317 of binding sites (N_{max}) available for GFP-CBM64, GFP-CBM64-CBM64, and GFP-CBM17. Nevertheless,
318 there was a ~3-fold reduction in N_{max} for the tandem CBM3a compared to the single CBM3a. However, the
319 binding affinity trend towards PASC was found to be similar for single and tandem CBM3a. GFP-CBM17
320 was found to have the maximum binding affinity towards PASC among all other CBMs, consistent with
321 previous studies (McLean *et al.*, 2002).

322 Furthermore, the partition coefficients of single CBM3a against cellulose-I and PASC were nearly 7- and 2-
323 fold higher than the tandem CBM3a. This reduction in partition coefficient for tandem CBMs on PASC is
324 probably due to the significant steric clashes it encounters with the non-native surface of amorphous
325 cellulose. The uneven topology of the hydrophobic binding face probably impairs the accessibility for
326 tandem CBMs (Chundawat *et al.*, 2021; Nemmaru *et al.*, 2021). On the other hand, the GFP-CBM64 has a
327 2- and a 4-fold reduction in partition coefficient compared to tandem CBM64. GFP-CBM17 is highly specific
328 to amorphous cellulose and showed a significant reduction (~5-fold) in binding affinity and an increased
329 partition coefficient towards cellulose-I.

330 The reversibility of tandem CBMs has not been studied in detail in the past, and most of the reported binding
331 reversibility data for single CBMs has been contradictory (Lim *et al.*, 2014; Møller *et al.*, 2021). Here, we
332 showed that the binding of single and tandem GFP-CBM3a is reversible on both cellulose-I and PASC
333 ruling out the possibility of protein structural deformation on the cellulose surface (Figure 2b; Figure S6).
334 However, GFP-CBM64 is irreversible on both cellulose-I and PASC, similar to our previous studies
335 (Nemmaru *et al.*, 2021). Tandem CBM64 is irreversible only towards PASC and not towards cellulose-I
336 (Figure S7). Conversely, GFP-CBM17 is reversible on both PASC and cellulose-I, respectively (Boraston
337 *et al.*, 2003; Blake *et al.*, 2006).



338
 339 **Figure 2.** (a) Binding curve of GFP-CBM3a towards cellulose-I using Langmuir-type adsorption model, (b)
 340 Binding reversibility of GFP-CBM3a binding towards cellulose-I; Black and red dots represent original and
 341 re-equilibrated data points; Red line represents linear fit joining original data points.

Construct	Cellulose-I			PASC		
	N_{max} ($\mu\text{mol/g}$ cellulose)	K_d (μM)	N_{max}/K_d (L/g cellulose)	N_{max} ($\mu\text{mol/g}$ cellulose)	K_d (μM)	N_{max}/K_d (L/g cellulose)
GFP-CBM3a	0.14±0.00	0.28±0.01	0.51±0.09	0.17±0.00	0.30±0.02	0.57±0.04
GFP-CBM3a-CBM3a	0.09±0.00	1.29±0.26	0.07±0.02	0.06±0.00	0.29±0.00	0.21±0.10
GFP-CBM64	0.02±0.00	0.29±0.03	0.07±0.03	0.13±0.01	2.12±0.53	0.06±0.02
GFP-CBM64-CBM64	0.04±0.00	0.30±0.03	0.13±0.10	0.11±0.00	0.48±0.04	0.23±0.06
GFP-CBM17	0.09±0.00	5.20±0.35	1.71±0.18	0.10±0.02	0.93±0.27	0.11±0.07

342
 343 **Table 1.** Binding parameters N_{max} , K_d for GFP-CBM3a, GFP-CBM3a-CBM3a, GFP-CBM64, GFP-CBM64-
 344 CBM64, and GFP-CBM17 obtained from Langmuir one-site model fitted to full-scale binding assay data.
 345 Errors are standard deviations of the mean obtained from the fitting analysis. Each experiment was
 346 performed using six replicates for each protein concentration.

347 **QCM-D assays show a reduced binding behavior for tandem CBMs**

348 Methods describing cellulose film preparation, QCM-D binding assays, and data analysis are discussed in
 349 detail in the 'Materials and Methods' section. To measure the binding kinetics of single and tandem CBMs
 350 on nanocrystalline cellulose, QCM-D binding assays were prepared (Figure 3a). Briefly, the number of
 351 protein molecules bound to the cellulose film was calculated by converting the frequency data obtained
 352 from the binding and unbinding of proteins (Brunecky *et al.*, 2020). Sauerbrey equation was used to obtain

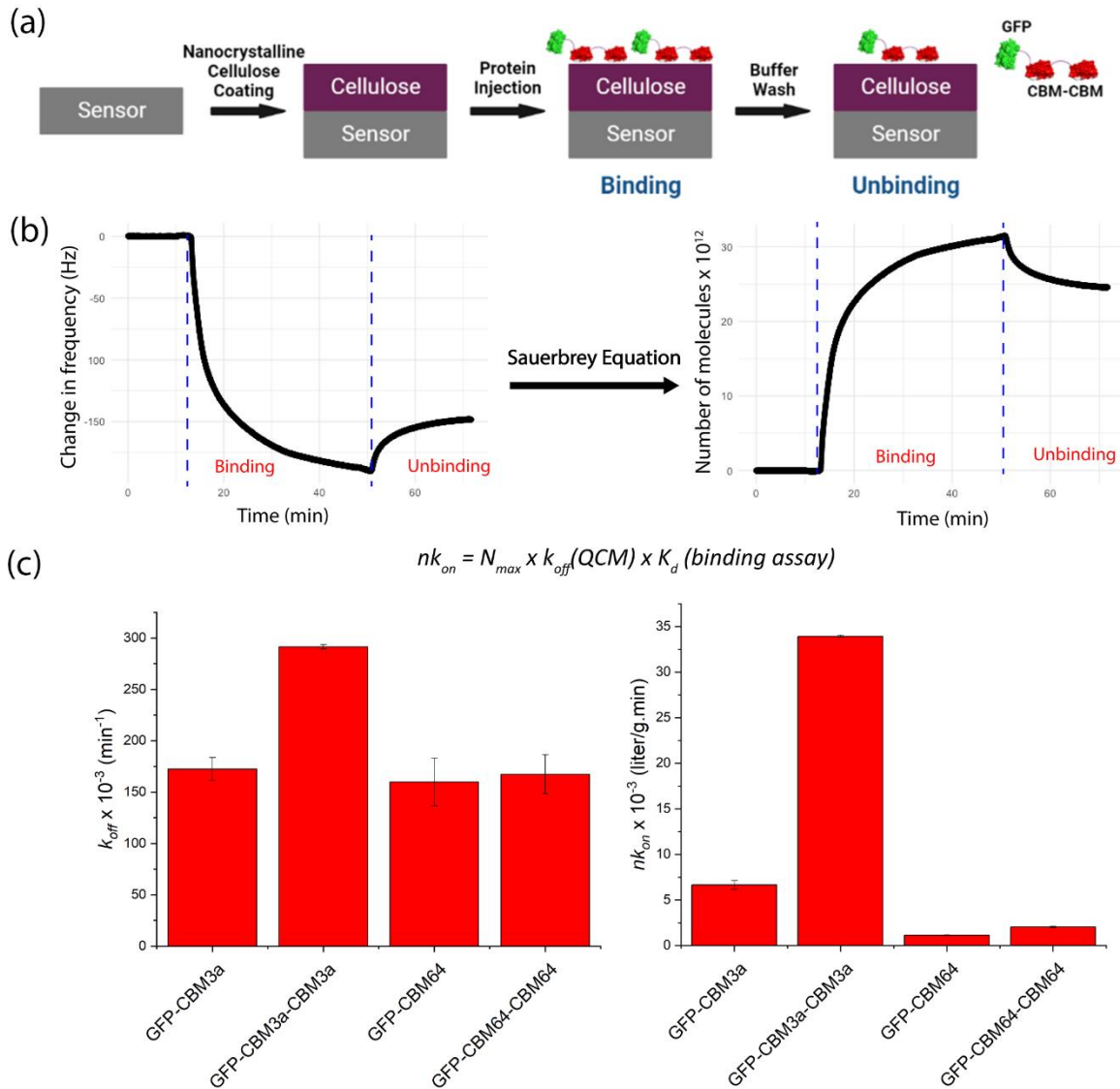
353 the mass of adsorbed protein on cellulose film using the frequency change (Kankare, 2002). In particular,
 354 the unbinding regime was fitted to an exponential decay to obtain the true desorption rate (k_{off}) as shown in
 355 Figure 3b. The effective adsorption rate constant (nk_{on}) was measured using desorption rate k_{off} from the
 356 QCM-D analysis and N_{max} , K_d obtained from the equilibrium binding assay results. These kinetic parameters
 357 were chosen since they were found to be an integral part of the kinetic models that we had reported
 358 previously (Gao *et al.*, 2013; Nemmaru *et al.*, 2021). QCM-D analysis shows a ~1.7-fold increase in k_{off} for
 359 GFP-CBM3a-CBM3a compared to GFP-CBM3a, indicating a reduced binding affinity for tandem CBM3a
 360 compared to the single one. However, the nk_{on} value for tandem CBM3a increased 5-fold compared to the
 361 single CBM3a. Conversely, a similar trend was observed for the k_{off} values between tandem and single
 362 versions of CBM64 towards cellulose-I nanocrystals (Figure 3c). To account for any effects in k_{off} and nk_{on}
 363 values arising from GFP being fused with CBM3a, we also prepared tandem CBM3a and single CBM3a
 364 without GFP. Interestingly, the non-GFP version of CBM3a-CBM3a showed a reduction in the k_{off} value
 365 (nearly 5-fold) compared to the GFP version (Figure S8; Table 2), suggesting that GFP fusion may increase
 366 the off-rate for the tandem CBM3a dimer. On the other hand, CBM3a seemed to have an increased off-rate
 367 when GFP was removed.

368 In summary, GFP-CBM3a showed a reduced nk_{on} compared to GFP-CBM3a-CBM3a. Removing GFP
 369 reduced the off-rate drastically for tandem CBM3a but increased it for single CBM3a. On the other hand,
 370 CBM64 didn't have a significant difference in their k_{off} values. These results suggest that there is no
 371 significant difference in binding affinity for single and tandem versions of CBM64 towards cellulose-I. Also,
 372 a weaker binding is probably more prevalent in the case of GFP tandem CBM3a towards cellulose-I, which
 373 could be determined from a higher desorption rate which is overcome by removing GFP. The k_{off} , nk_{on} , and
 374 number of bound molecules are summarized in table 2.

	$k_{off} \times 10^{-3} (min^{-1})$	$nk_{on} \times 10^{-3} (L.g^{-1}.min^{-1})$	$A (x10^{-12} \text{ bound molecules})$
GFP-CBM3a	172.45±11.11	6.65±0.47	27.91±2.26
GFP-CBM3a-CBM3a	291.56±2.02	33.92±0.00	18.20±2.55
GFP-CBM64	159.86±23.24	1.12±0.00	17.89±1.70
GFP-CBM64-CBM64	167.35±19.02	2.03±0.00	16.32±0.00
CBM3a	223.12±1.86	n.a.	23.90±0.42
CBM3a-CBM3a	64.24±16.84	n.a.	30.15±0.70

375 **Table 2.** Kinetic rate constants for GFP-CBM3a, GFP-CBM3a-CBM3a, GFP-CBM64, GFP-CBM64-
 376 CBM64, CBM3a, and CBM-CBM3a adsorption and desorption toward nanocrystalline cellulose allomorphs
 377 estimated using QCM-D-based binding assay data. n.a. – Not available.

378



379

380 **Figure 3.** (a) Schematic representation for QCM-D-based tandem CBM - nanocrystalline cellulose-based
 381 binding assay. (b) Frequency (Hz) versus time (min) data for a representative protein (GFP-CBM3a) was
 382 converted to a sensorgram (number of protein molecules $\times 10^{12}$ versus time (mins)) using the Sauerbrey
 383 equation. The binding and unbinding data in the plot on the right were then fitted to an exponential rise and
 384 decay function, respectively, as described in detail under the methods section. (c) nK_{on} (right) was
 385 calculated using the formula stated, and nK_{on} , and K_{off} (left) for GFP-CBM3a, GFP-CBM3a-CBM3a, GFP-
 386 CBM64, and GFP-CBM64-CBM64 toward cellulose-I.

387 QCM-D has been utilized as a tool to study the binding of full-length cellulases to different types of cellulose
388 (Brunecky *et al.*, 2020), lignin, and pretreated biomass (Kumagai *et al.*, 2014; Haarmeyer *et al.*, 2017).
389 QCM-D has also been used to study CBMs associated with oxidases (Mollerup *et al.*, 2016) and cellulase-
390 hemicellulase complexes (Freelove *et al.*, 2001a). These reports were mostly confined to single wild-type
391 and engineered CBMs associated with cellulases and other hydrolytic enzymes. We have previously
392 examined the viscoelastic properties of CBM-cellulose binding using QCM-D analysis (Chundawat *et al.*,
393 2021; Nemmaru *et al.*, 2021). Here, we explored the kinetic constants of tandem CBMs using QCM-D to
394 identify if they have an improved binding affinity and if they could be used as a weak or tight-binding imaging
395 probe in plant cells. A similar trend in adsorption and desorption rate constants was observed in CBM64
396 for both tandem and single versions suggesting no apparent avidity advantage for tandem CBM64. Similar
397 behavior was observed in SusF, a starch-utilizing system with tandem CBMs which didn't show an improved
398 affinity compared to its single and mutated versions (Cameron *et al.*, 2012). In contrast, GFP tandem
399 CBM3a showed a higher desorption rate and, ultimately, a lower affinity towards Avicel cellulose-I. This
400 reduced affinity towards crystalline cellulose could be attributed to the linker length that connects the GFP
401 to the CBM3a as well as between the two CBMs present in tandem CBM constructs. The native linker used
402 in this study is 42 amino acids long and has been found to be highly flexible in our previous study (Bandi *et al.*,
403 2020). Compared to shorter linker lengths, the inclusion of a flexible linker resulted in a reduced activity
404 when fused to a cellulase (CelE from *Clostridium thermocellum*). Similar behavior was observed when
405 double CBMs were fused to cellulases, Cel6A and Cel7A (from *Trichoderma reesei*). The 48 aa-linker
406 decreased the binding affinity and capacity of the tandem CBM compared to reduced linker lengths (Arola
407 & Linder, 2016).

408 ***Fluorescence microscopy for detecting cellulose fibrils during protoplast cell wall regeneration***

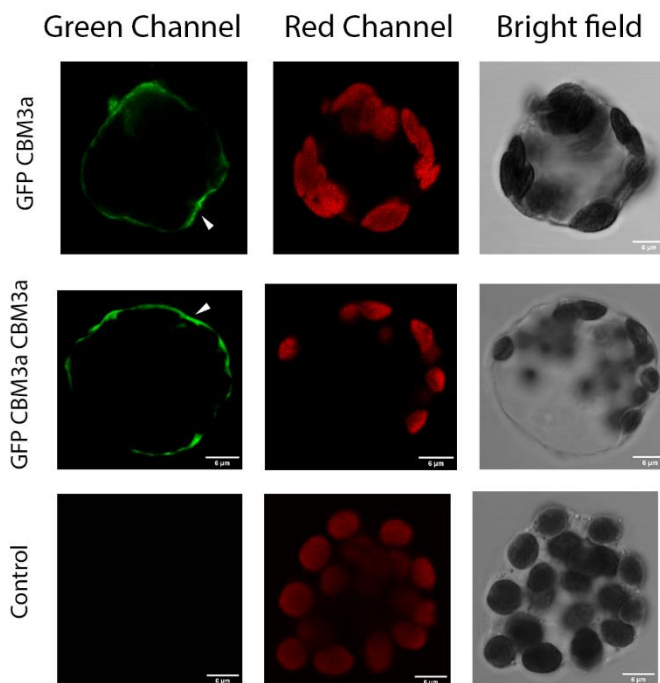
409 Procedures for preparing Arabidopsis wild-type mesophyll protoplasts and cell wall regeneration conditions
410 have been described in detail under the 'Materials and Methods' section. The isolated protoplasts were
411 healthy and relatively homogeneous in size, as observed under a simple light microscope (Figure 1). After
412 incubation of protoplasts in regeneration media (WI+2M2), the cells were fixed and stained with calcofluor,
413 a β -glucan-specific dye. Calcofluor white has been traditionally used to obtain high-resolution fluorescent
414 images of cellulose without autofluorescence (Kuki *et al.*, 2017). After 17 h of incubation in the regeneration
415 media, the cell wall network was spread over the entire surface of the plant protoplast (Figure S9a).
416 However, calcofluor is also known to bind to callose - a β -1,3 glucan present on the plant surface. Besides
417 non-specific binding targets, calcofluor is also known to affect the *in vivo* assembly of cellulose microfibrils
418 (Haigler *et al.*, 1980). Hence, the addition of calcofluor to the cell wall regeneration media could potentially
419 induce cytotoxicity for plant cells and preclude its use in monitoring cell wall regeneration in live protoplasts.

420 To overcome the cytotoxicity, improve the specificity, and predominantly visualize cellulose network around
421 regenerated plant protoplasts, engineered CBMs could be ideal probes to be employed for imaging. 100
422 nM GFP-CBM3a was used to visualize the regenerated cell wall under fluorescence microscopy. Both

423 calcofluor and GFP-CBM3a stained cells showed the presence of fibrous structures on the surface of
424 protoplasts after overnight incubation in regeneration media. The developed fibers labeled by GFP-CBM3a
425 were distinct and clear when observed under the green emission channel using a fluorescence microscope
426 (Figure S9b).

427 ***Type-A CBMs facilitate Arabidopsis thaliana cell wall visualization***

428 To explore further the general significance of the data mentioned above, the other type-A CBMs analyzed
429 biochemically were also used to image regenerated cell walls of plant protoplasts. All the CBMs used bound
430 to the cell wall surface of protoplasts, typical of type-A CBMs. Confocal laser scanning microscopy (CLSM)
431 was used to acquire images with high-resolution and across multiple focal planes. CLSM images showed
432 that the regenerated plant protoplasts had a pronounced accumulation of fluorescence along the edges of
433 the cell. Both GFP-CBM3a and GFP-CBM3a-CBM3a bound to fibers on the cell wall compared to the control
434 with no CBMs (Figure 4; Supplementary videos 1-4). To check if other type-A tandem CBMs behave in a
435 similar fashion, GFP-CBM64-CBM64 was used to image the regenerated plant protoplasts. Cell wall fibers
436 were observed throughout the edges of the protoplast (Figure S10; Supplementary videos 5). Alternatively,
437 the DS-red versions of CBM3a also showed similar probing characteristics (Figure S11; Supplementary
438 videos 6-7). However, the autofluorescence arising from plant chloroplasts under the red channel could
439 also contribute to the signals arising from DS-red fluorescence (Krause & Weis, 1991). Hence, GFP-CBMs
440 were used for further experiments.



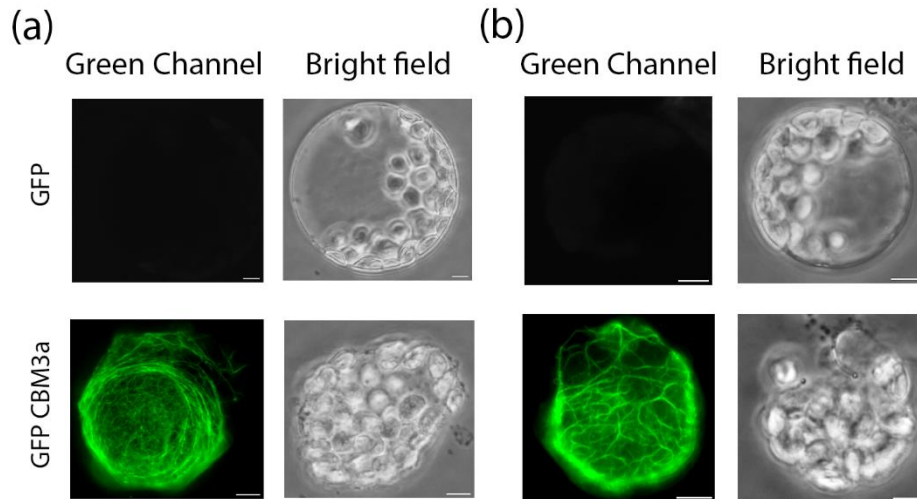
441

442 **Figure 4.** Confocal laser scanning microscopy images of GFP-CBM3a and GFP-CBM3a-CBM3a bound to
443 regenerated plant cell walls of Arabidopsis mesophyll protoplasts. The cell wall labeled with GFP-CBM is
444 visible under the green channel (left), and chloroplasts are visible under the red channel due to
445 autofluorescence (middle) and the bright field image of the intact protoplasts (right). Control had no GFP-
446 CBM added before imaging. White arrows indicate the GFP-CBMs binding to regenerated cell walls. (Scale
447 bars, 6 μm .)

448 ***GFP-CBM3a probe potentiates real-time imaging of cellulose synthesized by plant protoplasts***

449 Previous studies on imaging of plant cell walls were performed using immunocytochemistry and indirect
450 immunofluorescence (McCartney *et al.*, 2006; Knox, 2012). In particular, the time course of imaging of plant
451 cell walls was performed by isolating regenerated protoplasts at specific time intervals and labeling them
452 with calcofluor (Kuki *et al.*, 2017) or by directly visualizing plant roots using S4B dye (Anderson *et al.*, 2010).
453 These studies fail to capture the continuous dynamics of cellulose growth and movement on the surface of
454 regenerating plant protoplasts. These dyes are also toxic and cannot be added to the regeneration media
455 but only to fully developed cells. The absence of molecular probes that could be readily added to the
456 regeneration media is one of the primary reasons for the lack of complete understanding of cellulose
457 regeneration kinetics and other dynamic properties. Conversely, type-A CBMs are specific to cellulose and
458 are non-toxic to plant cells which makes them attractive to overcome this bottleneck.

459 To examine if these CBMs are compatible with live protoplasts and thus may not interfere with the cell wall
460 regeneration process, we added 100 nM of GFP-CBM3a to the regeneration media (WI+M2) from the outset
461 of the regeneration process. For controls, we had protoplasts in regeneration media with 100 nM GFP and
462 no GFP or GFP-CBM3a. Firstly, as mentioned previously, the regeneration of protoplasts was confirmed
463 by adding GFP-CBM3a to the regenerated protoplasts. Cell wall fibers were clearly distinguishable and
464 visible under fluorescence microscopy (Figure 5a). Similarly, protoplasts incubated in regeneration media,
465 along with GFP-CBM3a, also exhibited a uniform distribution of cellulose at the cell surface compared to
466 cells incubated with only GFP (Figure 5b). Taken together, these results show that GFP-CBMs could be
467 readily added to the regeneration media to continuously monitor the growing cellulose chains. Such a
468 system should now provide the tools to monitor cell wall growth and dynamics in real-time with live cells.



469

470 **Figure 5.** Fluorescence microscopy images of cell wall regenerated protoplasts. (a) GFP and GFP-CBM3a
471 were added to the regenerated plant cell walls of *Arabidopsis* mesophyll protoplasts after 18 h incubation.
472 (b) GFP and GFP-CBM3a were added along with the regeneration media at the outset of the 18 h
473 incubation. Regenerated cellulose fibers labeled with GFP-CBM3a are visible under the green channel (left)
474 and the bright field image of the intact protoplasts (right). (Scale bars, 6 μm.)

475 Discussion

476 Carbohydrate binding modules (CBMs) are crucial in targeting plant cell wall polysaccharides in
477 coordination with glycosyl hydrolases (Hervé *et al.*, 2010; Fox *et al.*, 2013). Several studies show the
478 presence of tandem CBMs coexisting with these modular microbial plant cell wall glycosyl hydrolases
479 (Freelove *et al.*, 2001b; McCartney *et al.*, 2006; Møller *et al.*, 2021). Some of these tandem CBMs exhibit
480 a synergistic effect with distinct specificities, and others exhibit an improved affinity towards
481 polysaccharides. Unlike other plant polysaccharides that have an extensive repertoire of mAbs readily
482 available, no stable mAb exists that can be used to visualize cellulose (Rydahl *et al.*, 2018). Type-A CBMs
483 bind specifically to cellulose, and particularly, CBM3a has a high affinity towards cellulose making it an
484 appropriate probe to detect plant cellulose (Chowdhury *et al.*, 2014; Johnsen *et al.*, 2015). However, CBM3a
485 could also binds non-specifically to other cell wall polysaccharides like xyloglucan (Hernandez-Gomez *et al.*,
486 2015).

487 To address this issue and to increase the odds of CBM3a binding to cellulose with high affinity, we prepared
488 tandem CBM3a fused to GFP to visualize the plant protoplasts directly. The ability of this tandem CBM3a
489 to recognize isolated cellulose both *in vitro* and in regenerated plant cell walls was assessed. Interestingly,
490 the tandem type-A CBMs – GFP-CBM3a and GFP-CBM64, showed reduced and equal affinity compared
491 to their single domain counterparts. This reduction in binding affinity could be attributed to the native 42-aa
492 linker that is highly flexible and might affect the binding of tandem CBMs. Furthermore, CBM3a exists as a

493 dimer in nature and might dimerize because of this highly flexible linker leading to an increased desorption
494 rate (Tormo *et al.*, 1996; Arola & Linder, 2016). Interestingly, removing GFP seemed to decrease off-rate
495 for tandem CBM3a drastically, suggesting that it might be more suitable for indirect immunofluorescence
496 imaging purposes that involve a continuous and dynamic movement of the substrate in an aqueous
497 environment. Additionally, these non-GFP-CBM versions could be conjugated to various amine-reactive
498 fluorophores providing a library of CBMs across different wavelengths. The resulting fluorophore
499 conjugated CBM proteins may exhibit greater photostability and brighter fluorescence than other
500 fluorochromes (Moran-Mirabal *et al.*, 2009).

501 In summary, these results provide qualitative and quantitative facets of the single and tandem CBMs used
502 to imaging plant cellulose fibers. Although single CBMs show a reduced desorption rate from QCM-D
503 assays, confocal imaging of plant protoplasts shows distinct cell walls with both single and tandem-CBMs.
504 Linker length could play a critical role in binding irreversibly to the cellulose fibers, and usage of shorter and
505 less flexible linkers might improve the binding affinity of tandem CBMs (Bandi *et al.*, 2020). To unravel its
506 complete potential, a more detailed study pertaining to the linker length of tandem CBMs for improving the
507 binding affinity and specificity to cellulose in the regenerated plant cell wall is required. Furthermore, the
508 plant cell wall is a natural amalgam of multiple polysaccharides intertwined with each other. Mimicking the
509 exact composition of complex plant cell wall polysaccharides *in vitro* might be unfeasible. In order to
510 overcome the non-specific binding nature exhibited by CBMs, techniques like phage display and directed
511 evolution might be needed to engineer proteinaceous probes with greater specificity and affinity for imaging
512 cellulose or other glycans synthesized and assembled in plant cell walls (DeVree *et al.*, 2021). Lastly, an
513 important characteristic of the GFP-CBM approach for detection and monitoring of cellulose fibers in plant
514 cells is its apparent non-toxic nature. This property opens the possibility for direct visualization of the kinetic
515 and spatial properties during the process of cell wall synthesis by using these protein-based labels for
516 cellulose.

517 **Acknowledgements:** The research was supported by the U.S. Department of Energy (Award Number:
518 DE-SC0019313) and Rutgers University.

519 **Competing Interest Statement:** The authors declare no competing interests.

520 **Author Contributions:** D.J. and S.P.S.C. designed the research. D.J., P.S., M.I., and J.S. conducted
521 research. D.J. and P.S. analyzed the data. D.J. and S.P.S.C. wrote the manuscript with input from all
522 authors.

523 **Data availability:** All raw data used to support the findings of this study are available upon request from
524 the corresponding author (Dr. Shishir P.S. Chundawat, Rutgers University,
525 shishir.chundawat@rutgers.edu).

526 **References**

- 527 **Anderson CT, Carroll A, Akhmetova L, Somerville C. 2010.** Real-time imaging of cellulose
528 reorientation during cell wall expansion in Arabidopsis roots. *Plant Physiology* **152**: 787–796.
- 529 **Arola S, Linder MB. 2016.** Binding of cellulose binding modules reveal differences between cellulose
530 substrates. *Scientific Reports* **6**: 1–9.
- 531 **Bandi CK, Goncalves A, Pingali SV, Chundawat SPS. 2020.** Carbohydrate-binding domains facilitate
532 efficient oligosaccharides synthesis by enhancing mutant catalytic domain transglycosylation activity.
533 *Biotechnology and Bioengineering* **117**: 2944–2956.
- 534 **Beckham GT, Ståhlberg J, Knott BC, Himmel ME, Crowley MF, Sandgren M, Sørlie M, Payne CM.**
535 **2014.** Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases. *Current*
536 *Opinion in Biotechnology* **27**: 96–106.
- 537 **Blake AW, McCartney L, Flint JE, Bolam DN, Boraston AB, Gilbert HJ, Knox JP. 2006.**
538 Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules
539 in prokaryotic enzymes. *Journal of Biological Chemistry* **281**: 29321–29329.
- 540 **Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. 2004.** Carbohydrate-binding modules: Fine-tuning
541 polysaccharide recognition. *Biochemical Journal* **382**: 769–781.
- 542 **Boraston AB, Kwan E, Chiu P, Warren RAJ, Kilburn DG. 2003.** Recognition and hydrolysis of
543 noncrystalline cellulose. *Journal of Biological Chemistry* **278**: 6120–6127.
- 544 **Brady SK, Sreelatha S, Feng Y, Chundawat SPS, Lang MJ. 2015.** Cellobiohydrolase 1 from
545 *Trichoderma reesei* degrades cellulose in single cellobiose steps. *Nature Communications* **6**.
- 546 **Brunecky R, Subramanian V, Yarbrough JM, Donohoe BS, Vinzant TB, Vanderwall TA, Knott BC,**
547 **Chaudhari YB, Bomble YJ, Himmel ME, et al. 2020.** Synthetic fungal multifunctional cellulases for
548 enhanced biomass conversion. *Green Chemistry* **22**: 478–489.
- 549 **Cameron EA, Maynard MA, Smith CJ, Smith TJ, Koropatkin NM, Martens EC. 2012.** Multidomain
550 carbohydrate-binding proteins involved in *Bacteroides thetaiotaomicron* starch metabolism. *Journal of*
551 *Biological Chemistry* **287**: 34614–34625.
- 552 **Cantarel BI, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009.** The Carbohydrate-
553 Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research* **37**:
554 233–238.
- 555 **Chowdhury J, Henderson M, Schweizer P, Burton RA, Fincher GB, Little A. 2014.** Differential

- 556 accumulation of callose, arabinoxylan and cellulose in nonpenetrated versus penetrated papillae on
557 leaves of barley infected with *Blumeria graminis* f. sp. *hordei*. *New Phytologist* **204**: 650–660.
- 558 **Chundawat SPS, Nemmaru B, Hackl M, Brady SK, Hilton MA, Johnson MM, Chang S, Lang MJ,**
559 **Huh H, Lee SH, et al. 2021.** Molecular origins of reduced activity and binding commitment of processive
560 cellulases and associated carbohydrate-binding proteins to cellulose III. *Journal of Biological Chemistry*
561 **296**: 100431.
- 562 **DeVree BT, Steiner LM, Głazowska S, Ruhnnow F, Herburger K, Persson S, Mravec J. 2021.** Current
563 and future advances in fluorescence-based visualization of plant cell wall components and cell wall
564 biosynthetic machineries. *Biotechnology for Biofuels* **14**: 1–26.
- 565 **Fox JM, Jess P, Jambusaria RB, Moo GM, Liphardt J, Clark DS, Blanch HW. 2013.** A single-molecule
566 analysis reveals morphological targets for cellulase synergy. *Nature Chemical Biology* **9**: 356–361.
- 567 **Freelove ACJ, Bolam DN, White P, Hazlewood GP, Gilbert HJ. 2001a.** A Novel Carbohydrate-binding
568 Protein Is a Component of the Plant Cell Wall-degrading Complex of *Piromyces equi*. *Journal of*
569 *Biological Chemistry* **276**: 43010–43017.
- 570 **Freelove ACJ, Bolam DN, White P, Hazlewood GP, Gilbert HJ. 2001b.** A Novel Carbohydrate-binding
571 Protein Is a Component of the Plant Cell Wall-degrading Complex of *Piromyces equi*. *Journal of*
572 *Biological Chemistry* **276**: 43010–43017.
- 573 **Gao D, Chundawat SPS, Sethi A, Balan V, Gnanakaran S, Dale BE. 2013.** Increased enzyme binding
574 to substrate is not necessary for more efficient cellulose hydrolysis. *Proceedings of the National Academy*
575 *of Sciences of the United States of America* **110**: 10922–10927.
- 576 **Guillén D, Sánchez S, Rodríguez-Sanoja R. 2010.** Carbohydrate-binding domains: Multiplicity of
577 biological roles. *Applied Microbiology and Biotechnology* **85**: 1241–1249.
- 578 **Haarmeyer CN, Smith MD, Chundawat SPS, Sammond D, Whitehead TA. 2017.** Insights into
579 cellulase-lignin non-specific binding revealed by computational redesign of the surface of green
580 fluorescent protein. *Biotechnology and Bioengineering* **114**: 740–750.
- 581 **Haigler CH, Brown RM, Benziman M. 1980.** Calcofluor white ST alters the in vivo assembly of cellulose
582 microfibrils. *Science* **210**: 903–906.
- 583 **Hashimoto H. 2006.** Recent structural studies of carbohydrate-binding modules. *Cellular and Molecular*
584 *Life Sciences* **63**: 2954–2967.
- 585 **Hernandez-Gomez MC, Rydahl MG, Rogowski A, Morland C, Cartmell A, Crouch L, Labourel A,**
586 **Fontes CMGA, Willats WGT, Gilbert HJ, et al. 2015.** Recognition of xyloglucan by the crystalline

- 587 cellulose-binding site of a family 3a carbohydrate-binding module. *FEBS Letters* **589**: 2297–2303.
- 588 **Hervé C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ, Knoxa JP. 2010.** Carbohydrate-binding
589 modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects.
590 *Proceedings of the National Academy of Sciences of the United States of America* **107**: 15293–15298.
- 591 **Johnsen HR, Striberny B, Olsen S, Vidal-Melgosa S, Fangel JU, Willats WGT, Rose JKC, Krause K.**
592 **2015.** Cell wall composition profiling of parasitic giant dodder (*Cuscuta reflexa*) and its hosts: A priori
593 differences and induced changes. *New Phytologist* **207**: 805–816.
- 594 **Kankare J. 2002.** Sauerbrey equation of quartz crystal microbalance in liquid medium. *Langmuir* **18**:
595 7092–7094.
- 596 **Kellett LE, Poole DM, Ferreira LMA, Durrant AJ, Hazlewood GP, Gilbert HJ. 1990.** Xylanase B and an
597 arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding
598 domains and are encoded by adjacent genes. *Biochemical Journal* **272**: 369–376.
- 599 **Knox JP. 2008.** Revealing the structural and functional diversity of plant cell walls. *Current Opinion in*
600 *Plant Biology* **11**: 308–313.
- 601 **Knox JP. 2012.** *In situ detection of cellulose with carbohydrate-binding modules.* Elsevier Inc.
- 602 **Kognole AA, Payne CM. 2018.** Cellulose-specific Type B carbohydrate binding modules: Understanding
603 oligomeric and non-crystalline substrate recognition mechanisms. *Biotechnology for Biofuels* **11**: 1–16.
- 604 **Krause GH, Weis E. 1991.** Annu. Rev. Plant Physiol. Plant Mol. Biol. 1991. 42:313-49 Copyright ©
605 1991 by Annual Reviews Inc. All rights reserved CHLOROPHYLL FLUORESCENCE AND
606 PHOTOSYNTHESIS. *Annu. Rev. Plant Physiol. Plant Mol. Biol* **42**: 313–49.
- 607 **Kuki H, Higaki T, Yokoyama R, Kuroha T, Shinohara N, Hasezawa S, Nishitani K. 2017.** Quantitative
608 confocal imaging method for analyzing cellulose dynamics during cell wall regeneration in *Arabidopsis*
609 mesophyll protoplasts. *Plant Direct* **1**.
- 610 **Kumagai A, Iwamoto S, Lee SH, Endo T. 2014.** Quartz crystal microbalance with dissipation monitoring
611 of the enzymatic hydrolysis of steam-treated lignocellulosic nanofibrils. *Cellulose* **21**: 2433–2444.
- 612 **Lehtiö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT. 2003.** The binding specificity
613 and affinity determinants of family 1 and family 3 cellulose binding modules. *Proceedings of the National*
614 *Academy of Sciences of the United States of America* **100**: 484–489.
- 615 **Levine SE, Fox JM, Blanch HW, Clark DS. 2010.** A mechanistic model of the enzymatic hydrolysis of
616 cellulose. *Biotechnology and Bioengineering* **107**: 37–51.

- 617 **Lim S, Chundawat SPS, Fox BG. 2014.** Expression, purification and characterization of a functional
618 carbohydrate-binding module from *Streptomyces* sp. SirexAA-E. *Protein Expression and Purification* **98**:
619 1–9.
- 620 **McCartney L, Blake AW, Flint J, Bolam DN, Boraston AB, Gilbert HJ, Knox JP. 2006.** Differential
621 recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. *Proceedings of*
622 *the National Academy of Sciences of the United States of America* **103**: 4765–4770.
- 623 **McKie VA, Vinoken JP, Voragen AGJ, Van Den Broek LAM, Stimson E, Gilbert HJ. 2001.** A new
624 family of rhamnogalacturonan lyases contains an enzyme that binds to cellulose. *Biochemical Journal*
625 **355**: 167–177.
- 626 **McLean BW, Boraston AB, Brouwer D, Sanaie N, Fyfe CA, Warren RAJ, Kilburn DG, Haynes CA.**
627 **2002.** Carbohydrate-binding modules recognize fine substructures of cellulose. *Journal of Biological*
628 *Chemistry* **277**: 50245–50254.
- 629 **Møller MS, Bouaballati S EI, Henrissat B, Svensson B. 2021.** Functional diversity of three tandem C-
630 terminal carbohydrate-binding modules of a β -mannanase. *Journal of Biological Chemistry* **296**: 1–13.
- 631 **Møllerup F, Parikka K, Vuong T V., Tenkanen M, Master E. 2016.** Influence of a family 29 carbohydrate
632 binding module on the activity of galactose oxidase from *Fusarium graminearum*. *Biochimica et*
633 *Biophysica Acta - General Subjects* **1860**: 354–362.
- 634 **Moran-Mirabal JM, Corgie SC, Bolewski JC, Smith HM, Cipriany BR, Craighead HG, Walker LP.**
635 **2009.** Labeling and purification of cellulose-binding proteins for high resolution fluorescence applications.
636 *Analytical Chemistry* **81**: 7981–7987.
- 637 **Nemmaru B, Ramirez N, Farino CJ, Yarbrough JM, Kravchenko N, Chundawat SPS. 2021.** Reduced
638 type-A carbohydrate-binding module interactions to cellulose I leads to improved endocellulase activity.
639 *Biotechnology and Bioengineering* **118**: 1141–1151.
- 640 **Pires VMR, Pereira PMM, Brás JLA, Correia M, Cardoso V, Bule P, Alves VD, Najmudin S, Venditto**
641 **I, Ferreira LMA, et al. 2017.** Stability and ligand promiscuity of type A carbohydrate-binding modules are
642 illustrated by the structure of *Spirochaeta thermophila* StCBM64C. *Journal of Biological Chemistry* **292**:
643 4847–4860.
- 644 **Ruel K, Nishiyama Y, Joseleau JP. 2012.** Crystalline and amorphous cellulose in the secondary walls of
645 *Arabidopsis*. *Plant Science* **193–194**: 48–61.
- 646 **Rydahl MG, Hansen AR, Kračun SK, Mravec J. 2018.** Report on the current inventory of the toolbox for
647 plant cell wall analysis: Proteinaceous and small molecular probes. *Frontiers in Plant Science* **9**: 1–20.

- 648 **Schiefner A, Angelov A, Liebl W, Skerra A. 2016.** Structural basis for cellulose binding by the type A
649 carbohydrate-binding module 64 of *Spirochaeta thermophila*. *Proteins: Structure, Function and*
650 *Bioinformatics* **84**: 855–858.
- 651 **Shang BZ, Chang R, Chu JW. 2013.** Systems-level modeling with molecular resolution elucidates the
652 rate-limiting mechanisms of cellulose decomposition by cellobiohydrolases. *Journal of Biological*
653 *Chemistry* **288**: 29081–29089.
- 654 **Stevenson J, Krycer JR, Phan L, Brown AJ. 2013.** A practical comparison of ligation-independent
655 cloning techniques. *PLoS ONE* **8**: 8–14.
- 656 **Talamantes D, Biabini N, Dang H, Abdoun K, Berlemont R. 2016.** Natural diversity of cellulases,
657 xylanases, and chitinases in bacteria. *Biotechnology for Biofuels* **9**: 1–11.
- 658 **Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y, Steitz TA. 1996.** Crystal structure of a
659 bacterial family-III cellulose-binding domain: A general mechanism for attachment to cellulose. *EMBO*
660 *Journal* **15**: 5739–5751.
- 661 **Vermaas J V., Kont R, Beckham GT, Crowley MF, Gudmundsson M, Sandgren M, Ståhlberg J,**
662 **Väljamäe P, Knott BC. 2019.** The dissociation mechanism of processive cellulases. *Proceedings of the*
663 *National Academy of Sciences of the United States of America* **116**: 23061–23067.
- 664 **Whitehead TA, Bandi CK, Berger M, Park J, Chundawat SPS. 2017.** Negatively Supercharging
665 Cellulases Render Them Lignin-Resistant. *ACS Sustainable Chemistry and Engineering* **5**: 6247–6252.
- 666 **Xu Y, Li R, Luo H, Wang Z, Li MW, Lam HM, Huang C. 2022.** Protoplasts: small cells with big roles in
667 plant biology. *Trends in Plant Science* **27**: 828–829.
- 668 **Y.-H. Percival Zhang JC, Lee R. Lynd LRK. 2006.** A Transition from Cellulose Swelling to Cellulose
669 Dissolution by o-Phosphoric Acid: Evidence from Enzymatic Hydrolysis and Supramolecular Structure. :
670 644–648.
- 671 **Yoo SD, Cho YH, Sheen J. 2007.** Arabidopsis mesophyll protoplasts: A versatile cell system for transient
672 gene expression analysis. *Nature Protocols* **2**: 1565–1572.
- 673
- 674

675 **Supporting Information**

676 **Fig. S1:** Plasmid maps of vectors and crystal structures of CBMs.

677 **Fig. S2:** Molecular architectures of CBM constructs.

678 **Fig. S3:** SDS-PAGE gel of purified CBMs and protein constructs.

679 **Fig. S4:** Langmuir-type adsorption model fits for different CBMs towards Cellulose-I.

680 **Fig. S5:** Langmuir-type adsorption model fits for different CBMs towards PASC.

681 **Fig. S6:** Binding Reversibility of CBMs towards Cellulose-I.

682 **Fig. S7:** Binding Reversibility of CBMs towards PASC.

683 **Fig. S8:** K_{off} for non-GFP versions of CBM3a and CBM3a-CBM3a towards Cellulose-I.

684 **Fig. S9:** SDS-PAGE gel of purified CBMs and protein constructs.

685 **Fig. S10:** Fluorescence microscopy images of calcofluor white and GFP-CBM3a labeled plant protoplasts.

686 **Fig. S11:** Confocal laser scanning microscopy images of GFP-CBM64-CBM64 labeled plant protoplasts.

687 **Fig. S12:** Confocal laser scanning microscopy images of DSR-CBM labeled plant protoplasts.

688

689 **Supplementary video 1:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
690 PBS buffer (no M2 media) with no GFP-CBM.

691 **Supplementary video 2:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
692 M2 media with no GFP-CBM.

693 **Supplementary video 3:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
694 M2 media labeled with GFP-CBM3a.

695 **Supplementary video 4:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
696 M2 media labeled with GFP-CBM3a-CBM3a.

697 **Supplementary video 5:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
698 M2 media labeled with GFP-CBM64-CBM64.

699 **Supplementary video 6:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
700 M2 media labeled with DSR-CBM3a.

701 **Supplementary video 7:** Confocal microscopy movie of regenerated Arabidopsis mesophyll protoplast in
702 M2 media labeled with DSR-CBM3a-CBM3a.

703

704 **Table S1:** Primer sequences used for cloning and sequencing of all constructs.

705 **Table S2:** Molecular weight and extinction coefficients of all CBMs used in this study.

706

707 **Supplementary Text S1:** Protein sequences of all major constructs used in this study.