

# Evidence for Plant-Conserved Region Mediated Trimeric CESAs in Plant Cellulose Synthase Complexes

*Juan Du<sup>1,2</sup>, Venu Gopal Vandavasi<sup>3</sup>, Kelly R. Molloy<sup>4</sup>, Hui Yang<sup>5</sup>, Lynnica Massenbourg<sup>2</sup>, Abhishek Singh<sup>6</sup>,  
Albert L. Kwansa<sup>6</sup>, Yaroslava G. Yingling<sup>6</sup>, Hugh O'Neill<sup>3</sup>, Brian T. Chait<sup>4</sup>, Manish Kumar<sup>7</sup>, B. Tracy Nixon<sup>2,\*</sup>*

<sup>1</sup>Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences,  
Sichuan University, Chengdu, 610064 China;

<sup>2</sup>Dept of Biochemistry & Molecular Biology, The Pennsylvania State University, University Park, PA 16802;

<sup>3</sup>Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831;

<sup>4</sup>Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY  
10065, USA

<sup>5</sup>Dept of Biology, The Pennsylvania State University, University Park, PA 16802.

<sup>6</sup>Dept of Materials Science and Engineering, North Carolina State University, Raleigh, NC 27695

<sup>7</sup>Dept of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802

\*Corresponding author: B.T.N. (btn1@psu.edu)

## **ABSTRACT**

Higher plants synthesize cellulose using membrane-bound, six-lobed cellulose synthase complexes (CSCs), each lobe containing trimeric cellulose synthases (CESAs). Although molecular biology reports support heteromeric trimers composed of different isoforms, homomeric trimer was reported for in-vitro studies of the catalytic domain of CESA1 of Arabidopsis (AtCESA1CatD) and confirmed in cryoEM structures of full-length CESA8 and CESA7 of poplar and cotton, respectively. In both structures a small portion of the P-CR region forms the only

26 contacts between catalytic domains of the monomers. We report inter-subunit lysine-crosslinks that localize to  
27 the small P-CR region, negative-stain EM structure, and modeling data for homotrimers of AtCESA1CatD.  
28 Molecular dynamics simulations for AtCESA1CatD trimers based on the CESA8 cryoEM structure were stable  
29 and depended upon a small set of residue contacts. The results suggest homomeric CESA trimers may be  
30 important for the synthesis of primary and secondary cell walls, and identify key residues for future mutagenic  
31 studies.

## 32 **KEYWORDS**

33 cellulose synthase complex, structure, plant cell wall, Arabidopsis

## 34 **INTRODUCTION**

35  
36 Cellulose is the world's most abundant organic compound and is used as raw material for the industrial production  
37 of paper, cardboard, wood, clothing, and medical supplies. Cellulose exists in plant cell walls as  $\beta$ -1,4 glucan  
38 chains, with primary walls existing around growing cells and relatively thick secondary walls surrounding non-  
39 growing cells. Increasingly, attention is focused on using fibers as feedstock for sustainable biofuel production as  
40 a renewable alternative to fossil fuels <sup>1-4</sup>.

41  
42 Cellulose is made by cellulose synthases (CESAs) that form cellulose synthase complexes (CSCs) in higher  
43 plants, bryophytes and algae. The size, pattern, and microfibril association of CSCs in the plasma membrane of  
44 moss were captured using freeze-fracture transmission electron microscopy <sup>5-9</sup>. In the model plant Arabidopsis,  
45 genetic evidence suggests three distinct isoforms are required for cellulose synthesis, in which AtCESA1,  
46 AtCESA3 and AtCESA6 or AtCESA6-like form CSCs for making primary walls, and AtCESA4, AtCESA7 and  
47 AtCESA8 form CSCs for making secondary walls <sup>10-14</sup>. Many observations suggest that the plant CSC is a multi-  
48 subunit heteromeric complex that produces cellulose microfibrils *in vivo*. Among these are similar sub-cellular  
49 locations, dynamic movement in plasma membranes for CESAs in the same group, loss of partner CESAs upon  
50 one's removal, protein interaction confirmed by *in vivo* analysis <sup>11-13, 15-16</sup>, and 1:1:1 stoichiometric presence in  
51 primary (AtCESA1, AtCESA3 and AtCESA6) and secondary (AtCESA4, AtCESA7 and AtCESA8) wall CSCs  
52 <sup>17-18</sup>. Until a few years ago it was believed that each CSC contained 36 CESA protomers, but now several lines  
53 of evidence support there being 18 protomers organized in 6 trimeric lobes, which together make an 18-chain

54 cellulose microfibril (reviewed by Polko and Kieber <sup>19</sup>). Recent studies of CSC in cotton shows that the  
55 oligomerization state of CSCs and chain number in one microfibril can vary in different species and  
56 developmental stages <sup>20</sup>.

57 We note that: a) CESA ratios different from 1:1:1 have been seen in poplar <sup>21</sup>; b) CESA5 of *Physcomitrium*  
58 *patens* functions as homomeric CSC *in vivo* <sup>22</sup>; and c) *in vitro* studies of *Pichia*-expressed poplar CESA8 and  
59 moss CESA5 revealed that a single CESA isoform can form oligomers and synthesize cellulose microfibrils *in*  
60 *vitro* <sup>23-24</sup>. The nature of the *in vitro* fibers is not fully determined, but cryo-electron tomography and solid state  
61 NMR studies show that they favor 18 chains and are similar to CMF in plant cell walls <sup>25</sup>. CryoEM structures of  
62 homomeric trimers of secondary wall CESAs PttCESA8 and GhCESA7 have been solved <sup>26-27</sup>. A trimeric crystal  
63 structure of OsCESA8 P-CR fragment is nearly identical to the P-CR based trimers in the cryoEM structures for  
64 secondary cell wall CESAs <sup>28</sup>.

65 For CESAs making primary walls there are reports of a homomeric dimer structure for the catalytic domain of  
66 AtCESA3 <sup>29</sup> and the formation of homomeric trimers for the catalytic domain of AtCESA1 <sup>30</sup>. Here we report  
67 negative staining TEM structure of the AtCESA1CatD monomer and trimer and describe a set of inter-subunit  
68 cross-linking lysine residues that cluster in a loop and helix of the P-CR of AtCESA1CatD. Since the cross linking  
69 lysine residues are highly conserved among CESA proteins, similar cross-linking potential is predicted for other  
70 lobes of trimeric CESAs and found to be so for the crystal and cryoEM structures. By considering goodness of fit  
71 to the published SAXS data and 3D reconstruction, and (most impactfully) match between observed and predicted  
72 cross-links, we show that among tens of thousands of computational models of P-CR based trimers of  
73 AtCESA1CatD the most likely models possess a trimerization motif like those in the crystal lattice of OsCESA8  
74 P-CR and cryoEM structures of PttCESA8 and GhCESA7. The crosslink data are not consistent with the recently  
75 reported crystal structure dimer of AtCESA3CatD <sup>29</sup>. MD simulations showed that the homotrimer of  
76 AtCESA1CatD is highly stable, with somewhat less stability present in modeled heterotrimers – in the latter,  
77 interactions between the TM helices of different protomers are likely essential for overall trimer stability. We  
78 thus propose that the P-CR based trimerization motif is common to plant CESAs regardless of whether they make  
79 primary or secondary cell walls and present a new model for AtCESA1 based upon structure of PttCESA8.

80 Comparing this model to bacterial cellulose synthase prompts us to suggest that substitution of the regulatory PilZ  
81 domain with the P-CR motif allowed plants to evolve oligomeric CESAs. This also raises the possibility that the  
82 P-CR trimer motif mediates allosteric communication between protomers to regulate synthesis of  $\beta$ -glucan chains.

## 84 **EXPERIMENTAL SECTION**

### 85 **Protein expression and purification**

86 The strain harboring AtCESA1CatD (amino acids 341 to 845) was cultured as previous described <sup>30</sup>. The  
87 monomeric AtCESA1CatD was extracted from inclusion bodies and purified to homogeneity as described, and  
88 analyzed by SDS page gel electrophoresis and dynamic light scattering. Previously it was shown by numerous  
89 methods that the protein is well folded and monomeric in the buffer despite the presence of sodium lauryl  
90 sarcosine. The hydrodynamic radius of the protein we purified was approximately 5 nm, consistent with the  
91 published dimensions of the monomeric protein.

92 To optimize oligomerization of AtCESA1CatD, a two-step procedure was used. Firstly, the monomer sample  
93 was dialyzed for 48 h at 4°C against the buffer containing 25 mM HEPES (pH 8.5), 50 mM NaCl, 10% glycerol  
94 and 5 mM DTT in a 10 kD cutoff dialysis cassette (slide-A-lyzer, Thermo Scientific Inc.), which allowed the  
95 gentle reduction (but not total loss) of the detergent. Dialysis was continued overnight in the same buffer except  
96 the pH was changed to 7.8 until the value of hydrodynamic radius was about 8 nm, which was previously reported  
97 to be true for the trimers characterized by small angle X-ray and neutron scattering <sup>30</sup>.

98 To generate more homogeneous sample for EM characterization, the dialyzed sample was applied to a size  
99 exclusion chromatography column (Superose 6 increase, 10/300 GL, GE healthcare) pre-equilibrated with 25 mM  
100 HEPES (pH7.8), 100 mM NaCl, 0.1% sodium lauroyl sarcosine and 0.5 mM tris(2-carboxyethyl)phosphine  
101 (TCEP). The fractions from the first shoulder of the peak were pooled and concentrated using a 100 kD cutoff  
102 centrifugal concentrator (Vivaspin 500, GE healthcare) to about 150  $\mu$ l and the gel filtration step described above  
103 was repeated. The sample corresponding to the peak fraction that eluted at 12.5 ml was used for EM analysis.

### 105 **Negative staining EM**

106 Trimer samples were prepared by the following procedure. Grids used for negative staining EM were prepared as  
107 previously described<sup>31</sup>. Briefly, 400-mesh copper grids (TED PELLA) were floating on a thin film of collodion.  
108 After drying overnight, the film side of grid was coated by a thin layer of continuous carbon (Denton Vacuum  
109 model DV-502B). Three and one half microlitres of protein sample (0.005 mg/ml) was applied onto glow  
110 discharged grids (Pelco easiGlow). Following absorbing for 1 min, the grid was washed/stained in 10 drops of  
111 1% (wt/vol) freshly made uranyl formate. Then the grid was stained for 20 sec before quick and complete air-  
112 drying. Negatively stained grids were imaged at  $20 \text{ e}^-/\text{\AA}^2$  on a Tecnai T12 transmission electron microscope  
113 operated at 120 keV with a defocus range of 0.5-2.5  $\mu\text{m}$ . Micrographs were captured on a 4K x 4K charge-coupled  
114 device (CCD) camera (Eagle, FEI) at a nominal magnification of 68,000 x (1.5  $\text{\AA}/\text{pixel}$  at specimen level).

115 Monomer samples were prepared similarly to trimers except for washes with 7 drops of filtered MilliQ water  
116 followed by 3 drops of 0.75% (wt/vol) freshly made uranyl formate. The monomer grid was stained for 45 sec  
117 before quick and complete air drying. Negatively stained grids were imaged on a Tecnai T12 transmission electron  
118 microscope operated at 120 keV with a defocus range of 0.5-2.5  $\mu\text{m}$ . Monomer micrographs were captured on a  
119 2K x 2K charge-coupled device (CCD) camera (Orius, FEI) at a nominal magnification of 98,000 x (3.0  $\text{\AA}/\text{pixel}$   
120 at specimen level).

## 121 **Image processing**

122 EMAN2<sup>32</sup> was used to process monomer images using e2ctf.py for CTF correction. Manually picked particles  
123 were extracted using an 84 x 84-pixel box and pruned to 3836 particles by 2D class averaging. Initial models  
124 were generated and used for refinement without symmetry using e2initialmodel.py and e2refine.py, respectively.

125 The evaluation of trimer micrographs was also carried out using e2evalimage.py function in EMAN2<sup>32</sup>. All  
126 the images displaying no significant astigmatism and well-spread particles were used for further processing. The  
127 contrast transfer function of each micrograph was determined using CTFFind3<sup>33</sup>. Manually picked trimer  
128 particles (~4000) were extracted using a 168 pixel x 168 pixel box and subjected to 2D-reference free  
129 classification after normalization. Classes generating blurry averages and containing heterogenous particles were  
130 discarded. Remaining particles were then subject to Iterative Stable Alignment and Classification using ISAC2.<sup>34</sup>  
131 Initial models were reconstructed from remaining good class averages using e2initialmodel.py with either no

132 symmetry or 3-fold symmetry imposed. Eventually, one initial model with 3-fold symmetry imposed was applied  
133 in subsequent analyses. 3D reconstruction and refinement were performed by using the program of  
134 e2refine\_easy.py in EMAN2. Meanwhile, particles were imported into RELION 2.0 for 3D classification and  
135 refinement<sup>35</sup>. Density maps were validated using Tilt-Pair Validation as described below. Rigid body docking  
136 between EM map and computational model was performed in UCSF Chimera<sup>36</sup>.

### 137 **Tilt-pair validation**

138 To evaluate the quality of density map of monomer and trimer, a set of image pairs were recorded for the same  
139 specimen field at two tilt angles (untilted and 10° tilted) and processed as described<sup>37</sup>. The analysis was  
140 performed using program EMAN2's function e2tiltvalidate.py and TiltStats<sup>37-38</sup>.

### 141 **Generation of heavy lysine labeled dimer and trimer protein**

142 To identify the interaction interface between monomers in trimer complex, we generated chimeric trimer complex  
143 consisting regular monomer (<sup>12</sup>C and <sup>14</sup>N) and heavy lysine labeled monomer (<sup>13</sup>C and <sup>15</sup>N). Given the cross links  
144 between regular lysine and heavy lysine must happen between distinct monomers (light monomer and heavy  
145 monomer, in this case), it is expected that fragments bearing cross links from the interaction interfaces will be  
146 increased 8 Daltons per lysine residue in MS spectra, which allowed us to distinguish intra-cross linking in  
147 monomer and inter-cross linking between monomers. To produce heavy lysine labeled AtCESA1CatD trimer  
148 complex, modified minimal M9 broth with lower glucose (0.2%) and lacking NH<sub>4</sub><sup>+</sup> was used to culture the strain.  
149 The heavy lysine labeled monomer (CatD\_H) and regular monomer (CatD\_L) were mixed in 1:1 ratio and dialyze  
150 to form trimer as the procedure mentioned above.

### 151 **Chemical cross linking and mass spectrometry**

152 To obtain the structural information of protein folding in monomer and interaction interface in trimer, we applied  
153 chemical cross linking followed by mass spectrometry to analyze the distance constrains between reactive residue  
154 pairs. The purified protein was diluted to 2 mg/ml using buffer containing 20 mM HEPES (pH7.9). Immediately  
155 before use, DSS (a homobifunctional amine-reactive cross linker, spacer arm = 11.4 Å, Thermo Scientific) was  
156 dissolved in DMSO to make 50 mM stock. To optimize the cross linking of trimer complex, 20-fold excess of  
157 cross linker (0.66 mM) was added to protein sample and followed by incubation for 1 hr at 25°C with gentle

158 agitation during incubation. Then the reaction was quenched by adding 50 mM Tris to a final concentration and  
159 subsequently incubated at room temperature for 15 min. The cross linked product was precipitated by adding cold  
160 methanol to a final concentration of 90% and incubating at -80°C overnight. After spinning at 14,000 rpm at 4°C  
161 for 15 min, the pellet was washed once by 90% cold methanol and finally dissolved in 50  $\mu$ l 1 x NuPAGE LDS  
162 loading buffer (Invitrogen, USA). Then protein sample was subjected to disulfide reduction and cysteine  
163 alkylation. After incubation for 15 min at 70°C, sample was resolved in 4-12% gradient SDS-PAGE gel  
164 (GenScript) in Tris-MOPS buffer, which resulted in good separation of cross linked monomer, dimer and trimer  
165 bands. The targeted gel bands were excised, subjected to in-gel trypsin digestion and subjected to LC-MS to  
166 identify cross links (Figure S1, Table S1). Cross-linked peptides were identified by comparing experimental and  
167 theoretical MS2 fragmentation spectra to find the closest matches. These are indicated as having direct evidence  
168 in Table S1. We suspected that in some cases, inter-subunit crosslinked peptides were present in the sample but  
169 not identified by MS2 because they generated low-quality fragmentation spectra or were not selected for MS2.  
170 To find such peptides, we searched the MS1 spectra for peaks matching their calculated masses. When we  
171 observed such a peak at the same LC retention time as the corresponding light-light and heavy-heavy peptides,  
172 and at least one of those peptides had been identified by MS2 fragmentation, we considered this indirect evidence  
173 of an inter-subunit cross-linked peptide as indicated in Table S1.

## 174 **Molecular Dynamics Simulations**

175 Molecular dynamics (MD) simulations were conducted using AMBER 2018<sup>39</sup>. An AtCESA1CatD trimer model  
176 (Figure S2) was assembled using the PttCESA8 trimer structure (PDB ID: 6WLB)<sup>27</sup> as the template, with the  
177 AtCESA1CatD monomer model generated using the SwissModel program<sup>40</sup> in which a major portion of the CSR  
178 region (THR649 – MET713) was predicted as unstructured (Figure S3). AlphaFold2 was not available at the time  
179 of performing the MD simulations. The SwissModel and AlphaFold2 predicted monomers are nearly identical  
180 (RMSD, 0.713 Å). Three AtCESA1CatD monomers formed a triangular P-CR interface. The CSR region was  
181 positioned at the other side of the GT domain, opposite the P-CR region. The unstructured CSR regions were  
182 excluded so that the simulations could finish in a reasonable period of time. Counterions (0.10 M NaCl) were  
183 added which was solvated with TIP3P water molecules<sup>41</sup>. The size of the initial system was 166  $\times$  177  $\times$  108 Å<sup>3</sup>

184 and contained ~255,000 atoms. Minimization and equilibration stages were conducted by gradually reducing  
185 restraints on the protein. First, a 1,000-step minimization was performed consisting of 400 steps of steepest  
186 descent and 600 steps of conjugate gradient minimization with protein positions harmonically restrained using a  
187 force constant of 50 kcal/mol/Å<sup>2</sup>. After minimization, the system temperature was increased to 300 K through  
188 two sequential runs, with 10 and 2 kcal/mol/Å<sup>2</sup> restraints placed on the positions of the protein, respectively. First,  
189 the system was heated to 100 K for 20 ps in an NVT ensemble, and then it was slowly heated to 300 K for 100 ps  
190 at 1 atm in an NPT ensemble, with a 2-fs time step, 10-Å nonbonded interaction cutoff, and SHAKE-constrained  
191 hydrogen bonds. A short 500-ps NPT simulation was then performed with no restraints prior to production  
192 simulations. Conventional MD simulations were conducted on the models for 1,000 ns. In all stages, MD  
193 simulations were performed with periodic boundary conditions, a temperature of 300 K, a pressure of 1 atm, 2-fs  
194 time step, 10-Å nonbonded interaction cutoff, SHAKE-constrained hydrogen bonds, and particle-mesh Ewald  
195 (PME) and Lennard-Jones correction for long-range Coulombic and Lennard-Jones interactions, respectively. In  
196 all simulations, the protein was described by the ff14SB force field <sup>42</sup>, and the water molecules and counterions  
197 were described by the TIP3P model <sup>41</sup> and Joung-Cheatham monovalent ions for TIP3P, respectively <sup>43</sup>.

198 The root-mean-square deviation (RMSD) of the trimer demonstrated that the system reached equilibrium after  
199 about 150 ns. A 1,000-ns-long production run was performed on the theoretical model of AtCESA1CatD trimer.  
200 The root-mean-square fluctuation (RMSF) profile of the trimer residues was calculated based on the whole 1,000-  
201 ns-long production run. One hundred frames taken from the last 100 ns production run were used to calculate the  
202 binding free energies between AtCESA1CatD monomers, using the MM-PBSA approach in the AMBER 2018  
203 package. Per-residue decomposition calculations were performed to decompose the free energy contribution to  
204 the binding free energy between AtCESA1CatD monomers.

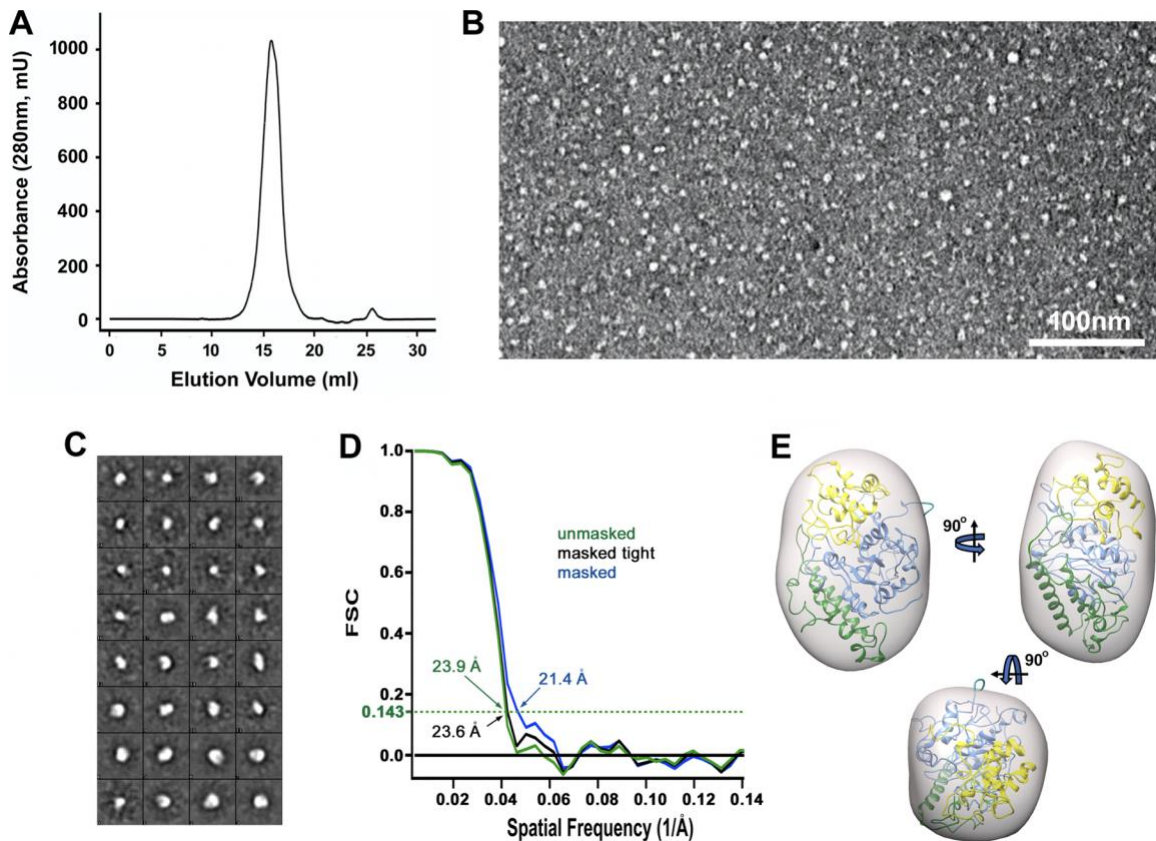
205 A newly developed protein structure prediction program, trRosetta <sup>44</sup>, was also utilized to generate AtCESA1,  
206 AtCESA3, and AtCESA6 catalytic domain models with the corresponding sequences listed in Table S2. As shown  
207 in Figure S3, the AtCESA1CatD model obtained from the trRosetta program superimposed tightly with the model  
208 predicted by the Swiss Model program, except for the CSR regions. The CSR region was predicted as unstructured  
209 in the SwissModel. In the trRosetta predicted structure, the CSR region is composed of four short helices

210 connected by short loops. Similar CSR structures were also obtained in the AtCESA3CatD, and AtCESA6CatD  
211 models, positioned at the other side of the GT domain, opposite the P-CR region, as shown in Figures S4-S5. Two  
212 heterotrimer models were then assembled by replacing two AtCESA1CatD units in the homotrimer as described  
213 before. One is in the counterclockwise arrangement (Figure S4), and the other is in the clockwise arrangement  
214 (Figure S5). Production runs of 500 ns were performed using the methods described above. The RMSD profile  
215 of each system was calculated based on the whole 500-ns-long production run. Distances between the mass center  
216 of each subunit (labeled as whole-#-# in Figure S6), and distances between the mass center of each P-CR domain  
217 (marked as PCR-#-# in Figure S6) were measured.

## 219 RESULTS

### 220 *TEM structure validation of AtCESA1CatD monomer model*

221 For the approach used in this study, the computational structure of AtCESA1CatD monomer forms the basis of  
222 understanding the structure and architecture of higher-order oligomer. A model of AtCESA1CatD was generated  
223 by threading its amino acid sequence (residues 341-845, RYD-GRL of AtCESA1) onto the recently published  
224 computational model of GhCESA1<sup>45</sup>. The overall size and shape of the model was validated by 3D  
225 reconstructions of purified monomer protein from negative stain TEM using EMAN2<sup>32</sup>. Size exclusion  
226 chromatography (SEC) of purified monomer yielded a single symmetric peak (Figure 1A). Negatively stained  
227 images (example presented in Figure 1B) showed well dispersed particles that were readily picked into a dataset  
228 of 7513 particles. All particles in this dataset were sorted by 2D classification (Figure 1C) into 32 classes and  
229 pruned via two rounds of refinement to 3975 particles. 3D reconstruction from these particles yielded a density  
230 map that is superimposed with the computational model in Figure 1D-E (21.4-23.9 Å resolution; FitMap CC  
231 0.92). In this optimal fit, the P-CR, CSR and catalytic core are well enclosed by the envelope. Since we are  
232 screening for trimers mediated by P-CR interactions, we conclude that the relevant parts of the computational  
233 model of monomer are consistent with the density map and that it is reasonable to use the model to build a library  
234 of putative P-CR/P-CR based trimer models.

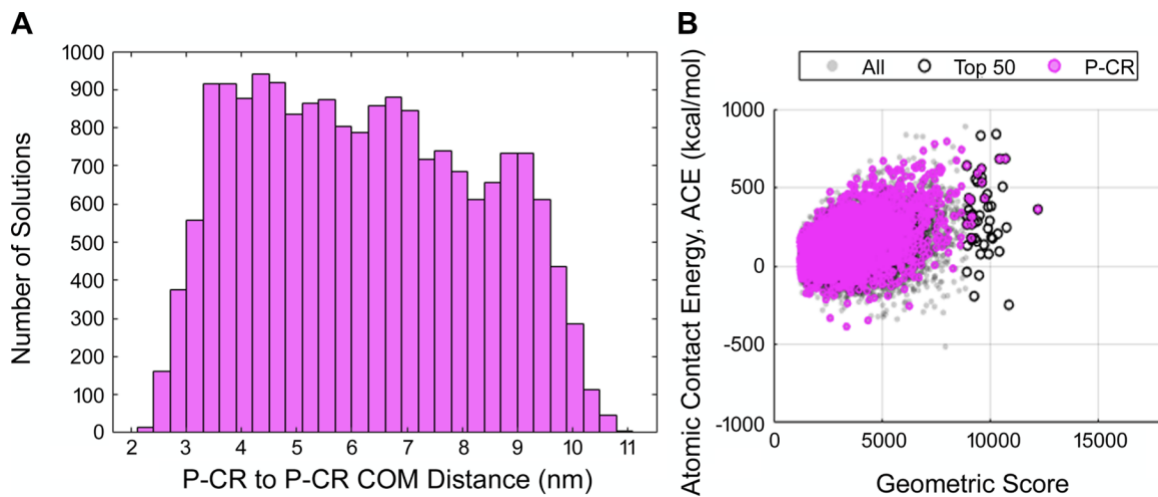


**Figure 1. Fit of computational model of AtCESA1CatD monomer into negative stain EM density map. A)** Size exclusion chromatogram of AtCESA1CatD monomer. **B)** Representative negative stain image of particles eluted from the main peak fraction in (A). **C)** 2D class averages of AtCESA1CatD monomer. **D)** Gold-standard Fourier shell correlation (FSC) plot vs spatial frequency. The density map is resolved to 21.4-23.9 Å according to the FSC value of 0.143. **E)** Model of AtCESA1CatD monomer superimposed on EM density map (catalytic core, P-CR, and CSR are represented by ribbons in blue, green and yellow, respectively).

### *Building trimer models of AtCESA1CatD with focus on optimizing P-CR regions at the monomer-monomer interfaces*

As suggested by previous studies, the P-CR and CSR regions may be involved in the higher-order oligomer assembly of CESA. New computational models have been built for trimers of GhCESA1 with focus on optimizing N-terminal (NT), CSR or P-CR regions at the monomer-monomer interfaces<sup>45</sup>. In that study, 20,488 trimer models were reduced to a “P-CR subset” of 367 models by accepting only those with a transmembrane helix (TMH) tilt angle less than or equal to 60 degrees and less than or equal to 5 nm distance between the centers of

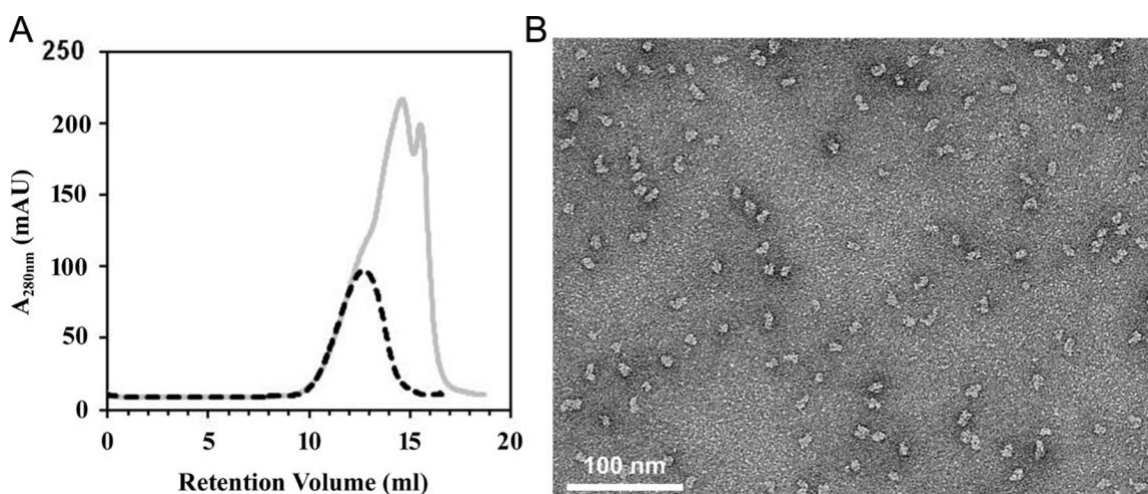
250 mass (COM) of adjacent P-CR regions. Omitting the TMH tilt angle criterion and just applying the P-CR/P-CR  
251 distance constraint (adjusted to less than or equal to 4 nm) yielded 2,651 models. For the current study, a similar  
252 library of 18,792 modeled trimers was generated using the AtCESA1CatD monomer structure and SymmDock  
253 <sup>46-47</sup>. Within the library, there were 3,251 models in which the P-CR elements' COM were not beyond 4 nm  
254 (Figure 2A; note that lacking the TMH regions in the models, there was no screening for TMH tilt angle). A  
255 scatter plot of SymmDock atomic contact energies vs geometric shape complementarity scores (geometric scores)  
256 for these models is shown in Figure 2B. Many of the top 50 scoring models, based on geometric score, were  
257 rejected by the P-CR to P-CR COM screening. Additional information is needed to screen through this P-CR/P-  
258 CR biased library to identify those that are reasonable based on biochemical observations. Small angle X-ray  
259 scattering data have been published for the AtCESA1CatD trimeric structure <sup>30</sup>. To supplement those data, we  
260 also reconstructed the trimer's 3D structure from negatively stained images and identified subunit-subunit cross  
261 links between lysine residues using the reagent disuccinimidyl suberate (DSS). As presented below, the latter  
262 information proved most informative.



264 **Figure 2. Library of trimers of AtCESA1CatD.** **A)** Number of solutions is plotted vs distance between P-CR  
265 centers of mass. Models with COM distances  $\leq 4$  nm were combined to make the P-CR subset. **B)** Scatter plot of  
266 atomic contact energy vs geometric score for the entire library (gray), the P-CR subset (magenta), and the top 50  
267 geometric-scoring models (black circles).

268  
269 **Purification of AtCESA1CatD trimers to homogeneity by sequential size exclusion chromatography**

270 We enriched the AtCESA1CatD trimers prepared for the SAXS studies for imaging by TEM. Initial attempts to  
271 pass trimer samples through a Superose 6 Increase SEC column failed, with the sample never entering the matrix.  
272 We thus identified a minimal concentration of sodium lauryl sarcosine to add back to the dialyzed sample that  
273 allowed material to enter the size exclusion matrix. Fractions containing protein of size expected for trimer were  
274 present in a shoulder followed by two peaks of smaller material (Figure 3A, gray). Analysis using negative  
275 staining EM on the shoulder fraction showed heterogeneous particles of variable size. To enrich trimer complex  
276 and minimize contaminants from larger and smaller material, we pooled the appropriate shoulder fractions and  
277 performed a second round of gel filtration, which resulted in a sharper, more symmetric major peak that showed  
278 no evidence of instability during the time needed for concentration and chromatography (~60 min; Figure 3A,  
279 black). Inspection using negative staining EM revealed improved particle integrity and homogeneity in the peak  
280 fraction of the second pass through the column (Figure 3B)

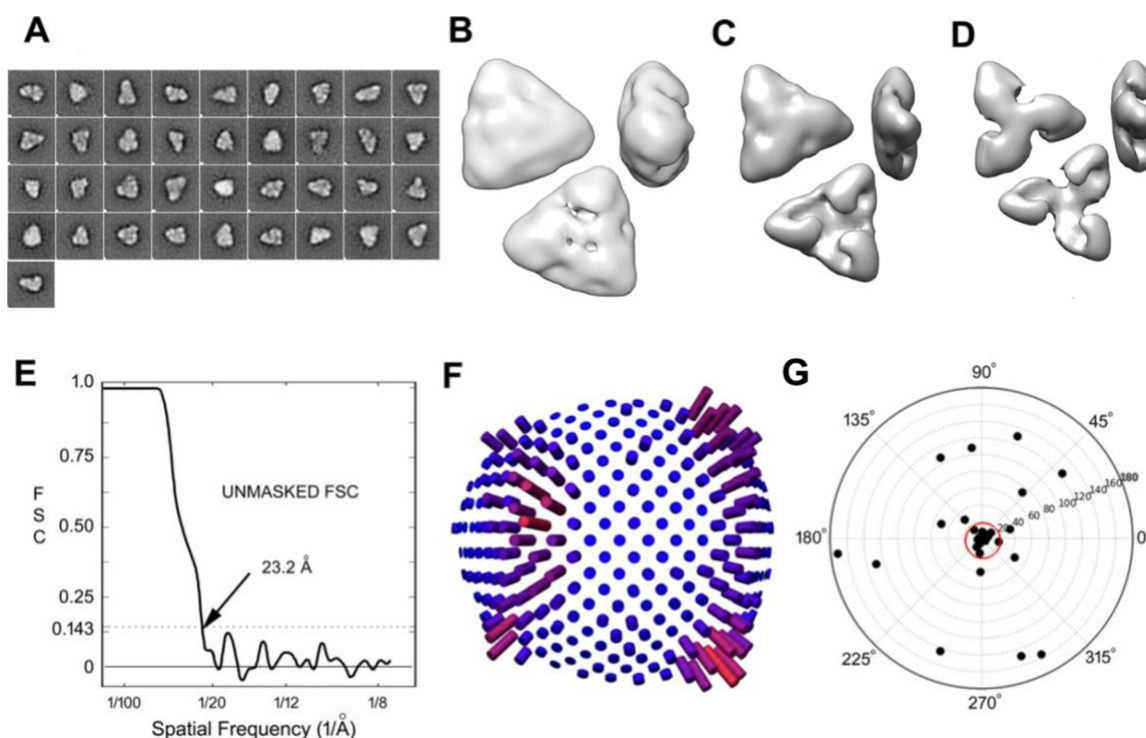


281  
282 **Figure 3. Purification and visualization of AtCESA1CatD trimers.** **A)** Size exclusion chromatography after  
283 dialysis (first run, gray line) and after pooling, concentrating and re-fractionating (second run, black dashed line).  
284 The elution peak of the second run (12.5 ml) is enriched with trimers. **B)** Negative stain image of particles at the  
285 elution peak of the second SEC (black dashed line in (A)).  
286  
287

288 **3D reconstruction of *AtCESA1CatD* trimer by negative staining EM**

289 Freshly eluted protein samples were then used to optimize grid preparation for negative staining EM. Final images  
290 showed monodispersed and evenly distributed particles in different orientations (Figure 3B). We manually picked  
291 ~4000 particles, which were reduced to 3417 particles after 2D reference free classification using the program  
292 EMAN2, and then to 3298 from 2D alignment and classification using ISAC2<sup>34</sup>. A final set of 37 stable averages  
293 are shown in Figure 4A, with a majority exhibiting triangular shapes in distinct views that could be matched with  
294 2D projections from prior published trimer models based on small angle x-ray scattering<sup>30</sup>.

295 We used RELION (Figure 4B-4F) and EMAN2 (Figure S7) to determine 3D structure of *AtCESA1CatD*. The  
296 particles were imported into RELION 2.0 and used to generate a C3 symmetric initial model (Figure 4B). After  
297 refinement, 3D classification gave only a single class (Figure 4C and 4D) resolved to 23.2 Å (Figure 4E). There  
298 was some preferred orientation, but all perspectives were observed (Figure 4F). Tilt pair validation (Figure 4G)  
299 gave a tilt angle mean of 5.9° (± 4.9°) that was consistent with the set angle of 10 degrees, and the number of tilt  
300 pairs in the cluster was 31 out of 44 (68%). TiltStats analysis with N = 44 and Winsorizing parameter = 10% gave  
301 31 tilt pairs with >99% probability of being non-randomly distributed (R/R<sub>0</sub> = 2.8) in a cluster of 8.3° radius about  
302 a geometric mean of 3.3° with angular accuracy of 6.3° and κ of 83.



304 **Figure 4. Negative stain structure of AtCESA1CatD using RELION. A)** Stable 2D classes from ISAC2. **B)**  
305 Initial model generated using RELION. Refined model rendered as surface levels 0.11 **C)** and 0.12 **D)**,  
306 respectively. **E)** Unmasked FSC plot. **F)** Particle distribution plot (only 1/3 shown since refinement used C3  
307 symmetry). The number of particles found in given orientations is shown as the cylinders of varying height. **G)**  
308 Tilt pair validation. Black dots represent the tilt axis and tilt angle for particle pairs in polar coordinates. The red  
309 circle is centered around the expected tilt angle 10°. The outer radius of the plot is 180°.

### 310 *Identification of inter-subunit chemical cross links in AtCESA1CatD trimers*

312 We supplemented the limited resolution achieved by negative staining EM by coupling disuccinimidyl suberate  
313 (DSS) chemical cross linking of AtCESA1CatD with high resolution mass spectrometry (MS). Despite the fact  
314 that application of cross linking MS in mapping interaction interfaces in protein complexes is well established,  
315 the analysis on homo-oligomers is challenging. To clearly distinguish intra-cross linking and inter-cross linking,  
316 we generated heavy lysine labeled AtCESA1CatD monomer (CatD\_H) by feeding *E. coli* expressing the protein  
317 with modified minimal M9 broth containing heavy lysine ( $^{13}\text{C}_6\text{H}_{14}^{15}\text{N}_2\text{O}_2\cdot\text{HCl}$ ), which resulted in increased  
318 molecular mass of 8 Da per lysine residue. As indicated in Figure 5A, light (CatD\_L) and heavy monomer  
319 (CatD\_H) were mixed in 1:1 ratio to form chimeric trimer. Assuming that the oligomers were stable over the 30  
320 min cross linking reaction, which was seen for the size-exclusion chromatography reported in Figure 3, the light-  
321 heavy crosslinks clearly occurred between two subunits, while the light-light or heavy-heavy ones could be intra-  
322 or inter-subunit cross links. After cross-linking, protein was separated by SDS-PAGE, excised as trimer,  
323 proteolyzed with trypsin, and subjected to LC-MS to identify cross links (Table S1).

324 In total, 18/32 cross links observed in trimers and 7/19 cross links observed in dimers were intersubunit (i.e.,  
325 they were light-heavy). Expectedly, 0/15 crosslinks observed in monomers were light-heavy. Six of the seven  
326 inter-subunit cross links seen in dimers were also present in trimers. Of the 18 inter-subunit cross links in trimers,  
327 13 reside between lysine residues of P-CR and P-CR, while 5 are between P-CR and CSR, and none are between  
328 CSR and CSR (Figure 5B). The crosslinks are formed between a total of nine lysine residues: six within a small  
329 segment of the P-CR region, one just downstream of it, and two within a predicted disordered region near the  
330 beginning of the CSR (Figure 5C).

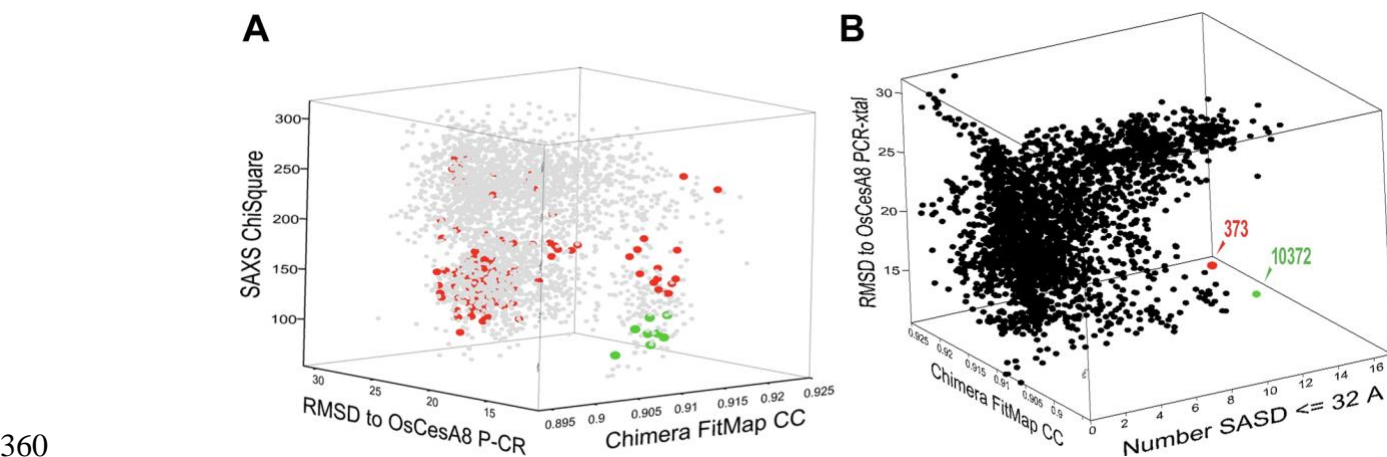


342 trimerizing portion of the P-CR (below the aligned P-CR sequences the asterisks identify the residues of the  
343 OsCESA8 and PttCESA8 P-CRs that come within 4 Å of a neighboring chain's P-CR in structures 5NJP (RCSB)  
344 and 21820 (EMDB). The secondary structures from crystal and cryoEM structures are also shown. The  
345 underlined, blue-shaded portion of PttCESA8 was present but not modeled in the cryoEM structure, and the  
346 sequence of the protein differs from the published one for PttCESA8 by substitutions P370A and S622P (red  
347 typeface) and A947T (not shown). The alignment was made using Clustal-Omega for AtCESA1CatD (from  
348 NP\_194967.1), OsCESA8 (Q84ZN6.1 GI:75149238), GhCESA7 (GI:376315426) and PttCESA8 (AAT09896.2).

349

### 350 *Correlations between Biochemical Data and AtCESA1CatD Trimer Models*

351 We used scripts and the programs of Chimera, Swiss PDB Viewer and Crysol\_30 to score each of the 3,251 trimer  
352 models reported above (see Figure 2) for cross correlation of fit into the negative stain density map for  
353 AtCESA1CatD trimer, root-mean-square deviation (RMSD) of fit between the P-CR regions and the P-CR trimer  
354 present in the crystal lattice of OsCESA8, and ChiSquare of fit to SAXS data. The highest scoring models fall  
355 into the lower right portion of the 3D scatter plot (Figure 6A). We also used the programs Jwalk and FreeSaSa to  
356 identify the number of reasonable solvent accessible pathways for lysine-DSS-lysine cross links. Those models  
357 that had 12 or more reasonable cross links are colored red or green in Figure 6A. The green colored ones fit best  
358 to the SAXS profile and P-CR structure. Figure 6B shows the distribution of trimer models ranked for the number  
359 of reasonable cross links (SASD  $\leq 32$  Å), fit to negative stain density and RMSD to OsCESA8 P-CR structure.

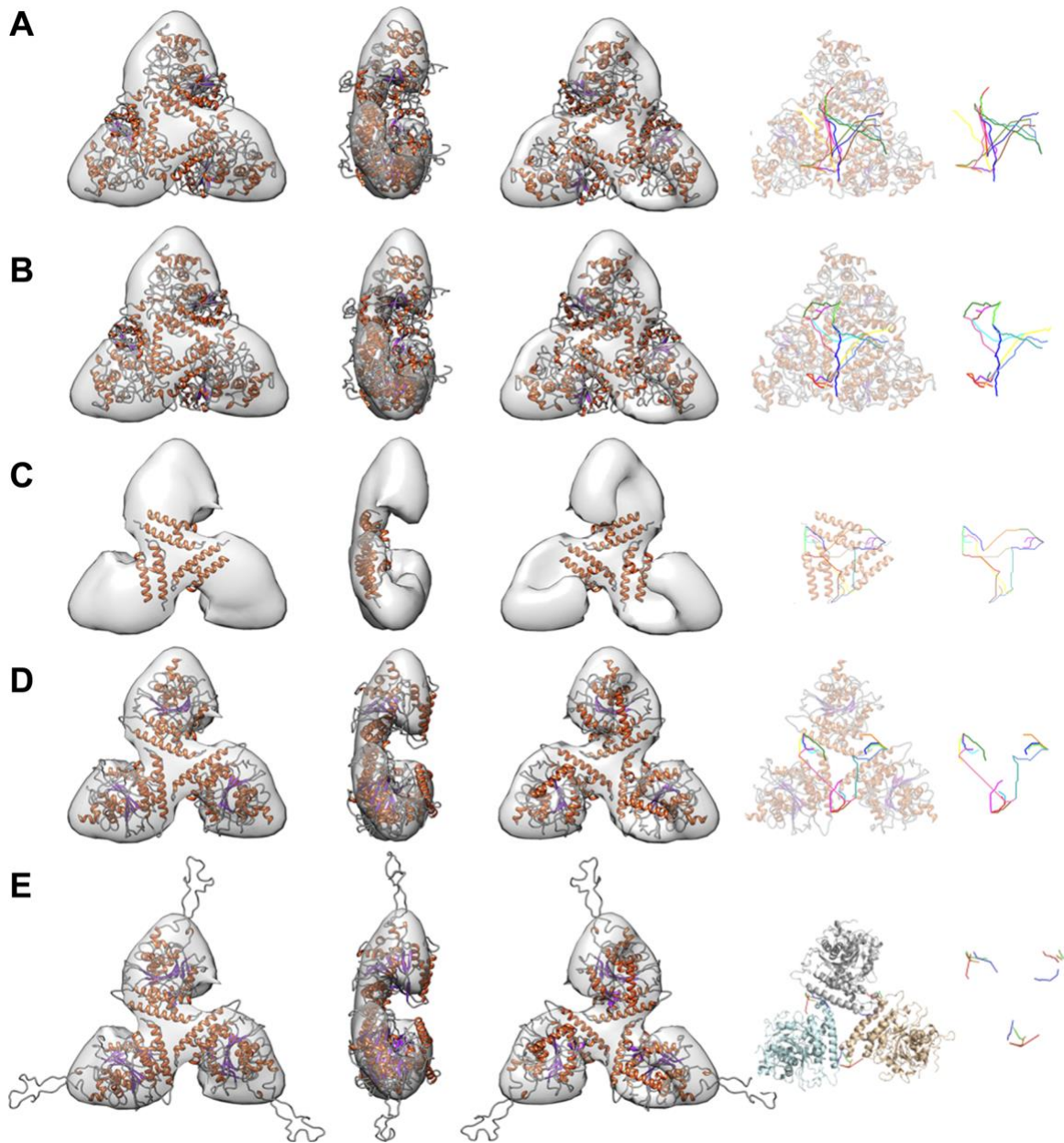


361 **Figure 6. Distribution of Scored AtCESA1CatD Trimer Models.** A) Ranked for fit to SAXS data, fit to the  
362 negative stain structure of AtCESA1CatD trimer, and RMSD to trimeric crystal structure of the OsCESA8 P-CR.

363 All computational trimer models of AtCESA1CatD trimers (gray); those with 12 or more SASD  $\leq 32$  Å (red or  
364 green); those with 12 or more SASD  $\leq 32$  Å, SAXS ChiSquare  $< 126$ , and RMSD  $< 13$  Å (green). **B)** Same  
365 ranking of the library of trimer models but with the parameter of 'SAXS ChiSquare' replaced with the number of  
366 reasonable cross links. The two models most consistent with the cross linking data are 373 (red) and 10372  
367 (green).

368

369 Two models stand out: #373 (red) and #10372 (green). They had 13 or 14 reasonable cross links (among the  
370 highest observed), lowest RMSD's to the P-CR crystal structure (14.9 Å and 12.5 Å, respectively), and displayed  
371 reasonable fits to the low-resolution EM density of AtCESA1CatD trimer (cross correlations of 0.92 and 0.91,  
372 respectively). Three other models scored closely to these two but supported fewer reasonable cross links. The  
373 negative stain density map is overlain with the two best models next to a mapping of their reasonable cross links  
374 in Figure 7A and 7B.



375

376

377

378

379

380

381

382

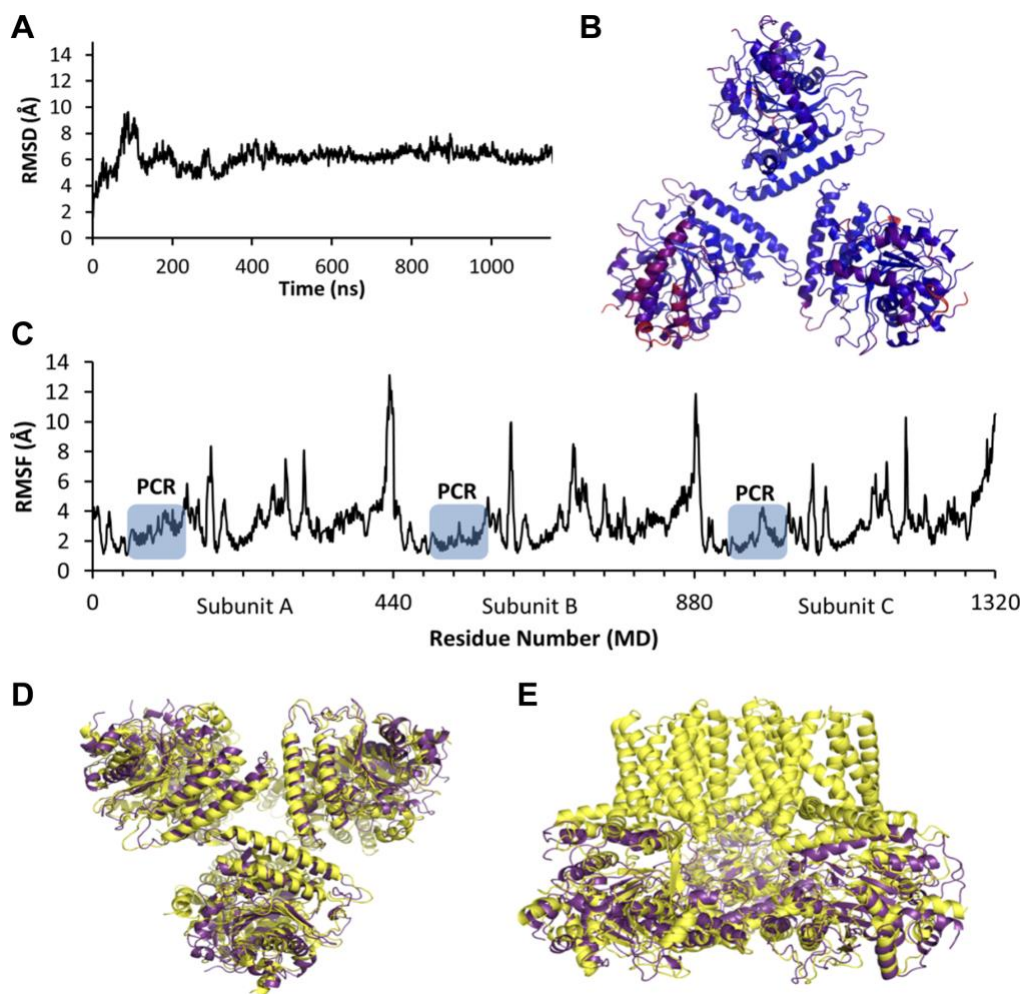
383

**Figure 7. Overlay of AtCESA1CatD negative stain density with best computational trimers, OsCESA8 P-CR and PttCESA8.** Columns 1-3 show top, side and bottom views of AtCESA1CatD negative stain density overlain with computational models **A)** #10372, **B)** #373 or **C)** OsCESA8 P-CR crystal structure and **D)** PttCESA8CatD cryoEM structure. Columns 4 and 5 show the reasonable DSS cross links for each model. For simplicity we only show the CatD portions of PttCESA8 in (D). Identical cross links were seen for the full PttCESA8 structure. **E)** Energy minimized SwissModel for AtCESA1 using PttCESA8 structure. The unstructured loop is part of the CSR.

384 In both ‘best’ model trimers, the P-CR helices bearing cross linkable lysines form a triangular contact surface  
385 very similar to the one seen in the OsCESA8 P-CR crystal structure and the PttCESA8 cryoEM structure. The  
386 cross linking lysine residues found in AtCESA1CatD are highly conserved (among most if not all CESA proteins)  
387 and present in both of these structures. When screened for capacity to support reasonable DSS cross links, 14 of  
388 those we observed in AtCESA1CatD were found in the poplar trimer structure (the number did not change if  
389 membrane spanning sequences were present or absent) and 15 were seen in the rice P-CR trimer (Figure 7C and  
390 7D). Encouraged by this agreement between trimer structures of PttCESA8 and AtCESA1, we used SwissModel  
391 <sup>39</sup>, a fully automated protein structure homology-modeling server, to generate a final model of AtCESA1 trimers  
392 based on the cryoEM structure – the CatD regions are shown in Figure 7E. Note that the loop extending beyond  
393 the negative stain density is a portion of the CSR that was not present in the cryoEM structure.

#### 394 *Molecular dynamics simulations of the AtCESA1CatD trimer*

395 To investigate stability of the AtCESA1CatD trimer, we performed conventional molecular dynamics (MD)  
396 simulations on a theoretical model of the AtCESA1CatD trimer. After the trimer system equilibrated to a stable  
397 RMSD, in about 150 ns (Figure 8A and 8B), a subsequent 1,000-ns-long production run was performed. The P-  
398 CR domains (ALA398-LYS475) were among a few regions showing the least conformational fluctuation (2 to 4  
399 angstroms) as measured by the atomic root mean squared fluctuations (RMSF; Figure 8C). The stability of the P-  
400 CR domains maintained the triangular interface between three AtCESA1CatD monomers during the simulation  
401 (Figure 8D and 8E). The most dynamic regions showing higher RMSF (>6 angstroms) were the loops connected  
402 to the truncated CSR region and the loops connected to the truncated transmembrane region.

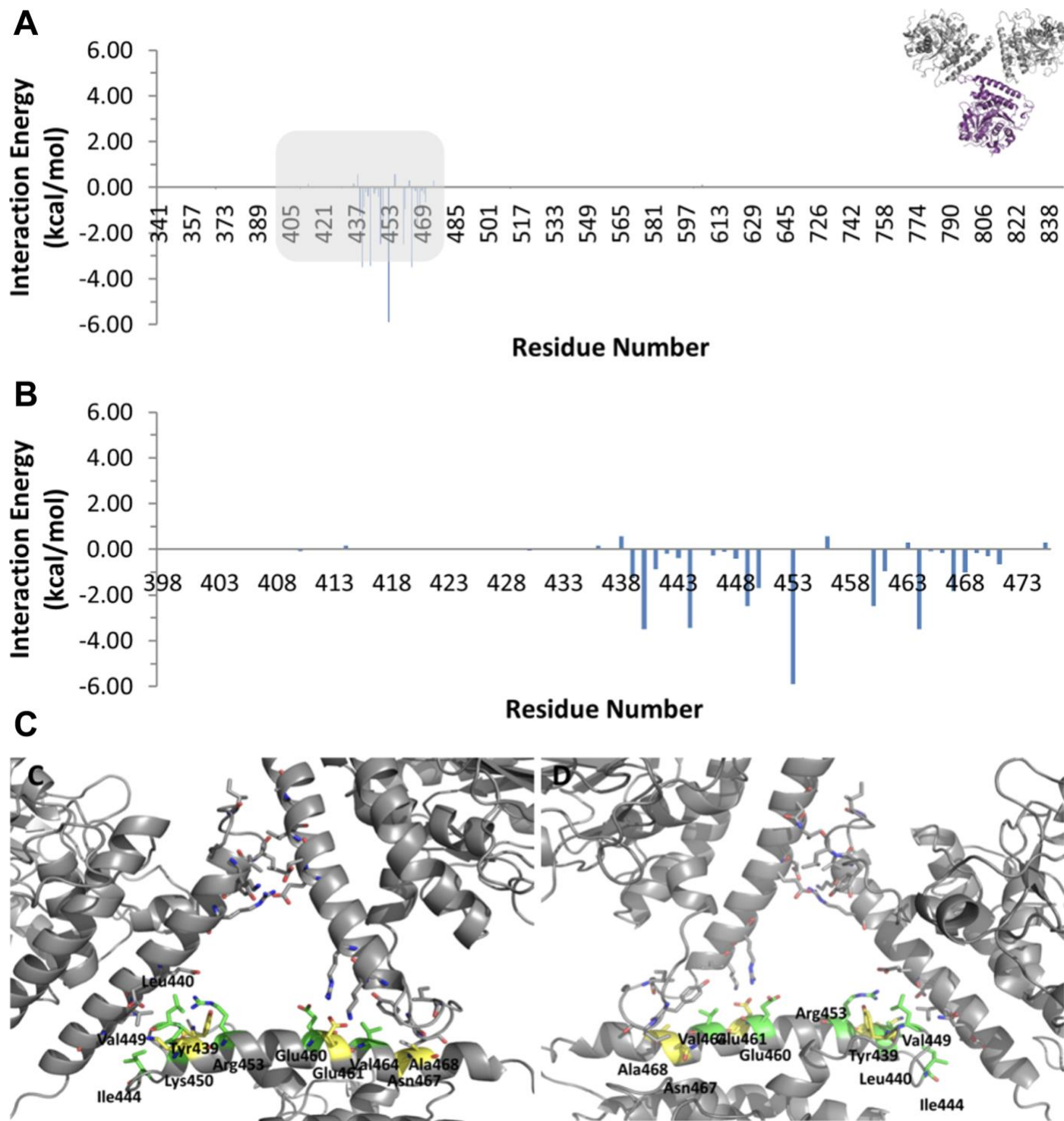


**Figure 8.** 1,000-ns-long MD simulation performed on the theoretical model of the AtCESA1CatD trimer. **A)** The RMSD profile of the AtCESA1CatD trimer during the 1,000-ns MDS. **B)** The final structure of the AtCESA1CatD trimer. The color represents the conformational fluctuations of the AtCESA1CatD protein measured as the RMSF, ranging from values less than 1 Å (blue) to values greater than 6 Å (red). **C)** RMSF profile of the AtCESA1CatD trimer residues during the 1,000-ns-long production MDS run. **D)** The bottom view and **E)** the side view of the superimposition of the CatD domain of the AtCESA1CatD trimer final structure and the cryoEM structure of the PttCESA8 trimer (PDB ID: 6WLB), shown in purple and yellow, respectively.

The binding free energies between one AtCESA1 monomer to the other two is  $\sim 56.4$  kcal/mol, as calculated using the MM-PBSA approach (Figure 9). Eleven residues that contributed significant ( $< -1$  kcal/mol) binding free energies are all located in the P-CR region. R453 contributed the most,  $\sim -5.9$  kcal/mol, to the binding free energy. L440, I444, and V464 each contributed  $\sim -3.5$  kcal/mol to the binding free energy. V449 and E460 each

416 contributed  $\sim$ -2.5 kcal/mol to the binding free energy. N467, K450, Y439, E461, and A468 each contributed -1~  
417 2 kcal/mol.

418



419

420 **Figure 9. Binding free energies between one AtCESA1CatD (colored in purple) to the other two**  
421 **AtCESA1CatD monomers (colored in grey). A) Per-residue decomposition binding free energies for residues**  
422 **of the AtCESA1CatD colored in purple. The shaded region represents the P-CR domain. B) A zoom-in view of**

423 the per-residue decomposition binding free energies for residues in the P-CR domain. **C)** A bottom view of the  
424 triangular shape interface formed by P-CR domains in the AtCESA1CatD trimer final structure. **D)** A top view  
425 of the triangular shape interface formed by P-CR domains in the AtCESA1CatD trimer final structure. All 11  
426 residues that contributed significant binding free energies ( $< -1$  kcal/mol) are shown as sticks. Hydrogen atoms  
427 are not shown. Residues contributed binding free energies greater than 2 kcal/mol are colored in green. Residues  
428 contributed binding free energies between 1 to 2 kcal/mol are colored in yellow.

429  
430 To investigate potential effects of the CSR region on the AtCESA1CatD trimer, we replaced the AtCESA1  
431 catalytic domains with an AtCESA1CatD model containing the CSR region obtained from the program trRosetta,  
432 a newly developed web-based homology modeling server based on deep learning and Rosetta<sup>44</sup> that predicted  
433 more structure for the CSR than predicted by the SwissModel program. Both the above trimer and this new one  
434 have similar RMSD profiles, as shown in Figure S6E and S6G. Figure S6F and S6H showed that both trimers  
435 preserved relative constant distances between each P-CR domain ( $\sim 30$  Å) and between each AtCESA1CatD unit  
436 ( $\sim 60$  Å and  $\sim 70$  Å, respectively), maintaining an equilateral triangle shape during the MD simulations. As shown  
437 in Figure S2, during 500 ns MD simulations, the P-CR domains remained in an equilateral triangle shape, and the  
438 catalytic domains largely maintained the parallel direction. Thus, the CSR region did not demonstrate a dramatic  
439 effect on the stability of the AtCESA1CatD homotrimer.

#### 440 **MD simulations computed stability of heterologous trimers of AtCESA1,3,6**

441 Because there is significant *in vivo* evidence that plants use heterotrimers of CESAs to make CSCs for synthesis  
442 of primary and secondary walls, we replaced two of the AtCESA1 catalytic domains with models of AtCESA3  
443 and AtCESA6, forming two possible heterotrimers with counterclockwise and clockwise distributions of the  
444 isomers. MD simulations of those two heterotrimers showed that the P-CR/P-CR contacts were stable. Figure  
445 S6B and S6D showed that the distances between each P-CR domain are  $\sim 30$  Å, maintaining an equilateral triangle  
446 shape during the MD simulations. However, the RMSD profiles (Figure S6A and S6C) showed that both  
447 heterotrimers shifted dramatically from their original conformation. Comparing structures obtained from the MD  
448 simulations (Figures S4 and S5), we found that during 500 ns MD simulations, the P-CR domains remained in an

449 equilateral triangle shape. In contrast, the catalytic domains did not maintain the original parallel directions. As  
450 shown in Figure S4, AtCESA3CatD tilted away by ~45 degrees in the counterclockwise model, whereas  
451 AtCESA1CatD and AtCESA6CatD leaned towards each other by ~15 degrees. These movements resulted in  
452 decreased distance between AtCESA1CatD and AtCESA6CatD, and increased distances between AtCESA3CatD  
453 and AtCESA1CatD, and AtCESA3CatD and AtCESA6CatD (Figure S6B). In the clockwise model, the catalytic  
454 domains of AtCesA6 tilted by only ~20 degrees (Figure S5) and the distances between each catalytic domain did  
455 not change dramatically (Figure S6D).

456 Theoretical DSS cross-links were calculated using Jwalk with SASD for ten representative structures obtained  
457 from the 1 $\mu$ s MDS on AtCESA1CatD trimer (Table S2). We found that of the total 13 inter-subunit cross links  
458 between P-CR and P-CR (Table S1 and Figure 5B) identified by the DSS chemical cross linking MS method, 11  
459 of them were confirmed by this final model of AtCESA1 trimers, shown in Figure S8. As shown in Figure S8A,  
460 K421-K472, K441-K472, K443-K472, and K450-K472 cross-links identified by the MS method were also found  
461 in the AtCESA1CatD trimer model, with SASD less than 32 Å, connecting two adjacent P-CRs at the vertices of  
462 the P-CR/P-CR/P-CR triangle (Table S3). To investigate if the dynamics of one subunit will account for the other  
463 cross-links between P-CR and P-CR, one subunit was removed from the trimer model and docked back with  
464 various different orientations. As shown in Figure S8B, the K421-K441, K421-K443, and K421-K450 cross-  
465 links, and Figure S8C, the K421-K537, K441-K537, K443-K537, and K450-K537 cross-links, were also found  
466 in the AtCESA1CatD trimer models after docking a green subunit tilted by ~90 and ~130 degrees, respectively,  
467 while preserving the P-CR contacts.

468 Of the 18 inter-subunit DSS chemical cross links (Table S1) 5 reside between lysine residues of P-CR and  
469 CSR. Since the AtCESA1CatD trimer was constructed via the P-CR/P-CR/P-CR contacts, the 5 P-CR/CSR cross  
470 links probably indicated inter-trimer interactions. Consistent with that possibility, docking one trimer to a second  
471 one rotated by ~60 degrees, as shown in Figure S9, resulted in a dimer of AtCESA1CatD trimers containing all 5  
472 P-CR/CSR cross links (SASD < 32 Å).

## 474 DISCUSSION

## **Evidence for a P-CR-based trimerization subdomain in primary and secondary cellulose synthases**

Cellulose is synthesized in single cells and multicellular life forms. The basic structure for catalysis appears common to the GT2 family of synthases, however the oligomeric state of the synthases appears to vary depending on whether the organism is making non-crystalline or crystalline cellulose. Most bacteria make non-crystalline cellulose and appear to possess monomeric synthases. Algae, plants and a few bacteria and animals make crystalline cellulose, and they appear to have higher ordered assemblies of synthases<sup>6, 48-51</sup>. It is widely believed that this ordered juxtaposition of individual enzymes is crucial for facilitating alignment and coalescence of emerging  $\beta$ -glucan chains into crystalline cellulose<sup>30, 52-53</sup>. In particular, plants use several distinct isoforms of synthase to build cellulose synthase complexes that appear to be hexamers of trimeric lobes of CESA proteins. Lobes comprised of CESA1/CESA3/CESA6 or CESA6-like contribute to synthesis of relatively thin, primary cell walls and those made of CESA4/CESA7/CESA8 are used for making thicker secondary cell walls. Here we provide evidence that a portion of the plant conserved region (P-CR) of AtCESA1 self assembles to form trimers *in vitro*, which may contribute to a trimeric assembly of full-length synthases as present in the lobes of primary cellulose synthase complexes.

Previously we reported analysis of freeze-fracture transmission electron micrographs that was best reconciled with trimeric lobes of CESA proteins in cellulose synthase complexes<sup>7</sup>. At about the same time, we reported that the catalytic domain of AtCESA1 studied here forms trimers even when expressed and purified in the absence of the membrane spanning helices and other domains of AtCESA1<sup>30</sup>. At that time, only computational models of synthase monomers and subsequently built trimers were available. In 2017 Rushton et al. published a crystal structure of the OsCESA8 P-CR fragment in isolation of the rest of the protein<sup>28</sup>. The asymmetric unit for that structure was dimeric, but in the crystal lattice one saw trimers of the P-CR. Based on size exclusion chromatography and SAXS data for isolated catalytic domain of rice CESA8 and then current computational models of CESA, the authors suggested that the dimer form was biologically relevant and that the trimeric form was an artifact. Based on that proposition, it was noted that while the trimeric form could explain trimerization of the AtCESA1CatD domain, it was likely an artifact of isolating the domain from the rest of the protein.

500 Here we propose that the trimer form seen in AtCESA1CatD is likely to be a functionally relevant form.  
501 Computational models have improved and now possess P-CR regions similar to those seen in the crystal structure  
502 <sup>7,53</sup>. These were used to generate a library of trimer models, from which to pull candidates that favor trimerization  
503 via CSR or P-CR domains <sup>44</sup>. To supplement the reported physical restraints used to screen the library, we applied  
504 four biochemical constraints to screen a similar library of AtCESA1CatD models. The effort brings into focus a  
505 small set of models that turn out to also be consistent with the new cryoEM structures for homotrimers of poplar  
506 CESA8 and cotton CESA7. The most discriminating information is in the set of DSS cross linking data. The  
507 SAXS data and negative stain structure envelopes provide some discrimination, but simply lack sufficient detail  
508 to be very selective. However, all of the models previously considered for AtCESA1CatD <sup>30</sup> supported zero  
509 reasonable solvent accessible pathways for the lysine-DSS-lysine cross links. Analysis of the recent set of cotton  
510 CESA1 models <sup>45</sup> for reasonable cross links points to models showing the same trimerization motif we see for  
511 AtCESA1CatD (data not shown). Among all the thousands of models screened, the ones providing the largest  
512 number of reasonable cross links are the ones most similar to the structures determined independently for rice,  
513 poplar and cotton CESA proteins. Finally, a microsecond MD simulation confirmed that AtCesA1CatD forms a  
514 stable homotrimer. P-CR domains, especially residues R453, L440, I444, V464, V449, E460, N467, K450, Y49,  
515 E461, and A468, played a crucial role in the stability of the homotrimer. MD simulations also found a stable  
516 heterotrimer of AtCESA1, AtCESA3 and AtCESA6 catalytic domains arranged in the clockwise and counter-  
517 clockwise direction, demonstrating that P-CR/P-CR contacts were stable but the rest of the catalytic domains  
518 were less stable than the homotrimer of AtCESA1CatD. We thus propose that self-association by the P-CR  
519 element contributes to trimerization of plant CESAs, both in those making primary (AtCESA1) and secondary  
520 (OsCESA8, PttCESA8, GhCESA7) walls. We note that the two cryoEM structures of homotrimers of CESAs  
521 show interaction between TMH 7 of two subunits. Our observation that the heterotrimer of AtCESA1, AtCESA3  
522 and AtCESA6 catalytic domains was less stable than the homotrimer of AtCESA1CatD supports the idea that the  
523 TM helix interaction is crucial for stability of heterotrimers of CESAs.

524 Recently, crystallographic data revealed that AtCESA3CatD can form homodimers in solution <sup>29</sup>. The interface  
525 between subunits in the dimer structure requires altering the structural context of residues that form a linkage

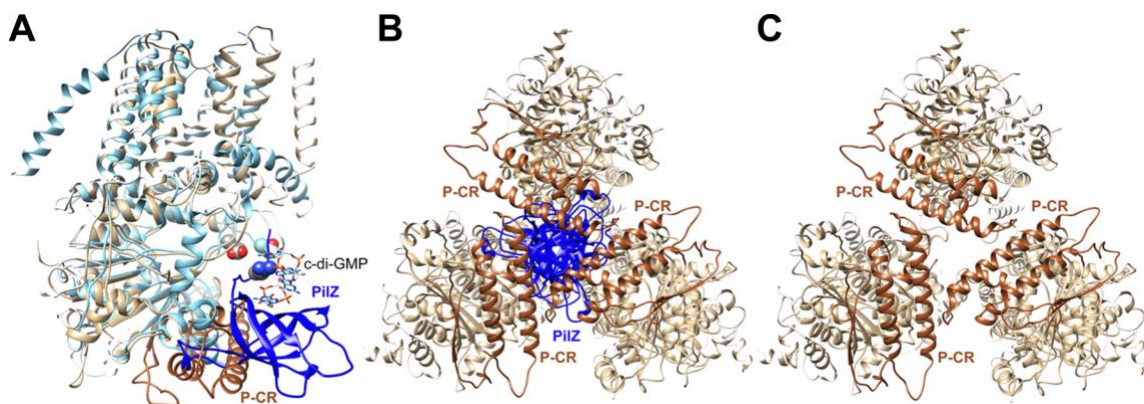
526 between the finger helix of the catalytic domain and the helix interfacing with the transmembrane helices in the  
527 poplar CESA8 and cotton CESA7 cryoEM structures. The same study showed *in vivo* fluorescence data consistent  
528 with oligomers – dimers or larger. Mutations that altered residues critical for the dimerization interface also  
529 disrupted the *in vivo* oligomerization, causing the authors to hypothesize a dimer phase in the assembly of CESAs.  
530 This is consistent with prior work from which an assembly model was proposed for making secondary CSCs, in  
531 which homodimers of isoforms CESA4, CESA7 and CESA8 assemble and then further form tetramers and  
532 hexamers to make CSCs<sup>16</sup>. However, it is not known if the reported AtCESA3CatD dimer structure is present in  
533 the dimers seen for AtCESA7. The latter were shown to be partially sensitive to reducing agents, implying that  
534 disulfide crosslinks were important for their stability in SDS-PAGE. The subunit-subunit, DSS crosslinks reported  
535 here for trimeric and dimeric forms of the AtCESA1CatD that were stable in SDS-PAGE are not plausible in the  
536 dimeric CESA3CatD crystal structure, but they are plausible in the trimeric structures of full-length forms of  
537 poplar CESA8 and cotton CESA7. This can be directly assessed because the lysine residues involved in DSS  
538 crosslinks are conserved in AtCESA3CatD (and other CESAs). An alternative hypothesis explaining the  
539 formation of reduction-sensitive dimers of CESA7 is that cysteines in the carboxy terminus, located in the  
540 apoplast outside of the cell plasma membrane, form disulfide bonds joining two monomers in a trimeric lobe.  
541 Detergent purification of those crosslinked dimers could have given rise to the CESA7 observations.

542 Since the subunit-subunit DSS crosslinks seen in SDS-PAGE stable dimers reported here are consistent with  
543 the P-CR based trimer structure, it follows that if there were dimers similar to those present in the AtCESA3CatD  
544 crystal structure they were present in amounts below that needed for detection in the mass spectrometry analysis.  
545 The sequence of AtCESA3CatD and AtCESA1CatD are quite similar but not identical. AtCESA1CatD does  
546 contain the residues of the dimerization motif described for AtCESA3CatD but a vector-derived N-terminal tag  
547 is in the AtCESA1CatD protein that is not present in the AtCESA3CatD protein. There was clearly sub-trimer  
548 state material present in the first gel filtration steps we took to purify trimers of AtCESA1CatD, but we removed  
549 much of the smaller material and did not study it further. Further analysis of the discarded material might reveal  
550 dimers similar to the crystal structure of AtCESA3CatD.

We also observed some inter-subunit crosslinks between P-CR and CSR regions, but these were not supported by any of the best ranking trimer models nor the cryoEM structures for PttCESA7 and GhCESA8 for which the CSR is largely ‘unstructured’. Because the P-CR/CSR crosslinks could be reasonably modeled between closely apposed trimers of AtCESA1CatD we tentatively suggest that they arose from trimer-trimer interactions and could be *in vitro* artifacts of the N-terminal and TM-domain truncations. Further studies are necessary to determine if assembly and delivery of CESA trimers to cellulose synthase complexes involves a dimeric phase, and to define the juxtaposition of P-CR and CSR regions in CSCs.

### The trimerizing motif of the P-CR replaces regulatory PilZ domain of bacterial CESA

An overlay of the bacterial cellulose synthase structure with the cryoEM structure of poplar CESA8 reveals that the P-CR trimerization motif physically occupies the same space next to the catalytic fold as does the regulatory PilZ domain of bacterial synthases (Figure 10).



**Figure 10. Similar juxtapositions of PilZ and P-CR within bacterial and plant cellulose synthases. A)** Overlay of BcsA (light and dark blue; PDB 4P02) and poplar CesA8 (tan and brown; PDB 6WLB). Dark colors highlight bacterial PilZ and plant P-CR regions. **B)** Overlay of three BcsA subunits on three CESA8 subunits in the homotrimer structure, with only PilZ domains (dark blue) rendered for the bacterial enzymes. **C)** Same as B, with PilZ domains removed to highlight P-CR interactions.

It has also been pointed out that putative bacterial synthases may have PilZ, replace it with another domain, or simply lack anything in its place<sup>55</sup> (e.g., a Pfam search of *Acaryochloris marina* WP\_012165112.1 reveals a putative cellulose synthase containing a region homologous to the Pfam domain “histidine kinase-like ATPase”).

572 These observations raise the possibility that in the evolution of plants, this malleable spatial region of the enzyme  
573 fold has become specialized to facilitate trimerization. Moreover, since a small portion of the P-CR is the major  
574 place within the cytosolic domains of the trimer where monomers make direct contact, it may also be a conduit  
575 for dynamic communication between subunits. Such communication could certainly contribute to regulation of  
576 catalytic activity within a lobe, perhaps synchronizing chain synthesis and translocation across membranes in a  
577 ‘coalescence-friendly’ manner. Replacing subunits with P-CR regions of different isoforms could fine tune such  
578 communication, which could also be influenced by contact between neighboring lobes, perhaps mediated by CSR  
579 or amino-terminal domain interactions. As pointed out by Zhong et al., a previously studied mutant of AtCESA8,  
580 *fra6*, maps to P-CR (R362K, same position as 453 of AtCESA1CatD in this study) <sup>56</sup>. Notably, residue R453 of  
581 AtCESA1CatD contributed the most stabilizing energy for maintaining P-CR based trimers during MD  
582 simulations.

## 583

## 584 CONCLUSIONS

585 Recent evidence suggests that plant CESAs self-assemble into trimeric lobes that further assemble into a larger  
586 complex containing 18 monomers called the CSC or rosette. We report purification to homogeneity of a self-  
587 assembled trimeric complex of the catalytic domain of AtCESA1. 3D-structure resolved to 23.2 Å was obtained  
588 via reconstruction of negatively stained single particles imaged by TEM. Further chemical cross linking followed  
589 by high resolution mass spectrometry was performed to identify the interaction interface between monomers,  
590 which revealed a set of lysine residues that cluster in a loop and helix of the P-CR domain play a crucial role in  
591 the stabilization of homotrimer. Finally, MD simulations verified that 11 residues R453, L440, I444, V464, V449,  
592 E460, N467, K450, Y49, E461, and A468 in P-CR domains contribute significantly to the binding free energy.  
593 MD simulations also found two stable heterotrimers of AtCESA1, AtCESA3 and AtCESA6 catalytic domains  
594 with counterclockwise and clockwise organization of monomers, respectively. These results provide important  
595 experimental evidence to support the role of P-CR domain in the trimeric assembly of CESAs and lay the  
596 foundation for the future study on the function of important residues by mutagenesis.

598 **ASSOCIATED CONTENT**

599 **Supporting Information**

- 600 1. Figure S1 - Fragmentation spectra of inter-subunit crosslinked peptides identified from trimers by pLink
- 601 2. Table S1 - DSS Cross links in trimer of AtCESA1CatD.
- 602 3. Figure S2 - Representative structures of the homotrimer of AtCESA1 catalytic domain.
- 603 4. Table S2 - Sequence alignment of AtCesA1, AtCesA3 and AtCesA6.
- 604 5. Figure S3 - Structural alignment of AtCESA1 models obtained from Swiss Model and trRosetta.
- 605 6. Figures S4 - Representative MD simulations structures of catalytic domain heterotrimers of AtCESA1,
- 606 AtCESA3 and AtCESA6 arranged counterclockwise.
- 607 7. Figures S5 - Representative MD simulations structures of catalytic domain heterotrimers of AtCESA1,
- 608 AtCESA3 and AtCESA6 arranged clockwise.
- 609 8. Figure S6 - RMSD profiles, distances between subunits and between P-CR domains during MDS.
- 610 9. Figure S7 – 3D reconstruction of AtCESA1CatD trimer by software suite EMAN2.
- 611 10. Figure S8 - Inter-subunit chemical cross links between lysine residues of P-CR and P-CR.
- 612 11. Table S3 - DSS Cross links in trimer of AtCESA1CatD MD models.
- 613 12. Figure S9 - Inter-subunit chemical cross links identified by the MS method suggesting a P-CR/CSR
- 614 dimerization between two AtCESA1CatD trimers.

615 **EMDB Submissions**

617 Negative stain reconstructions for monomer and trimer of AtCESA1CatD have been deposited to the EMDB as

618 D\_1000260420 and D\_1000260421.

619 **AUTHOR INFORMATION**

621 **Corresponding Author**

622 B. Tracy Nixon – Department of Biochemistry and Molecular Biology, Pennsylvania State University,

623 University Park, PA 16802; orcid.org/0000-0002-3568-5703; Email: btn1@psu.edu

624 **Authors**

625 **Juan Du** - Department of Biochemistry and Molecular Biology, Pennsylvania State University, University  
626 Park, PA 16802. Current address: Key Laboratory of Bio-Resource and Eco-Environment of Ministry of  
627 Education, College of Life Sciences, Sichuan University, Chengdu, 610064 China; orcid.org/0000-0001-  
628 9380-8831.

629 **Venu Gopal Vandavasi** - Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge,  
630 Tennessee 37831. Current address: Department of Chemistry, Princeton University, Princeton, NJ 08544;  
631 orcid.org/0000-0002-8894-1395.

632 **Kelly R. Molloy** - Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University,  
633 New York, NY 10065.

634 **Hui Yang** - Department of Biology, The Pennsylvania State University, University Park, PA 16802. Current  
635 address: Schrödinger, Inc., New York City, New York 10036; orcid.org/0000-0001-6331-6801

636 **Lynnica Massenburg** - Department of Biochemistry and Molecular Biology, Pennsylvania State University,  
637 University Park, PA 16802; orcid.org/0000-0002-6590-273X.

638 **Abhishek Singh** - Department of Materials Science and Engineering, North Carolina State University,  
639 Raleigh, NC 27695. Current address: Computational Biology, Digitalization in Bioscience R&D, BASF  
640 Corporation, Morrisville, NC 27560; orcid.org/0000-0002-9450-7229.

641 **Albert L. Kwansa** - Department of Materials Science and Engineering, North Carolina State University,  
642 Raleigh, NC 27695

643 **Yaroslava G. Yingling** - Department of Materials Science and Engineering, North Carolina State University,  
644 Raleigh, NC 27695; orcid.org/0000-0002-8557-9992.

645 **Hugh O'Neill** - Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831;  
646 orcid.org/0000-0003-2966-5527.

647 **Brian T. Chait** - Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University,  
648 New York, NY 10065; orcid.org/0000-0003-3524-557X.

649 **Manish Kumar** - Department of Chemical Engineering, The Pennsylvania State University, University Park,  
650 PA 16802. Current address: Dept of Civil, Architectural and Environmental Engineering, University of  
651 Texas at Austin, Austin, TX, 78712; orcid.org/0000-0001-5545-3793.

### 653 **Author Contributions**

654 The manuscript was written through contributions of all authors. All authors have given approval to the final  
655 version of the manuscript. J.D. conceived the original research plans; B.T.N., M.K., B.T.C., H.O. and Y.Y.  
656 supervised the experiments; J.D. performed most of the experiments with protein initially provided by V.G.V.;  
657 subsequent TEM experiments were performed by L.M.; mass spectrometry performed by K.R.M.; molecular  
658 dynamic simulations were performed by H.Y.; J.D., L.M., H.Y., B.T.N., B.T.C., K.R.M., Y.Y, A.S. and A.L.K.  
659 designed the experiments and analyzed the data; J.D. wrote the article with contributions of all the authors; B.T.N.  
660 supervised and complemented the writing.

### 662 **Notes**

663 The authors declare no competing financial interests.

### 665 **ACKNOWLEDGMENTS**

666 This work was supported in part by the Center for Lignocellulose Structure and Formation, an Energy Frontier  
667 Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences (grant-no:  
668 DE-SC0001090), and P41 GM109824 and P41 GM103314 (to B.T.C.). We thank Sung Hyun Cho (The  
669 Pennsylvania State University) and Chaowen Xiao (Sichuan University) for insightful discussions.  
670 Transmission electron microscopic experiments were performed at the Huck Institutes of the Life Sciences  
671 Microscopy Facility at Penn State – University Park, PA, USA.

### 673 **ABBREVIATIONS**

674 2D – two dimensional

675	3D	– three dimensional
676	At	– <i>Arabidopsis thaliana</i>
677	CatD	– catalytic domain
678	CatD_H	– catalytic domain containing heavy isotope of lysine
679	CatD_L	– catalytic domain containing light isotope of lysine
680	CCD	– charge coupled device electron detector
681	CESA#	the number refers to a particular isoform of CESA
682	CESAs	cellulose synthases
683	CMF	cellulose microfibrils
684	COM	center of mass
685	cryoEM	cryo-electron microscopy
686	CSCs	cellulose synthase complexes
687	CSR	hypervariable or class-specific region of plant CESA
688	CTF	contrast transfer function
689	DMSO	dimethyl sulfoxide
690	DSS	disuccinimidyl suberate
691	DTT	dithiotheitol
692	EMDB	Electron Microscopy Data Bank
693	FSC	Fourier shell correlation
694	Gh	<i>Gossypium hirsutum</i> (cotton)
695	GT	glycosyltransferase
696	GT2	glycosyltransferase family 2
697	HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
698	ISAC	iterative stable alignment and clustering
699	LC	liquid chromatography
700	LC-MS	liquid chromatography-mass spectrometry

701	LDS	lauryl dodecyl sulfate
702	MD	molecular dynamics
703	MOPS	3-(N-morpholino)propanesulfonic acid
704	MS1, MS2	first and second mass spectrum in tandem mass spectrometry
705	NH <sub>4</sub> <sup>+</sup>	ammonium ion
706	NMR	nuclear magnetic resonance
707	NPT	isothermal-isobaric ensemble for MD simulations
708	NVT	canonical ensemble for MD simulations
709	Os	<i>Oryza sativa</i> (rice)
710	PDB	Protein Data Bank
711	P-CR	plant conserved region of CESA
712	Ptt	<i>Populus tremula x tremuloides</i> – poplar
713	RMSD	root mean squared deviation
714	RMSF	root mean squared fluctuation
715	SASD	surface accessible solvent distance
716	SAXS	small angle X-ray scattering
717	SDS	sodium dodecyl sulfate
718	SDS-PAGE	polyacrylamide gel electrophoresis with SDS
719	SEC	size exclusion chromatography
720	TCEP	Tris(2-carboxyethyl)phosphine
721	TEM	transmission electron microscopy
722	TIP3P	water molecule for MD simulations
723	TMH	transmembrane helix
724		
725		
726		

727 **REFERENCES**

- 728 (1) Anderson, W. F.; Akin, D. E., Structural and chemical properties of grass lignocelluloses related to  
729 conversion for biofuels. *J. Ind. Microbiol. Biotechnol.* **2008**, *35* (5), 355-366.
- 730 (2) Nookaraju, A.; Pandey, S. K.; Bae, H. J.; Joshi, C. P., Designing cell walls for improved bioenergy  
731 production. *Mol. Plant* **2013**, *6* (1), 8-10.
- 732 (3) Pauly, M.; Keegstra, K., Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant*  
733 *Journal* **2008**, *54*, 559-568.
- 734 (4) Zhu, H.; Luo, W.; Ciesielski, P. N.; Fang, Z.; Zhu, J. Y.; Henriksson, G.; Himmel, M. E.; Hu, L., Wood-  
735 Derived Materials for Green Electronics, Biological Devices, and Energy Applications. *Chem. Rev.* **2016**, *116*,  
736 9305-9374.
- 737 (5) Mueller, S. C.; Brown, R. M., Jr.; Scott, T. K., Cellulosic microfibrils: nascent stages of synthesis in a  
738 higher plant cell. *Science* **1976**, *194* (4268), 949-951.
- 739 (6) Kimura, S.; Laosinchai, W.; Itoh, T.; Cui, X.; Linder, C. R.; Brown, R. M., Jr., Immunogold labeling of  
740 rosette terminal cellulose-synthesizing complexes in the vascular plant vigna angularis. *Plant Cell* **1999**, *11*  
741 (11), 2075-2086.
- 742 (7) Nixon, B. T.; Mansouri, K.; Singh, A.; Du, J.; Davis, J. K.; Lee, J. G.; Slabaugh, E.; Vandavasi, V. G.;  
743 O'Neill, H.; Roberts, E. M.; Roberts, A. W.; Yingling, Y. G.; Haigler, C. H., Comparative Structural and  
744 Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Sci.*  
745 *Rep.* **2016**, *6*, 28696. <https://doi.org/10.1038/srep28696>.
- 746 (8) Brown, R. M. J.; Montezinos, D., Cellulose microfibrils: visualization of biosynthetic and orienting  
747 complexes in association with the plasma membrane. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73* (1), 143-147.
- 748 (9) Mueller, S. C.; Brown, R. M. J., Evidence for an intramembrane component associated with a cellulose  
749 microfibril-synthesizing complex in higher plants. *J. Cell Biol.* **1980**, *84* (2), 315-326.
- 750 (10) Persson, S.; Paredez, A.; Carroll, A.; Palsdottir, H.; Doblin, M.; Poindexter, P.; Khitrov, N.; Auer, M.;  
751 Somerville, C. R., Genetic evidence for three unique components in primary cell-wall cellulose synthase  
752 complexes in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (39), 15566-15571.
- 753 (11) Taylor, N. G.; Laurie, S.; Turner, S. R., Multiple cellulose synthase catalytic subunits are required for  
754 cellulose synthesis in Arabidopsis. *Plant Cell* **2000**, *12* (12), 2529-2540.
- 755 (12) Desprez, T.; Juraniec, M.; Crowell, E. F.; Jouy, H.; Pochylova, Z.; Parcy, F.; Hofte, H.; Gonneau, M.;  
756 Vernhettes, S., Organization of cellulose synthase complexes involved in primary cell wall synthesis in  
757 Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (39), 15572-15577.
- 758 (13) Taylor, N. G.; Howells, R. M.; Huttly, A. K.; Vickers, K.; Turner, S. R., Interactions among three distinct  
759 CesA proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (3), 1450-1455.
- 760 (14) Gardiner, J. C.; Taylor, N. G.; Turner, S. R., Control of cellulose synthase complex localization in  
761 developing xylem. *Plant Cell* **2003**, *15* (8), 1740-1748.

- 762 (15) Wang, J.; Elliott, J. E.; Williamson, R. E., Features of the primary wall CESA complex in wild type and  
763 cellulose-deficient mutants of *Arabidopsis thaliana*. *J. Exp. Bot.* **2008**, *59* (10), 2627-2637.
- 764 (16) Atanassov, II; Pittman, J. K.; Turner, S. R., Elucidating the mechanisms of assembly and subunit  
765 interaction of the cellulose synthase complex of *Arabidopsis* secondary cell walls. *J. Biol. Chem.* **2009**, *284* (6),  
766 3833-3841.
- 767 (17) Hill, J. L., Jr.; Hammudi, M. B.; Tien, M., The *Arabidopsis* cellulose synthase complex: a proposed  
768 hexamer of CESA trimers in an equimolar stoichiometry. *Plant Cell* **2014**, *26* (12), 4834-4842.
- 769 (18) Gonneau, M.; Desprez, T.; Guillot, A.; Vernhettes, S.; Höfte, H., Catalytic subunit stoichiometry within the  
770 cellulose synthase complex. *Plant Physiol.* **2014**, *166* (4), 1709-12.
- 771 (19) Polko, J. K.; Kieber, J. J., The Regulation of Cellulose Biosynthesis in Plants. *Plant Cell* **2019**, *31* (2), 282-  
772 296.
- 773 (20) Wen, X.; Zhai, Y.; Zhang, L.; Chen, Y.; Zhu, Z.; Chen, G.; Wang, K.; Zhu, Y. (2022) Molecular studies of  
774 cellulose synthase supercomplex from cotton fiber reveals its unique biochemical properties. *Sci. China Life*  
775 *Sci.* (2022). <https://doi.org/10.1007/s11427-022-2083-9>.
- 776 (21) Zhang, X.; Dominguez, P. G.; Kumar, M.; Bygdell, J.; Miroshnichenko, S.; Sundberg, B.; Wingsle, G.;  
777 Niittyla, T., Cellulose Synthase Stoichiometry in Aspen Differs from *Arabidopsis* and Norway Spruce. *Plant*  
778 *Physiol.* **2018**, *177* (3), 1096-1107.
- 779 (22) Li, X.; Chaves, A. M.; Dees, D. C. T.; Mansoori, N.; Yuan, K.; Speicher, T. L.; Norris, J. H.; Wallace, I.  
780 S.; Trindade, L. M.; Roberts, A. W., Cellulose synthesis complexes are homo-oligomeric and hetero-oligomeric  
781 in *Physcomitrium patens*. *Plant Physiol.* **2022**, *188*(4), 2115–2130.
- 782 (23) Purushotham, P.; Cho, S. H.; Diaz-Moreno, S. M.; Kumar, M.; Nixon, B. T.; Bulone, V.; Zimmer, J., A  
783 single heterologously expressed plant cellulose synthase isoform is sufficient for cellulose microfibril formation  
784 in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (40), 11360-11365.
- 785 (24) Cho, S. H.; Purushotham, P.; Fang, C.; Maranas, C.; Diaz-Moreno, S. M.; Bulone, V.; Zimmer, J.; Kumar,  
786 M.; Nixon, B. T., Synthesis and Self-Assembly of Cellulose Microfibrils from Reconstituted Cellulose  
787 Synthase. *Plant Physiol.* **2017**, *175* (1), 146-156.
- 788 (25) Deligey, F.; Frank, M. A.; Cho, S. H.; Kirui, A.; Mentink-Vigier, F.; Swulius, M. T.; Nixon, B. T.; Wang,  
789 T., Structure of In Vitro-Synthesized Cellulose Fibrils Viewed by Cryo-Electron Tomography and (13)C  
790 Natural-Abundance Dynamic Nuclear Polarization Solid-State NMR. *Biomacromolecules* **2022**, *23*(6), 2290–  
791 2301. [doi 10.1021/acs.biomac.1c01674](https://doi.org/10.1021/acs.biomac.1c01674).
- 792 (26) Zhang, X. N.; Xue, Y.; Guan, Z. Y.; Zhou, C.; Nie, Y. F.; Men, S.; Wang, Q.; Shen, C. C.; Zhang, D. L.;  
793 Jin, S. X.; Tu, L. L.; Yin, P.; Zhang, X. L., Structural insights into homotrimeric assembly of cellulose synthase  
794 CesA7 from *Gossypium hirsutum*. *Plant Biotechnol. J.* **2021**, *19*(8), 1579–1587. doi: 10.1111/pbi.13571.
- 795 (27) Purushotham, P.; Ho, R.; Zimmer, J., Architecture of a catalytically active homotrimeric plant cellulose  
796 synthase complex. *Science* **2020**, *369* (6507), 1089-1094. <https://doi.org/10.1126/science.abb2978>.

797 (28) Rushton, P. S.; Olek, A. T.; Makowski, L.; Badger, J.; Steussy, C. N.; Carpita, N. C.; Stauffacher, C. V.,  
798 Rice Cellulose SynthaseA8 Plant-Conserved Region Is a Coiled-Coil at the Catalytic Core Entrance. *Plant*  
799 *Physiol.* **2017**, *173* (1), 482-494.

800 (29) Qiao, Z.; Lampugnani, E. R.; Yan, X. F.; Khan, G. A.; Saw, W. G.; Hannah, P.; Qian, F.; Calabria, J.;  
801 Miao, Y.; Gruber, G.; Persson, S.; Gao, Y. G., Structure of Arabidopsis CESA3 catalytic domain with its  
802 substrate UDP-glucose provides insight into the mechanism of cellulose synthesis. *Proc. Natl. Acad. Sci. U. S.*  
803 *A.* **2021**, *118* (11), e2024015118.

804 (30) Vandavasi, V. G.; Putnam, D. K.; Zhang, Q.; Petridis, L.; Heller, W. T.; Nixon, B. T.; Haigler, C. H.;  
805 Kalluri, U.; Coates, L.; Langan, P.; Smith, J. C.; Meiler, J.; O'Neill, H., A Structural Study of CESA1 Catalytic  
806 Domain of Arabidopsis Cellulose Synthase Complex: Evidence for CESA Trimers. *Plant Physiol.* **2016**, *170*  
807 (1), 123-135.

808 (31) Du, J.; Vepachedu, V.; Cho, S. H.; Kumar, M.; Nixon, B. T., Structure of the Cellulose Synthase Complex  
809 of *Gluconacetobacter hansenii* at 23.4 Å Resolution. *PLoS One* **2016**, *11* (5), e0155886.

810 (32) Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J., EMAN2: an extensible  
811 image processing suite for electron microscopy. *J. Struct. Biol.* **2007**, *157* (1), 38-46.

812 (33) Mindell, J. A.; Grigorieff, N., Accurate determination of local defocus and specimen tilt in electron  
813 microscopy. *J. Struct. Biol.* **2003**, *142* (3), 334-347.

814 (34) Yang, Z.; Fang, J.; Chittuluru, J.; Asturias, F. J.; Penczek, P. A., Iterative stable alignment and clustering of  
815 2D transmission electron microscope images. *Structure.* **2012**, *20*(2), 237–247.  
816 doi.org/10.1016/j.str.2011.12.007

817 (35) Scheres, S. H., RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J.*  
818 *Struct. Biol.* **2012**, *180* (3), 519-530.

819 (36) Goddard, T. D.; Huang, C. C.; Ferrin, T. E., Visualizing density maps with UCSF Chimera. *J. Struct. Biol.*  
820 **2007**, *157* (1), 281-287.

821 (37) Rosenthal, P. B.; Henderson, R., Optimal determination of particle orientation, absolute hand, and contrast  
822 loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **2003**, *333* (4), 721-745.

823 (38) Russo, C. J.; Passmore, L. A., Robust evaluation of 3D electron cryomicroscopy data using tilt-pairs. *J.*  
824 *Struct. Biol.* **2014**, *187* (2), 112-118.

825 (39) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, I., T.E. ; Cruzeiro, V. W. D.;  
826 Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.;  
827 Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu,  
828 J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.;  
829 Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.;  
830 Simmerling, C. L.; Smith, J.; SalomonFerrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu,  
831 X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2018*, University of California, San Francisco: 2018.

832 (40) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T.  
833 A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein  
834 structures and complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296-W303.

835 (41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple  
836 potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926-935.

837 (42) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB:  
838 Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory*  
839 *Comput.* **2015**, *11* (8), 3696-3713.

840 (43) Joung, I. S.; Cheatham, T. E., 3rd, Determination of alkali and halide monovalent ion parameters for use in  
841 explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020-9041.

842 (44) Yang, J. Y.; Anishchenko, I.; Park, H.; Peng, Z. L.; Ovchinnikov, S.; Baker, D., Improved protein structure  
843 prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (3), 1496-1503.

844 (45) Singh, A.; Kwansa, A. L.; Kim, H. S.; Williams, J. T.; Yang, H.; Li, N. K.; Kubicki, J. D.; Roberts, A.;  
845 Haigler, C. H.; Yingling, Y. G., In silico prediction of full-length integral membrane protein structure, cotton  
846 cellulose synthase (GhCESA1), and its hierarchical complexes. *Cellulose* **2020**, *27*, 5597-5616.

847 (46) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J., PatchDock and SymmDock: servers for  
848 rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33* (suppl\_2), W363-W367.

849 (47) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J., Geometry-based flexible and  
850 symmetric protein docking. *Proteins.* **2005**, *60* (2), 224-231.

851 (48) Keshk, S. M. A. S., Bacterial Cellulose Production and its Industrial Applications. *J. Bioprocess. Biotech.*  
852 **2014**, *4* (2), 1-10.

853 (49) Giddings, T. H., Jr.; Brower, D. L.; Staehelin, L. A., Visualization of particle complexes in the plasma  
854 membrane of *Micrasterias denticulata* associated with the formation of cellulose fibrils in primary and  
855 secondary cell walls. *J. Cell Biol.* **1980**, *84* (2), 327-339.

856 (50) Grimson, M. J.; Haigler, C. H.; Blanton, R. L., Cellulose microfibrils, cell motility, and plasma membrane  
857 protein organization change in parallel during culmination in *Dictyostelium discoideum*. *J. Cell Sci.* **1996**, *109*  
858 3079-3087.

859 (51) Kimura, S.; Itoh, T., New cellulose synthesizing complexes (terminal complexes) involved in animal  
860 cellulose biosynthesis in the tunicate *Metandrocarpa uedai*. *Protoplasma.* **1996**, *194*, 151-163.

861 (52) Nicolas, W. J.; Ghosal, D.; Tocheva, E. I.; Meyerowitz, E. M.; Jensen, G. J., Structure of the Bacterial  
862 Cellulose Ribbon and Its Assembly-Guiding Cytoskeleton by Electron Cryotomography. *J. Bacteriol.* **2021**, *203*  
863 (3):e00371-20. doi: 10.1128/JB.00371-20.

864 (53) Sethaphong, L.; Haigler, C. H.; Kubicki, J. D.; Zimmer, J.; Bonetta, D.; DeBolt, S.; Yingling, Y. G.,  
865 Tertiary model of a plant cellulose synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (18), 7512-7517.

- 866 (54) Sethaphong, L.; Davis, J. K.; Slabaugh, E.; Singh, A.; Haigler, C. H.; Yingling, Y. G., Prediction of the  
867 structures of the plant-specific regions of vascular plant cellulose synthases and correlated functional analysis.  
868 *Cellulose* **2016**, 23, 145–161.
- 869 (55) Kawano, Y.; Saotome, T.; Ochiai, Y.; Katayama, M.; Narikawa, R.; Ikeuchi, M., Cellulose accumulation  
870 and a cellulose synthase gene are responsible for cell aggregation in the cyanobacterium *Thermosynechococcus*  
871 *vulcanus* RKN. *Plant Cell Physiol.* **2011**, 52 (6), 957-966.
- 872 (56) Zhong, R.; Morrison, W. H., 3rd; Freshour, G. D.; Hahn, M. G.; Ye, Z. H., Expression of a mutant form of  
873 cellulose synthase *AtCesA7* causes dominant negative effect on cellulose biosynthesis. *Plant Physiol.* **2003**, 132  
874 (2), 786-795.