

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

Juan Du, Venu Gopal Vandavasi, Kelly R. Molloy, Hui Yang, Lynnica N. Massenburg, Abhishek Singh, Albert L. Kwansa, Yaroslava G. Yingling, Hugh O'Neill, Brian T. Chait, Manish Kumar, and B. Tracy Nixon*



Cite This: <https://doi.org/10.1021/acs.biomac.2c00550>



Read Online

ACCESS |



Metrics & More

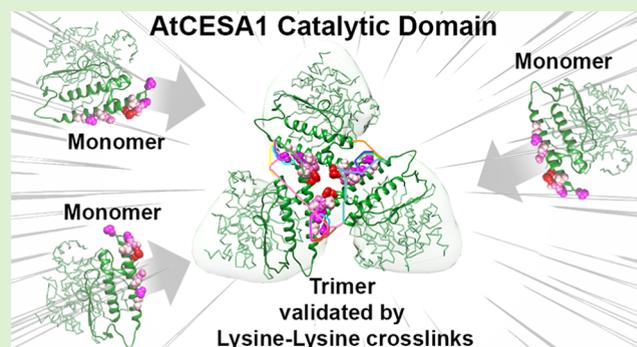


Article Recommendations



Supporting Information

ABSTRACT: Higher plants synthesize cellulose using membrane-bound, six-lobed cellulose synthase complexes, each lobe containing trimeric cellulose synthases (CESAs). Although molecular biology reports support heteromeric trimers composed of different isoforms, a 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 plant-conserved region (P-CR) forms the only contacts between catalytic domains of the monomers. We report inter-subunit lysine-crosslinks that localize to the small P-CR, negative-stain EM structure, and modeling data for homotrimers of AtCESA1CatD. Molecular dynamics simulations for AtCESA1CatD trimers based on the CESA8 cryoEM structure were stable and dependent upon a small set of residue contacts. The results suggest that homomeric CESA trimers may be important for the synthesis of primary and secondary cell walls and identify key residues for future mutagenic studies.



INTRODUCTION

Cellulose is the world's most abundant organic compound and is used as a raw material for the industrial production of paper, cardboard, wood, clothing, and medical supplies. Cellulose exists in plant cell walls as β -1,4 glucan chains, with primary walls existing around growing cells and relatively thick secondary walls surrounding non-growing cells. Increasing attention is focused on using fibers as feedstock for sustainable biofuel production as a renewable alternative to fossil fuels.^{1–4}

Cellulose is made by cellulose synthases (CESAs) that form cellulose synthase complexes (CSCs) in higher plants, bryophytes, and algae. The size, pattern, and microfibril association of CSCs in the plasma membrane of moss were captured using freeze-fracture transmission electron microscopy.^{5–9} In the model plant *Arabidopsis*, genetic evidence suggests that three distinct isoforms are required for cellulose synthesis, in which AtCESA1, AtCESA3, and AtCESA6 or AtCESA6-like form CSCs for making primary walls, and AtCESA4, AtCESA7, and AtCESA8 form CSCs for making secondary walls.^{10–14} Many observations suggest that the plant CSC is a multi-subunit heteromeric complex that produces cellulose microfibrils in vivo. Among these, are similar sub-cellular locations, dynamic movement in plasma membranes for CESAs in the same group, loss of partner CESAs upon one's removal, protein interaction confirmed by in vivo

analysis,^{11–13,15,16} and 1:1:1 stoichiometric presence in primary (AtCESA1, AtCESA3, and AtCESA6) and secondary (AtCESA4, AtCESA7, and AtCESA8) wall CSCs.^{17,18} Until a few years ago, it was believed that each CSC contained 36 CESA protomers, but now several pieces of evidence indicate that there are 18 protomers organized in 6 trimeric lobes, which together make an 18-chain cellulose microfibril (reviewed by Polko and Kieber¹⁹). Recent studies of CSCs in cotton show that the oligomerization state of CSCs and the chain number in one microfibril can vary in different species and developmental stages.²⁰

We note that: (a) CESA ratios different from 1:1:1 have been seen in poplar;²¹ (b) CESA5 of *Physcomitrium patens* functions as a homomeric CSC in vivo;²² and (c) in vitro studies of *Pichia*-expressed poplar CESA8 and moss CESAs revealed that a single CESA isoform can form oligomers and synthesize cellulose microfibrils in vitro.^{23,24} The nature of the in vitro fibers is not fully determined, but cryo-electron

Received: April 29, 2022

Revised: July 29, 2022

tomography and solid state NMR studies show that they favor 18 chains and are similar to CMF in plant cell walls.²⁵ CryoEM structures of homomeric trimers of secondary wall CESAs PttCESA8 and GhCESA7 have been solved.^{26,27} A trimeric crystal structure of the OsCESA8 plant-conserved region (P-CR) fragment is nearly identical to the P-CR-based trimers in the cryoEM structures for secondary cell wall CESAs.²⁸

For CESAs making primary walls, there are reports of a homomeric dimer structure for the catalytic domain of AtCESA3²⁹ and the formation of homomeric trimers for the catalytic domain of AtCESA1.³⁰ Here, we report the negative staining transmission electron microscopy (TEM) structure of the AtCESA1CatD monomer and trimer and describe a set of inter-subunit cross-linking lysine residues that cluster in a loop and helix of the P-CR of AtCESA1CatD. Since the cross linking lysine residues are highly conserved among CESA proteins, a similar cross-linking potential is predicted for other lobes of trimeric CESAs and found to be the same for the crystal and cryoEM structures. By considering goodness of fit to the published SAXS data and 3D reconstruction and (most impactfully) match between observed and predicted cross-links, we show that among tens of thousands of computational models of P-CR-based trimers of AtCESA1CatD, the most likely models possess a trimerization motif like those in the crystal lattice of OsCESA8 P-CR and cryoEM structures of PttCESA8 and GhCESA7. The crosslink data are not consistent with the recently reported crystal structure dimer of AtCESA3CatD.²⁹ Molecular dynamics (MD) simulations showed that the homotrimer of AtCESA1CatD is highly stable, with somewhat less stability present in modeled heterotrimers—in the latter, interactions between the TM helices of different protomers are likely essential for the overall trimer stability. We thus propose that the P-CR-based trimerization motif is common to plant CESAs, regardless of whether they make primary or secondary cell walls and present a new model for AtCESA1, based upon the structure of PttCESA8. Comparing this model to bacterial cellulose synthase prompts us to suggest that substitution of the regulatory PilZ domain with the P-CR motif allowed plants to evolve oligomeric CESAs. This also raises the possibility that the P-CR trimer motif mediates allosteric communication between protomers to regulate synthesis of β -glucan chains.

EXPERIMENTAL SECTION

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

To optimize oligomerization of AtCESA1CatD, a two-step procedure was used. First, the monomer sample was dialyzed for 48 h at 4 °C, against the buffer containing 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 8.5), 50 mM NaCl, 10% glycerol, and 5 mM dithiothreitol in a 10 kD cutoff dialysis cassette (slide-A-lyzer, Thermo Scientific Inc.), which allowed the gentle reduction (but not total loss) of the detergent. Dialysis was continued overnight in the same buffer, except that the pH was changed to 7.8, until the value of the hydrodynamic radius was about 8 nm, which was previously reported to be true for the trimers characterized by small angle X-ray and neutron scattering.³⁰

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

Negative Staining EM. Trimer samples were prepared by the following procedure. Grids used for negative staining EM were prepared as previously described.³¹ Briefly, 400-mesh copper grids (TED PELLA) were floating on a thin film of collodion. After drying overnight, the film side of grid was coated by a thin layer of continuous carbon (Denton Vacuum model DV-502B). Three and one half microliters of protein sample (0.005 mg/mL) was applied onto glow-discharged grids (Pelco easiGlow). Following absorbing for 1 min, the grid was washed/stained in 10 drops of 1% (wt/vol) freshly made uranyl formate. Then, the grid was stained for 20 s before quick and complete air-drying. Negatively stained grids were imaged at 20 e⁻/Å² on the Tecnai T12 transmission electron microscope operated at 120 keV with a defocus range of 0.5–2.5 μ m. Micrographs were captured on a 4 K \times 4 K charge-coupled device (CCD) camera (Eagle, FEI) at a nominal magnification of 68,000 \times (1.5 Å/pixel at specimen level).

Monomer samples were prepared similar to trimers, except for washes with 7 drops of filtered Milli-Q water, followed by 3 drops of 0.75% (wt/vol) freshly made uranyl formate. The monomer grid was stained for 45 s before quick and complete air drying. Negatively stained grids were imaged on the Tecnai T12 transmission electron microscope operated at 120 keV, with a defocus range of 0.5–2.5 μ m. Monomer micrographs were captured on a 2 K \times 2 K CCD camera (Orion, FEI) at a nominal magnification of 98,000 \times (3.0 Å/pixel at specimen level).

Image Processing. EMAN2³² was used to process monomer images using e2ctf.py for CTF correction. Manually picked particles were extracted using an 84 \times 84-pixel box and pruned to 3836 particles by 2D class averaging. Initial models were generated and used for refinement without symmetry using e2initialmodel.py and e2refine.py, respectively.

The evaluation of trimer micrographs was also carried out using e2evalimage.py function in EMAN2.³² All the images displaying no significant astigmatism and well-spread particles were used for further processing. The contrast transfer function of each micrograph was determined using CTFFind3.³³ Manually picked trimer particles (~4000) were extracted using a 168 pixel \times 168 pixel box and subjected to 2D-reference free classification, after normalization. Classes generating blurred averages and containing heterogeneous particles were discarded. Remaining particles were then subjected to iterative stable alignment and classification, using ISAC2.³⁴ Initial models were reconstructed from remaining good class averages, using e2initialmodel.py with either no symmetry or 3-fold symmetry imposed. Eventually, one initial model with 3-fold symmetry imposed was applied in subsequent analyses. 3D reconstruction and refinement were performed by using the program of e2refine_easy.py in EMAN2. Meanwhile, particles were imported into RELION 2.0 for 3D classification and refinement.³⁵ Density maps were validated using tilt-pair validation as described below. Rigid body docking between the EM map and computational model was performed in UCSF Chimera.³⁶

Tilt-Pair Validation. To evaluate the quality of the density map of a monomer and trimer, a set of image pairs were recorded for the same specimen field at two tilt angles (untilted and 10° tilted) and processed as described.³⁷ The analysis was performed using program EMAN2's function e2tiltvalidate.py and TiltStats.^{37,38}

Generation of a Heavy Lysine-Labeled Dimer and Trimer Protein. To identify the interaction interface between monomers in a trimer complex, we generated a chimeric trimer complex consisting of a regular monomer (¹²C and ¹⁴N) and heavy lysine-labeled monomer

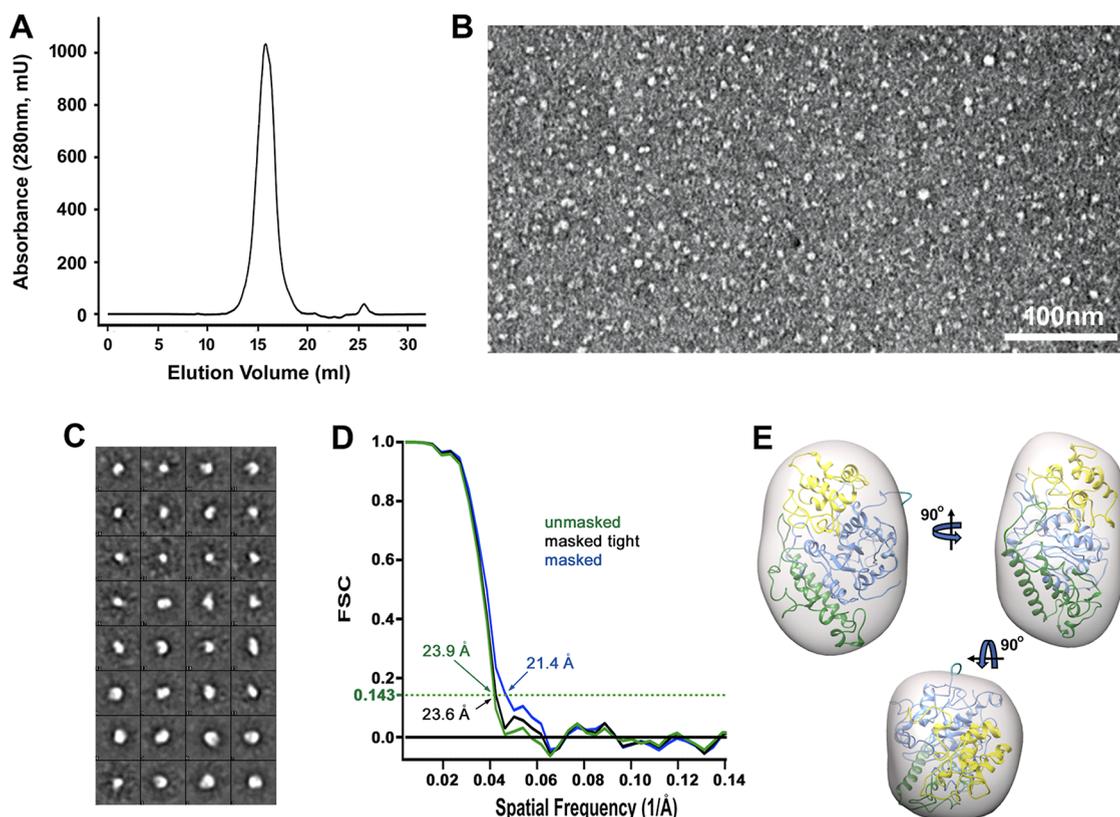


Figure 1. Fit of the computational model of the AtCESA1CatD monomer into the negatively stained EM density map. (A) Size exclusion chromatogram of the AtCESA1CatD monomer. (B) Representative negatively stained image of particles eluted from the main peak fraction in (A). (C) 2D class averages of the AtCESA1CatD monomer. (D) Gold-standard Fourier shell correlation (FSC) plot vs the spatial frequency. The density map is resolved to 21.4–23.9 Å, according to the FSC value of 0.143. (E) Model of the AtCESA1CatD monomer superimposed on the 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 the P-CR at the monomer–monomer interfaces.

(^{13}C and ^{15}N). Given the cross links between regular lysine and heavy lysine must happen between distinct monomers (light monomer and heavy monomer, in this case), it is expected that fragments bearing cross links from the interaction interfaces will be increased 8 Da per lysine residue in MS spectra, which allowed us to distinguish intra-cross linking in the monomer and inter-cross linking between monomers. To produce a heavy lysine-labeled AtCESA1CatD trimer complex, modified minimal M9 broth with lower glucose (0.2%) and lacking NH_4^+ was used to culture the strain. The heavy lysine-labeled monomer (CatD_H) and regular monomer (CatD_L) were mixed in a 1:1 ratio and dialyze to form trimer, as mentioned in the procedure above.

Chemical Cross Linking and Mass Spectrometry. To obtain the structural information of protein folding in a monomer and interaction interface in a trimer, we applied chemical cross linking, followed by mass spectrometry (MS) to analyze the distance constraints between reactive residue pairs. The purified protein was diluted to 2 mg/ml using buffer containing 20 mM HEPES (pH7.9). Immediately before use, disuccinimidyl suberate (DSS) (a homobifunctional amine-reactive cross linker, spacer arm = 11.4 Å, Thermo Scientific) was dissolved in dimethyl sulfoxide to make 50 mM stock. To optimize the cross linking of the trimer complex, 20-fold excess of the cross linker (0.66 mM) was added to the protein sample and followed by incubation for 1 h at 25 °C with gentle agitation during incubation. Then, the reaction was quenched by adding 50 mM Tris to a final concentration and subsequently incubated at room temperature for 15 min. The cross-linked product was precipitated by adding cold methanol to a final concentration of 90% and incubating at -80 °C overnight. After spinning at 14,000 rpm at 4 °C for 15 min, the pellet was washed once by 90% cold methanol and finally dissolved in 50 μL 1 \times NuPAGE LDS loading buffer

(Invitrogen, USA). Then, the protein sample was subjected to disulfide reduction and cysteine alkylation. After incubation for 15 min at 70 °C, the sample was resolved in 4–12% gradient SDS-PAGE gel (GenScript) in Tris-MOPS buffer, which resulted in good separation of cross-linked monomer, dimer and trimer bands. The targeted gel bands were excised, subjected to in-gel trypsin digestion, and subjected to liquid chromatography–MS (LC–MS) to identify cross links (Figure S1 and Table S1). Cross-linked peptides were identified by comparing experimental and theoretical MS2 fragmentation spectra to find the closest matches. These are indicated as having direct evidence in Table S1. We suspected that in some cases, inter-subunit cross-linked peptides were present in the sample but not identified by MS2 because they generated low-quality fragmentation spectra or were not selected for MS2. To find such peptides, we searched the MS1 spectra for peaks matching their calculated masses. When we observed such a peak at the same LC retention time as the corresponding light–light and heavy–heavy peptides, and at least one of those peptides had been identified by MS2 fragmentation, we considered this indirect evidence of an inter-subunit cross-linked peptide, as indicated in Table S1.

MD Simulations. MD simulations were conducted using AMBER 2018.³⁹ An AtCESA1CatD trimer model (Figure S2) was assembled using the PtCESA8 trimer structure (PDB ID: 6WLB)²⁷ as the template, with the AtCESA1CatD monomer model generated using the SWISS-MODEL program⁴⁰ in which a major portion of the CSR (THR649–MET713) was predicted as unstructured (Figure S3). AlphaFold2 was not available at the time of performing the MD simulations. The SWISS-MODEL and AlphaFold2-predicted monomers are nearly identical (rmsd, 0.713 Å). Three AtCESA1CatD monomers formed a triangular P-CR interface. The CSR was positioned at the other side of the GT domain, opposite to the P-

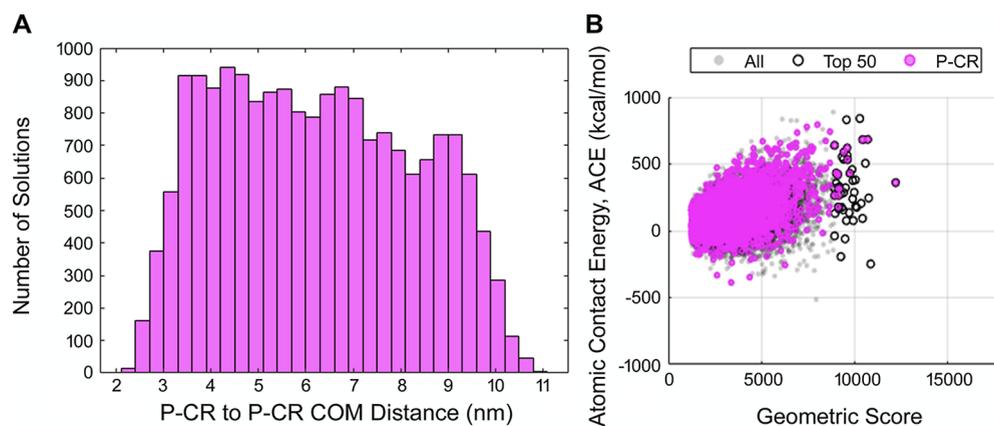


Figure 2. Library of trimers of AtCESA1CatD. (A) Number of solutions vs distance is plotted between P-CR COM. Models with COM distances ≤ 4 nm were combined to make the P-CR subset. (B) Scatter plot of atomic contact energy vs geometric score for the entire library (gray), the P-CR subset (magenta), and the top 50 geometric-scoring models (black circles). Purification of AtCESA1CatD trimers to homogeneity by sequential SEC.

CR. The unstructured CSRs were excluded so that the simulations could finish in a reasonable period of time. Counterions (0.10 M NaCl) were added which was solvated with TIP3P water molecules.⁴¹ The size of the initial system was $166 \times 177 \times 108 \text{ \AA}^3$ and contained $\sim 255,000$ atoms. Minimization and equilibration stages were conducted by gradually reducing restraints on the protein. First, a 1000-step minimization was performed, consisting of 400 steps of the steepest descent and 600 steps of conjugate gradient minimization with protein positions harmonically restrained using a force constant of $50 \text{ kcal/mol/\AA}^2$. After minimization, the system temperature was increased to 300 K through two sequential runs, with 10 and 2 kcal/mol/ \AA^2 restraints placed on the positions of the protein, respectively. First, 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 at 1 atm in an NPT ensemble, with a 2-fs time step, 10- \AA nonbonded interaction cutoff, and SHAKE-constrained hydrogen bonds. A short 500-ps NPT simulation was then performed with no restraints, prior to production simulations. Conventional MD simulations were conducted on the models for 1000 ns. In all stages, MD simulations were performed with periodic boundary conditions, a temperature of 300 K, a pressure of 1 atm, 2 fs time step, 10 \AA nonbonded interaction cutoff, SHAKE-constrained hydrogen bonds, and particle-mesh Ewald and Lennard–Jones correction for long-range Coulombic and Lennard–Jones interactions, respectively. In all simulations, the protein was described by the ff14SB force field,⁴² and the water molecules and counterions were described by the TIP3P model⁴¹ and Joung–Cheatham monovalent ions for TIP3P, respectively.⁴³

The root-mean-square deviation (rmsd) of the trimer demonstrated that the system reached equilibrium after about 150 ns. A 1000 ns-long production run was performed on the theoretical model of the AtCESA1CatD trimer. The root-mean-square fluctuation (RMSF) profile of the trimer residues was calculated, based on the whole 1000 ns-long production run. 100 frames taken from the last 100 ns production run were used to calculate the binding free energies between AtCESA1CatD monomers, using the MM-PBSA approach in the AMBER 2018 package. Per-residue decomposition calculations were performed to decompose the free energy contribution to the binding free energy between AtCESA1CatD monomers.

A newly developed protein structure prediction program, trRosetta,⁴⁴ was also utilized to generate AtCESA1, AtCESA3, and AtCESA6 catalytic domain models with the corresponding sequences listed in Table S2. As shown in Figure S3, the AtCESA1CatD model obtained from the trRosetta program superimposed tightly with the model predicted by the SWISS-MODEL program, except for the CSR. The CSR was predicted as unstructured in the SWISS-MODEL. In the trRosetta-predicted structure, the CSR is composed of four short helices connected by short loops. Similar CSR structures were also obtained in the AtCESA3CatD and AtCESA6CatD models,

positioned at the other side of the GT domain, opposite to the P-CR, as shown in Figures S4 and S5. Two heterotrimer models were then assembled by replacing two AtCESA1CatD units in the homotrimer as described before. One is in the counterclockwise arrangement (Figure S4), and the other is in the clockwise arrangement (Figure S5). Production runs of 500 ns were performed using the methods described above. The rmsd profile of each system was calculated based on the whole 500 ns-long production run. Distances between the mass center of each subunit (labeled as whole-## in Figure S6) and distances between the mass center of each P-CR domain (marked as PCR-## in Figure S6) were measured.

RESULTS

TEM Structure Validation of the AtCESA1CatD Monomer Model. For the approach used in this study, the computational structure of the AtCESA1CatD monomer forms the basis of understanding the structure and architecture of the higher-order oligomer. A model of AtCESA1CatD was generated by threading its amino acid sequence (residues 341–845, RYD-GRL of AtCESA1) onto the recently published computational model of GhCESA1.⁴⁵ The overall size and shape of the model was validated by 3D reconstruction of the purified monomer protein from negative stain TEM using EMAN2.³² SEC of the purified monomer yielded a single symmetric peak (Figure 1A). Negatively stained images (example presented in Figure 1B) showed well-dispersed particles that were readily picked into a dataset of 7513 particles. All particles in this dataset were sorted by 2D classification (Figure 1C) into 32 classes and pruned via 2 rounds of refinement to 3836 particles. 3D reconstruction from these particles yielded a density map that is superimposed with the computational model in Figure 1D–E (21.4–23.9 \AA resolution; FitMap CC 0.92). In this optimal fit, the P-CR, CSR, and catalytic core are well enclosed by the envelope. Since we are screening for trimers mediated by P-CR interactions, we conclude that the relevant parts of the computational model of the monomer are consistent with the density map and that it is reasonable to use the model to build a library of putative P-CR/P-CR-based trimer models.

As suggested by previous studies, the P-CR and CSR 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 the N-terminal (NT), CSR, or P-CR at the monomer–monomer interfaces.⁴⁵ In this

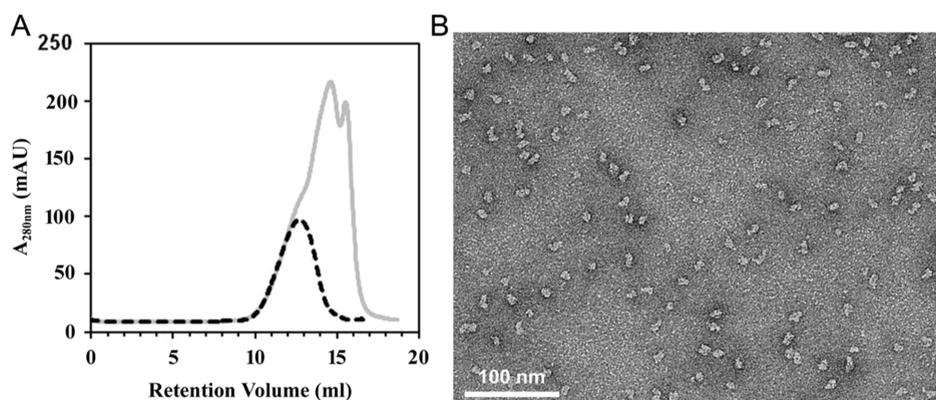


Figure 3. Purification and visualization of AtCESA1CatD trimers. (A) SEC after dialysis (first run, gray line) and after pooling, concentrating, and re-fractionating (second run, black dashed line). The elution peak of the second run (12.5 mL) is enriched with trimers. (B) Negatively stained image of particles at the elution peak of the second SEC [black dashed line in (A)].

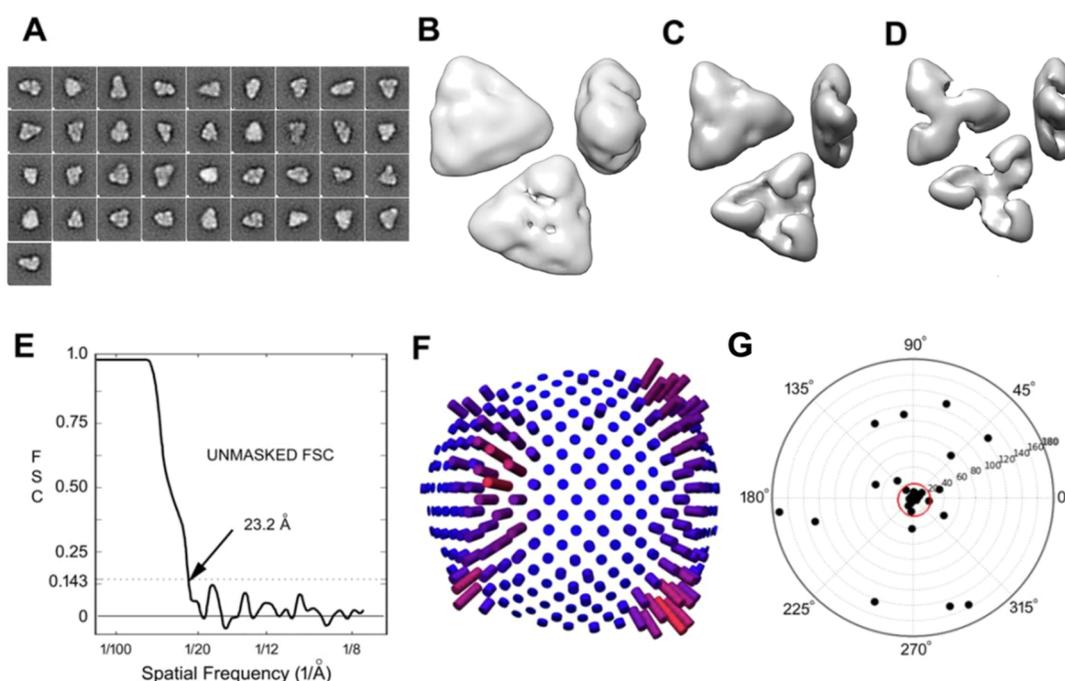


Figure 4. Negatively stained structure of AtCESA1CatD using RELION. (A) Stable 2D classes from ISAC2. (B) Initial model generated using RELION. Refined model rendered surface levels as 0.11 (C) and 0.12 (D), respectively. (E) Unmasked FSC plot. (F) Particle distribution plot (only 1/3 shown since refinement used C3 symmetry). The number of particles found in a given orientations is shown as the cylinders of varying heights. (G) Tilt pair validation. Black dots represent the tilt axis and tilt angle for particle pairs in polar coordinates. The red circle is centered around the expected tilt angle of 10°. The outer radius of the plot is 180°. Identification of inter-subunit chemical cross links in AtCESA1CatD trimers.

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° and less than or equal to 5 nm distance between the center of mass (COM) of the adjacent P-CR. Omitting the TMH tilt angle criterion and just applying the P-CR/P-CR distance constraint (adjusted to less than or equal to 4 nm) yielded 2651 models. For the current study, a similar library of 18,792 modeled trimers was generated using the AtCESA1CatD monomer structure and SymmDock.^{46,47} Within the library, there were 3251 models in which the P-CR elements’ COM were not beyond 4 nm (Figure 2A; note that lacking the TMH regions in the models, there was no screening for the TMH tilt angle). A scatter plot of SymmDock atomic contact energies versus geometric shape complementarity scores (geometric scores)

for these models is shown in Figure 2B. Many of the top 50 scoring models, based on the geometric score, were rejected by P-CR to P-CR COM screening. Additional information is needed to screen through this P-CR/P-CR-biased library to identify those that are reasonable, based on biochemical observations. Small angle X-ray scattering data have been published for the AtCESA1CatD trimeric structure.³⁰ To supplement these data, we also reconstructed the trimer 3D structure from negatively stained images and identified subunit–subunit cross links between lysine residues using the reagent DSS. As presented below, the latter information proved the most informative.

We enriched the AtCESA1CatD trimers prepared for the SAXS studies for imaging by TEM. Initial attempts to pass trimer samples through a Superose 6 increase SEC column

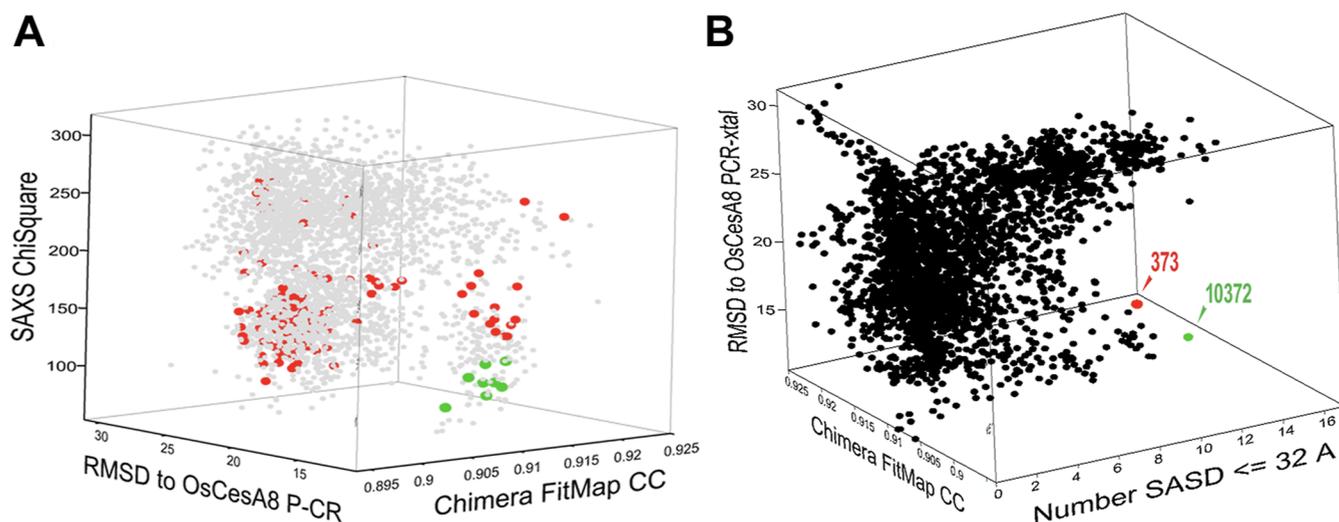


Figure 6. Distribution of scored AtCESA1CatD trimer models. (A) Ranked for fit to SAXS data, fit to the negative stain structure of the AtCESA1CatD trimer, and rmsd to the trimeric crystal structure of OsCESA8 P-CR. All computational trimer models of AtCESA1CatD trimers (gray); those with 12 or more SASD ≤ 32 Å (red or green); those with 12 or more SASD ≤ 32 Å, SAXS ChiSquare < 126 , and rmsd < 13 Å (green). (B) Same ranking of the library of trimer models but with the parameter of “SAXS ChiSquare” replaced with the number of reasonable cross links. The two models most consistent with the cross linking data are 373 (red) and 10,372 (green).

shoulder fractions and performed a second round of gel filtration, which resulted in a sharper, more symmetric major peak that showed no evidence of instability during the time needed for the concentration and chromatography (~ 60 min; Figure 3A, black). Inspection using negative staining EM revealed improved particle integrity and homogeneity in the peak fraction of the second pass through the column (Figure 3B).

3D Reconstruction of the AtCESA1CatD Trimer by Negative Staining EM. Freshly eluted protein samples were then used to optimize grid preparation for negative staining EM. Final images showed monodispersed and evenly distributed particles in different orientations (Figure 3B). We manually picked ~ 4000 particles, which were reduced to 3417 particles after 2D reference free classification using the program EMAN2 and then to 3298 from 2D alignment and classification using ISAC2.³⁴ A final set of 37 stable averages is shown in Figure 4A, with a majority exhibiting triangular shapes in distinct views that could be matched with 2D projections from prior published trimer models, based on small angle X-ray scattering.⁵⁰

We used RELION (Figure 4B–F) and EMAN2 (Figure S7) to determine the 3D structure of AtCESA1CatD. The particles were imported into RELION 2.0 and used to generate a C3 symmetric initial model (Figure 4B). After refinement, 3D classification gave only a single class (Figure 4C,D) resolved to 23.2 Å (Figure 4E). There were some preferred orientations, but all perspectives were observed (Figure 4F). Tilt pair validation (Figure 4G) gave a tilt angle mean of 5.9° ($\pm 4.9^\circ$) that was consistent with the set angle of 10° , and the number of tilt pairs in the cluster was 31 out of 44 (68%). TiltStats analysis with $N = 44$ and Winsorizing parameter = 10% gave 31 tilt pairs with $>99\%$ probability of being non-randomly distributed ($R/R_0 = 2.8$) in a cluster of 8.3° radius about a geometric mean of 3.3° with an angular accuracy of 6.3° and κ of 83.

We supplemented the limited resolution achieved by negative staining EM by coupling DSS chemical cross linking of AtCESA1CatD with high resolution MS. Despite the fact

that application of cross linking MS in mapping interaction interfaces in protein complexes is well established, the analysis on homo-oligomers is challenging. To clearly distinguish intra-cross linking and inter-cross linking, we generated a heavy lysine-labeled AtCESA1CatD monomer (CatD_H) by feeding *Escherichia coli* expressing the protein 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 an increased molecular mass of 8 Da per lysine residue. As indicated in Figure 5A, a light (CatD_L) and heavy monomer (CatD_H) were mixed in a 1:1 ratio to form a chimeric trimer. Assuming that the oligomers were stable over the 60 min cross linking reaction, which was seen for size-exclusion chromatography reported in Figure 3, the light–heavy crosslinks clearly occurred between two subunits, while the light–light or heavy–heavy ones could be intra- or inter-subunit cross links. After cross-linking, the protein was separated by SDS-PAGE, excised as a trimer, proteolyzed with trypsin, and subjected to LC–MS to identify cross links (Table S1).

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

We used scripts and the programs of Chimera, Swiss PDB viewer, and Crysol_30 to score each of the 3251 trimer models reported above (see Figure 2) for cross correlation of fit into the negative stain density map for the AtCESA1CatD trimer, rmsd of fit between the P-CR and the P-CR trimer present in the crystal lattice of OsCESA8, and ChiSquare of fit to SAXS data. The highest scoring models fall into the lower right portion of the 3D scatter plot (Figure 6A). We also used the

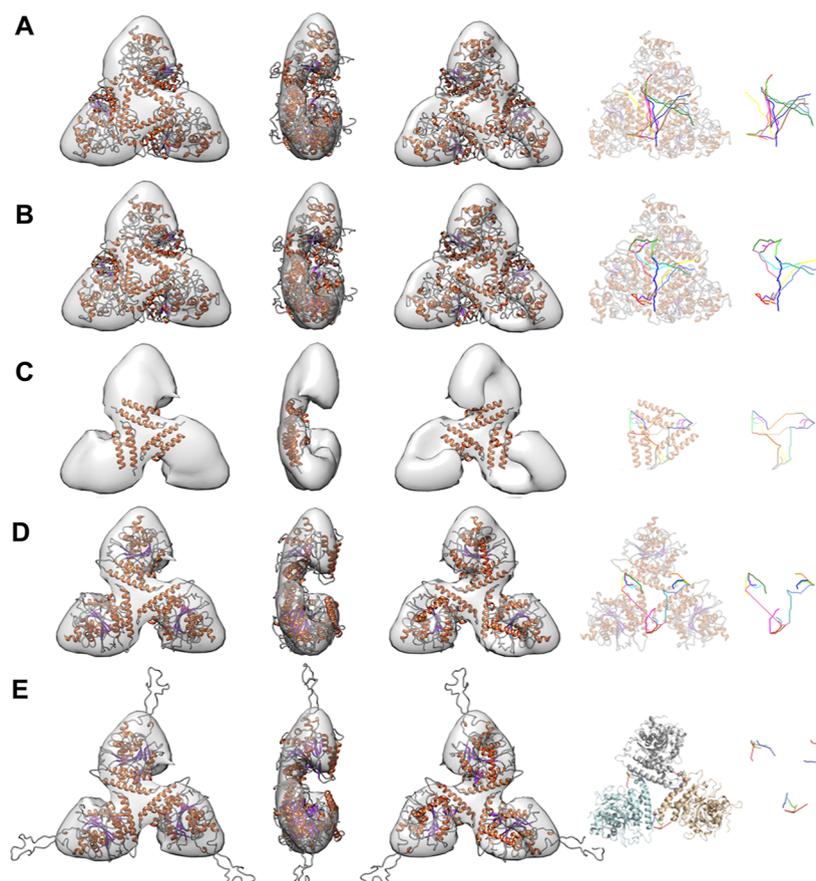


Figure 7. Overlay of the AtCESA1CatD negative stain density with best computational trimers, OsCESA8 P-CR and PttCESA8. Columns 1–3 show top, side, and bottom views of the AtCESA1CatD negative stain density overlain with computational models (A) #10372, (B) #373, or (C) OsCESA8 P-CR crystal structures 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 SWISS-MODEL for AtCESA1, using the PttCESA8 structure. The unstructured loop is part of the CSR.

programs JWalk and FreeSaSa to identify the number of reasonable solvent accessible pathways for lysine-DSS-lysine cross links. The models that had 12 or more reasonable cross links are colored red or green, as shown in Figure 6A. The green colored ones fit the best to the SAXS profile and P-CR structure. Figure 6B shows the distribution of trimer models ranked for the number of reasonable cross links [surface-accessible solvent distance (SASD) ≤ 32 Å], fit to the negative stain density, and rmsd to the OsCESA8 P-CR structure.

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

In both “best” model trimers, the P-CR helices bearing cross linkable lysines form a triangular contact surface very similar to the one seen in the OsCESA8 P-CR crystal structure and the PttCESA8 cryoEM structure. The cross linking lysine residues found in AtCESA1CatD are highly conserved (among most if not all CESA proteins) and present in both of these structures. When screened for capacity to support reasonable DSS cross

links, 14 of these we observed in AtCESA1CatD were found in the poplar trimer structure (the number did not change if membrane spanning sequences were present or absent) and 15 were seen in the rice P-CR trimer (Figure 7C,D). Encouraged by this agreement between trimer structures of PttCESA8 and AtCESA1, we used SWISS-MODEL,⁴⁰ a fully automated protein structure homology-modeling server, to generate a final model of AtCESA1 trimers based on the cryoEM structure—the CatD regions are shown in Figure 7E. Note that the loop extending beyond the negative stain density is a portion of the CSR that was not present in the cryoEM structure.

MD Simulations of the AtCESA1CatD Trimer. To investigate stability of the AtCESA1CatD trimer, we performed conventional MD simulations on a theoretical model of the AtCESA1CatD trimer. After the trimer system equilibrated to a stable rmsd, in about 150 ns (Figure 8A,B), a subsequent 1000 ns-long production run was performed. The P-CR domains (ALA398-LYS475) were among a few regions showing the least conformational fluctuation (2 to 4 Å), as measured by the atomic RMSF (Figure 8C). The stability of the P-CR domains maintained the triangular interface among three AtCESA1CatD monomers during the simulation (Figure 8D,E). The most dynamic regions showing higher RMSF (>6 Å) were the loops connected to the truncated CSR and the loops connected to the truncated transmembrane region.

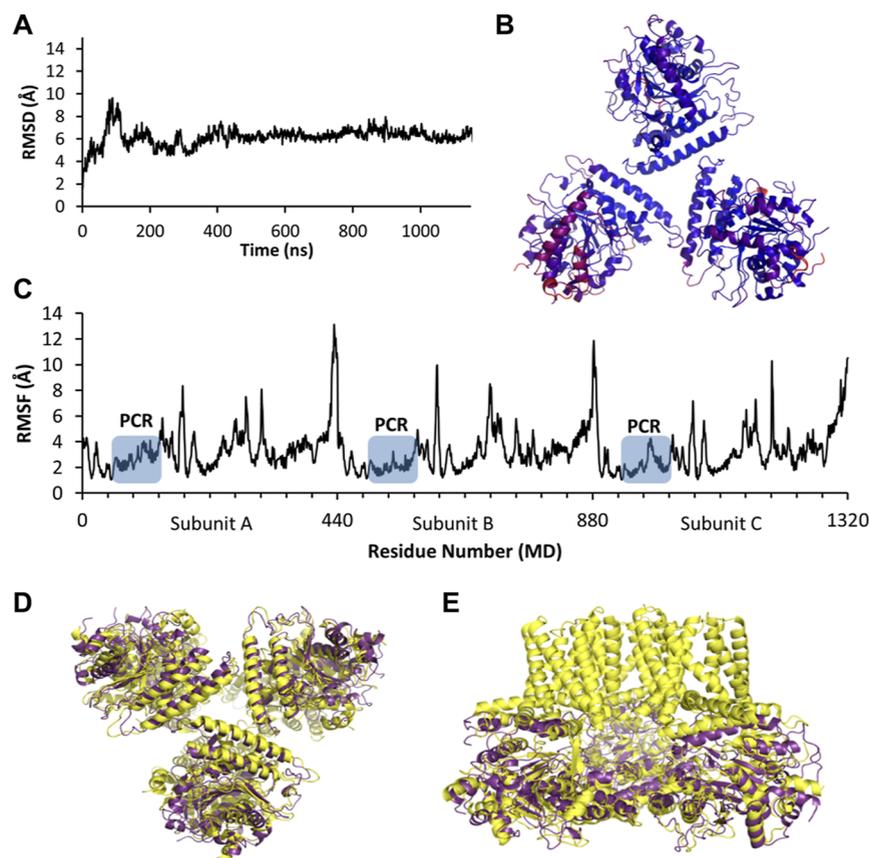


Figure 8. 1000 ns-long MD simulation performed on the theoretical model of the AtCESA1CatD trimer. (A) rmsd profile of the AtCESA1CatD trimer during the 1000 ns MDS. (B) Final structure of the AtCESA1CatD trimer. The color represents the conformational fluctuations of the AtCESA1CatD protein measured 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 1000 ns-long production MDS run. (D) Bottom view and (E) 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 ~ 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. R453 contributed the most, ~ -5.9 kcal/mol, to the binding free energy. L440, I444, and V464 each contributed ~ -3.5 kcal/mol to the binding free energy. V449 and E460 each contributed ~ -2.5 kcal/mol to the binding free energy. N467, K450, Y439, E461, and A468 each contributed -1 to -2 kcal/mol.

To investigate potential effects of the CSR region on the AtCESA1CatD trimer, we replaced the AtCESA1 catalytic domains with an AtCESA1CatD model containing the CSR obtained from the program trRosetta, a newly developed web-based homology modeling server based on deep learning and Rosetta⁴⁴ that predicted more structures for the CSR than those predicted by the SWISS-MODEL program. Both the above trimer and this new one have similar rmsd profiles, as shown in Figure S6E,G. Figure S6F,H showed that both trimers preserved relative constant distances between each P-CR domain (~ 30 Å) and between each AtCESA1CatD unit (~ 60 and ~ 70 Å, respectively), maintaining an equilateral triangle shape during the MD simulations. As shown in Figure S2, during 500 ns MD simulations, the P-CR domains remained in an equilateral triangle shape, and the catalytic domains largely maintained the parallel direction. Thus, the

CSR did not demonstrate a dramatic effect on the stability of the AtCESA1CatD homotrimer.

MD Simulations Computed Stability of Heterologous Trimers of AtCESA1,3,6.

Because there is significant in vivo evidence that plants use heterotrimers of CESAs to make CSCs for synthesis of primary and secondary walls, we replaced two of the AtCESA1 catalytic domains with models of AtCESA3 and AtCESA6, forming two possible heterotrimers with counterclockwise and clockwise distributions of the isomers. MD simulations of these two heterotrimers showed that the P-CR/P-CR contacts were stable. Figure S6B,D showed that the distances between each P-CR domain are ~ 30 Å, maintaining an equilateral triangle shape during the MD simulations. However, the rmsd profiles (Figure S6A,C) showed that both heterotrimers shifted dramatically from their original conformation. Comparing structures obtained from the MD simulations (Figures S4 and S5), we found that during 500 ns MD simulations, the P-CR domains remained in an equilateral triangle shape. In contrast, the catalytic domains did not maintain the original parallel directions. As shown in Figure S4, AtCESA3CatD tilted away by $\sim 45^\circ$ in the counterclockwise model, whereas AtCESA1CatD and AtCESA6CatD leaned toward each other by $\sim 15^\circ$. These movements resulted in decreased distances between AtCESA1CatD and AtCESA6CatD and increased distances between AtCESA3CatD and AtCESA1CatD and AtCESA3CatD and

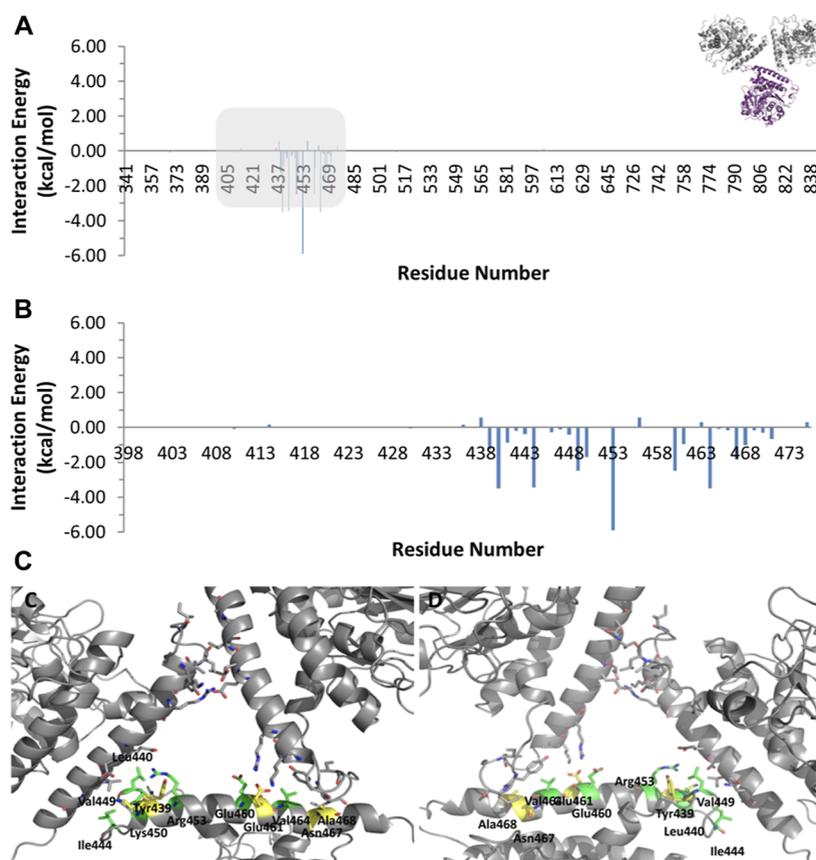


Figure 9. Binding free energies between one AtCESA1CatD (colored in purple) to the other two AtCESA1CatD monomers (colored in grey). (A) Per-residue decomposition binding free energies for residues of AtCESA1CatD colored in purple. The shaded region represents the P-CR domain. (B) Zoom-in view of the per-residue decomposition binding free energies for residues in the P-CR domain. (C) Bottom view of the triangular shape interface formed by P-CR domains in the AtCESA1CatD trimer final structure. (D) Top view of the triangular shape interface formed by P-CR domains in the AtCESA1CatD trimer final structure. All 11 residues that contributed significant binding free energies (<-1 kcal/mol) are shown as sticks. Hydrogen atoms are not shown. Residues which contributed binding free energies greater than 2 kcal/mol are colored in green. Residues which contributed binding free energies between 1 and 2 kcal/mol are colored in yellow.

AtCESA6CatD (Figure S6B). In the clockwise model, the catalytic domains of AtCesA6 tilted by only $\sim 20^\circ$ (Figure S5), and the distances between each catalytic domain did not change dramatically (Figure S6D).

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

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

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.^{6,48–51} It is widely believed that this ordered juxtaposition of individual enzymes is crucial for facilitating alignment and coalescence of emerging β -glucan chains into crystalline cellulose.^{30,52,53} In particular, plants use

several distinct isoforms of synthase to build CSCs that appear to be hexamers of trimeric lobes of CESA proteins. Lobes comprising 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 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 CSCs.

Previously, we reported analysis of freeze-fracture transmission electron micrographs, which was best reconciled with trimeric lobes of CESA proteins in CSCs.⁷ 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 AtCESA1.³⁰ 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.²⁸ The asymmetric unit for this structure was dimeric, but in the crystal lattice, one saw trimers of the P-CR. Based on SEC and SAXS data for the 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 for isolating the domain from the rest of the protein.

Here, we propose that the trimer form seen in AtCESA1CatD is likely to be a functionally relevant form. Computational models have improved and now possess P-CRs similar to those seen in the crystal structure.^{7,54} These were used to generate a library of trimer models, from which to pull candidates that favor trimerization via CSR or P-CR domains.⁴⁴ To supplement the reported physical restraints used to screen the library, we applied four biochemical constraints to screen a similar library of AtCESA1CatD models. The effort brings into focus a small set of models that turn out to also be consistent with the new cryoEM structures for homotrimers of poplar CESA8 and cotton CESA7. The most discriminating information is in the set of DSS cross linking data. The SAXS data and negative stain structure envelopes provide some discrimination but simply lack sufficient detail to be very selective. However, all of the models previously considered for AtCESA1CatD³⁰ supported zero reasonable solvent accessible pathways for the lysine–DSS–lysine cross links. Analysis of the recent set of cotton CESA1 models⁴⁵ for reasonable cross links points to models showing the same trimerization motif that we see for AtCESA1CatD (data not shown). Among all the thousands of models screened, the ones providing the largest number of reasonable cross links are the ones most similar to the structures determined independently for rice, poplar, and cotton CESA proteins. Finally, a microsecond MD simulation confirmed that AtCESA1CatD forms a stable homotrimer. P-CR domains, especially residues R453, L440, I444, V464, V449, E460, N467, K450, Y49, E461, and A468, played a crucial role in the stability of the homotrimer. MD simulations also found a stable heterotrimer of AtCESA1, AtCESA3, and AtCESA6 catalytic domains arranged in the clockwise and counter-clockwise direction, demonstrating that P-CR/P-CR contacts were stable, but the rest of the catalytic domains were

less stable than the homotrimer of AtCESA1CatD. We thus propose that self-association by the P-CR element contributes to trimerization of plant CESAs, both in those making primary (AtCESA1) and secondary (OsCESA8, PttCESA8, and GhCESA7) walls. We note that the two cryoEM structures of homotrimers of CESAs show interaction between TMH 7 of two subunits. Our observation that the heterotrimer of AtCESA1, AtCESA3, and AtCESA6 catalytic domains was less stable than the homotrimer of AtCESA1CatD supports the idea that the TM helix interaction is crucial for stability of heterotrimers of CESAs.

Recently, crystallographic data revealed that AtCESA3CatD can form homodimers in solution.²⁹ The interface between subunits in the dimer structure requires altering the structural context of residues that form a linkage between the finger helix of the catalytic domain and the helix interfacing with the transmembrane helices in the poplar CESA8 and cotton CESA7 cryoEM structures. The same study showed *in vivo* fluorescence data consistent with oligomers—dimers or larger. Mutations that altered residues critical for the dimerization interface also disrupted the *in vivo* oligomerization, causing the authors to hypothesize a dimer phase in the assembly of CESAs. This is consistent with prior work from which an assembly model was proposed for making secondary CSCs, in which homodimers of isoforms CESA4, CESA7, and CESA8 assemble and then further form tetramers and hexamers to make CSCs.¹⁶ However, it is not known if the reported AtCESA3CatD dimer structure is present in the dimers seen for AtCESA7. The latter were shown to be partially sensitive to reducing agents, implying that disulfide crosslinks were important for their stability in SDS-PAGE. The subunit–subunit DSS crosslinks reported here for trimeric and dimeric forms of the AtCESA1CatD that were stable in SDS-PAGE are not plausible in the dimeric CESA3CatD crystal structure, but they are plausible in the trimeric structures of full-length forms of poplar CESA8 and cotton CESA7. This can be directly assessed because the lysine residues involved in DSS crosslinks are conserved in AtCESA3CatD (and other CESAs). An alternative hypothesis explaining the formation of reduction-sensitive dimers of CESA7 is that cysteines in the carboxy terminus, located in the apoplast outside of the cell plasma membrane, form disulfide bonds joining two monomers in a trimeric lobe. Detergent purification of these crosslinked dimers could have given rise to the CESA7 observations.

Since the subunit–subunit DSS crosslinks seen in SDS-PAGE stable dimers reported here are consistent with the P-CR-based trimer structure, it follows that if there were dimers similar to those present in the AtCESA3CatD crystal structure, they were present in amounts below that needed for detection in the MS analysis. The sequence of AtCESA3CatD and AtCESA1CatD is quite similar but not identical. AtCESA1-CatD does contain the residues of the dimerization motif described for AtCESA3CatD, but a vector-derived NT tag is in the AtCESA1CatD protein that is not present in the AtCESA3CatD protein. There was clearly a sub-trimer state material present in the first gel filtration steps we performed to purify trimers of AtCESA1CatD, but we removed much of the smaller material and did not study it further. Further analysis of the discarded material might reveal dimers similar to the crystal structure of AtCESA3CatD.

We also observed some inter-subunit crosslinks between the P-CR and CSR, but these were not supported by any of the best ranking trimer models or the cryoEM structures for

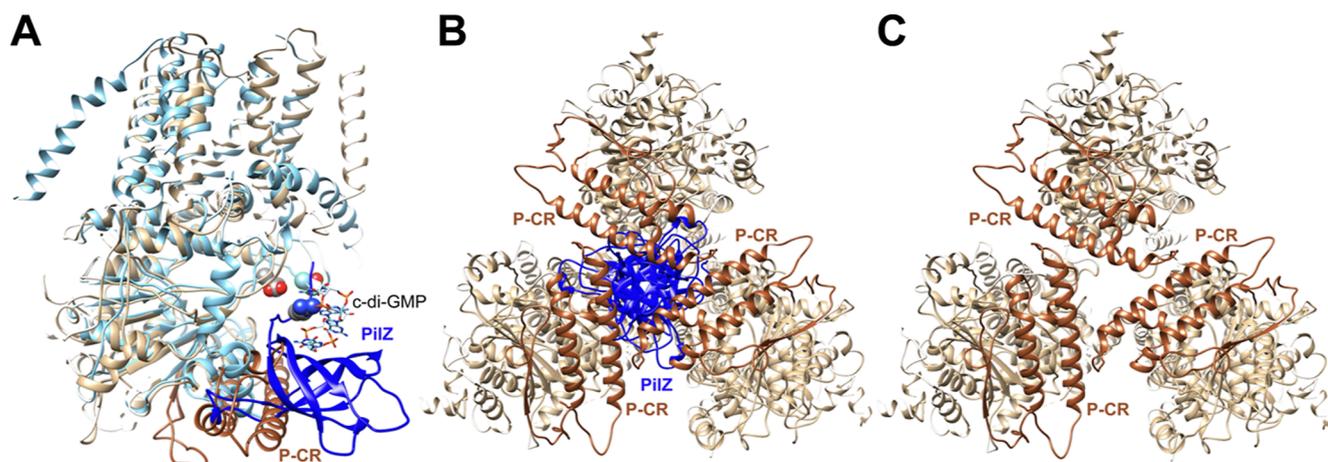


Figure 10. Similar juxtapositions of PilZ and P-CR within bacterial and plant cellulose synthases. (A) Overlay of BcsA (light and dark blue; PDB 4P02) and poplar CesA8 (tan and brown; PDB 6WLB). Dark colors highlight bacterial PilZ and plant P-CR. (B) Overlay of three BcsA subunits on three CesA8 subunits 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.

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 NT and TM-domain truncations. Further studies are necessary to determine if assembly and delivery of CESA trimers to CSCs involve a dimeric phase and to define the juxtaposition of the P-CR and CSR in CSCs.

Trimerizing Motif of the P-CR Replaces the 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 done by the regulatory PilZ domain of bacterial synthases (Figure 10).

It has also been pointed out that putative bacterial synthases may have PilZ, replace it with another domain, or simply lack anything in its place⁵⁵ (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”). These observations raise the possibility that in the evolution of plants, this malleable spatial region of the enzyme fold has become specialized to facilitate trimerization. Moreover, since a small portion of the P-CR is the major place within the cytosolic domains of the trimer where monomers make direct contact, it may also be a conduit for dynamic communication between subunits. Such a communication could certainly contribute to regulation of catalytic activity within a lobe, perhaps synchronizing chain synthesis and translocation across membranes in a “coalescence-friendly” manner. Replacing subunits with the P-CR of different isoforms could fine tune such a communication, which could also be influenced by contact between neighboring lobes, perhaps mediated by CSR or amino-terminal domain interactions. As pointed out by Zhong et al., a previously studied mutant of AtCESA8, *fra6*, maps to P-CR (R362K, same position as 453 of AtCESA1CatD in this study).⁵⁶ Notably, residue R453 of AtCESA1CatD contributed the most stabilizing energy for maintaining P-CR-based trimers during MD simulations.

CONCLUSIONS

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

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biomac.2c00550>.

Fragmentation spectra of inter-subunit crosslinked peptides identified from trimers by pLink, DSS cross links in the trimer of AtCESA1CatD, representative structures of the homotrimer of the AtCESA1 catalytic domain, sequence alignment of AtCesA1, AtCesA3, and AtCesA6, structural alignment of AtCESA1 models obtained from SWISS-MODEL and trRosetta, representative MD simulation structures of catalytic domain heterotrimers of AtCESA1, AtCESA3, and AtCESA6 arranged counterclockwise, representative MD simulation structures of catalytic domain heterotrimers of AtCESA1, AtCESA3, and AtCESA6 arranged clockwise,

rmsd profiles, distances between subunits and between P-CR domains during MDS,3D reconstruction of the AtCESA1CatD trimer by software suite EMAN2, inter-subunit chemical cross links between lysine residues of P-CR and P-CR, DSS cross links in the trimer of AtCESA1CatD MD models, and inter-subunit chemical cross links identified by the MS method, suggesting a P-CR/CSR dimerization between two AtCESA1CatD trimers (PDF)

AUTHOR INFORMATION

Corresponding Author

B. Tracy Nixon – Department of Biochemistry & Molecular Biology, The Pennsylvania State University, State College, Pennsylvania 16802, United States; orcid.org/0000-0002-3568-5703; Email: btn1@psu.edu

Authors

Juan Du – Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610064, China; Department of Biochemistry & Molecular Biology, The Pennsylvania State University, State College, Pennsylvania 16802, United States

Venu Gopal Vandavasi – Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States

Kelly R. Molloy – Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York 10065, United States

Hui Yang – Department of Biology, The Pennsylvania State University, State College, Pennsylvania 16802, United States; orcid.org/0000-0001-6331-6801

Lynnicia N. Massenburg – Department of Biochemistry & Molecular Biology, The Pennsylvania State University, State College, Pennsylvania 16802, United States

Abhishek Singh – Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, United States

Albert L. Kwansa – Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, United States

Yaroslava G. Yingling – Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695, United States; orcid.org/0000-0002-8557-9992

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

Brian T. Chait – Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York 10065, United States

Manish Kumar – Department of Chemical Engineering, The Pennsylvania State University, State College, Pennsylvania 16802, United States; orcid.org/0000-0001-5545-3793

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.biomac.2c00550>

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. J.D. conceived the original research plans;

B.T.N., M.K., B.T.C., H.O., and Y.G.Y. supervised the experiments; J.D. performed most of the experiments with protein initially provided by V.G.V.; subsequent TEM experiments were performed by L.M.; mass spectrometry was performed by K.R.M.; molecular dynamics simulations were performed by H.Y.; J.D., L.M., H.Y., B.T.N., B.T.C., K.R.M., Y.G.Y., A.S., and A.L.K. designed the experiments and analyzed the data; J.D. wrote the article with contributions of all the authors; B.T.N. supervised and complemented the writing.

Notes

The authors declare no competing financial interest.

EMDB Submissions: negative stain reconstructions for the monomer and trimer of AtCESA1CatD have been deposited to the EMDB as D_1000260420 and D_1000260421.

ACKNOWLEDGMENTS

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

ABBREVIATIONS

2D	two dimensional
3D	three dimensional
At	<i>Arabidopsis thaliana</i>
CatD	catalytic domain
CatD_H	catalytic domain containing heavy isotope of lysine
CatD_L	catalytic domain containing light isotope of lysine
CCD	charge coupled device electron detector
CESA#	the number refers to a particular isoform of CESA
CESAs	cellulose synthases
CMF	cellulose microfibrils
COM	center of mass
cryoEM	cryo-electron microscopy
CSCs	cellulose synthase complexes
CSR	hypervariable or class-specific region of plant CESA
CTF	contrast transfer function
DMSO	dimethyl sulfoxide
DSS	disuccinimidyl suberate
DTT	dithiothreitol
EMDB	Electron Microscopy Data Bank
FSC	Fourier shell correlation
Gh	<i>Gossypium hirsutum</i> (cotton)
GT	glycosyltransferase
GT2	glycosyltransferase family 2
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ISAC	iterative stable alignment and clustering
LC	liquid chromatography
LC-MS	liquid chromatography-mass spectrometry
LDS	lauryl dodecyl sulfate
MD	molecular dynamics
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid

MS1, MS2	first and second mass spectrum in tandem mass spectrometry
NH ₄ ⁺	ammonium ion
NMR	nuclear magnetic resonance
NPT	isothermal-isobaric ensemble for MD simulations
NVT	canonical ensemble for MD simulations
Os	<i>Oryza sativa</i> (rice)
PDB	Protein Data Bank
P-CR	plant conserved region of CESA
Ptt	<i>Populus tremula</i> × <i>tremuloides</i> —poplar
rmsd	root mean squared deviation
RMSF	root mean squared fluctuation
SASD	surface accessible solvent distance
SAXS	small angle X-ray scattering
SDS	sodium dodecyl sulfate
SDS-PAGE	polyacrylamide gel electrophoresis with SDS
SEC	size exclusion chromatography
TCEP	tris(2-carboxyethyl)phosphine
TEM	transmission electron microscopy
TIP3P	water molecule for MD simulations
TMH	transmembrane helix

REFERENCES

- Anderson, W. F.; Akin, D. E. Structural and chemical properties of grass lignocelluloses related to conversion for biofuels. *J. Ind. Microbiol. Biotechnol.* **2008**, *35*, 355–366.
- Nookaraju, A.; Pandey, S. K.; Bae, H. J.; Joshi, C. P. Designing cell walls for improved bioenergy production. *Mol. Plant* **2013**, *6*, 8–10.
- Pauly, M.; Keegstra, K. Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* **2008**, *54*, 559–568.
- Zhu, H.; Luo, W.; Ciesielski, P. N.; Fang, Z.; Zhu, J. Y.; Henriksson, G.; Himmel, M. E.; Hu, L. Wood-Derived Materials for Green Electronics, Biological Devices, and Energy Applications. *Chem. Rev.* **2016**, *116*, 9305–9374.
- Mueller, S. C.; Brown, R. M., Jr.; Scott, T. K. Cellulosic microfibrils: nascent stages of synthesis in a higher plant cell. *Science* **1976**, *194*, 949–951.
- Kimura, S.; Laosinchai, W.; Itoh, T.; Cui, X.; Linder, C. R.; Brown, R. M., Jr. Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *vigna angularis*. *Plant Cell* **1999**, *11*, 2075–2086.
- Nixon, B. T.; Mansouri, K.; Singh, A.; Du, J.; Davis, J. K.; Lee, J. G.; Slabaugh, E.; Vandavasi, V. G.; O'Neill, H.; Roberts, E. M.; Roberts, A. W.; Yingling, Y. G.; Haigler, C. H. Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Sci. Rep.* **2016**, *6*, 28696.
- Brown, R. M. J.; Montezinos, D. Cellulose microfibrils: visualization of biosynthetic and orienting complexes in association with the plasma membrane. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 143–147.
- Mueller, S. C.; Brown, R. M. J. Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. *J. Cell Biol.* **1980**, *84*, 315–326.
- Persson, S.; Paredes, A.; Carroll, A.; Palsdottir, H.; Doblin, M.; Poindexter, P.; Khitrov, N.; Auer, M.; Somerville, C. R. Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15566–15571.
- Taylor, N. G.; Laurie, S.; Turner, S. R. Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in *Arabidopsis*. *Plant Cell* **2000**, *12*, 2529–2540.
- Desprez, T.; Juraniec, M.; Crowell, E. F.; Jouy, H.; Pochylova, Z.; Parcy, F.; Höfte, H.; Gonneau, M.; Vernhettes, S. Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15572–15577.
- Taylor, N. G.; Howells, R. M.; Huttly, A. K.; Vickers, K.; Turner, S. R. Interactions among three distinct CesaA proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1450–1455.
- Gardiner, J. C.; Taylor, N. G.; Turner, S. R. Control of cellulose synthase complex localization in developing xylem. *Plant Cell* **2003**, *15*, 1740–1748.
- Wang, J.; Elliott, J. E.; Williamson, R. E. Features of the primary wall CESA complex in wild type and cellulose-deficient mutants of *Arabidopsis thaliana*. *J. Exp. Bot.* **2008**, *59*, 2627–2637.
- Atanassov, I. I.; Pittman, J. K.; Turner, S. R. Elucidating the mechanisms of assembly and subunit interaction of the cellulose synthase complex of *Arabidopsis* secondary cell walls. *J. Biol. Chem.* **2009**, *284*, 3833–3841.
- Hill, J. L., Jr.; Hammudi, M. B.; Tien, M. The *Arabidopsis* cellulose synthase complex: a proposed hexamer of CESA trimers in an equimolar stoichiometry. *Plant Cell* **2014**, *26*, 4834–4842.
- Gonneau, M.; Desprez, T.; Guillot, A.; Vernhettes, S.; Höfte, H. Catalytic subunit stoichiometry within the cellulose synthase complex. *Plant Physiol.* **2014**, *166*, 1709–1712.
- Polko, J. K.; Kieber, J. J. The Regulation of Cellulose Biosynthesis in Plants. *Plant Cell* **2019**, *31*, 282–296.
- Wen, X.; Zhai, Y.; Zhang, L.; Chen, Y.; Zhu, Z.; Chen, G.; Wang, K.; Zhu, Y. Molecular studies of cellulose synthase super-complex from cotton fiber reveal its unique biochemical properties. *Sci. China: Life Sci.* **2022**, DOI: 10.1007/s11427-022-2083-9.
- Zhang, X.; Dominguez, P. G.; Kumar, M.; Bygdell, J.; Miroshnichenko, S.; Sundberg, B.; Wingsle, G.; Niittylä, T. Cellulose Synthase Stoichiometry in Aspen Differs from *Arabidopsis* and Norway Spruce. *Plant Physiol.* **2018**, *177*, 1096–1107.
- Li, X.; Chaves, A. M.; Dees, D. C. T.; Mansoori, N.; Yuan, K.; Speicher, T. L.; Norris, J. H.; Wallace, I. S.; Trindade, L. M.; Roberts, A. W. Cellulose synthesis complexes are homo-oligomeric and hetero-oligomeric in *Physcomitrium patens*. *Plant Physiol.* **2022**, *188*, 2115–2130.
- Purushotham, P.; Cho, S. H.; Díaz-Moreno, S. M.; Kumar, M.; Nixon, B. T.; Bulone, V.; Zimmer, J. A single heterologously expressed plant cellulose synthase isoform is sufficient for cellulose microfibril formation in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 11360–11365.
- Cho, S. H.; Purushotham, P.; Fang, C.; Maranas, C.; Díaz-Moreno, S. M.; Bulone, V.; Zimmer, J.; Kumar, M.; Nixon, B. T. Synthesis and Self-Assembly of Cellulose Microfibrils from Reconstituted Cellulose Synthase. *Plant Physiol.* **2017**, *175*, 146–156.
- Deligey, F.; Frank, M. A.; Cho, S. H.; Kirui, A.; Mentink-Vigier, F.; Swulius, M. T.; Nixon, B. T.; Wang, T. Structure of In Vitro-Synthesized Cellulose Fibrils Viewed by Cryo-Electron Tomography and (13)C Natural-Abundance Dynamic Nuclear Polarization Solid-State NMR. *Biomacromolecules* **2022**, *23*, 2290–2301.
- Zhang, X.; Xue, Y.; Guan, Z.; Zhou, C.; Nie, Y.; Men, S.; Wang, Q.; Shen, C.; Zhang, D.; Jin, S.; Tu, L.; Yin, P.; Zhang, X. Structural insights into homotrimeric assembly of cellulose synthase CesaA7 from *Gossypium hirsutum*. *Plant Biotechnol. J.* **2021**, *19*, 1579–1587.
- Purushotham, P.; Ho, R.; Zimmer, J. Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science* **2020**, *369*, 1089–1094.
- Rushton, P. S.; Olek, A. T.; Makowski, L.; Badger, J.; Steussy, C. N.; Carpita, N. C.; Stauffacher, C. V. Rice Cellulose SynthaseA8 Plant-Conserved Region Is a Coiled-Coil at the Catalytic Core Entrance. *Plant Physiol.* **2017**, *173*, 482–494.
- Qiao, Z.; Lampugnani, E. R.; Yan, X. F.; Khan, G. A.; Saw, W. G.; Hannah, P.; Qian, F.; Calabria, J.; Miao, Y.; Gruber, G.; Persson, S.; Gao, Y. G. Structure of *Arabidopsis* CESA3 catalytic domain with its substrate UDP-glucose provides insight into the mechanism of cellulose synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2024015118.
- Vandavasi, V. G.; Putnam, D. K.; Zhang, Q.; Petridis, L.; Heller, W. T.; Nixon, B. T.; Haigler, C. H.; Kalluri, U.; Coates, L.; Langan, P.; Smith, J. C.; Meiler, J.; O'Neill, H. A Structural Study of CESA1

- Catalytic Domain of Arabidopsis Cellulose Synthesis Complex: Evidence for CESA Trimers. *Plant Physiol.* **2016**, *170*, 123–135.
- (1) Du, J.; Vepachedu, V.; Cho, S. H.; Kumar, M.; Nixon, B. T. Structure of the Cellulose Synthase Complex of *Gluconacetobacter hansenii* at 23.4 Å Resolution. *PLoS One* **2016**, *11*, No. e0155886.
- (2) Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **2007**, *157*, 38–46.
- (3) Mindell, J. A.; Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **2003**, *142*, 334–347.
- (4) Yang, Z.; Fang, J.; Chittuluru, J.; Asturias, F. J.; Penczek, P. A. Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* **2012**, *20*, 237–247.
- (5) Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **2012**, *180*, 519–530.
- (6) Goddard, T. D.; Huang, C. C.; Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **2007**, *157*, 281–287.
- (7) Rosenthal, P. B.; Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **2003**, *333*, 721–745.
- (8) Russo, C. J.; Passmore, L. A. Robust evaluation of 3D electron cryomicroscopy data using tilt-pairs. *J. Struct. Biol.* **2014**, *187*, 112–118.
- (9) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; SalomonFerrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2018*; University of California: San Francisco, 2018.
- (10) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303.
- (11) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (12) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (13) Joung, I. S.; Cheatham, T. E., 3rd Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.
- (14) Yang, J. Y.; Anishchenko, I.; Park, H.; Peng, Z.; Ovchinnikov, S.; Baker, D. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 1496–1503.
- (15) Singh, A.; Kwansa, A. L.; Kim, H. S.; Williams, J. T.; Yang, H.; Li, N. K.; Kubicki, J. D.; Roberts, A. W.; Haigler, C. H.; Yingling, Y. G. In silico structure prediction of full-length cotton cellulose synthase protein (GhCESA1) and its hierarchical complexes. *Cellulose* **2020**, *27*, 5597–5616.
- (16) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33*, W363–W367.
- (17) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. Geometry-based flexible and symmetric protein docking. *Proteins* **2005**, *60*, 224–231.
- (18) Keshk, S. M. Bacterial Cellulose Production and its Industrial Applications. *J. Bioprocess. Biotech.* **2014**, *4*, 150.
- (19) Giddings, T. H., Jr.; Brower, D. L.; Staehelin, L. A. Visualization of particle complexes in the plasma membrane of *Microsterias denticulata* associated with the formation of cellulose fibrils in primary and secondary cell walls. *J. Cell Biol.* **1980**, *84*, 327–339.
- (20) Grimson, M. J.; Haigler, C. H.; Blanton, R. L. Cellulose microfibrils, cell motility, and plasma membrane protein organization change in parallel during culmination in *Dictyostelium discoideum*. *J. Cell Sci.* **1996**, *109*, 3079–3087.
- (21) Kimura, S.; Itoh, T. New cellulose synthesizing complexes (terminal complexes) involved in animal cellulose biosynthesis in the tunicate *Metandrocarpa uedai*. *Protoplasma* **1996**, *194*, 151–163.
- (22) Nicolas, W. J.; Ghosal, D.; Tocheva, E. I.; Meyerowitz, E. M.; Jensen, G. J. Structure of the Bacterial Cellulose Ribbon and Its Assembly-Guiding Cytoskeleton by Electron Cryotomography. *J. Bacteriol.* **2021**, *203*, No. e00371.
- (23) Sethaphong, L.; Haigler, C. H.; Kubicki, J. D.; Zimmer, J.; Bonetta, D.; DeBolt, S.; Yingling, Y. G. Tertiary model of a plant cellulose synthase. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 7512–7517.
- (24) Sethaphong, L.; Davis, J. K.; Slabaugh, E.; Singh, A.; Haigler, C. H.; Yingling, Y. G. Prediction of the structures of the plant-specific regions of vascular plant cellulose synthases and correlated functional analysis. *Cellulose* **2016**, *23*, 145–161.
- (25) Kawano, Y.; Saotome, T.; Ochiai, Y.; Katayama, M.; Narikawa, R.; Ikeuchi, M. Cellulose accumulation and a cellulose synthase gene are responsible for cell aggregation in the cyanobacterium *Thermosynechococcus vulcanus* RKN. *Plant Cell Physiol.* **2011**, *52*, 957–966.
- (26) Zhong, R.; Morrison, W. H., 3rd; Freshour, G. D.; Hahn, M. G.; Ye, Z. H. Expression of a mutant form of cellulose synthase *AtCesA7* causes dominant negative effect on cellulose biosynthesis. *Plant Physiol.* **2003**, *132*, 786–795.