



Small subunit isoform diversity underlies structural heterogeneity in native plant Rubisco

Thomas Reynolds^{a,1}, Zheming Zhang^{b,1} , Dušan Živković^{a,c}, Steven Kelly^a , and Jani R. Bolla^{a,c,2}

Affiliations are included on p. 7.

Edited by Krishna K. Niyogi, University of California, Berkeley, CA; received July 25, 2025; accepted March 19, 2026

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), the most abundant protein on Earth, catalyzes the fixation of CO₂ in photosynthesis. In terrestrial plants, Rubisco assembles as a hexadecameric complex (L₈S₈), comprising eight large subunits (LSu) and eight small subunits (SSu). While LSu is encoded by a single chloroplastic gene, a nuclear multigene family results in diverse SSu protein isoforms, but structural evidence is currently lacking for Rubisco holoenzyme SSu heterogeneity. In this study, utilizing native Rubisco purified from *Arabidopsis thaliana*, we employed high-resolution mass spectrometry and cryo-electron microscopy to demonstrate that multiple SSu isoforms can co-assemble within individual Rubisco complexes. We unambiguously identify the composition of these mixed-isoform complexes and elucidate isoform-specific structural interactions at near-atomic resolution. The structural heterogeneity in plant Rubisco established here will underpin future research to establish the impact of SSu diversity on kinetic and functional plasticity of Rubisco activity under diverse environmental conditions.

Rubisco | cryo-EM | native mass spectrometry | *Arabidopsis thaliana* | structural heterogeneity

Photosynthesis underpins almost all life on Earth, converting carbon dioxide (CO₂) and water into sugars and oxygen using sunlight. Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) catalyzes the first step of CO₂ fixation (1). Rubisco is both essential and inefficient; its slow turnover and competing oxygenase reaction limit photosynthetic output, but its abundance compensates for these limitations (2, 3). In land plants, Rubisco is a hexadecameric (L₈S₈) complex composed of eight chloroplast-encoded large subunits (LSu) and eight nucleus-encoded small subunits (SSu). Nucleus-encoded subunits are translated in the cytosol as preproteins with an N-terminal chloroplast-targeting sequence, called a transit peptide, which is cleaved and degraded after importing preproteins into the chloroplast through translocons of the outer and inner chloroplast membrane (TOC and TIC, respectively) (4, 5). Once folded in the stroma, the SSu is complexed with folded LSu, assisted by chaperones, to form the functional holoenzyme, L₈S₈ (SI Appendix, Fig. S1A). The large subunits form the catalytic core, whereas the small subunits cap its ends and are required for correct folding, assembly, and function. LSu is strikingly conserved among higher plants, but the small subunits are not. Most species express multiple SSu isoforms, ranging from a few to more than twenty (6). Differential expression of these isoforms across tissues and environmental conditions has been reported in *Arabidopsis*, tomato, tobacco, rice, and maize (6), suggesting a regulatory complexity whose mechanistic basis remains poorly understood.

Although SSu does not directly participate in catalysis, it plays an essential role in Rubisco biogenesis, structural stability, and catalytic efficiency. SSu contributes to enzyme assembly, enhances CO₂ binding efficiency, and may serve as a reservoir for CO₂ (6–8). In several cases, replacing the native SSu with a heterologous isoform has been sufficient to alter carboxylation rates and specificity (8–12). However, the structural determinants of this functional variability remain obscure, largely due to an incomplete understanding of how different SSu isoforms interact with LSu and influence enzyme activity. A fundamental question also remains unresolved: Do individual L₈S₈ complexes in planta contain a single small-subunit isoform, or can multiple isoforms coexist within the same particle?

Over the past decades, increasing efforts have focused on Rubisco biogenesis, structural stability, and engineering, motivated by the urgent need to improve photosynthetic efficiency in crops under rising global demands and climate stress (6, 7, 13). To guide these investigations, several Rubisco structures have been solved at near-atomic resolution, revealing molecular details of protein–protein interactions within Rubisco. However, nearly all available models contain only a single SSu isoform, likely due to crystal packing

Significance

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), the key enzyme in photosynthetic CO₂ fixation, has long challenged efforts to understand its assembly, regulation, and potential for enhancement. Here, we show that multiple small subunit (SSu) isoforms can coexist within individual Rubisco complexes in *Arabidopsis thaliana*. Combining native protein extraction, high-resolution mass spectrometry, and cryo-electron microscopy, we define these mixed-isoform assemblies and reveal how different SSu variants shape local structural environments. Previous studies have demonstrated that plants maintain diverse SSu gene families and suggest that isoform composition might fine-tune Rubisco activity in response to developmental and environmental changes. Our results provide fundamental insights into Rubisco composition at the holoenzyme level and a structural framework for interpreting differences in Rubisco activity arising from varying isoform compositions.

Competing interest statement: S.K. is a co-founder of Wild Bioscience Ltd. and an employee of Ellison Institute of Technology, Oxford Limited.

This article is a PNAS Direct Submission.

Copyright © 2026 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

¹T.R. and Z.Z. contributed equally to this work.

²To whom correspondence may be addressed. Email: jani.bolla@bioch.ox.ac.uk.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2519949123/-/DCSupplemental>.

Published April 15, 2026.

constraints. Only two studies to date report any heterogeneity within the L_8S_8 complex (14, 15), but no heterogeneous structures have been deposited in the protein databank (PDB), reinforcing the prevailing assumption that Rubisco incorporates only one SSu isoform per assembly, despite clear evidence that multiple isoforms are expressed simultaneously in most plants and their relative abundance shifts with changing environmental and developmental context (6, 16). Under such circumstances, it is not obvious why every Rubisco complex should be compositionally identical (17, 18). Until very recently, the available methodological toolkit was not well-suited to answering this question decisively.

Clarifying the true composition of plant Rubisco is important for both fundamental biology and practical applications. Most engineering strategies, whether importing faster algal Rubiscos into crops (several crop species also express various SSu isoforms, see *SI Appendix, Fig. S2* for a few examples) or redesigning plant small subunits to improve CO_2 specificity, focus on introducing a single “superior” SSu into plants. However, it is not yet known if native Rubisco typically incorporates a mixture of SSu isoforms with distinct structural roles, and the potential structure–function consequences of mixed populations have not yet been considered. For example, it is not known whether natively produced heterogeneous Rubisco would be beneficial, detrimental, or neutral for Rubisco assembly, stability, function, or regulation. To explore the extent and structural ramifications of heterogeneous Rubisco, we combined high-resolution cryo–electron microscopy (cryo-EM) with mass spectrometry (MS) to examine Rubisco purified directly from the model plant, *Arabidopsis thaliana*. MS unambiguously identifies the presence of multiple SSu isoforms, while cryo-EM resolves their spatial arrangement and isoform-specific interactions within the holoenzyme. These findings challenge the long-standing model of isoform exclusivity and establish that mixed assemblies occur *in planta*.

Results

Isolation of Intact Rubisco from Plant Leaves. To investigate the native assembly of Rubisco and assess whether distinct SSu isoforms can coexist within a single holoenzyme, we purified Rubisco from *A. thaliana*. Chloroplasts were isolated from 14-d-old plants and lysed using a hypotonic solution. The stroma of lysed chloroplasts

was obtained using previously established fractionation methods (19). The stromal soluble fraction was further concentrated using a 100 kDa molecular weight concentrator to enrich large complexes. SDS–PAGE analysis showed that the preparation was dominated by Rubisco, with prominent bands at ~55 kDa and ~15 kDa corresponding to the LSu and SSu bands (*Fig. 1A* and *SI Appendix, Fig. S3*). To verify that the purified enzyme remained intact and correctly assembled, we performed mass photometry, a label-free single-molecule technique that enables stoichiometric estimation of protein complexes in solution (*SI Appendix, Fig. S4*). The resulting histogram revealed a major symmetric peak centered at approximately 541 kDa, consistent with the theoretical molecular mass of an L_8S_8 complex. These results indicate that the isolation of intact Rubisco from plant leaves preserved Rubisco’s native architecture.

Isolated Rubisco Contains Multiple SSu Isoforms. *A. thaliana* expresses four SSu isoforms (SSu1A, SSu1B, SSu2B, and SSu3B) (*SI Appendix, Fig. S1B*). The mature forms of SSu2B and SSu3B are identical (henceforth denoted SSu2B/3B in this paper). To identify which of these isoforms are present in our Rubisco preparation, we performed bottom–up proteomics. This approach allowed us to detect isoform-specific peptides of SSu (*SI Appendix, Fig. S3*), and their relative quantification is shown in *Fig. 1B*. When considered semiquantitatively, the relative intensities of isoform-specific peptides are broadly consistent with the *rbcS* expression patterns observed across 21,154 publicly available RNA-seq libraries (with SSu2B and SSu3B aggregated because they are indistinguishable at the protein level). After aggregating the nearly identical SSu2B and SSu3B transcripts, expression profiles revealed that SSu2B/3B dominate in most photosynthetic tissues, whereas SSu1A and SSu1B are more variably expressed across developmental stages and environmental conditions (*Fig. 1C*). The close concordance between transcript abundance and proteoform representation strongly suggests that the detected mixture of SSu isoforms reflects bona fide *in planta* diversity rather than adventitious association or contamination. Together, these findings indicate that *A. thaliana* Rubisco does not exist as a homogeneous population of identical L_8S_8 complexes. Instead, it comprises a heterogeneous ensemble of assemblies that incorporate various combinations of SSu isoforms, potentially expanding the enzyme’s functional versatility *in planta*.

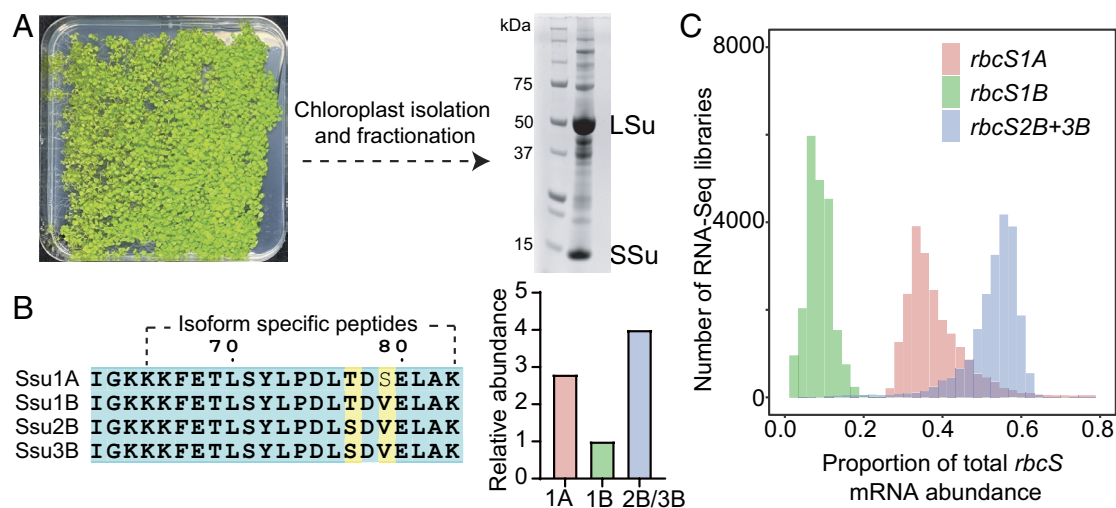


Fig. 1. Native Rubisco contains multiple SSu isoforms. (A) Leaves from 14-d-old *A. thaliana* plants were used for chloroplast isolation, followed by fractionation to obtain the soluble stromal fraction. SDS–PAGE (Right) shows that Rubisco is abundant in this stromal fraction. Notably, this workflow was designed to enrich the chloroplast stroma rather than to purify Rubisco to homogeneity, thereby reducing the risk of losing protein subpopulations during more stringent purification. (B) Regions of the SSu sequence containing isoform-specific peptides identified by proteomics, and their relative quantification (bar graph, Right) is consistent with the *rbcS* expression profile across 21,154 publicly available RNA-seq libraries (C).

SSu1A and SSu2B/3B Are the Two Stably Binding Isoforms. To assess whether the above-identified SSu isoform ratios reflect stable incorporation into L_8S_8 complexes, rather than weak or transient association, we turned our attention to high-resolution native MS coupled with in-source dissociation and top-down tandem fragmentation (MS/MS). This approach allows detection of intact proteoforms, their posttranslational modifications, and sequence variations without denaturation or proteolytic digestion. We first analyzed the above purified stromal soluble fraction, which predominantly contained Rubisco, by native MS, and the resultant spectrum revealed a major charge-state series whose mass corresponded to the intact L_8S_8 complex (Fig. 2). Next, we applied in-source activation energy to release intact SSu monomers, namely m/z corresponding to 2,453 and 2,469, which were subsequently isolated and fragmented using higher-energy collisional dissociation (HCD). The resulting spectra contained a plethora of fragment ions that yielded unambiguous sequence coverage for SSu1A and SSu2B/3B proteoforms (Fig. 2 and *SI Appendix, Fig. S5*), with weaker signals from SSu1B. Of note, all “B” isoforms (1B, 2B, 3B) of SSu share the same molecular weight. The relative signal intensities of these species suggest a nonstoichiometric but reproducible representation of multiple isoforms within the same preparation. Interestingly, all detected SSu proteoforms displayed an N-terminal methylation, a conserved modification previously reported in spinach and tobacco Rubisco, that has been proposed to enhance holoenzyme stability (20).

Spatial Distribution of SSu Isoforms around the L_8 Core Is Disordered. To directly visualize how different SSu isoforms are arranged within the holoenzyme, we subjected the same Rubisco preparations from above to cryo-EM analysis. Over 250,000 particle images were collected and processed using single-particle reconstruction workflows, yielding two major structural states (state I and II) (Fig. 3 and *SI Appendix, Fig. S6 and Table S1*), each

characterized by a distinct composition and spatial distribution of SSu isoforms. In both states, the overall architecture resembles previously reported structures (16). Briefly, the large subunits form an L_8 core made up of a tetramer of dimers, with four small subunits forming cap-like coverings at the top and four capping the bottom of this barrel-shaped LSu core. Focused refinement of the small-subunit regions, however, revealed subtle but reproducible density differences consistent with compositional heterogeneity. In State I, each end of the core is capped by one SSu1A and three SSu2B/3B subunits, while in State II, the upper cap contains two SSu1A and two SSu2B/3B subunits, and the lower cap consists exclusively of four SSu2B/3B (Fig. 3). These observations are fully consistent with the relative isoform abundances observed by MS, providing orthogonal validation of mixed-isoform assembly.

Isoform-Specific Residues Influence Rubisco Core Stability. SSu1A and SSu2B/3B differ in seven positions: Q2K, T22S, S24V, I34L, S58T, N96G, and G124E (*SI Appendix, Fig. S1B*). Our cryo-EM density maps enabled confident assignment of side-chain identities at six of the seven positions distinguishing SSu1A from the B-type isoforms (Q2K, T22S, S24V, I34L, S58T, N96G) (*SI Appendix, Fig. S7*). The substitutions cluster at two interfaces: near the C-terminal β -barrel of SSu and the $\alpha 8$ - αK helices of LSu, which together shape the catalytic pocket (Fig. 4 and *SI Appendix, Fig. S8*). Notably, residue 2 in SSu (Q in SSu1A, K in SSu2B/3B) sits within 5.5 Å of LSu helix $\alpha 8$ and within 3.9 to 4.6 Å of E454 on helix αK . The lysine present in SSu2B/3B is therefore well positioned to form a stabilizing salt bridge with E454, which could, in turn, influence the conformational dynamics of loop 6 during CO_2 fixation (21). Given the central role of loop 6 in carboxylation chemistry, even such minor alterations may modulate catalytic turnover or substrate affinity.

The conservative serine-to-threonine substitution at position 58 occurs on the βA - βB loop that projects into the solvent

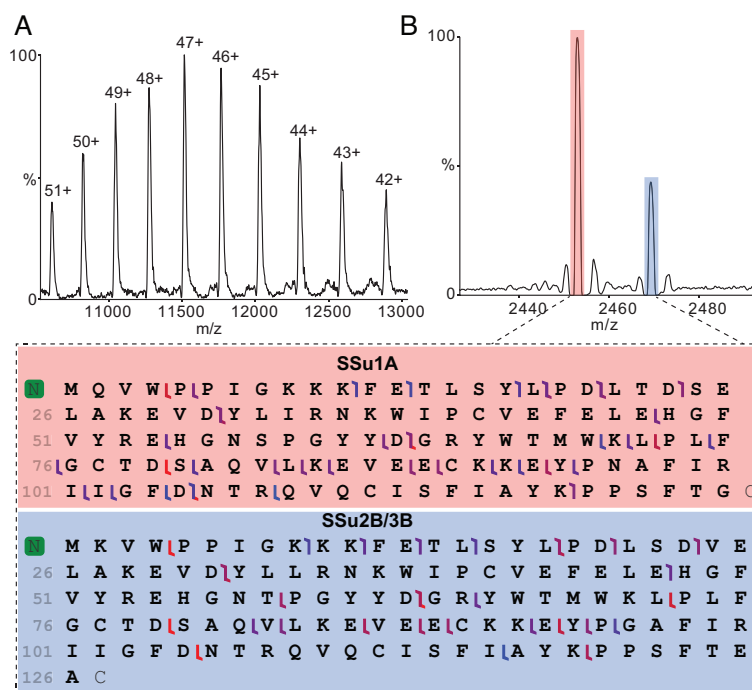


Fig. 2. Top-down MS reveals the cooccurrence of two SSu isoforms in the native Rubisco population. (A) Native MS of Rubisco yields a charge-state envelope that deconvolutes to a measured mass of $541,220 \pm 19$ Da, consistent with the intact L_8S_8 holoenzyme, confirming that the complex remains assembled during analysis. Charge states are labeled for clarity. (B) Increasing collisional activation dissociates SSu subunits from the holoenzyme, yielding two SSu signals at $\sim 2,453$ and $\sim 2,469$ m/z . Top-down fragmentation (using HCD) of these SSu ions generates isoform-specific fragment ions that uniquely identify SSu1A and SSu2B/3B. Together, these data show that the native Rubisco population contains more than one SSu isoform.

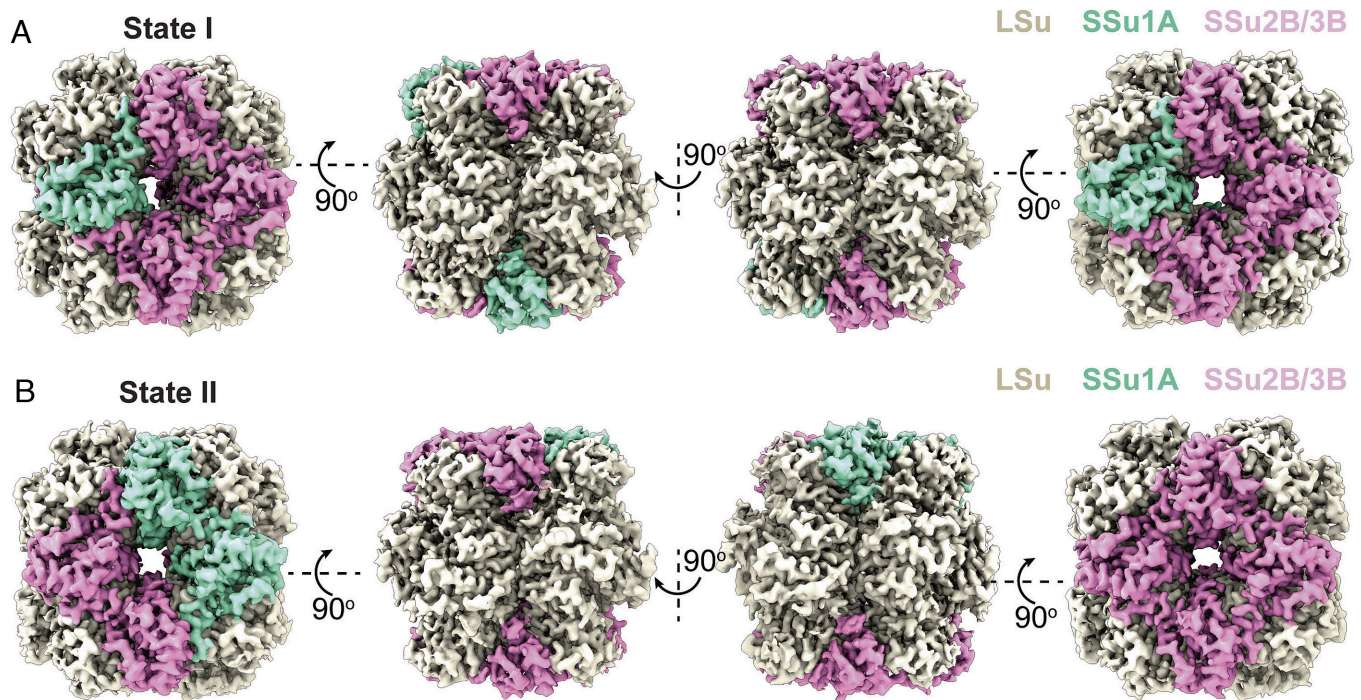


Fig. 3. Cryo-EM analysis reveals two distinct heterogeneous architectures of Rubisco. Cryo-EM reconstructions of the native enzyme identify two structural states, State I (A) and State II (B), each represented as surface views from multiple orientations (Top view—Left, Bottom view—Right). LSu are shown in beige, SSu1A in green, and SSu2B/3B in magenta. The two states differ in both the composition and spatial arrangement of SSu1A and SSu2B/3B around the L_8 core, demonstrating that native Rubisco exists as a mixture of distinct isoform-containing assemblies.

channel. Although the polarity is retained, the additional methyl group in threonine may reduce flexibility, thereby modulating access to the active-site cavity. Similarly, the N96G substitution shortens the side chain in the α B– β C loop, possibly affecting intersubunit packing or allosteric communication between the small and large subunits (Fig. 4 and SI Appendix, Fig. S8). The terminal SSu residues (124 and 125) remain unresolved in the current structural model, indicating inherent flexibility within this region. Collectively, these findings reveal that even minimal sequence divergence among SSu isoforms can exert localized structural effects that propagate to the LSu core, underscoring the evolutionary fine-tuning of Rubisco stability and activity.

Comparison with Reported Structures Reveals Substantial Conformational Flexibility. To assess whether isoform-specific arrangements influence the global conformation of Rubisco, we performed pairwise superimpositions of State I and State II structures with three reference structures, 5IU0 (an *A. thaliana* crystal structure), 9MUR, and 9N37 (two structures reported in the accompanying paper, with 9MUR containing all eight SSu2B/3B and 9N37 containing all eight SSu1A) yielding rmsd values ranging from 0.7 to 1.6 Å (Fig. 5). When aligned on one face of the holoenzyme, the superimposed side shows good structural agreement (gray). In contrast, the opposite face exhibits greater positional shifts in both LSu and SSu regions. Local structural comparisons reveal discrete conformational differences in SSu1A (magenta) and SSu2B/3B (cyan/green) isoforms, as well as in specific large subunit domains (blue), with notable structural deviations. On the heterogeneous SSu side, both State I and State II adopt more extended conformations relative to 5IU0, yet are notably more compact than 9MUR and 9N37.

Importantly, States I and II represent the first Rubisco structures from *A. thaliana* with heterogeneous small-subunit composition, incorporating both SSu1A and SSu2B/3B isoforms. Notably, State II exhibits an unprecedented asymmetric distribution of SSu

isoforms (2 SSu1A+2 SSu2B/3B on one face; 4 SSu2B/3B on the opposite face), a structural arrangement that has not been documented in any previous Rubisco structures. All prior *Arabidopsis* Rubisco structures, including 5IU0, contained homogeneous small-subunit populations, leading to speculation that heterogeneous populations cannot be accommodated in ordered crystals. Our cryo-EM structures definitively demonstrate that *Arabidopsis* Rubisco can assemble with mixed SSu isoforms in planta, resolving this long-standing question and establishing a unique architectural arrangement that distinguishes these states from all previously characterized Rubisco structures.

Discussion

Our findings indicate that Rubisco in *A. thaliana* is not a uniform L_8S_8 complex but a mixed assembly where multiple SSu isoforms co-assemble in specific spatial arrangements. This challenges the longstanding assumption that each holoenzyme contains eight identical small subunits, instead revealing a structurally diverse complex. The most straightforward interpretation is that this compositional heterogeneity provides an additional layer of flexibility, allowing Rubisco to adapt its properties to different physiological or environmental conditions. For instance, *Arabidopsis* SSu “B” isoforms are upregulated at elevated temperatures (18), consistent with a role in enhancing stability and performance under heat stress.

While the majority of complexes detected by native MS/MS and resolved by cryo-EM exhibited mixed SSu composition, we cannot exclude the presence of a minor population of homogeneous L_8S_8 assemblies below the detection limit of our methods. In this context, the previously reported *A. thaliana* Rubisco crystal structure 5IU0, purified from a native source and interpreted as containing eight copies of a single low-abundance SSu isoform, likely represents one end of a continuum of possible SSu compositions under different physiological or experimental conditions. Our data suggest that

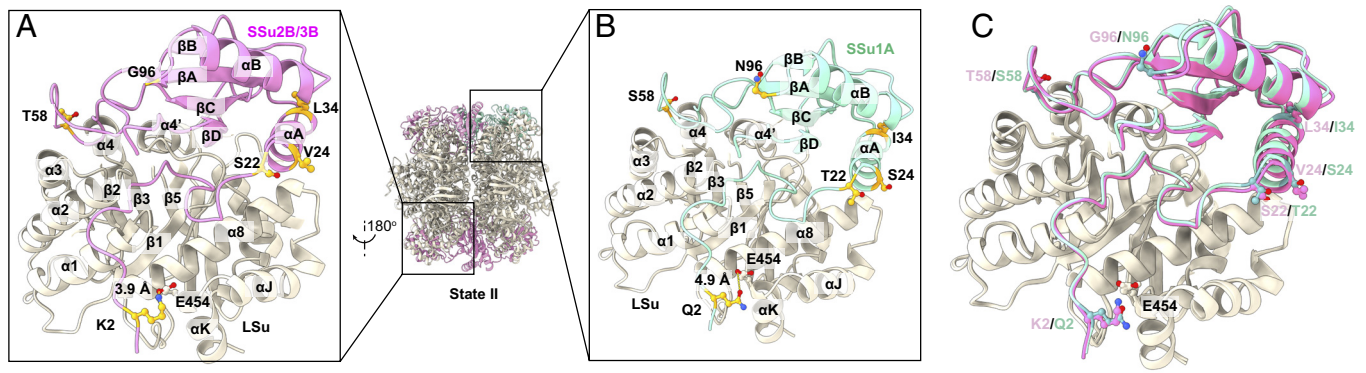


Fig. 4. Isoform-specific interactions at the Rubisco LSU-SSu interface in State II. (A) Zoomed-in view of the interface between the LSU (beige) and SSu2B/3B (magenta), highlighting isoform-specific contacts at the N terminus and β -barrel of the SSu. Selected interacting residues are shown as sticks, with putative hydrogen bonds/salt bridges indicated by dashed lines and distances. (B) Corresponding view of the LSU-SSu1A interface (LSU in beige, SSu1A in teal), illustrating alternative side-chain identities and geometries at analogous positions that give rise to a distinct interaction network. (C) Overlay of the LSU-SSu2B/3B (magenta) and LSU-SSu1A (teal) complexes, showing that the LSU backbone is largely conserved while the different SSu isoforms adopt subtly shifted orientations and side-chain conformations at the interface. Rubisco secondary-structure elements are labeled according to the established convention (19).

heterogeneous assemblies predominate under the growth and purification conditions used here, but homogeneous complexes may also occur transiently or under alternative environmental regimes. Compositional heterogeneity of Rubisco appears pronounced in *Arabidopsis* and in several other species with diversified RbcS/SSu families (for example, the examples shown in *SI Appendix, Fig. S2*) but may be minimal in species where multiple *rbcS* genes encode identical mature SSus, for example, rice.

Recognizing that native Rubisco is built from a mixture of SSu isoforms may also help explain why some Rubisco-engineering efforts have produced confounding results. Many studies have reconstructed the enzyme with a single SSu variant or replaced plant small subunits with those from faster algal Rubiscos (6, 7). In most cases, these complexes do not perform as well as the native enzyme. Our findings suggest that these strategies might have missed a key point: In planta, Rubisco does not depend on a single “best” small subunit but on a carefully balanced ensemble. The unique kinetic effect of heterogeneity is demonstrated in our companion paper, which finds that compositionally heterogeneous Rubisco has an intermediate turnover rate but decreased substrate and inhibitor affinities relative to its pure counterparts (22). Therefore, engineering efforts must consider potential heterogeneity-related kinetic effects when introducing nonnative SSus.

The hypotheses arising from this work are now directly testable. Genome editing allows the deletion of specific *rbcS* genes, creation of defined isoform combinations, or swapping of promoters to modify their expression profiles. Examining the effects on Rubisco kinetics, assembly efficiency, and in planta photosynthetic performance would establish a direct link between structural heterogeneity and function. Complementary reconstitution in manageable hosts such as *Nicotiana* or *Chlamydomonas*, including existing hybrid Rubisco systems (23), could aid in distinguishing effects on biogenesis from those on catalysis and in determining the extent to which native-like mixtures are necessary to restore full activity.

Methodologically, the combined use of native top-down MS and cryo-EM provides a versatile approach for examining multi-isoform protein complexes in plants. Top-down MS identifies which proteoforms are present within intact assemblies, while cryo-EM shows how these variants are spatially organized in 3D. Together, these methods reveal structural heterogeneity that would otherwise be averaged out in bulk biochemical analyses or hidden during crystallization. The same framework should be applicable to other paralog-rich chloroplast machinery where isoform-specific functions remain poorly understood, including Rubisco activase, the Cpn60-RbcL chaperone system, and components of the TOC-TIC protein import pathway (4, 23–27).

More broadly, our findings suggest a conceptual shift. Enhancing enzyme function in plants may not be achieved by perfecting a single protein sequence, but by maintaining a group of closely related isoforms that collectively balance rigidity and flexibility. Such compositional plasticity would enable quick adjustment of enzymatic properties in response to changing environments, without the need for extensive rewiring of transcriptional programs or metabolic networks. From this perspective, Rubisco heterogeneity is not a defect to be eliminated but a feature that evolution has repeatedly favored and preserved.

Materials and Methods

Rubisco Sample Purification. *A. thaliana* Col-0 chloroplast stroma was extracted by existing methods (19), and Rubisco complexes further enriched by concentrating with MWCO 100,000 kDa filter.

MS.

Bottom-up MS. Bottom-up proteomics analysis was performed using an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific). Data were acquired in a data-dependent acquisition mode employing a TopN method with HCD fragmentation and a 1 h elution gradient HPLC method, as previously described (28). Raw MS data were processed and analyzed using MaxQuant (29). Peptide intensities were normalized across runs to account for technical variability.

Native MS. Purified Rubisco was buffer exchanged into MS Buffer (200 mM ammonium acetate pH 8.0) using a centrifugal buffer exchange device (Micro Bio-Spin 6, Bio-Rad). Samples were introduced directly into the mass spectrometer using gold-coated capillary needles (prepared in-house). Data were collected on a Q-Exactive UHMR mass spectrometer (Thermo Fisher Scientific). The instrument parameters were as follows: capillary voltage 1.1 kV, S-lens RF 100%, quadrupole selection from 1,000 to 20,000 m/z range, collisional activation in the HCD cell 100 to 200 V, trapping gas pressure setting 7.5, temperature 200 °C, and resolution of the instrument 12,500. The noise level was set at 3 rather than the default value of 4.64. No in-source dissociation was applied. Data were analyzed using Xcalibur 4.2 (Thermo Scientific) and UniDec software packages.

Top-down MS. The top-down analysis was also performed on the same instrument. The typical MS settings were spray voltage of 1.1 kV, source temperature of 200 °C. The SSu subunits were dissociated from the complex with 100 V using in-source fragmentation. The dissociated subunit peaks were selected using ion trap with a window of m/z 10. The selected ions were then further fragmented using HCD energy. The fragment ions were detected in the Orbitrap. The top-down mass spectra were manually validated in TDValidator (Proteinaceous) (30).

Cryo-EM Sample Preparation. For imaging Rubisco, a 2.5 μ l of 5 mg/ml stroma sample was directly applied to glow-discharged holey carbon grids (Quantifoil Cu R1.2/1.3, 300 mesh), blotted for 18 s and then plunge-frozen in liquid ethane

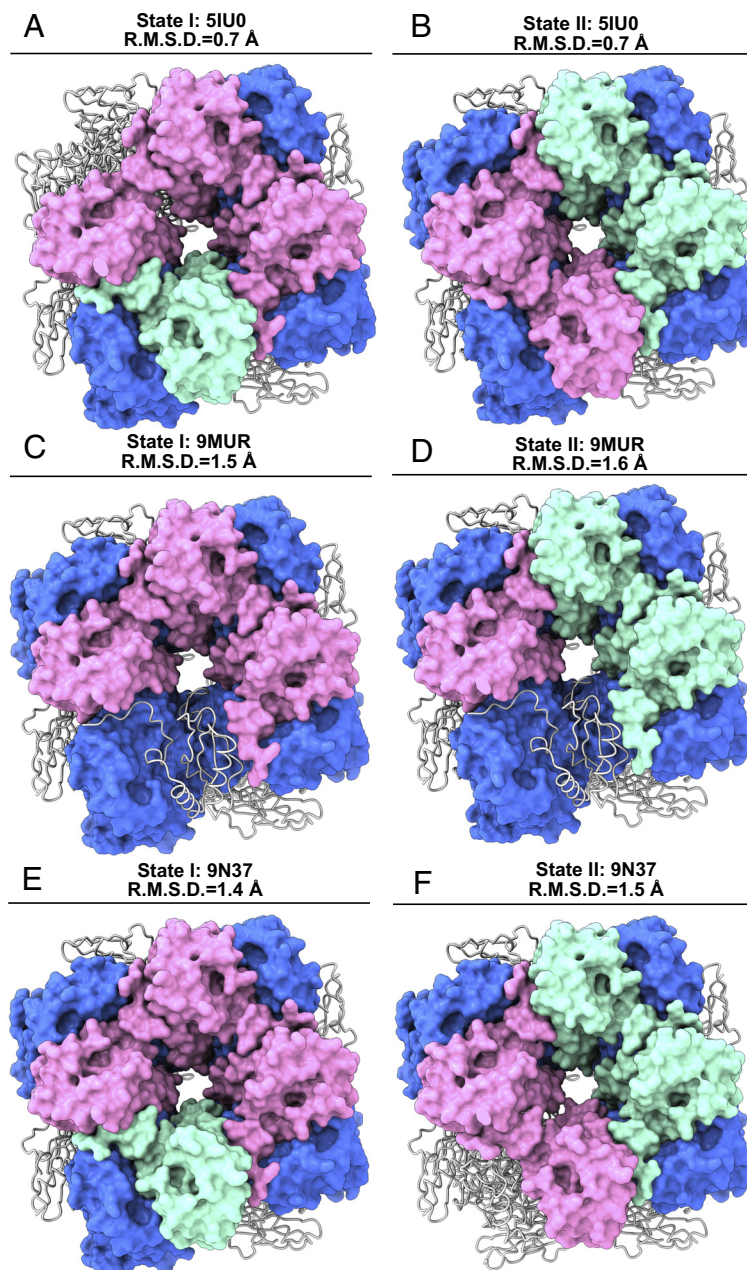


Fig. 5. Structural comparison of State I and State II Rubisco with previously reported structures. Superimpositions of the cryo-EM structures of *A. thaliana* Rubisco State I and State II with available Rubisco models. In all panels, structurally identical regions are shown in a gray string. Variable regions are highlighted in color surface: SSu1A in green, SSu2B/3B in magenta, and LSu in blue. (A and B) Overlay of State I (A) and State II (B) with the *A. thaliana* Rubisco crystal structure (PDB ID: 5IU0, rmsd = 0.7 Å). (C and D) Overlay of State I (C) and State II (D) with the recently reported Rubisco structure 9MUR (rmsd = 1.5 Å and 1.6 Å, respectively). (E and F) Overlay of State I (E) and State II (F) with Rubisco structure 9N37 (rmsd = 1.4 Å and 1.5 Å, respectively).

using a Vitrobot (Thermo Fisher). For imaging, all grids were then transferred into cartridges prior to data collection.

Cryo-EM Data Collection. The images were collected in superresolution mode at 81 K magnification on a Titan Krios equipped with a K3 direct electron detector (Gatan). The physical pixel size was 1.07 Å/pix (superresolution of 0.535 Å/pix). Each micrograph was exposed to a total dose of 37.8 e⁻/Å² for 3.5 s and 40 frames were captured using SerialEM (31).

Cryo-EM Data Processing. The super-resolution image stack was aligned and binned by 2 using patch motion. The contrast transfer function (CTF) was estimated using patch CTF in cryoSPARC (32). A procedure for blob picker followed by 2D classification were applied to generate templates for automated template picking. Initially, 1,659,413 particles were selected after autopicking in cryoSPARC. Several iterative rounds of 2D classifications followed by ab initio

and heterogeneous 3D classifications were performed to remove false picks and classes with unclear features, ice contamination, or carbon. The 3D classification analysis was then employed, resulting in two distinct classes of Rubisco images. Non-uniform refinement followed by local refinement with nonuniform sampling resulted in 2.87 Å and 2.97 Å resolution cryo-EM maps for two forms Rubisco based on the gold standard Fourier shell correlation (FSC 0.143).

Model Building and Refinement. Model buildings of Rubisco state I and state II were based on the cryo-EM maps, respectively. A predicted Rubisco structure using AlphaFold (33) was used and fitted into the corresponding density maps using Chimera (34). The subsequent model rebuilding was performed using Coot (35). Structural refinements were performed using the phenix.real_space_refine program (36) from the PHENIX suite (37). The final atomic model was evaluated using MolProbity (38). The statistics associated with data collection, 3D reconstruction, and model refinement are included in [SI Appendix, Table S1](#).

Data, Materials, and Software Availability. Cryo-EM maps and atomic coordinates data have been deposited in EMDB and PDB [EMD-70094 (39) and 904C (40) for state I and EMD-70090 (41) and 9049 (42) for state II].

ACKNOWLEDGMENTS. This research was supported by the Royal Society (URF/R1\211567 awarded to J.R.B.); UK Research and Innovation Frontier Research Guarantee Grant (EP/Y036158/1 awarded to J.R.B.) and the Wellcome Trust (226598/Z/22/Z awarded to S.K.). D.Ž. is supported by the European Research Executive Agency under the HORIZON TMA Marie Skłodowska-Curie grant agreement (101205973). We are grateful to the Cryo-Electron Microscopy Core at the

CWRU School of Medicine, as well as to Dr. Corinne Lutomski and Jack Bennett, for their insightful discussions.

Author affiliations: ^aDepartment of Biology, University of Oxford, Oxford OX1 3RB, United Kingdom; ^bDepartment of Pharmacology, Case Western Reserve University, Cleveland, OH 44106; and ^cDepartment of Biochemistry, University of Oxford, Oxford, OX1 3QU, United Kingdom

Author contributions: T.R., Z.Z., and J.R.B. designed research; T.R., Z.Z., D.Ž., S.K., and J.R.B. performed research; J.R.B. contributed new reagents/analytic tools; T.R., Z.Z., D.Ž., S.K., and J.R.B. analyzed data; and T.R., Z.Z., and J.R.B. wrote the paper.

1. J. A. Raven, Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. *Aquat. Microb. Ecol.* **56**, 177–192 (2009).
2. A. I. Flamholz *et al.*, Revisiting trade-offs between Rubisco kinetic parameters. *Biochemistry* **58**, 3365–3376 (2019).
3. Y. M. Bar-On, R. Milo, The global mass and average rate of Rubisco. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4738–4743 (2019).
4. Y. Sun, R. P. Jarvis, Chloroplast proteostasis: Import, sorting, ubiquitination, and proteolysis. *Annu. Rev. Plant Biol.* **74**, 259–283 (2023).
5. S. Ramundo *et al.*, Coexpressed subunits of dual genetic origin define a conserved supercomplex mediating essential protein import into chloroplasts. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 32739–32749 (2020).
6. Y. W. Mao *et al.*, The small subunit of Rubisco and its potential as an engineering target. *J. Exp. Bot.* **74**, 543–561 (2023).
7. N. Prywes, N. R. Phillips, O. T. Tuck, L. E. Valentin-Alvarado, D. F. Savage, Rubisco function, evolution, and engineering. *Annu. Rev. Biochem.* **92**, 385–410 (2023).
8. R. J. Spreitzer, Role of the small subunit in ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.* **414**, 141–149 (2003).
9. H. Matsumura *et al.*, Hybrid rubisco with complete replacement of rice rubisco small subunits by sorghum counterparts confers C(4) plant-like high catalytic activity. *Mol. Plant.* **13**, 1570–1581 (2020).
10. E. Martin-Avila *et al.*, Modifying plant photosynthesis and growth via simultaneous chloroplast transformation of Rubisco large and small subunits. *Plant Cell* **32**, 2898–2916 (2020).
11. S. Karkehabadi *et al.*, Chimeric small subunits influence catalysis without causing global conformational changes in the crystal structure of ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Biochemistry* **44**, 9851–9861 (2005).
12. T. Genkov, R. J. Spreitzer, Highly conserved small subunit residues influence Rubisco large subunit catalysis. *J. Biol. Chem.* **284**, 30105–30112 (2009).
13. M. Gionfriddo, K. Zang, M. Hayer-Hartl, The challenge of engineering Rubisco for improving photosynthesis. *FEBS Lett.* **597**, 1679–1680 (2023).
14. N. Shibata *et al.*, Orderly disposition of heterogeneous small subunits in D-ribulose-1, 5-bisphosphate carboxylase/oxygenase from spinach. *J. Biol. Chem.* **271**, 26449–26452 (1996).
15. P. C. Loewen *et al.*, Structure of *Pisum sativum* Rubisco with bound ribulose 1, 5-bisphosphate. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **69**, 10–14 (2013).
16. K. Valegard, D. Hasse, I. Andersson, L. H. Gunn, Structure of Rubisco from *Arabidopsis thaliana* in complex with 2-carboxyarabinitol-1, 5-bisphosphate. *Acta Crystallogr. D Struct. Biol.* **74**, 1–9 (2018).
17. M. Buti *et al.*, Transcriptome profiling of short-term response to chilling stress in tolerant and sensitive *Oryza sativa* ssp. Japonica seedlings. *Funct. Integr. Genomics* **18**, 627–644 (2018).
18. A. P. Cavanagh, R. Slattery, D. S. Kubien, Temperature-induced changes in *Arabidopsis* Rubisco activity and isoform expression. *J. Exp. Bot.* **74**, 651–663 (2023).
19. U. Flores-Pérez, P. Jarvis, Isolation and suborganelle fractionation of *Arabidopsis* chloroplasts. *Methods Mol. Biol.* **1511**, 45–60 (2017).
20. Z. Ying, R. M. Mulligan, N. Janney, R. L. Houtz, Rubisco small and large subunit N-methyltransferases. Bi- and mono-functional methyltransferases that methylate the small and large subunits of Rubisco. *J. Biol. Chem.* **274**, 36750–36756 (1999).
21. S. Knight, I. Andersson, C. I. Branden, Crystallographic analysis of ribulose 1, 5-bisphosphate carboxylase from spinach at 2.4 Å resolution. Subunit interactions and active site. *J. Mol. Biol.* **215**, 113–160 (1990).
22. L. Gunn *et al.*, Rubisco kinetic acclimation at the holoenzyme level. *Proc. Natl. Acad. Sci. U.S.A.* **123**, e2519914123. (2026).
23. P. Khumsupan *et al.*, Generating and characterizing single- and multigene mutants of the Rubisco small subunit family in *Arabidopsis*. *J. Exp. Bot.* **71**, 5963–5975 (2020).
24. M. J. Blayney, S. M. Whitney, J. L. Beck, NanoESI mass spectrometry of Rubisco and Rubisco activase structures and their interactions with nucleotides and sugar phosphates. *J. Am. Soc. Mass Spectrom.* **22**, 1588–1601 (2011).
25. J. Y. Bhat *et al.*, Mechanism of enzyme repair by the AAA(+) chaperone Rubisco activase. *Mol. Cell* **67**, 744–756.e6 (2017).
26. I. Sparrow-Munoz, T. C. Chen, S. J. Burgess, Recent developments in the engineering of Rubisco activase for enhanced crop yield. *Biochem. Soc. Trans.* **51**, 627–637 (2023).
27. A. Gjindali *et al.*, Two cowpea Rubisco activase isoforms for crop thermotolerance. *New Phytol.* **247**, 1199–1217 (2025).
28. C. A. Lutomski *et al.*, Defining proteoform-specific interactions for drug targeting in a native cell signalling environment. *Nat. Chem.* **17**, 204–214 (2025). <https://doi.org/10.1038/s41557-024-01711-w>.
29. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008). <https://doi.org/10.1038/nbt.1511>.
30. L. Fornelli *et al.*, Accurate sequence analysis of a monoclonal antibody by top-down and middle-down orbitrap mass spectrometry applying multiple ion activation techniques. *Anal. Chem.* **90**, 8421–8429 (2018).
31. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
32. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
33. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
34. E. F. Pettersen *et al.*, UCSF chimera-A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
35. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
36. P. V. Afonine *et al.*, Real-space refinement in PHENIX for cryo-em and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
37. P. D. Adams *et al.*, PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
38. V. B. Chen *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
39. Z. Zhang, J. Bolla, Data title "Cryo-EM structure of Rubisco with hetero small subunit 1A3B form II." Electron Microscopy Data Bank. <https://www.ebi.ac.uk/emdb/EMD-70094>. Deposited 4 August 2025.
40. Z. Zhang, J. Bolla, Data title "Cryo-EM structure of Rubisco with hetero small subunit 1A3B form II." Protein Data Bank. <https://www.rcsb.org/structure/904C>. Deposited 4 August 2025.
41. Z. Zhang, J. Bolla, Data title "Cryo-EM structure of Rubisco with hetero small subunit 1A3B form I." Electron Microscopy Data Bank. <https://www.ebi.ac.uk/emdb/EMD-70090>. Deposited 4 August 2025.
42. Z. Zhang, J. Bolla, Data title "Cryo-EM structure of Rubisco with hetero small subunit 1A3B form I." Protein Data Bank. <https://www.rcsb.org/structure/9049>. Deposited 4 August 2025.