RESEARCH ARTICLE

BIOCHEMISTRY

Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2

H. Aigner,* R. H. Wilson,* A. Bracher, L. Calisse, J. Y. Bhat, F. U. Hartl, M. Hayer-Hartl†

Plant RuBisCo, a complex of eight large and eight small subunits, catalyzes the fixation of CO_2 in photosynthesis. The low catalytic efficiency of RuBisCo provides strong motivation to reengineer the enzyme with the goal of increasing crop yields. However, genetic manipulation has been hampered by the failure to express plant RuBisCo in a bacterial host. We achieved the functional expression of *Arabidopsis thaliana* RuBisCo in *Escherichia coli* by coexpressing multiple chloroplast chaperones. These include the chaperonins Cpn60/Cpn20, RuBisCo accumulation factors 1 and 2, RbcX, and bundle-sheath defective-2 (BSD2). Our structural and functional analysis revealed the role of BSD2 in stabilizing an end-state assembly intermediate of eight RuBisCo large subunits until the small subunits become available. The ability to produce plant RuBisCo recombinantly will facilitate efforts to improve the enzyme through mutagenesis.

Form I ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo) is a ~540-kDa complex of eight large (RbcL, ~53 kDa) and eight small (RbcS, ~15 kDa) subunits (*1*, *2*). The RbcL subunits are arranged as a tetramer of antiparallel RbcL dimers (the $RbcL_8$ core), carrying the active sites. Four RbcS subunits each cap the top and bottom, forming the $RbcL_8S_8$ holoenzyme. RuBisCo is the most abundant enzyme in nature, owing in part to its low catalytic turnover rate and limited specificity for CO_2 versus O_2 (*3*). Thus, RuBisCo has long been a target for reengineering so as to increase crop yields (*4*, *5*). However, the plant enzyme has not been amenable to functional recombinant expression, suggesting that its biogenesis has a more extensive requirement for auxiliary factors than that of cyanobacterial $RbcL_8S_8$ RuBisCo (*2*). Furthermore, RuBisCo biogenesis in plants involves the need to coordinate the synthesis of the plastid-encoded RbcL and the nuclear-encoded RbcS subunits, which are imported into chloroplasts as an unfolded protein (*6*).

Folding of the RbcL subunit is mediated by the chaperonin, a cylindrical complex of two stacked heptameric rings of ~60-kDa subunits (*7*). Although RbcS is able to fold spontaneously in vitro, it may require chaperone assistance after its translocation into chloroplasts (*8*). The chloroplast chaperonin is a 1:1 hetero-oligomer of Cpn60α and Cpn60β subunits; it cooperates with the cofactors Cpn20 and Cpn10, which form the lid of the chaperonin folding cage (*2*, *9*). Assembly of the $RbcL_8$ core is mediated through specific assembly chaperones. As shown for the cyanobacterial enzyme, these include the proteins RbcX (a dimer of ~15-kDa subunits) (*10*–*19*) and RuBisCo accumulation factor1 (Raf1, a dimer of ~40-kDa subunits) (*20*–*23*), both of which are conserved in plants. An additional RuBisCo accumulation factor, Raf2 (a dimer of ~10- to 18-kDa subunits), is thought to function in assembly by interacting with RbcS (*24*, *25*). Furthermore, the chloroplast-specific protein bundle-sheath defective-2 (BSD2), a zinc-finger domain protein of ~8 to 10 kDa, has been implicated in plant RuBisCo biogenesis and translational

Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, Martinsried, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: mhartl@biochem.mpg.de (M.H.-H)

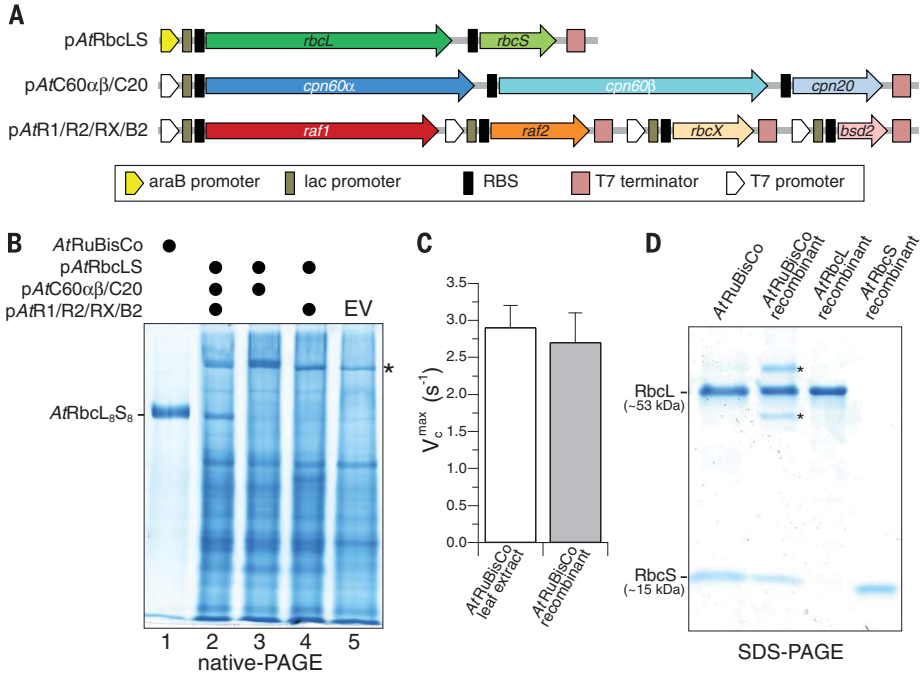

Fig. 1. Plant RuBisCo folding and assembly in *E. coli* requires coexpression of chloroplast chaperonin and auxiliary factors. (**A**) Operon organization of plasmids encoding *A. thaliana* RuBisCo (p*At*RbcLS); chloroplast chaperonin factors (p*At*C60αβ/C20); and predicted RuBisCo biogenesis factors Raf1, Raf2, RbcX, and BSD2 (p*At*R1/R2/RX/B2) (fig. S1A). RBS, ribosome binding site. (**B**) Native-PAGE analysis of cell extracts from *E. coli* cells expressing *At*RbcL and *At*RbcS with and without auxiliary factors, as indicated (lanes 2 to 5). RuBisCo holoenzyme from *A. thaliana* leaf extract (lane 1) was used as standard. EV, empty vector control. Asterisk marks the position of chloroplast or *E. coli* chaperonins. (**C**) RuBisCo synthesized in *A. thaliana* leaves and in *E. coli* show equivalent carboxylation rates (V_c^{max}). Data are averages ± SD from at least three independent experiments. (**D**) Analysis by means of SDS-PAGE of partially purified, recombinantly expressed *At*RuBisCo. Impurities are marked with asterisks. The enzyme purified from leaves as well as recombinant *At*RbcL and *At*RbcS were used as standards (fig. S1B).

regulation of RbcL (*20*, *26*, *27*), but its mechanism has remained elusive.

Functional expression of *At*RuBisCo in *E. coli*

Some cyanobacterial RuBisCos can be functionally expressed in *Escherichia coli* dependent only on GroEL/GroES, the bacterial chaperonin homolog, whereas others also require coexpression of RbcX or Raf1 for assembly (*12*, *14*, *21*, *22*, *28*). However, our preliminary attempts to express *Arabidopsis thaliana* RuBisCo (*At*RuBisCo) with coexpression of cognate RbcX and/or Raf1 failed to produce functional enzymes. This suggested that additional factors may be necessary for the biogenesis of plant RuBisCo, including Raf2 and BSD2, which is consistent with results of a screen of photosynthetic maize mutants (*20*). Plant RuBisCo may also have a specific requirement for the hetero-oligomeric chloroplast chaperonin (*2*, *9*).

We generated an *E. coli* strain containing three plasmids: one expressing *At*RbcL and *At*RbcS under control of the arabinose-regulated pBAD promoter (p*At*RbcLS); the second expressing the chloroplast chaperonin proteins (p*At*C60αβ/C20); and the third expressing Raf1, Raf2, RbcX, and BSD2 (p*At*R1/R2/RX/B2), both under the isopropyl-β-D-thiogalactopyranoside (IPTG)–inducible T7 promoter (Fig. 1A). Each coding sequence (without transit peptide) is preceded by a ribosome binding site. Induction with IPTG for 3 hours produced all auxiliary factors, as confirmed with mass spectrometry (MS) (fig. S1A). Subsequent induction of p*At*RbcLS with arabinose for ~18 hours at 23°C (in the absence of IPTG) resulted in the robust production of a protein complex migrating on native–polyacrylamide gel electrophoresis (PAGE) at the position of RuBisCo from *A. thaliana* leaves (Fig. 1B, lanes 1 and 2). This band was not observed in *E. coli* strains lacking either the chloroplast chaperonins (p*At*C60αβ/C20) or the auxiliary factors (p*At*R1/R2/RX/B2) (Fig. 1B, lanes 3 and 4). We quantified the recombinant RuBisCo through binding of the high-affinity, ^{14}C-labeled substrate analog carboxyarabinitol-1,5-bisphosphate (CABP). Activity assays showed a maximal carboxylation rate (V_c^{max}) and affinity for CO_2 (K_c^{air}) of the recombinant enzyme similar to *At*RuBisCo standard (Fig. 1C) (*23*, *29*). The recombinant holoenzyme contained both RbcL and RbcS (Fig. 1D), and MS analysis showed that the first two amino acids of RbcL were missing, as for authentic *At*RbcL (fig. S1B). The resulting N terminus was not acetylated, and no other posttranslational modifications were detected.

Requirement for chloroplast chaperones

The level of chloroplast chaperonin upon expression from pC60αβ/C20 was approximately fourfold higher than that of endogenous GroEL/GroES (fig. S1A). However, even when overexpressed, GroEL/GroES could not replace Cpn60αβ/Cpn20 for *At*RuBisCo production (Fig. 2, A and B, lanes 1 to 3). Both Cpn60α and Cpn60β were required for efficient expression of functional enzyme (Fig. 2, A and B, lanes 4 and 5). Cpn60β, which forms tetradecamer complexes on its own (*30*), mediated production of RuBisCo with low efficiency (Fig. 2, A and B, lane 4). Small amounts of active RuBisCo were also generated in the absence of Cpn20 (Fig. 2, A and B, lane 6), suggesting that *E. coli* GroES can replace Cpn20 as the cofactor of Cpn60. Indeed, overexpression of GroES supported RuBisCo production as efficiently as Cpn20 (Fig. 2, A and B, lane 9), which is consistent with previous findings that GroES can cooperate with plant chaperonin (*30*). Whereas GroES is a heptamer of 10-kDa subunits, Cpn20 is a tetramer of tandem repeat GroES-like domains (*31*). The GroES-like Cpn10 of chloroplasts not only failed to replace Cpn20 but also interfered with the function of Cpn20 (Fig. 2, A and B, lane 7 and 8). This may be explained by suboptimal relative expression levels of Cpn20 and Cpn10, which can form nonfunctional, mixed complexes (*32*).

To determine the requirement for the auxiliary factors, we stepwise deleted each factor and confirmed the expression of the remaining proteins (fig. S2). Deletion of Raf1, Raf2, or BSD2 each abolished RuBisCo holoenzyme production (Fig. 3, A and B, lanes 3, 4, and 6). Raf1, as shown for the cyanobacterial homolog, functions downstream of chaperonin in assembling RbcL subunits up to $RbcL_8$ (*21*, *22*). The function of Raf2 remains to be defined (*24*, *25*). BSD2 is thought to have homology to the zinc finger domain of the chaperone DnaJ and has been implicated in translational regulation of RbcL (*20*, *26*, *27*). Like Raf1, cyanobacterial RbcX mediates $RbcL_8$ core assembly, albeit with a distinct mechanism (*10*, *11*, *22*). However, in contrast to Raf1, deletion of RbcX resulted in only ~50 to 60% reduction of assembled RuBisCo (Fig. 3, A to C), indicating that RbcX is not essential but enhances recombinant *At*RuBisCo production. The loss of assembled RuBisCo in the deletion strains was mirrored by a decrease in soluble RbcL (Fig. 3C). This was apparently owing to aggregation because the overall expression of RuBisCo was similar in all strains (Fig. 3C).

To test whether the *A. thaliana* chaperones can also mediate the folding and assembly of a heterologous plant RuBisCo, we expressed the RbcL and RbcS from *Nicotiana tabacum*. Only a small amount of *Nt*RuBisCo was produced, migrating at the level of authentic *Nt*RuBisCo (Fig. 3D, lanes 1 and 3). The amount of recombinant enzyme increased when *At*Raf1 was replaced by *Nt*Raf1, and an increase in *Nt*RuBisCo activity was observed (Fig. 3, D and E, lanes 3 and 4). This is consistent with previous findings that foreign RuBisCo expression in chloroplasts is augmented by cognate Raf1 coexpression (*23*). Presumably, RuBisCo assembly is only efficient with all cognate auxiliary factors.

Function of the chloroplast-specific BSD2

Whereas Raf1, Raf2, and RbcX have homologs in cyanobacteria, BSD2

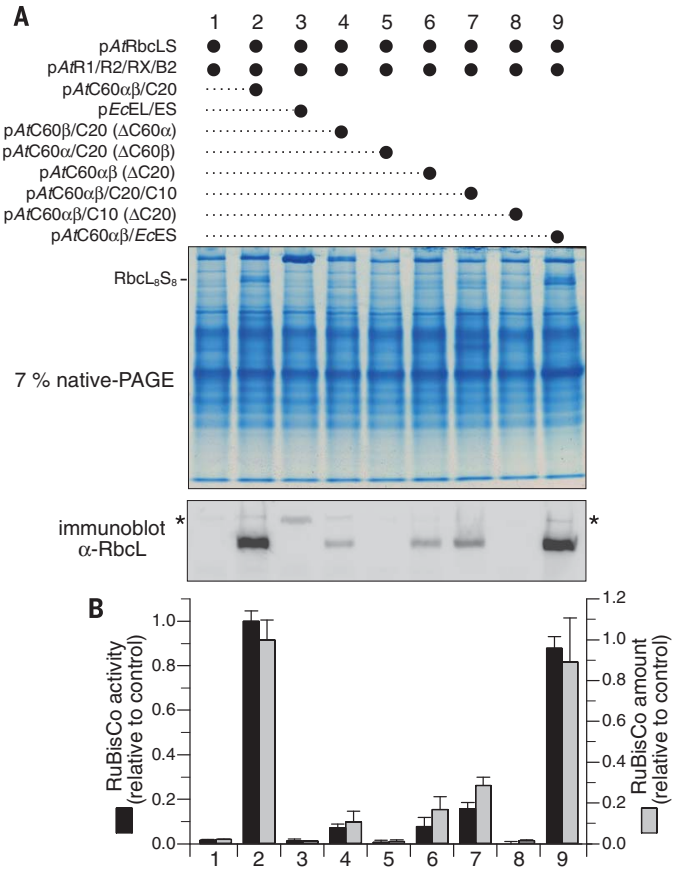

Fig. 2. Chaperonin dependence of *At*RbcL folding. (**A** and **B**) *At*RuBisCo production in *E. coli* strains expressing *A. thaliana* chloroplast chaperonin proteins *At*Cpn60α/β (*At*C60α/β), *At*Cpn20 (C20), and *At*Cpn10 (C10) or the *E. coli* chaperonin system GroEL/GroES (*Ec*EL/ES) in the combinations indicated. (A) Analysis of soluble cell lysates by means of native-PAGE and antibody-to-RbcL immunoblot. Native-PAGE samples were loaded on the basis of equal OD_{600} (optical density of a sample measured at a wavelength of 600 nm) of cells (~15 μg total protein per lane). Asterisk indicates chaperonin-bound RbcL. (B) RuBisCo activity in cell lysates through CO_2 fixation (black bars) and densitometry of native-PAGE immunoblot to quantify RuBisCo amount (gray bars). Data are averages ± SD from three independent experiments.

appears to be present only in green algae and plants, suggesting that it has evolved after the endosymbiotic event leading to the evolution of chloroplasts. The function of BSD2 may thus be related to the chloroplast-specific requirement for RbcS subunits to be imported into the organelle, possibly leading to a limited availability of RbcS. In some experiments, two distinct RbcL complexes were observed migrating above the *At*RbcL$_8$S$_8$ holoenzyme on native-PAGE (Fig. 4, A and B, lane 1). These bands presumably represented chaperone-bound RbcL$_8$ complexes. Immunoblot analysis showed that the top band contained BSD2, whereas the lower band contained BSD2 and RbcS (Fig. 4, B and C, lanes 1 to 3). Other auxiliary factors were not detected (fig. S3A). Upon coexpression of a second *rbcS* gene (p*At*C60αß/C20+p*At*RbcS) to increase RbcS production, only RbcL$_8$S$_8$ formed (Fig. 4, A and B, lane 2), suggesting that the BSD2-bound complexes accumulated owing to limited RbcS availability. We next expressed RbcL in the absence of RbcS, but with coexpression of BSD2 and all other chaperones. Only the top RbcL/BSD2 complex was observed (Fig. 4, A and B, lane 3). When RbcS and Raf2 were omitted, no RbcL/BSD2 complexes were detected, suggesting that Raf2 also interacts with RbcL during biogenesis (fig. S3B).

To further analyze the function of BSD2 in RbcL assembly, we used the RbcL from the cyanobacterium *Synechococcus elongatus* PCC7942. *Se*RbcL is highly homologous to *At*RbcL (~80% identity and ~90% similarity) but forms isolatable RbcL$_8$ core complexes upon recombinant expression in the absence of auxiliary factors (*12*). We purified *Se*RbcL$_8$ and confirmed its oligomeric state by means of native-MS (theoretical mass, 419576 Da) (Fig. 4D). Addition of increasing amounts of *At*BSD2 to *Se*RbcL$_8$ produced complexes with three to eight BSD2 molecules bound (theoretical mass of *Se*RbcL$_8$:*At*BSD2$_8$, 487744 Da) (Fig. 4D). BSD2 alone behaved mostly as a monomer (theoretical mass, 8565 Da) (Fig. 4E). Addition of purified *Se*RbcS to *Se*RbcL$_8$:*At*BSD2$_8$ resulted in the formation of *Se*RbcL$_8$S$_8$ holoenzyme (Fig. 4F), suggesting that the RbcL/BSD2 complex (Fig. 4, A to C) is a productive assembly intermediate. In the absence of RbcS, RbcL$_8$ accumulates as a stoichiometric complex with BSD2 (Fig. 4). Thus, BSD2 is critical at a late stage of RuBisCo biogenesis.

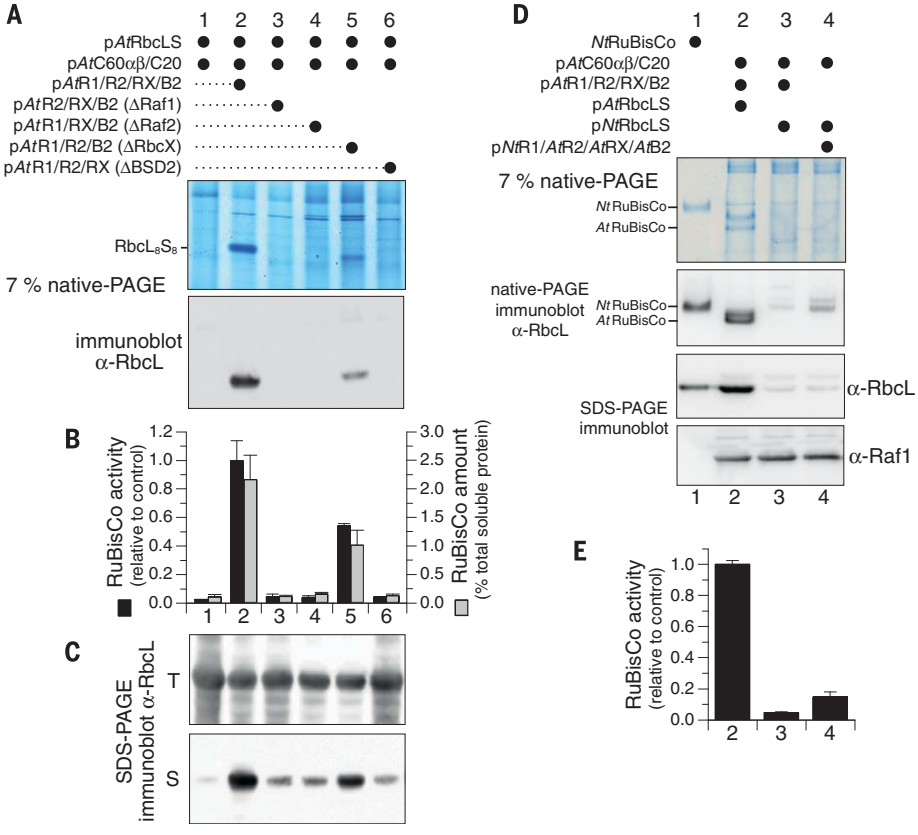

Fig. 3. Chaperone dependence of *At*RuBisCo assembly. (**A** to **C**) RuBisCo content analysis in *E. coli* strains upon deletion of specific auxiliary factors from p*At*R1/R2/RX/B2 (lanes 2 to 6) or containing empty pCDF-Duet vector (lane 1) (fig. S2). (A) Native-PAGE and antibody-to-RbcL immunoblot. (B) RuBisCo content in soluble lysates through CO_2 fixation (black bars) and [^{14}C]-CABP binding (gray bars). Amounts of RuBisCo are expressed as percent of total soluble protein. Data are averages ± SD from three independent experiments. (C) Soluble expression of RbcL by means of SDS-PAGE and antibody-to-RbcL immunoblotting. T, total lysate fractions; S, soluble lysate fractions obtained by means of centrifugation. Equivalent amounts of fractions were analyzed. (**D** and **E**) Recombinant expression and assembly efficiency of *N. tabacum* RuBisCo in the presence of *A. thaliana* chaperones or substitution of *N. tabacum* Raf1. (D) *Nt*RbcL and *Nt*RbcS were expressed in *E. coli* with coexpression of the *A. thaliana* chaperonin system and auxiliary factors (lane 3) or upon substitution of *At*Raf1 by *Nt*Raf1 (lane 4). *Nt*RuBisCo from tobacco leaves was analyzed as standard (lane 1), and *At*RuBisCo was expressed as control (lane 2). Native- and SDS-PAGE samples were loaded with equal amounts of protein (20 and 12 μg per lane, respectively). (E) RuBisCo carboxylation activity was analyzed as above. Data are averages ± SD from three independent experiments.

Structure of BSD2 and RbcL$_8$:BSD2 complex

To obtain insight into the BSD2 mechanism, we solved the crystal structure of *At*BSD2 (residues 57 to 136), lacking the chloroplast transit peptide, at 1.90-Å resolution by means of zinc-multiwavelength anomalous dispersion (Zn-MAD) (fig. S4A and table S1). The model comprises residues 68 to 129 (Fig. 5A and fig. S4B). Consistent with native-MS analysis (Fig. 4E), *At*BSD2 is monomeric in the crystal lattice. The elongated molecule is crescent-shaped, with dimensions of ~50 by 30 by 15 Å, and has a central groove. BSD2 has a hairpin architecture arranged around two Zn atoms, each coordinated by four cysteines (Fig. 5A). The chain termini (residues 56 to 67 and 130 to 136) are disordered. BSD2 has little regular secondary structure and only a limited hydrophobic core at the hairpin tip. The extensive surface of BSD2 comprises numerous hydrophobic and uncharged side chains (Fig. 5B). Close sequence homologs of BSD2 are found in plants and some green algae (Znf2 of *C. rheinhardtii* is more distantly related) (fig. S4, B and C). The concave surface of Zn center 2 exhibits the largest area of high surface conservation, including residues Trp^{108}, Leu^{109}, Arg^{111}, and Lys^{113}, all of which line the rim of the central groove. The opposite rim contains the highly conserved residues Leu^{117} and Gly^{119}. A second surface area, mainly conserved in plant BSD2 proteins, is located at the hairpin tip (residues Asp^{95}, Phe^{97}, Gln^{100}, Phe^{101}, and Lys^{102}) (Fig. 5C). The proposed similarity of BSD2 to the Zn-finger domain of Hsp40 chaperones (*26*) is limited to the overall hairpin architecture.

To obtain the crystal structure of BSD2 bound to the RbcL$_8$ core, we used the thermostable RbcL$_8$ from the cyanobacterium *Thermosynechococcus elongatus* BP-1. Mutations F345I and P415A [*Te*RbcL(IA)] were introduced to further increase *Te*RbcL stability (*33*, *34*). (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example,

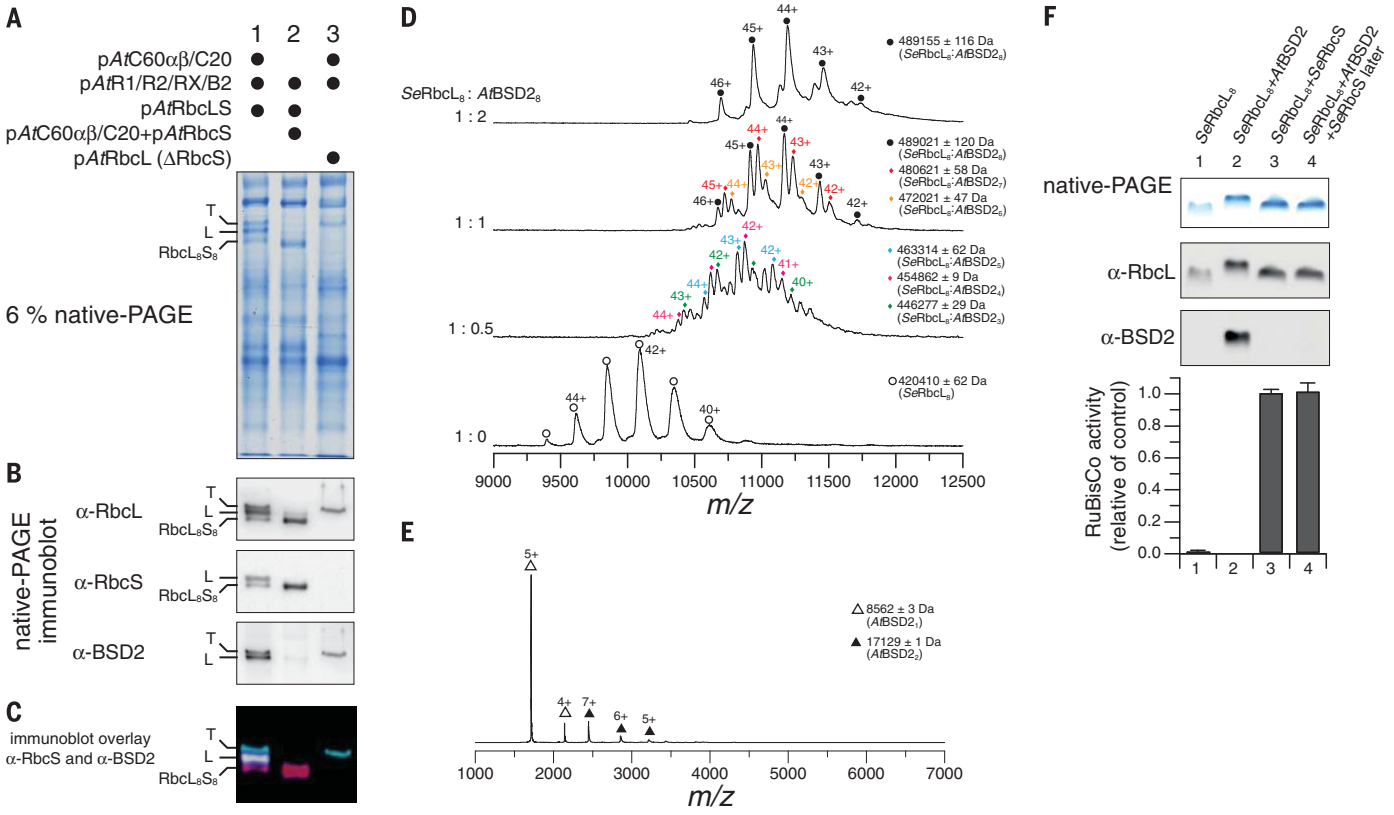

Fig. 4. Functional role of *A. thaliana* BSD2. (**A** to **C**) Formation of three distinct high-molecular-weight RbcL complexes at RbcS limiting conditions (6% native PAGE at 200 V for ~1.5 hours). Top band (T), RbcL/BSD2; lower band (L), RbcL/BSD2/RbcS. (A) Coomassie-stained native-PAGE. (B) Native-PAGE immunoblots against RbcL, RbcS, and BSD2. Expression of *At*RuBisCo with all auxiliary factors (lane 1), with additional RbcS (lane 2) or expression of RbcL in the absence of RbcS (lane 3). (C) Overlay of BSD2 and RbcS immunoblots, with RbcS in pink and BSD2 in cyan. The middle band (white) contains both BSD2 and RbcS. (**D**) Native-MS spectra of *Se*RbcL$_8$ complexes incubated with *At*BSD2 at increasing molar ratios (*Se*RbcL$_8$:*At*BSD2 1:0.5 to 1:2) for 15 min at 25°C. (**E**) Native-MS spectra of *At*BSD2. Charge-state distributions are shown with corresponding symbols for each RbcL:BSD2 complex population. The calculated mass around the *m/z* values of the respective protein complexes and the accuracy of mass values calculated from the different *m/z* peaks are indicated. (**F**) Native-PAGE and immunoblot analysis of *Se*RbcL$_8$ complexes alone (lane 1) and upon incubation for 30 min at 25°C with *At*BSD2 (lane 2), *Se*RbcS (lane 3), or upon incubation with *At*BSD2 for 15 min followed by addition of *Se*RbcS for 15 min (lane 4). The molar ratio of RbcL to BSD2 or RbcS was 1:2. Immunoblotting with antibodies to RbcL and BSD2. Relative CO_2 fixation activities of the reactions are indicated. Data are averages ± SE from three independent experiments.

F345I indicates that phenylalanine at position 345 was replaced by isoleucine.) The complex produced by coexpression of *Te*RbcL and *At*BSD2 consisted of a *Te*RbcL$_8$ core and eight *At*BSD2 (fig. S5A). The crystal structure of the complex was solved by means of molecular replacement at 2.63-Å resolution. The asymmetric unit contains 10 *Te*RbcL(IA)$_8$*At*BSD2$_8$ complexes, which are virtually identical in structure [root mean square deviation (RMSD) 0.23 to 0.57 Å, average RMSD 0.42 Å for Cα positions]. Two BSD2 molecules are bound per antiparallel RbcL$_2$ unit (Fig. 5D). BSD2 appears to join the RbcL units in the dimer, which is reminiscent of the interaction of RbcX with RbcL$_2$ (fig. S5B) (*10*). The highly conserved, concave surface of Zn center 2 (Fig. 5, A and C) of BSD2 forms the major interaction site and cradles the C-terminal domain of one RbcL subunit, whereas the surface of Zn center 1 (Fig. 5, A and C) binds to the N-terminal domain of the adjacent RbcL (Fig. 5D). Zn center 2 contacts helices α16 and α19 of RbcL, exhibiting considerable shape complementarity at the interface (Fig. 5E). The RbcL residue W411 hydrogen bonds with the backbone of BSD2 at W108, and the guanidinium group of R111 in BSD2 forms π-stacking interactions with the indole moiety of W411. Furthermore, the side chains of the conserved BSD2 residues F97, W108, and L109 (Fig. 5C) point into hydrophobic pockets along helix α19 (Fig. 5E). These interactions stabilize the C-terminal domain of RbcL, which is disordered in the cryo–electron microscopy structure of RbcL$_8$ alone (*11*). Zn center 1 of BSD2 contacts helix α2 and makes hydrogen bonds with the so-called 60s loop (residues 63 to 70) of RbcL (Fig. 5F), which forms part of the catalytic site (*35*) and is shifted and remodeled by BSD2 binding as compared with holoenzyme structures (Fig. 5G). In this topology, the 60s loop would clash with loop-CD (residues 92 to 95) of RbcS (Fig. 5G). The conserved residue L117 at the rim of the BSD2 groove forms hydrophobic contacts to V69 and W70 and to G408 in the other RbcL subunit (Fig. 5, E and F). Residues 130 to 136 of the C-terminal tail of BSD2, which are disordered in the structure of BSD2 alone (fig. S5C), contribute substantially to the interaction, occupying the position held by loop 6 of RbcL (residues 330 to 336) in the closed state of the holoenzyme (Fig. 5H) (*35*). As a result, the BSD2-bound complex assumes an open conformation of the catalytic site. In total, the BSD2-RbcL$_2$ interface covers ~970 Å^2 of accessible surface in one RbcL and ~750 Å^2 in the other. Although the binding sites of BSD2 and RbcS do not overlap (Fig. 5G), binding of RbcS, upon spontaneous dissociation of BSD2, would favor the canonical conformation of the 60s loop, resulting in the burial of part of the interface to BSD2 and preventing its rebinding.

On the basis of the crystal structure, we mutated BSD2 and analyzed the ability of the mutants to support *At*RuBisCo assembly in *E. coli*. The mutant proteins were expressed in soluble form together with the other auxiliary factors (Fig. 6A).

Mutation of residues lining either side of the groove (double mutants W108A/L109E and L117E/G119T) (Fig. 5C) caused the loss of BSD2 function (Fig. 6, B and C, lanes 6 and 8), which is consistent with the critical role of these residues in stabilizing the $RbcL_2$ unit (Fig. 5, E and F). Mutation of the two positively charged residues to glutamate (R111E/K113E) (Fig. 5C) also resulted in a substantial reduction of assembled *At*RuBisCo (Fig. 6, B and C, lane 5). By contrast, mutations in the conserved hairpin region of Zn center 2 (D95N/F97S and Q100E/F101A/K102A) (Fig. 5C) showed no substantial defect (Fig. 6, B and C, lanes 4 and 7), which is consistent with these residues being located at the periphery of the BSD2-$RbcL_2$

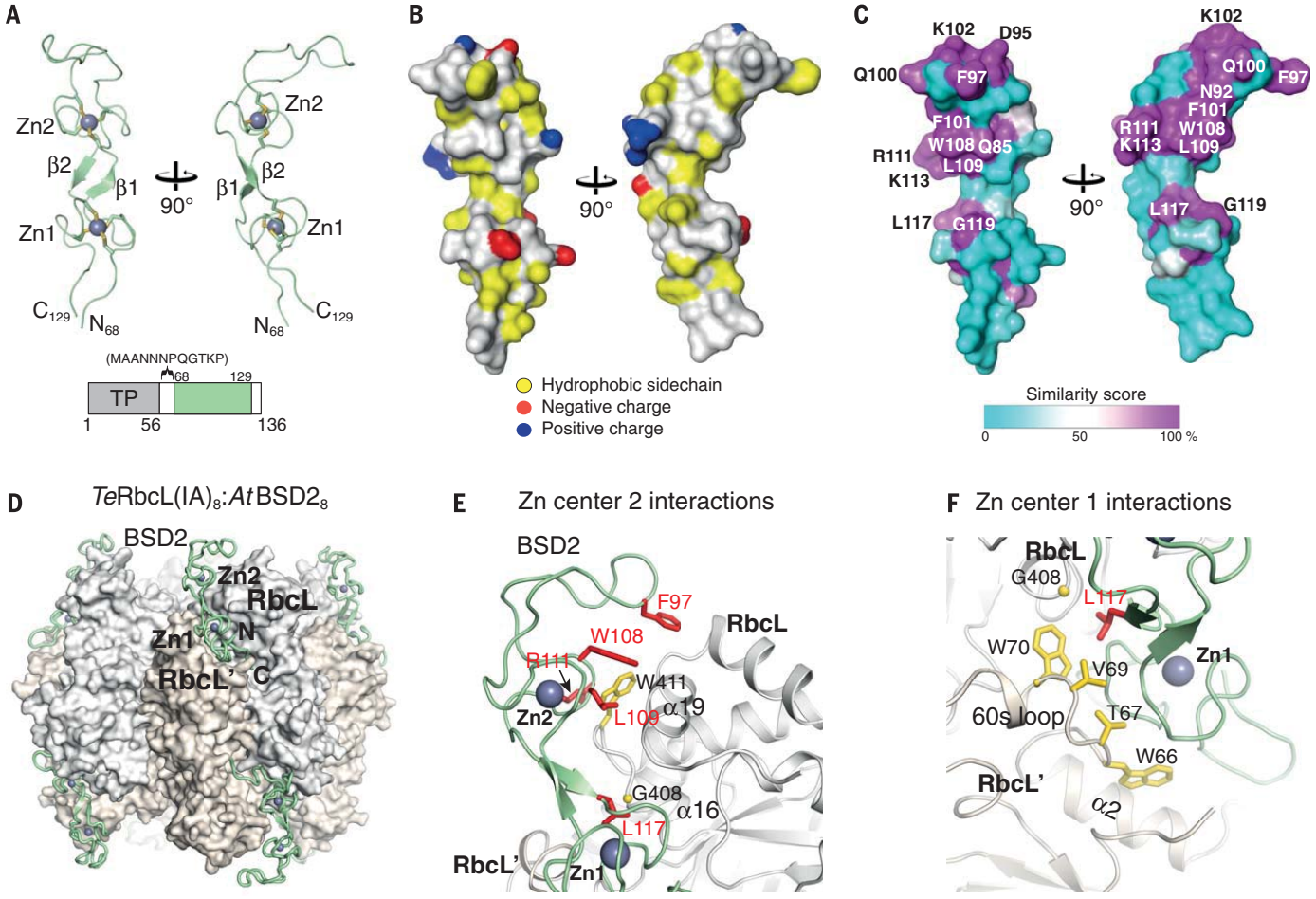

Fig. 5. Crystal structures of *At*BSD2 and heterologous *Te*$RbcL_8$:*At*$BSD2_8$ complex. (**A**) Ribbon representation of the *At*BSD2 crystal structure. Two perpendicular views are shown. The Zn centers and cysteine ligands are shown in space-filling and stick representation, respectively. The amino acid sequence of the crystallized *At*BSD2 construct is shown schematically. Green, sequence resolved in the structure; residues 57 to 68 and 130 to 136 are unstructured. TP, transit peptide. (**B**) Surface properties of *At*BSD2. Hydrophobic side chains are indicated in yellow. Red and blue represent negatively and positively charged groups, respectively. (**C**) Surface conservation of *At*BSD2. Color gradient from cyan to magenta represents increasing conservation, based on sequence alignment of BSD2 homologs (fig. S4B). The positions of residues chosen for mutational analysis are indicated. (**D**) Crystal structure of the *Te*$RbcL(IA)_8$:*At*$BSD2_8$ complex. BSD2 (green) is shown in ribbon representation. $RbcL_8$ is shown as surface with RbcL in white and RbcL' in light orange. (**E**) Interactions between Zn center 2 of BSD2 (green) and RbcL (white). Critical amino acid residues of BSD2 and RbcL are shown in red and yellow stick representation, respectively. (**F**) Interactions between Zn center 1 of BSD2 (green) and the $RbcL_2$ unit. Interacting residues are colored as in (E). (**G**) Rearrangement of the 60s loop in *Te*$RbcL(IA)_8$:*At*$BSD2_8$ complex (green) compared with apo-RuBisCo (DOI: 10.2210/pdb2ybv/pdb) and CABP-bound RuBisCo (DOI: 10.2210/pdb3zxw/pdb) (red and blue, respectively). RbcS and *At*BSD2 are shown for orientation. (**H**) Interaction of the C-terminal tail of BSD2 with the catalytic center of $RbcL_2$. BSD2 in green and the RbcL in the $RbcL_8$:$BSD2_8$ complex in white/light orange ribbon representation. The position of loop 6 in the CABP-bound RuBisCo is shown in cyan.

interface (Fig. 5, C and E). Thus, the mutational analysis validates the interaction of BSD2 with $RbcL_8$ in the crystal structure.

Discussion

The complex folding and assembly pathway of higher plant RuBisCos so far made it impossible to study these proteins outside closely related host chloroplasts (*36*). The ability to produce functional plant RuBisCo in *E. coli* now removes this limitation and will facilitate efforts to improve its catalytic properties through genetic engineering. Furthermore, understanding the assembly pathway of eukaryotic RuBisCo is expected to pave the way for heterologous RuBisCo expression in higher plants, leading to plant varieties with higher yield (*37*, *38*), improved water use efficiency (*39*), or enhanced temperature resistance (*40*)—properties of particular importance in light of future climate uncertainties and increasing water scarcity (*41*).

We used seven auxiliary proteins to express the functional RuBisCo enzyme of *A. thaliana* in *E. coli*. These are the chloroplast chaperonin subunits Cpn60α, Cpn60β, and Cpn20 as well as the auxiliary factors Raf1, Raf2, RbcX, and BSD2 (Fig. 6D). The chloroplast Cpn60 could not be replaced by the bacterial chaperonin GroEL, suggesting that the former is adapted to folding the plant RbcL subunits. By contrast, the Cpn20 cofactor is not essential for substrate specificity and could be replaced by the bacterial GroES. Raf1 and RbcX function downstream of chaperonin in mediating RbcL assembly, acting either sequentially or in parallel (Fig. 6D). The role of Raf2 remains to be clarified because our data are consistent with a function either downstream or upstream of chaperonin (Fig. 6D). Raf1, Raf2, and BSD2 have an essential role in recombinant RuBisCo biogenesis, with RbcX being required for efficiency.

Our analysis of BSD2 provides insight into the role of this chloroplast-specific protein as a late-stage assembly factor. Two complexes containing RbcL/BSD2 or RbcL/BSD2/RbcS were observed under conditions of limiting RbcS. The RbcL/BSD2 complex consists of the $RbcL_8$ core, with eight BSD2 bound. We suggest that this complex represents the end-state assembly intermediate from which BSD2 is displaced by RbcS (Fig. 6D). In contrast to the aggregation-prone $RbcL_8$ core, the $RbcL_8$:$BSD2_8$ complex appears more stable and so may limit RuBisCo aggregation. Indeed, in the crystal structure of the complex, BSD2 stabilizes the $RbcL_2$ units of the $RbcL_8$ core, in a manner similar to that described for cyanobacterial RbcX (*10*). It would then appear that in chloroplasts,

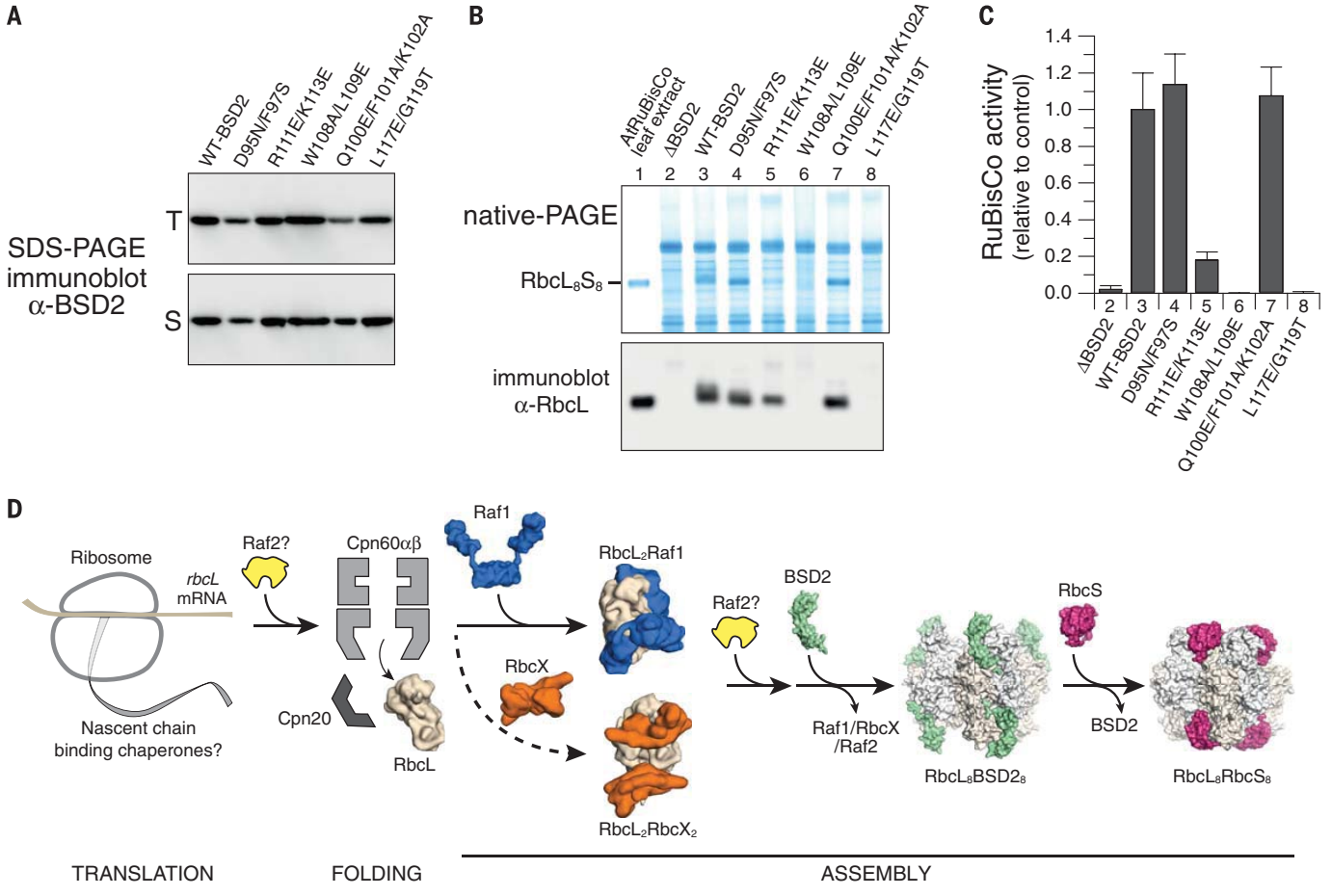

Fig. 6. Function of auxiliary factors in RuBisCo assembly.
(**A** to **C**) Mutational analysis of *At*BSD2. (A) Soluble expression of mutant proteins. Wild-type (WT)–*At*BDS2 in pR1/R2/RX/B2 was replaced with the mutant proteins indicated. Cell extracts were fractionated as in Fig. 3C, and total (T) and soluble (S) fractions were analyzed by means of antibody-to-*At*BSD2 immunoblotting. (B) Native-PAGE and immunoblot analysis of RuBisCo assembly in *E. coli* cells expressing mutant *At*BSD2 (lanes 4 to 9). Cells lacking *At*BSD2 (lane 2) or expressing WT-*At*BSD2 (lane 3) served as control. Purified *At*RuBisCo from leaf extracts is shown in lane 1. (C) Relative RuBisCo carboxylation activities in cell extracts as above. Data are averages ± SE from three independent experiments. (**D**) Model of chaperone-assisted folding and assembly of plant RuBisCo. Upon folding of newly synthesized RbcL subunits by the Cpn60αβ/Cpn20 chaperonin, RbcL assembly to RbcL dimers and higher oligomers is mediated by Raf1 and RbcX acting in cooperation or in parallel. Binding of BSD2 causes the displacement of these factors and stabilizes $RbcL_8$ cores in a state competent for association with RbcS. RbcS binding causes displacement of BSD2, forming the functional holoenzyme. Raf2 is essential for RuBisCo biogenesis and may act downstream or upstream of chaperonin.

BSD2 may have diminished the role of RbcX in RuBisCo assembly. How BSD2 may act as a negative regulator of RbcL transcription (*27*, *42*), in addition to its role in assembly, remains unclear.

REFERENCES AND NOTES

1. I. Andersson, A. Backlund, *Plant Physiol. Biochem.* **46**, 275–291 (2008).
2. A. Bracher, S. M. Whitney, F. U. Hartl, M. Hayer-Hartl, *Annu. Rev. Plant Biol.* **68**, 29–60 (2017).
3. R. H. Wilson, S. M. Whitney, in *Directed Enzyme Evolution: Advances and Applications.*, M. Alcalde, Ed. (Springer International Publishing, 2017), chap. 4, pp. 101–126.
4. R. T. Furbank, W. P. Quick, X. R. R. Sirault, *Field Crops Res.* **182**, 19–29 (2015).
5. T. J. Erb, J. Zarzycki, *Curr. Opin. Chem. Biol.* **34**, 72–79 (2016).
6. I. Sjuts, J. Soll, B. Bölter, *Front. Plant Sci.* **8**, 168 (2017).
7. M. Hayer-Hartl, A. Bracher, F. U. Hartl, *Trends Biochem. Sci.* **41**, 62–76 (2016).
8. R. Trösch, T. Mühlhaus, M. Schroda, F. Willmund, *Biochim. Biophys. Acta* **1847**, 872–888 (2015).
9. A. Vitlin Gruber, S. Nisemblat, A. Azem, C. Weiss, *Trends Plant Sci.* **18**, 688–694 (2013).
10. A. Bracher, A. Starling-Windhof, F. U. Hartl, M. Hayer-Hartl, *Nat. Struct. Mol. Biol.* **18**, 875–880 (2011).
11. C. Liu *et al.*, *Nature* **463**, 197–202 (2010).
12. S. Saschenbrecker *et al.*, *Cell* **129**, 1189–1200 (2007).
13. P. Kolesinski *et al.*, *Biochim. Biophys. Acta* **1830**, 2899–2906 (2013).
14. D. Emlyn-Jones, F. J. Woodger, G. D. Price, S. M. Whitney, *Plant Cell Physiol.* **47**, 1630–1640 (2006).
15. T. Onizuka *et al.*, *Plant Cell Physiol.* **45**, 1390–1395 (2004).
16. S. Tanaka, M. R. Sawaya, C. A. Kerfeld, T. O. Yeates, *Acta Crystallogr. D Biol. Crystallogr.* **63**, 1109–1112 (2007).
17. M. Tarnawski, B. Gubernator, P. Kolesinski, A. Szczepaniak, *Acta Biochim. Pol.* **55**, 777–785 (2008).
18. M. Tarnawski, S. Krzywda, W. Bialek, M. Jaskolski, A. Szczepaniak, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67**, 851–857 (2011).
19. A. Bracher, T. Hauser, C. Liu, F. U. Hartl, M. Hayer-Hartl, *PLOS ONE* **10**, e0135448 (2015).
20. L. Feiz *et al.*, *Plant Cell* **24**, 3435–3446 (2012).
21. P. Kolesinski, I. Belusiak, M. Czarnocki-Cieciura, A. Szczepaniak, *FEBS J.* **281**, 3920–3932 (2014).
22. T. Hauser *et al.*, *Nat. Struct. Mol. Biol.* **22**, 720–728 (2015).
23. S. M. Whitney, R. Birch, C. Kelso, J. L. Beck, M. V. Kapralov, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3564–3569 (2015).
24. L. Feiz *et al.*, *Plant J.* **80**, 862–869 (2014).
25. N. M. Wheatley, C. D. Sundberg, S. D. Gidaniyan, D. Cascio, T. O. Yeates, *J. Biol. Chem.* **289**, 7973–7981 (2014).
26. T. P. Brutnell, R. J. Sawers, A. Mant, J. A. Langdale, *Plant Cell* **11**, 849–864 (1999).
27. L. Doron, N. Segal, H. Gibori, M. Shapira, *Plant J.* **80**, 345–355 (2014).
28. P. Goloubinoff, A. A. Gatenby, G. H. Lorimer, *Nature* **337**, 44–47 (1989).
29. D. J. Orr *et al.*, *Plant Physiol.* **172**, 707–717 (2016).
30. L. P. Cloney, D. R. Bekkaoui, M. G. Wood, S. M. Hemmingsen, *J. Biol. Chem.* **267**, 23333–23336 (1992).
31. F. Baneyx *et al.*, *J. Biol. Chem.* **270**, 10695–10702 (1995).
32. Y.-C. C. Tsai, O. Mueller-Cajar, S. Saschenbrecker, F. U. Hartl, M. Hayer-Hartl, *J. Biol. Chem.* **287**, 20471–20481 (2012).
33. O. Mueller-Cajar, S. M. Whitney, *Biochem. J.* **414**, 205–214 (2008).
34. R. H. Wilson, E. Martin-Avila, C. Conlan, S. M. Whitney, *J. Biol. Chem.* jbc.M117.810861 (2017).
35. A. P. Duff, T. J. Andrews, P. M. Curmi, *J. Mol. Biol.* **298**, 903–916 (2000).
36. R. E. Sharwood, *New Phytol.* **213**, 494–510 (2017).
37. R. E. Sharwood, O. Ghannoum, S. M. Whitney, *Curr. Opin. Plant Biol.* **31**, 135–142 (2016).
38. S. P. Long, A. Marshall-Colon, X. G. Zhu, *Cell* **161**, 56–66 (2015).
39. J. R. Evans, *Plant Physiol.* **162**, 1780–1793 (2013).
40. W. Yamori, K. Hikosaka, D. A. Way, *Photosynth. Res.* **119**, 101–117 (2014).
41. K. A. Bishop, A. M. Betzelberger, S. P. Long, E. A. Ainsworth, *Plant Cell Environ.* **38**, 1765–1774 (2015).
42. K. Wostrikoff, D. Stern, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6466–6471 (2007).

ACKNOWLEDGMENTS

Expert technical assistance by S. Gärtner, A. Jungclaus, and A. Ries is acknowledged. Antibodies against Raf2 were a kind gift from S. Merchant (University of California, Los Angeles). We thank the immunization facility of the Max Planck Institute of Biochemistry (MPIB) for the generation of primary rabbit antibodies. We thank the staff at the MPIB Crystallization facility and at beamline X10SA of the Swiss Light Source in Villigen, Switzerland, and the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, for their support. M.H.-H. acknowledges funding by the Minerva Foundation of the Max Planck Gesellschaft and a grant of the Deutsche Forschungsgemeinschaft (SFB1035) to M.H.-H. and F.U.H. X-ray structure factors and refined coordinates have been deposited in the Protein Data Bank (www.wwpdb.org) under accession numbers 6EKB and 6EKC for *At*BSD2 and *Te*RbcL(IA)$_8$:*At*BSD2$_8$, respectively. H.A. generated the plasmids and the *E. coli* strain for *At*RuBisCo expression. H.A. and R.H.W designed and performed most of the in vivo and in vitro experiments. H.A. performed mass spectrometry proteomics and generated antibodies. R.H.W. kinetically characterized recombinant *At*RuBisCo, crystallized RuBisCo complexes, and performed native-MS with assistance from J.Y.B. L.C. assisted in BSD2 crystallization and mutagenic analysis. A.B. solved the crystal structures. F.U.H. and M.H.-H. supervised the experimental design and wrote the manuscript, with contributions from the other authors.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6368/1272/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S3
References (*43*–*59*)

12 September 2017; accepted 17 October 2017
10.1126/science.aap9221

Science

Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2

H. Aigner, R. H. Wilson, A. Bracher, L. Calisse, J. Y. Bhat, F. U. Hartl and M. Hayer-Hartl

Science **358** (6368), 1272-1278.
DOI: 10.1126/science.aap9221

A biotech tour de force

RuBisCo, the key enzyme of photosynthesis, is a complex of eight large and eight small subunits. It mediates the fixation of atmospheric CO_2 in the Calvin-Benson-Bassham cycle. In addition to being enzymatically inefficient, RuBisCo has a problem with distinguishing between CO_2 and O_2. The fixation of O_2 results in the energetically wasteful reaction of photorespiration. Thus, there is a strong incentive to improve RuBisCo's catalytic properties by engineering. However, for decades, it has been impossible to express the enzyme from plants in an easily manipulatable bacterial host. Aigner *et al.* succeeded in functionally expressing plant RuBisCo in *Escherichia coli* (see the Perspective by Yeates and Wheatley). This should allow for the systematic mutational analysis of RuBisCo and selection of favorable variants for improved crop yields.

Science, this issue p. 1272; see also p. 1253

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.