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


Plant RuBisCo, a complex of eight large and eight small subunits, catalyzes the fixation of CO₂ 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.

![Image](https://science.sciencemag.org/content/358/6363/650/F1.large.jpg)

**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 (pARbcLS); chloroplast chaperonin factors (pATC60uq/C20); and predicted RuBisCo biogenesis factors Raf1, Raf2, RbcX, and BSD2 (pAR1R2/RX/B2) (fig. S1A). RBS, ribosome binding site. (B) Native-PAGE analysis of cell extracts from *E. coli* cells expressing ARbcL and ARbcS 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 (V₅₀max). Data are averages ± SD from at least three independent experiments. (D) Analysis by means of SDS-PAGE of partially purified, recombinantly expressed ARubisCo. Impurities are marked with asterisks. The enzyme purified from leaves as well as recombinant ARbcL and ARbcS were used as standards (fig. S1B).
regulation of RbcL (20, 26, 27), but its mechanism has remained elusive.

Functional expression of AtrRuBisCo 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 Ralf for assembly (12, 14, 31, 22, 28). However, our preliminary attempts to express Arabidopsis thaliana RuBisCo (AtrRuBisCo) with coexpression of cognate RbcX and/or Ralf 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 AtrRbcL and AtrRbcS under control of the arabinose-regulated pBAD promoter (pAtrRbcLS); the second expressing the chloroplast chaperonin proteins (pAtC600/oβ/C20); and the third expressing Raf1, Raf2, RbcX, and BSD2 (pAtrR1/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 pAtrRbcLS 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-PAGE (PAGE) at the position of RuBisCo from Arabidopsis leaves (Fig. 1B, lanes 1 and 2). This band was not observed in E. coli strains lacking either the chloroplast chaperonins (pAtC600/oβ/C20) or the auxiliary factors (pAtrR1/R2/RX/B2) (Fig. 1B, lanes 3 and 4). We quantified the recombinant RuBisCo through binding of the high-affinity, 14C-labeled substrate analog carboxyarabinitol-1,5-bisphosphate (CABP). Activity assays showed a maximal carboxylation rate (Vmax) and affinity for CO2 (Km) of the recombinant enzyme similar to AtrRuBisCo 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 AtrRbcL (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 pC600/oβ/C20 was approximately fourfold higher than that of endogenous GroEL/GroES (fig. S1A). However, even when overexpressed, GroEL/GroES could not replace Cpn60oβ/Cpn20 for AtrRuBisCo production (Fig. 2, A and B, lanes 1 to 3). Both Cpn90oβ and Cpn60oβ were required for efficient expression of functional enzyme (Fig. 2, A and B, lanes 4 and 5). Cpn60oβ, 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 RbcLa (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 RbcLa 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 AtrRuBisCo 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 chaperonins 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 NtRuBisCo was produced, migrating at the level of authentic NtRuBisCo (Fig. 3D, lanes 1 and 3). The amount of recombinant enzyme increased when NtRuBisCo was replaced by NtRalf1, and an increase in NtRuBisCo 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...
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 AtRbcL8S8 holoenzyme on native-PAGE (Fig. 4, A and B, lane 1). These bands presumably represented chaperone-bound RbcL 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 (pCa060005/C20+pNtvRbcS) to increase RbcS production, only RbcL8S8 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. SeRbcL is highly homologous to AtRbcL (~80% identity and ~90% similarity) but forms isolatable RbcL8/α core complexes upon recombinant expression in the absence of auxiliary factors (22).

We purified SeRbcL and confirmed its oligomeric state by means of native-MS (theoretical mass, 419770 Da) (Fig. 4D). Addition of increasing amounts of AtBSD2 to SeRbcL produced complexes with three to eight BSD2 molecules bound (theoretical mass of SeRbcL-se BSD2n, 487744 Da) (Fig. 4D). BSD2 alone behaved mostly as a monomer (theoretical mass, 8565 Da) (Fig. 4E). Addition of purified SeRbcS to SeRbcL-se BSD2n resulted in the formation of SeRbcL8:S8 holoenzyme (Fig. 4F), suggesting that the RbcL/BSD2 complex (Fig. 4, A to C) is a productive assembly intermediate. In the absence of RbcS, RbcLα accumulates as a stoichiometric complex with BSD2 (Fig. 4). Thus, BSD2 is critical at a late stage of RuBisCo biogenesis.

Structure of BSD2 and RbcLα:BSD2 complex

To obtain insight into the BSD2 mechanism, we solved the crystal structure of AtBSD2 (residues 57 to 136), lacking the chloroplast transit peptide, at 1.90 Å resolution by means of single-wavelength 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), AtBSD2 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. reinhardtii 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 Trp33, Leu93, Arg143, and Lys153 all of which line the rim of the central groove. The opposite rim contains the highly conserved residues Leu97 and Gly102. A second surface area, mainly conserved in plant BSD2 proteins, is located at the hairpin tip (residues Asp37, Phe50, Glu52, Phe130, and Lys130) (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 RbcL8 core, we used the thermostable RbcL8 from the cyanobacterium Thermosynechococcus elongatus BP-1. Mutations F434S and P415A (TrRbcL1[A]) were introduced to further increase 7RbcL 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. 3. Chaperone dependence of AtRuBisCo assembly. (A to C) RuBisCo content analysis in E. coli strains upon deletion of specific auxiliary factors from pAR1/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 CO2 fixation (black bars) and [14C]-CABP binding (gray bars). Amounts of RuBisCo are expressed as percent of total soluble RuBisCo activity (relative to control).

Downloaded from http://science.sciencemag.org/ on December 27, 2018
F345I indicates that phenylalanine at position 345 was replaced by isoleucine. The complex produced by coexpression of RuBisCo and A. thaliana BSD2 consisted of a RuBisCo core and eight 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 TeRuBisCo 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 RbcLα unit (Fig. 5D). BSD2 appears to join the RuBisCo units in the dimer, which is reminiscent of the interaction of RbcX with RbcL (fig. S8B) (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 RuBisCo with the backbone of BSD2 at L117 at the rim of the BSD2 groove forms 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–RbcLα interface covers ~970 Å² of accessible surface in one RbcL and ~750 Å² 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 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, Band C, lanes 6 and 8), which is consistent with the critical role of these residues in stabilizing the RbcL2 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 AtRuBisCo (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-RbcL2 complex.
interface (Fig. 5, C and E). Thus, the mutational analysis validates the interaction of BSD2 with RbcLα 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 (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 RbcLα 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 RbcLα core, the RbcLα-BSD2α complex appears more stable and so may limit RuBisCo aggregation. Indeed, in the crystal structure of the complex, BSD2 stabilizes the RbcLα units of the RbcLα core, in a manner similar to that described for cyanobacterial RbcX (40). It would then appear that in chloroplasts,
BS2 may have diminished the role of RbcX in RuBisCo assembly. How BS2 may act as a negative regulator of RbcL transcription (27, 42), in addition to its role in assembly, remains unclear.

REFERENCES AND NOTES

ACKNOWLEDGMENTS
Expert technical assistance by S. Gätter, A. Jungclaus, and A. Ries is acknowledged. Antibodies against Ral2 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 MPB Crystalization facility and at beamline XDS 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.pdb.org) under accession numbers 6EKB and 6EKC for ABsd2 and TFbbsd1(IAA)4ABsd2a, respectively. H.A. generated the plasmids and the E. coli strain for AtRubsCo 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 AtRubsCo, crystallized RbsCo complexes, and performed native MS with assistance from J.Y.B. L.C. assisted in BS2 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
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