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.

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 (pAC60αβ/C20); and predicted RuBisCo biogenesis factors Raf1, Raf2, RbcX, and BSD2 (pAR1/R2/RX/B2) (fig. S1A). RBS, ribosome binding site. (B) Native-PAGE analysis of cell extracts from *E. coli* cells expressing pARbcL and pARbcs 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 (V₄₅₀°C). Data are averages ± SD from at least three independent experiments. (D) Analysis by means of SDS-PAGE of partially purified, recombinantly expressed pARbcLs. Impurities are marked with asterisks. The enzyme purified from leaves as well as recombinant pARbcL and pARbcs were used as standards (fig. S1B).
regulation of RbcL (20, 26, 27), but its mechanism has remained elusive.

Functional expression of AtRuBisCo in *E. coli*

Some cyanobacterial RuBisCo can be functionally expressed in *Escherichia coli* dependent only on GroEL/GroES, the bacterial chaperonin homolog, whereas others require coexpression of RbcX or Raf1 for assembly (12, 14, 31, 22, 28). However, our preliminary attempts to express *Arabidopsis thaliana* RuBisCo (AtRuBisCo) 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 AtRbcL and AtRbcS under control of the arabinose-regulated pBAD promoter (pAtRbcLS); the second expressing the chloroplast chaperonin proteins (pAtC60αβ/C20); and the third expressing Raf1, Raf2, RbcX, and BSD2 (pAtR1/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 pAtRbcLS 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 (pAtC60αβ/C20) or the auxiliary factors (pAtR1/R2/RX/B2) (Fig. 1B, lanes 3 and 4). We quantified the recombinant RuBisCo through coomassie staining of the high-affinity, 14C-labeled substrate analog carboxyarabinitol-1,5-bisphosphate (CABP). Activity assays showed a maximal carbonic anhydrase rate (Vmax) and affinity for CO2 (Km) of the recombinant enzyme similar to AtRuBisCo 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 AtRbcL (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 AtRuBisCo 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 RbcL4 (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 RbcL4 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 AtRuBisCo production. The loss of assembled RuBisCo in the deletion strains was mirrored by a decrease in soluble RuBisCo (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 NiRuBisCo was produced, migrating at the level of authentic NiRuBisCo (Fig. 3D, lanes 1 and 3). The amount of recombinant enzyme increased when AtRaf1 was replaced by NiRaf1, and an increase in NiRuBisCo activity was observed (Fig. 3, D and F, 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).

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 RbcLcore 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, 419,776 Da) (Fig. 4D). Addition of increasing amounts of *AtBSD2* to *SeRbcL* produced complexes with three to eight BSD2 molecules bound (theoretical mass of *SeRbcLα*~*2* *AtBSD2*~*3*, 487,744 Da) (Fig. 4D). BSD2 alone behaved mostly as a monomer (theoretical mass, 85,655 Da) (Fig. 4E). Addition of purified *SeRbcS* to *SeRbcL*: *AtBSD2* resulted in the formation of *SeRbcL*: *AtBSD2* complexes (Fig. 4). SDS-PAGE alone behaved mostly as a monomer (theoretical mass, 487,744 Da) (Fig. 4). *AtBSD2* behaves as a monomer (theoretical mass, 85,655 Da) (Fig. 4E). Addition of purified *SeRbcS* to *SeRbcL*: *AtBSD2* resulted in the formation of *SeRbcL*: *AtBSD2* complexes (Fig. 4). SDS-PAGE alone behaved mostly as a monomer (Fig. 4D). BSD2 alone behaved mostly as a monomer (theoretical mass, 85,655 Da) (Fig. 4E).

**Structure of BSD2 and RbcLcore: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 zinc-multilength 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 (*Zmf2* 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 TrpGluGluLeuArgLys, all of which line the rim of the central groove. The opposite rim contains the highly conserved residues Leu and Gly. A second surface area, mainly conserved in plant BSD2 proteins, is located at the hairpin tip (residues AspPhen Gly). 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 RbcLcore, we used the thermostable RbcLcore from the cyanobacterium *Thermosynechococcus elongatus* BP-1. Mutations F454I and P415A (TrRbcL1A) were introduced to further increase TrRbcL 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,
of RbcX with RbcL2 (Fig. S5B) (dimer, which is reminiscent of the interaction between BSD2 and RbcL2 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 incubation with BSD2 for 15 min followed by addition of SeRbcS, 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 CO₂ fixation activities of the reactions are indicated. Data are averages ± SE from three independent experiments.

Aigner et al., Science 358, 1272–1278 (2017) 8 December 2017
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.

Fig. 5. Crystal structures of AtBSD2 and heterologous TeRbcL8:AtBSD28 complex. (A) Ribbon representation of the AtBSD2 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 AtBSD2 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 AtBSD2. Hydrophobic side chains are indicated in yellow. Red and blue represent negatively and positively charged groups, respectively. (C) Surface conservation of AtBSD2. 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 TeRbcL(IA)8:AtBSD28 complex. BSD2 (green) is shown in ribbon representation. RbcL8 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 RbcL2 unit. Interacting residues are colored as in (E). (G) Rearrangement of the 60s loop in TeRbcL(IA)8:AtBSD28 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 AtBSD2 are shown for orientation. (H) Interaction of the C-terminal tail of BSD2 with the catalytic center of RbcL2. BSD2 in green and the RbcL in the RbcL8:BSD28 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 RbcL8 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. RuBisCo is essential for RuBisCo biogenesis and may act downstream or upstream of the bacterial 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 RbcL8 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 RbcL8 core, the RbcL8-BSD2 complex appears more stable and so may limit RuBisCo aggregation. Indeed, in the crystal structure of the complex, BSD2 stabilizes the RbcL8 units of the RbcL8 core, in a manner similar to that described for cyanobacterial RbcX (30). It would then appear that in chloroplasts,
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


ACKNOWLEDGMENTS

Expert technical assistance by S. Gärtner, 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 (MPB8) for the generation of primary rabbit antibodies. We thank the staff at the MPB Crystalization facility and at beamline XDOSA 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 6EXG and 6EC6 for ARBSD2 and T4RbsL(IA)8, respectively. H.A. generated the plasmids and the E. coli strain for ARBSD2 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 ARBSD2, 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

Figures 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 CO₂ in the Calvin-Benson-Bassham cycle. In addition to being enzymatically inefficient, RuBisCo has a problem with distinguishing between CO₂ and O₂. The fixation of O₂ 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

This article cites 58 articles, 13 of which you can access for free
http://science.sciencemag.org/content/358/6368/1272#BIBL

Use of this article is subject to the Terms of Service

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. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title Science is a registered trademark of AAAS.