BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Structure of *Arabidopsis thaliana* Rubisco activase

Dirk Hasse, Anna M. Larsson and Inger Andersson*

Laboratory of Molecular Biophysics, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, SE-75124 Uppsala, Sweden. *Correspondence e-mail: inger@xray.bmc.uu.se

Received 24 October 2014
Accepted 20 January 2015

Edited by Q. Hao, University of Hong Kong

Keywords: Rubisco activase; AAA+ protein; ATPase; sulfate ion binding; helical crystal packing.

PDB reference: Rubisco activase, 4w5w

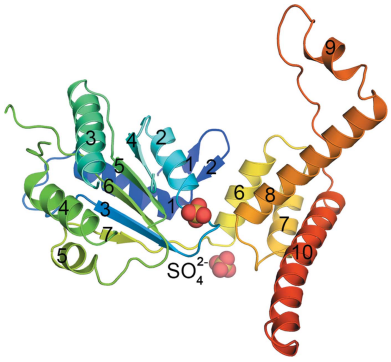
Supporting information: this article has supporting information at journals.iucr.org/d

The CO_2-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is inactivated by the formation of dead-end complexes with inhibitory sugar phosphates. In plants and green algae, the ATP-dependent motor protein Rubisco activase restores catalytic competence by facilitating conformational changes in Rubisco that promote the release of the inhibitory compounds from the active site. Here, the crystal structure of Rubisco activase from *Arabidopsis thaliana* is presented at 2.9 Å resolution. The structure reveals an AAA+ two-domain structure. More than 100 residues in the protein were not visible in the electron-density map owing to conformational disorder, but were verified to be present in the crystal by mass spectrometry. Two sulfate ions were found in the structure. One was bound in the loop formed by the Walker A motif at the interface of the domains. A second sulfate ion was bound at the N-terminal end of the first helix of the C-terminal domain. The protein packs in a helical fashion in the crystal, as observed previously for Rubisco activase, but differences in the helical pitch indicate flexibility in the packing of the protein.

1. Introduction

The activity of the CO_2-assimilating enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) depends on the stromal redox state and the ATP:ADP ratio in a light-mediated fashion (reviewed in Andersson, 2008). Light-driven electron transport leads to increases in the stromal pH and the stromal concentration of Mg^{2+}. These changes in turn promote the activation of Rubisco by the carbamylation of a specific lysine residue in the active site of the enzyme and stabilization of the carbamate by an Mg^{2+} ion (Lorimer & Miziorko, 1980), thereby creating a catalytic centre capable of catalysing the carboxylation of ribulose 1,5-bisphosphate (RuBP). Binding of inhibitory sugar phosphates to the decarbamylated enzyme locks the enzyme into a catalytically impaired dead-end complex. Inhibitors, for example xylulose 1,5-biphosphate or pentodiulose 1,5-bisphosphate, may be produced by enzymatic misfire reactions (Pearce & Andrews, 2003). Alternatively, the substrate itself, RuBP, can act as an inhibitor by binding to the decarbamylated enzyme. Such catalytically stalled complexes also accumulate during *in vitro* assays and lead to a decay in the measured catalytic activity, historically termed fallover (Edmondson *et al.*, 1990; Zhu & Jensen, 1991). In some plants, 2-carboxyarabinitol-1-phosphate accumulates in darkness or low light and binds to the carbamylated enzyme, blocking access by RuBP (Gutteridge *et al.*, 1986).

The motor protein Rubisco activase restores catalytic competence by facilitating conformational changes in Rubisco that promote the release of the inhibitory compounds from the active site (Portis, 2003; Mueller-Cajar *et al.*, 2014). Rubisco activase uses energy from ATP hydrolysis to drive inhibitor release (Wang *et al.*, 1992). The activase function depends on



the ambient ATP:ADP ratio of the stroma and in some plants also on the redox status of the chloroplast by co-expression of two isoforms of the activase (Zhang *et al.*, 2002; Portis *et al.*, 2008). The longer isoform, named α, is redox-sensitive owing to the presence of two C-terminal cysteine residues (Werneke *et al.*, 1989).

Rubisco activase belongs to the family of AAA+ ATPases (ATPases associated with diverse cellular activities; Neuwald *et al.*, 1999). AAA+ proteins are two-domain proteins with a nucleotide-binding site located between the domains that display considerable conformational flexibility. This conformational instability and the tendency to form large assemblies or aggregates have impeded attempts to determine the crystal structure of the protein. A successful strategy to achieve crystal formation has been to express truncated versions of the protein. This has resulted in crystal structures of the activase from *Nicotiana tabacum* lacking 90 amino acids (Stotz *et al.*, 2011) and of the C-terminal domain core of the activase from *Larrea tridentata* (Henderson *et al.*, 2011).

Here, we describe the crystal structure of Rubisco activase from *Arabidopsis thaliana* at 2.9 Å resolution. We show that although the crystals contain an almost complete version of the protein, a large part of the N-terminus is disordered in the crystal and is not visible in the electron-density maps. The refined structure shows a high resemblance to the structure of the tobacco activase. However, differences in the packing of the monomers in the crystal demonstrate the flexibility of the protein that presumably facilitates its function as a motor protein.

2. Materials and methods

2.1. DNA cloning

An N-terminally and C-terminally truncated version of the gene (UniProt ID P10896-2) coding for the short isoform of Rubisco activase (*RCA*) was amplified from *A. thaliana* cDNA (Columbia wild type) using Phusion DNA Polymerase Master Mix (Thermo Scientific) with the forward primer 5′-AGCGGCTCTTCAATGGCTGTGAAAGAAGACAAACAAA-3′ and the reverse primer 5′-AGCGGCTCTTCTCCCTTTTCCGTAGAAAGTTCCGCG-3′. The truncated version lacks the 58-amino-acid N-terminal transit peptide. At the C-terminal end, nine residues were replaced by an 11-residue long *Strep*-tag (GSAWSHPQFEK). In this sense, the truncated version is similar to a Rubisco activation-proficient deletion mutant (−35 del; Zhang & Portis, 1999), except that the two last amino acids, Lys and Gly, were replaced by the *Strep*-tag. The *RCA* gene was cloned into the vector pENTRY-IBA 51 using StarGate Entry Combi Cloning (IBA Life Sciences) and was confirmed by sequencing before it was transferred to the pASG-IBA3 expression vector using the StarGate Transfer Reagent Set (IBA Life Sciences).

2.2. Protein expression

The expression plasmid pASG-IBA3 (IBA Life Sciences) was used to express *A. thaliana* Rubisco activase (*At*Rca) in *Escherichia coli* Rosetta (DE3) pLys cells (Novagen). In order to prevent the formation of inclusion bodies, a combined heat-shock and salt-shock treatment was applied. The *E. coli* cells were grown in 8 l 2×YT medium (1.6% tryptone, 1% yeast extract, 0.5% NaCl) with 100 µg ml^{-1} ampicillin, 2 m*M* betaine (Sigma–Aldrich) and 2% NaCl. After cultivation at 310 K to an OD_{600} of 0.4, the cells were heat-shocked in a water bath (318 K) and further cultivated for 1 h at 318 K. The cells were cooled to 303 K and protein expression was induced with a final concentration of 0.1 µg ml^{-1} anhydrotetracycline. The cells were harvested after further incubation for 4 h at 303 K.

2.3. Protein purification

The protein was purified using a three-step protocol comprising affinity, ion-exchange and size-exclusion chromatography. The cells were resuspended in *Strep*-binding buffer [20 m*M* bis-tris propane pH 8, 200 m*M* KCl, 10% glycerol, 5 m*M* ATP, 5 m*M* $MgCl_2$, 5 m*M* tris(2-carboxyethyl)phosphine (TCEP)] and lysed by sonication. After clearing the lysate by centrifugation at 40 000*g* (SS-34, Sorvall), the supernatant was applied onto a *Strep*-Tactin Superflow column (IBA Life Sciences), washed thoroughly with *Strep*-binding buffer and eluted with the same buffer containing 2.5 m*M* D-desthiobiotin. After lowering the salt concentration by diluting the eluate in buffer *A* (20 m*M* bis-tris propane pH 8, 10% glycerol, 2.5 m*M* $MgCl_2$, 2 m*M* TCEP), the sample was applied onto a Mono Q column (GE Healthcare) and eluted with a 0–50% gradient of buffer *B* (20 m*M* bis-tris propane pH 8, 10% glycerol, 2.5 m*M* $MgCl_2$, 1 *M* KCl, 2 m*M* TCEP). Size-exclusion chromatography was carried out using a HiLoad 26/60 Superdex 200 column (GE Healthcare) equilibrated with buffer *C* (20 m*M* bis-tris propane pH 8, 50 m*M* KCl, 10% glycerol, 5 m*M* TCEP) at a flow rate of 2 ml min^{-1}. The protein corresponding to *At*Rca eluted as a broad peak, indicating that it was a mixture of several oligomeric states. Care was taken to use only the high-molecular-weight portion of the peak for further work. An ÄKTApurifier system (GE Healthcare) was used for all chromatography steps.

2.4. Crystallization and data collection

*At*Rca was concentrated to a final concentration of about 8 mg ml^{-1} (Vivaspin, Sartorius Stedim Biotech). Equal volumes of *At*Rca protein in buffer *C* and reservoir solution (100 m*M* Tris pH 7, 0.15 *M* ammonium sulfate, 13% PEG 3350, 5 m*M* TCEP) were mixed. Diffracting crystals were obtained by the hanging-drop vapour-diffusion method after streak-seeding and incubation for several weeks at 293 K. Prior to data collection, crystals were soaked in reservoir solution containing 20% glycerol, transferred into a nylon loop (Hampton Research) and flash-cooled in liquid nitrogen. X-ray diffraction data were collected at 100 K with a 225 mm MAR Mosaic detector on beamline ID23-2 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France.

Table 1
Data-collection and refinement statistics for *At*Rca.

Values in parentheses are for the outer shell.

Data collection	
Beamline	ID23-2, ESRF
Wavelength (Å)	0.8726
Space group	$P6_5$
Unit-cell parameters (Å)	$a = b = 83.9$, $c = 105.8$
Resolution range (Å)	41.93–2.90 (3.08–2.90)
No. of observations	41585 (6765)
No. of unique reflections	9423 (1509)
R_{meas}†	0.065 (0.579)
$\langle I/\sigma(I)\rangle$	9.5 (3.1)
Completeness (%)	99.9 (99.9)
Multiplicity	4.4 (4.5)
Refinement	
Resolution range (Å)	41.93–2.90 (3.24–2.90)
No. of reflections	9393 (2520)
R_{work}‡	0.196 (0.231)
R_{free}§	0.249 (0.303)
No. of atoms	
Protein	2154
Ligand	10
Waters	5
Average B values (Å^2)	
Estimated from Wilson plot	61.9
Protein	82.7
Ligands	122.4
Waters	60.0
R.m.s. deviations from ideal values	
Bond lengths (Å)	0.010
Bond angles (°)	1.19
Ramachandran analysis¶	
Outliers (%)	2.3

† $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$ (Evans, 2006; Evans & Murshudov, 2013). ‡ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. § R_{free} was calculated in the same way as R_{work} but from a randomly chosen 5% of all unique reflections. ¶ From *MolProbity* (Chen *et al.*, 2010).

2.5. Structure determination and refinement

The data were processed and scaled using *XDS* (Kabsch, 2010) and *AIMLESS* (Evans, 2006). Phases were obtained by molecular replacement in *Phaser* (McCoy *et al.*, 2007) using the X-ray crystal structure of *N. tabacum* Rubisco activase (PDB entry 3t15; Stotz *et al.*, 2011) as a search model. *Zanuda* (Lebedev & Isupov, 2014) in the *CCP*4 suite (Winn *et al.*, 2011) was used to verify the space group and to obtain an improved model for automated building of the *At*Rca structure in *Buccaneer* (Cowtan, 2006). Initial refinement was performed in *REFMAC*5 (Murshudov *et al.*, 2011). The final structure was obtained by alternating restrained refinement in *BUSTER* (Bricogne *et al.*, 2011) and manual rebuilding in *O* (Jones *et al.*, 1991). Waters were added using the water-insertion command in *BUSTER*. Attempts to improve the data by taking into account anisotropy were performed with the UCLA MBI *Diffraction Anisotropy Server* (Strong *et al.*, 2006).

2.6. Comparison with other Rubisco activase structures

*At*Rca was compared with other Rubisco activases in the PDB using the *DALI* server (Holm & Rosenström, 2010) and by using the least-squares alignment function in *O* (Jones *et al.*, 1991).

2.7. Mass spectrometry

Crystals were dissolved in water and the protein was assessed by SDS–PAGE and MALDI-TOF mass spectrometry. The protein was alkylated with iodoacetamide and digested with modified porcine trypsin (Promega). The extracted peptides were analysed on a Bruker Ultraflex TOF/TOF and the signals in the spectra corresponding to theoretical tryptic peptides from the target protein were identified.

2.8. Other software

The figures were prepared with *PyMOL* (v.1.6.0.0; Schrödinger).

3. Results and discussion

3.1. Data analysis and refinement

The heterologously expressed *At*Rca consists of a truncated version in which the nine amino acids at the C-terminus of the (short) β-isoform were replaced by a *Strep*-tag. *At*Rca crystallized in space group $P6_5$, with unit-cell parameters $a = b = 83.9$, $c = 105.8$ Å and one molecule in the asymmetric unit. The initial crude model was improved by using the phases obtained after refinement in *Zanuda* (Lebedev & Isupov, 2014) for automatic tracing in *Buccaneer* (Cowtan, 2006). Initial refinement of the traced model with *REFMAC*5 was not successful as judged by a slow gradual increase in R_{free} during the refinement. Refinement using *BUSTER* (Bricogne *et al.*, 2011) reduced the R_{free} as well as the gap between the R_{work} and R_{free}. The improved refinement may be explained by the capacity of *BUSTER* to handle the large number of amino acids present in the construct but missing from the structure (Blanc *et al.*, 2004).

The reciprocal-space correlation coefficient in *BUSTER* showed signs of anisotropy in the data. Based on the criteria of the correlation coefficient between random half data sets, $CC_{1/2} > 0.5$, and the average signal-to-noise ratio of the merged intensities as a function of resolution, $Mn(I/\sigma) > 2$ (Evans & Murshudov, 2013), the resolution cutoffs for scaling of the data in *AIMLESS* may be set to 3.1 Å for the *hk* plane and 2.5 Å along the *l* axis. The spread in the values of the three principal components that describe the directional dependence of the intensity falloff with resolution was 28.57 Å^2, which indicates strong anisotropy in the data according to the *Diffraction Anisotropy Server* at UCLA (Strong *et al.*, 2006). However, attempts to improve the data using this server did not produce a significant improvement in the map quality. The lower B factors that were observed after anisotropic scaling and negative isotropic B-factor correction cannot be compared with the B factors from the original data. Thus, data to 2.9 Å resolution were used for refinement in *BUSTER* without prior correction of the detected anisotropy. The final model has an R_{work} and R_{free} of 0.196 and 0.249, respectively. Statistics of data collection and refinement are presented in Table 1.

Table 2
Comparison of *At*Rca with other Rubisco activase structures.

Protein	Organism	PDB code	No. aa†	*Z*-score	R.m.s.d. (Å)	No.‡	No. id§
Rca	*Nicotiana tabacum*	3t15¶	293	27.9	1.587	232	185
p97	*Homo sapiens*	3hu2††	452	15.5	2.046	157	24
Rca	*Larrea tridentata*	3thg‡‡	107	12.8	1.031	82	63
CbbX	*Rhodobacter sphaeroides*	3syl§§	309	8.5	2.176	88	11

† Number of amino acids in the structure. ‡ Number of amino acids within the 3.8 Å cutoff. § Number of identical residues in the amino acids within the cutoff. ¶ Stotz *et al.* (2011). †† Tang *et al.* (2010). ‡‡ Henderson *et al.* (2011). §§ Mueller-Cajar *et al.* (2011).

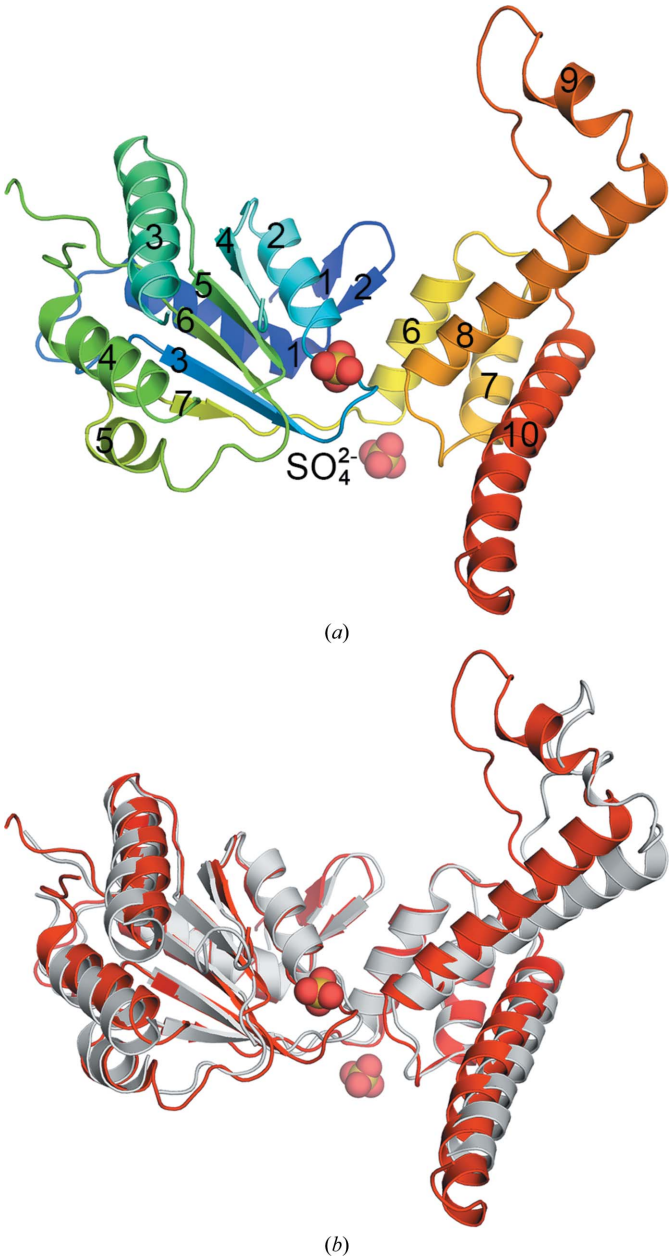


Figure 1
The crystal structure of Rubisco activase from *A. thaliana* (*At*Rca). (*a*) Ribbon representation of *At*Rca rainbow-coloured from blue at the N-terminus to red at the C-terminus. (*b*) Superposition of *At*Rca in red with the structure of Rubisco activase from *N. tabacum* (*Nt*Rca) in grey.

3.2. Overall structure

The expressed *At*Rca protein consisted of 379 amino acids (excluding the C-terminal *Strep*-tag), but only 267 amino acids could be located in the electron density. Dissolved *At*Rca crystals were analysed by mass spectrometry and 101 amino acids of the *At*Rca protein that were missing in the *At*Rca structure were verified to be present in the crystal. The presence of the N-terminal domain may be of significance because deletion of the first 50 amino acids almost eliminated the Rubisco activation activity (van de Loo & Salvucci, 1996). The refined *At*Rca structure consists of amino acids 65–362, with three absent loops within this range comprising residues 134–143, 172–80 and 203–213. The structure contains two sulfate ions and five water molecules. The numbering of the *At*Rca polypeptide starts with the first amino acid of the mature protein; the 58-amino-acid chloroplast-transit peptide is not included.

The structure of *At*Rca presented here has the overall fold of the AAA+ proteins with a two-domain architecture (Fig. 1). The N-terminal domain of the AAA+ module has an $\alpha\beta\alpha$ fold, with five β-strands forming a central β-sheet surrounded by helices H1 and H2 on one side and helices H3, H4 and H5 on the opposite side. Helices H1 and H2 face the C-terminal domain of the AAA+ module, which comprises two helical hairpins. The C-terminal hairpin contains two extended α-helices (H8 and H10). The loop region between the extended α-helices is preceded by a short α-helix (H9) and forms a paddle-like structure.

A total of six Ramachandran outliers were observed (2.3%). Four are close to either of the two sulfate ions. Lys113 and Asp171, which are key residues in the nucleotide-binding region (Walker A and B motifs, respectively, see below), are in the vicinity of the sulfate ion bound in the Walker A motif. Asn317 and Arg319 are located close to the second sulfate ion. The two remaining outliers may have structural roles: Pro96 is positioned at the tip of a loop extending toward an adjacent molecule in the helical spiral and Phe229 points towards the interior of the N-terminal domain close to the Walker B motif.

3.3. Structural comparison with other Rubisco activases

The major form I Rubisco enzymes belong to two phylogenetically distinct subgroups: the green-type enzymes, which are present in land plants, green algae and cyanobacteria, and the red-type enzymes, which are found in red algae, proteobacteria and phytoplankton. At present, crystal structures of two other plant (green-type) Rubisco activases are available in the Protein Data Bank (PDB). The crystal structure of Rubisco activase from tobacco (*N. tabacum*; *Nt*Rca) is from a fragment which lacks the first (N-terminal) 67 residues and the last (C-terminal) 23 residues of the mature protein (Stotz *et al.*, 2011). The resulting *Nt*Rca structure is highly similar to the structure of *At*Rca presented here; *Nt*Rca and *At*Rca have a sequence identity of 84%, excluding the signal peptide, and a *DALI* *Z*-score of 27.9. The structures superimpose with an r.m.s.d. of 1.59 Å, with 79% of the C^{α} atoms of *Nt*Rca falling within the default distance cutoff of 3.8 Å (Fig. 1*b*; Table 2).

The major structural differences, observed for example in helices 8 and 10 and in the paddle region, are owing to the individual movement of structural elements and not to rigid-body domain movements. The crystal structure of *L. tridentata* Rubisco activase (*Lt*Rca) is from a fragment containing 107 amino acids of the C-terminal domain (Henderson *et al.*, 2011). *Lt*Rca superimposes on *At*Rca with an r.m.s.d. of 1.03 Å for 82 C^{α} atoms. The *DALI Z*-score is 12.8.

No homologue has been found of the green-type Rubisco activase in organisms containing red-type Rubisco. Structural and functional studies of the product of the *cbbX* gene in *Rhodobacter sphaeroides* (Mueller-Cajar *et al.*, 2011) have revealed that the CbbX protein is an AAA+ protein and functions as a red-type Rubisco activase. The three-dimensional structure of the red-type Rubisco activase CbbX from *R. sphaeroides* diverges significantly from the extended AAA group into which *At*Rca is classified (Neuwald *et al.*, 1999), with only 28% of its C^{α} atoms superimposing with *At*Rca within the 3.8 Å cutoff and with an r.m.s.d of 2.18 Å (Table 2). Comparison with proteins in the PDB revealed that the CbbX structure is more similar to those from the helicases and clamp-loaders group, with the highest *Z*-score (16.6) to the Holliday junction DNA helicase RuvB from *Thermotoga maritima* (PDB entry 1in5; Putnam *et al.*, 2001).

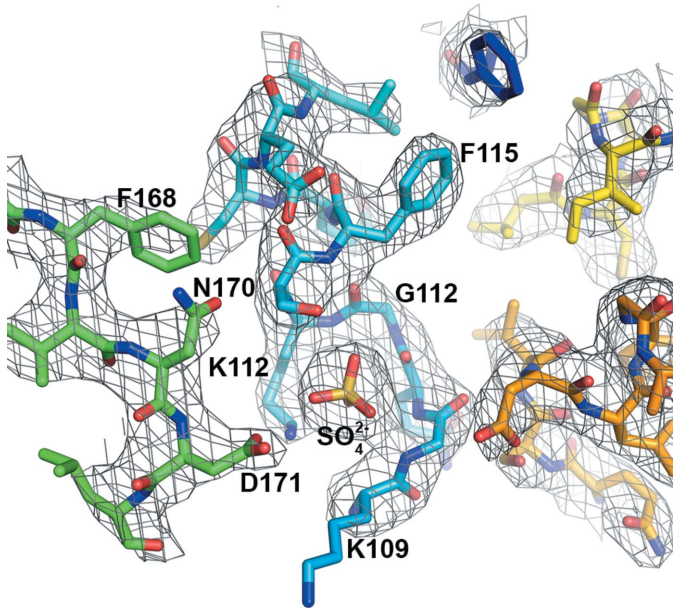


Figure 2
Electron-density map ($2mF_o - DF_c$) contoured at the 1σ level in the region around the sulfate ion bound in the Walker A motif.

3.4. ATP-binding motifs

A common feature of the diverse AAA+ protein group is the presence of Walker A and Walker B motifs. The Walker A motif, or P-loop, is involved in the binding of ATP and is found in many nucleotide-binding proteins. In the structure of *At*Rca, a sulfate ion is bound in the loop formed by the Walker A motif (Figs. 1 and 2). Although no structural information on nucleotide binding to Rubisco activase is available to date, comparison with other AAA+ proteins is revealing. A *DALI* structural homology search with the *At*Rca model as a query showed the highest score for the AAA+ protein p97 (Tang *et al.*, 2010) with ATPγS bound (*DALI Z*-score of 15.5; Table 2). The structure of p97 used for comparison (PDB entry 3hu2; Tang *et al.*, 2010) consists of one of the two AAA+ ATPase modules (named D1) and an N-terminal cofactor-binding domain. *At*Rca and p97 superimpose with an r.m.s.d of 2.05 Å for 157 amino acids with good overlap for the AAA+ domains (Fig. 3*a*). In particular, the Walker A motifs are highly similar and the sulfate

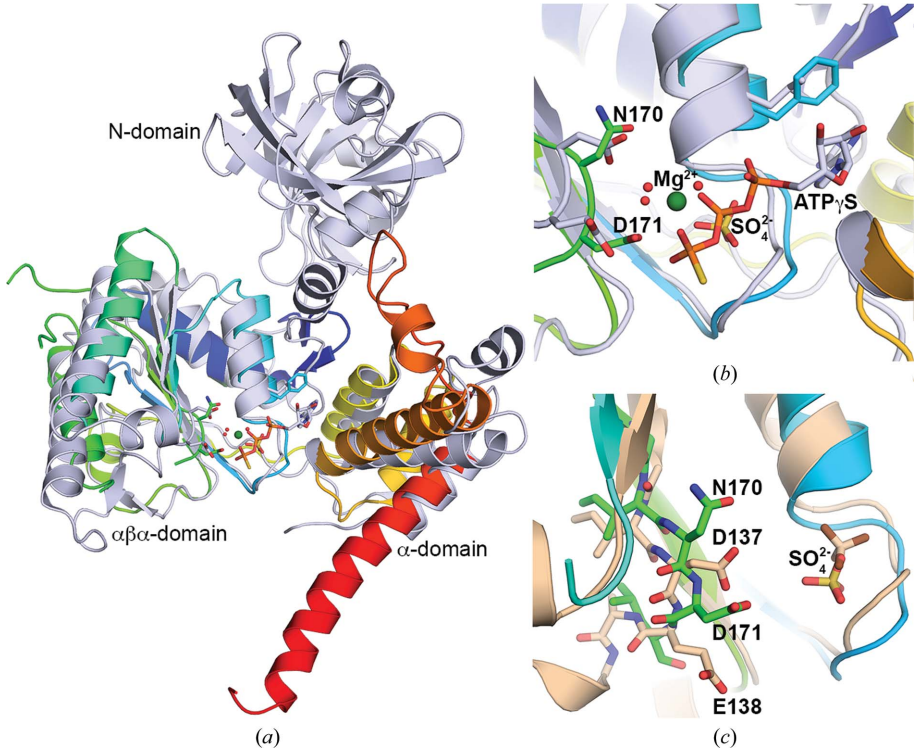


Figure 3
Nucleotide-binding site. (*a*) Superposition of *At*Rca (rainbow-coloured ribbon) with sulfate bound and p97 (light blue ribbon) in complex with ATPγS and Mg^{2+} (PDB entry 3hu2). (*b*) Comparison of the Walker A motifs of *At*Rca and p97 in the same orientation as in (*a*). The sulfate ion and ATPγS are depicted as sticks, the Mg^{2+} ion as a green sphere and water molecules as red spheres. (*c*) Comparison of the Walker B motifs of *At*Rca (rainbow-coloured) and *Rs*CbbX (beige). The Walker A and B motifs of the *At*Rca structure are depicted in blue and green, respectively. Note that although *Rs*CbbX features the common Walker B hhhhDE motif, VLFIDE, whereas *At*Rca displays ^{167}CLFINDL^{173}, the two aspartates occupy similar positions in space.

ion of *At*Rca superimposes with the β-phosphate of ATPγS in p97 (Fig. 3*b*).

A sulfate ion is also present in the corresponding position in the structure of the red-type Rubisco activase CbbX. Superposition of the two structures reveals similarities in the Walker A region (Fig. 3*c*). In contrast, there are significant differences in the Walker B regions involved in ATP hydrolysis and Mg^{2+} binding. In the red-type Rubisco activase the Walker B motifs display the normal consensus motif with a hydrophobic stretch followed by two negatively charged amino acids (hhhhDE; Hanson & Whiteheart, 2005). However, in the green-type Rubisco activase the Walker B motif features an asparagine in one of the positions of the negatively charged amino acids (the hhhhND stretch). Although the two aspartates in the different Rubisco activase Walker B motifs do not align in sequence, the side chains of the two aspartates are directed towards the sulfate ion and their positions superpose in the structures (Fig. 3*c*). The Asp171 carboxyl O atom is located at a distance of 4 Å from one of the sulfate O atoms. When the corresponding aspartate, Asp174, in *Nt*Rca is mutated to alanine the protein has a reduced tendency to bind ATP (van de Loo & Salvucci, 1998).

3.5. Oligomeric states of functionally active Rubisco activases

AAA+ proteins most often assemble into oligomeric structures of diverse sizes in the cell, and the functional units vary according to the function. Proteins involved in unfolding and degradation typically form closed hexameric rings, in which substrate processing is achieved by translocation of an unfolded polypeptide of the target protein through the central pore of the ring (reviewed in Sauer *et al.*, 2004). Initiator or clamp-loader proteins, on the other hand, form long helical fibrils or pentameric spirals (Bowman *et al.*, 2004; Erzberger *et al.*, 2006) that are appropriate for their role in DNA recognition. High concentration can drive oligomerization towards higher molecular-weight complexes, and changes in the oligomerization state are nucleotide-dependent (Monroe *et al.*, 2014; Keown *et al.*, 2013).

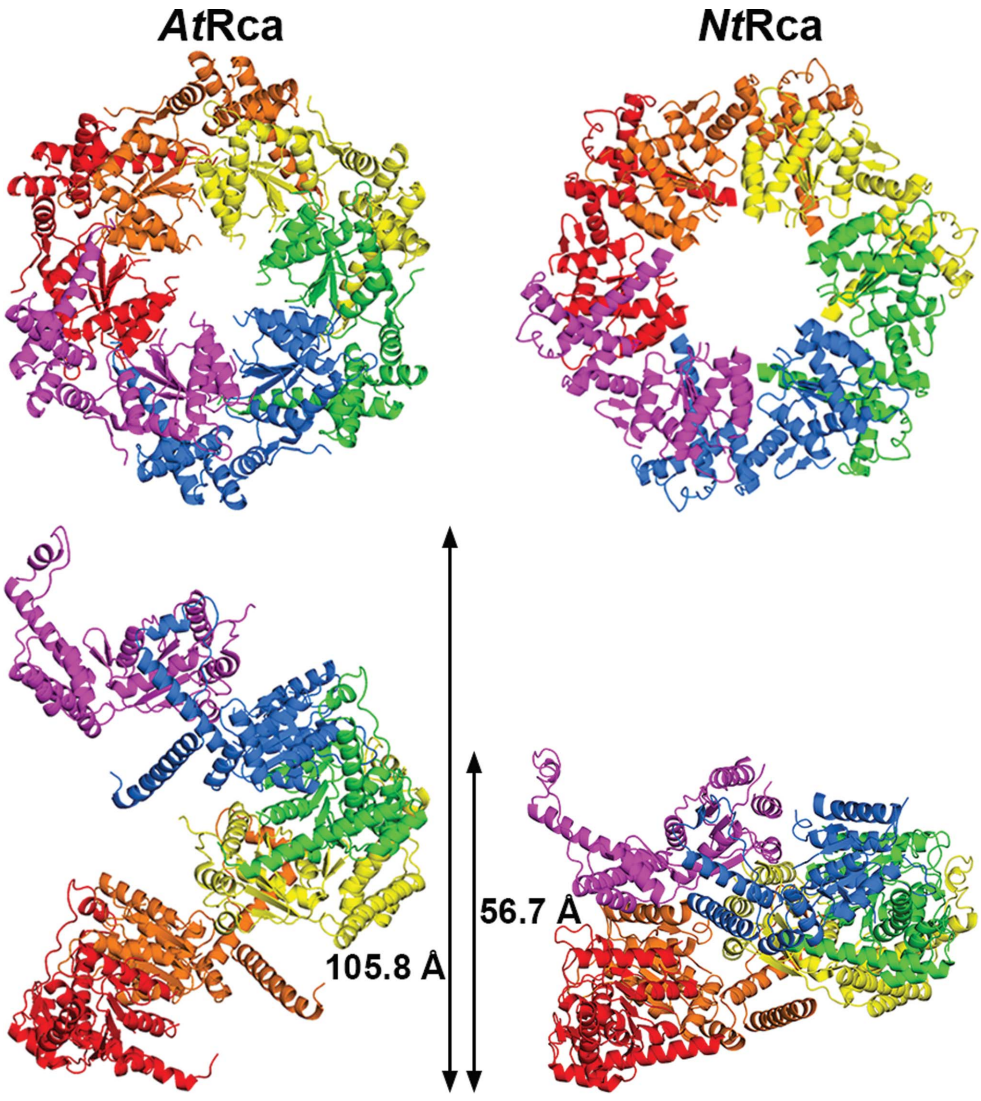


Figure 4
Differences in the crystal packing of *At*Rca and *Nt*Rca. Upper panel, view down the *z* axis; lower panel, view perpendicular to the *z* axis. The length of the sixfold screw axis is indicated. The diameter of the inner pore is narrower in *At*Rca than in *Nt*Rca.

The evidence for a functional unit of Rubisco activase is more diverse. The green-type activase is polydisperse in solution (Wang *et al.*, 1993; Blayney *et al.*, 2011; Chakraborty *et al.*, 2012; Henderson *et al.*, 2013), but no consensus structure has been found for the native protein. Studies of *Nt*Rca in solution have shown that oligomers shorter than hexamers (2–4 Rca units) are sufficient for full activity, while a helical packing was observed at higher *Nt*Rca concentrations (Keown *et al.*, 2013). The longer α-isoform of spinach activase has been observed to form closed hexamers upon binding of Mg·ATPγS, while the spinach β-isoform assembled in a range of oligomers, similar to the behaviour of the tobacco β-isoform (Keown & Pearce, 2014).

This points to a complex pattern of assembly of Rubisco activases. In particular, the oligomeric composition of the

activase is not known at the high protein concentration in the chloroplast (Portis, 2003), which is estimated to be ∼100 times the concentration used in the *in vitro* assays of Keown & Pearce (2014), and in the presence of high concentrations of Rubisco. The packing of plant Rubisco activases in the crystal shows that the proteins are able to form helical structures. The high protein concentration used during crystallization may explain the tendency to form helices in the crystal. This helical arrangement may be further stabilized by the presence of polyethylene glycol, which has been shown to stabilize the oligomeric state (Salvucci, 1992). *Nt*Rca and *At*Rca both pack as hexameric spirals in the crystal but with differences in the length of the pitch (corresponding to the sixfold *z* axis). The *z* axis of *Nt*Rca is 56.7 Å, while it is 105.8 Å in *At*Rca (Fig. 4). The interactions between molecules in the helical packing are mainly formed between the paddle region in the C-terminal domain of one molecule and the first α-helix and the following loop in the N-terminal domain of the adjacent molecule, and packing is promoted by hydrophobic patches (Stotz *et al.*, 2011). There are differences in the intermolecular contacts in the two Rca crystal structures indicating flexibility in the interface formation.

The presence of the flexible N-terminal domain in the *At*Rca crystals could possibly influence packing. Superposition of *At*Rca with p97 (Tang *et al.*, 2010) in complex with ATPγS (Fig. 3*a*) shows good overlap of the core of the $\alpha\beta\alpha$ domain of the AAA+ modules. The structural homology extends to the α-helical bundle domain, with the exception of helix 8 and the paddle, which rotate towards the $\alpha\beta\alpha$ domain in *At*Rca (Fig. 3*a*). The linker connecting the AAA+ module to the N-terminal domain occupies the same space but deviates in structure. The N-terminal domain in *At*Rca is disordered in the crystal, but inspection of the p97 structure may be informative. The N-terminal domains of p97 radiate from the hexameric ring formed by the AAA+ modules (Tang *et al.*, 2010), which may indicate an approximate position for the (smaller) N-terminal domain of *At*Rca at the periphery of the oligomer. Although this position is partly occupied by a symmetry-related molecule in the *At*Rca crystals, there appears to be ample space without defined electron density that may be able to accommodate the flexible N-terminal domain of *At*Rca. The activase N-terminal domain may be intrinsically flexible, awaiting the binding of ATP, its target Rubisco, or both. Mobility in the N-terminal domain is common among AAA+ proteins (see, for example, Ishikawa *et al.*, 2004; Tang *et al.*, 2010). Studies of p97 in solution and in the crystal have demonstrated large nucleotide-dependent conformational changes of the N-terminal domain, including a coil-to-helix transition in the linker region (Tang *et al.*, 2010). The highly mobile N-terminal domain of p97 interacts with the cofactor, and if translated to the activase structure one may speculate that the N-terminal domain of the activase may play a similar role, *i.e.* interaction with its target protein, Rubisco.

Another notable difference between the two structures that could affect the packing is found in the second helical hairpin motif in the C-terminal domain. The paddle region is shifted towards the N-terminal domain in the *At*Rca structure (Fig. 5). The change in the relative position of the paddle compared with its position in *Nt*Rca may be a result of the sulfate ion binding in the Walker A motif of *At*Rca. The carbonyl group of Gln111 in the Walker A motif of *At*Rca interacts with the sulfate ion, and the Gln111 side chain is shifted in position compared with the corresponding glutamine residue (Gln114) in *Nt*Rca. The shift in position permits Ile284 in *At*Rca, which is located at the bottom of the extended helix H8 forming the paddle, to take up a position that overlaps with the position of Gln114 of the Walker A motif in *Nt*Rca (Fig. 6).

In line with the notion that oligomerization is sensitive to small structural changes, tobacco activase containing substitutions of Arg294 was shown to form hexamers in solution in the presence of ATPγS (Blayney *et al.*, 2011; Stotz *et al.*, 2011). Arg294 is located at the subunit interface (Li *et al.*, 2006) and

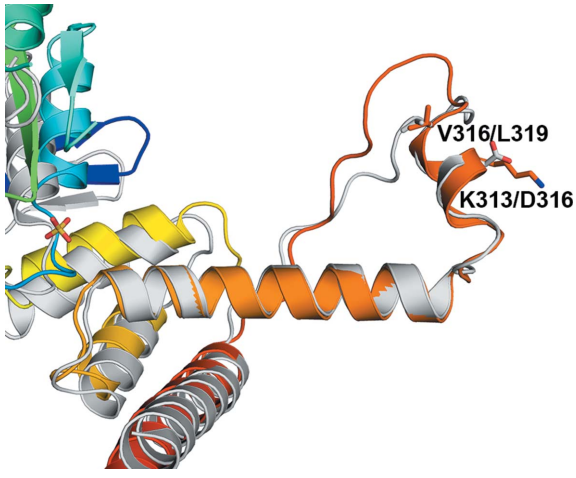


Figure 5
Comparison of the paddle regions of *At*Rca (rainbow-coloured) and *Nt*Rca (grey), with the amino acids that define substrate specificity for Solanaceae and non-Solanaceae plants presented as sticks.

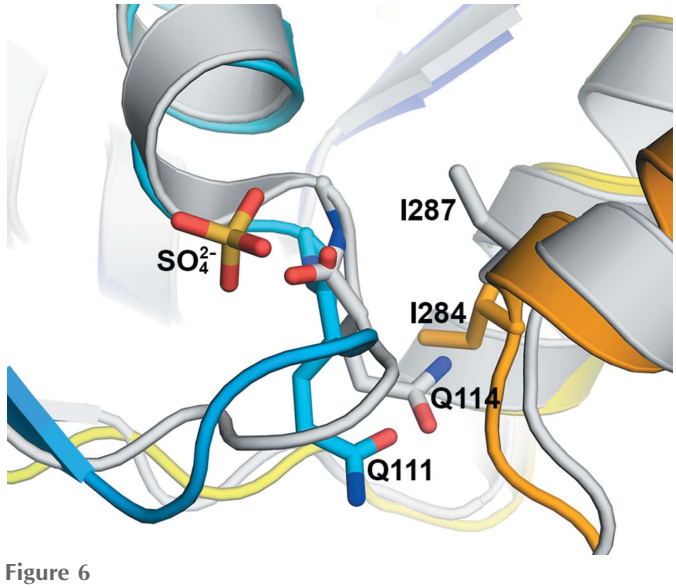


Figure 6
Superposition of the ATP-binding Walker A motifs of *At*Rca (rainbow-coloured) and *Nt*Rca (grey). Differences in the position of the glutamine residue of the Walker A motif may affect the position of the isoleucine residue in the extended helix H8 in the paddle motif.

the R294A substitution, which greatly diminishes Rubisco activation activity, may cause the disruption of a hydrogen bond between Arg294 and Asn99 in the interface between two molecules (Stotz *et al.*, 2011). Negative-stain electron microscopy showed that the activase-proficient R294V variant forms closed hexameric rings in the presence of ATP or ATPγS (Stotz *et al.*, 2011) in a manner similar to the red-type Rubisco activase CbbX, which forms hexameric rings in the presence of Mg·ATP and RuBP (Mueller-Cajar *et al.*, 2011).

Rubisco activase in higher plants displays differences in specificity between Solanaceae plants and non-Solanaceae plants (Wang *et al.*, 1992). Two amino acids in the short helix H9, Asp316 and Leu319 (using the numbering of the tobacco structure), have been identified to be essential for the specificity (Fig. 4; Li *et al.*, 2005). The corresponding residues in *At*Rca are Lys313 and Val316. In the helical packing of *At*Rca in the crystal the paddle region containing helices H8 and H9 interacts with the C-terminal domain of the adjacent molecule. This interaction results in the burial of Val316, which may block the proposed interaction with residues in Rubisco (Larson *et al.*, 1997; Ott *et al.*, 2000; Li *et al.*, 2005). However, if open-end spirals or a small oligonucleotide of 2–4 units is the active form of Rubisco activase (Chakraborty *et al.*, 2012; Keown *et al.*, 2013) interactions may occur between residues on Rubisco and Val316 of the paddle at the free end not involved in oligomerization. The details of such a mechanism for green-type Rubisco activase remain to be investigated.

Acknowledgements

We thank ESRF/EMBL, Grenoble for providing beam time and data-collection facilities. This work was supported by the Swedish Research Councils Formas and VR, the Knut and Alice Wallenberg Foundation, the European Research Council and the Röntgen–Ångström Cluster. We thank Åke Engström for assistance with mass-spectrometric analysis and Evalena Andersson for helpful discussions.

References

Andersson, I. (2008). *J. Exp. Bot.* **59**, 1555–1568.
Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). *Acta Cryst.* D**60**, 2210–2221.
Blayney, M. J., Whitney, S. M. & Beck, J. L. (2011). *J. Am. Soc. Mass Spectrom.* **22**, 1588–1601.
Bowman, G. D., O'Donnell, M. & Kuriyan, J. (2004). *Nature (London)*, **429**, 724–730.
Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. & Womack, T. O. (2011). Cambridge: Global Phasing Ltd.
Chakraborty, M., Kuriata, A. M., Henderson, J. N., Salvucci, M. E., Wachter, R. M. & Levitus, M. (2012). *Biophys. J.* **103**, 949–958.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Edmondson, D. L., Badger, M. R. & Andrews, T. J. (1990). *Plant Physiol.* **93**, 1390–1397.
Erzberger, J. P., Mott, M. L. & Berger, J. M. (2006). *Nature Struct. Mol. Biol.* **13**, 676–683.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Gutteridge, S., Parry, M. A. J., Burton, S., Keys, A. J., Mudd, A., Feeney, J., Servaites, J. C. & Pierce, J. (1986). *Nature (London)*, **324**, 274–276.
Hanson, P. I. & Whiteheart, S. W. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 519–529.
Henderson, J. N., Kuriata, A. M., Fromme, R., Salvucci, M. E. & Wachter, R. M. (2011). *J. Biol. Chem.* **286**, 35683–35688.
Henderson, J. N., Hazra, S., Dunkle, A. M., Salvucci, M. E. & Wachter, R. M. (2013). *Biochim. Biophys. Acta*, **1834**, 87–89.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Ishikawa, T., Maurizi, M. R. & Steven, A. C. (2004). *J. Struct. Biol.* **146**, 180–188.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Keown, J. R., Griffin, M. D., Mertens, H. D. & Pearce, F. G. (2013). *J. Biol. Chem.* **288**, 20607–20615.
Keown, J. R. & Pearce, F. G. (2014). *Biochem. J.* **464**, 413–423.
Larson, E. M., O'Brien, C. M., Zhu, G., Spreitzer, R. J. & Portis, A. R. Jr (1997). *J. Biol. Chem.* **272**, 17033–17037.
Lebedev, A. A. & Isupov, M. N. (2014). *Acta Cryst.* D**70**, 2430–2443.
Li, C., Salvucci, M. E. & Portis, A. R. Jr (2005). *J. Biol. Chem.* **280**, 24864–24869.
Li, C., Wang, D. & Portis, A. R. Jr (2006). *Arch. Biochem. Biophys.* **450**, 176–182.
Loo, F. J. van de & Salvucci, M. E. (1996). *Biochemistry*, **35**, 8143–8148.
Loo, F. J. van de & Salvucci, M. E. (1998). *Biochemistry*, **37**, 4621–4625.
Lorimer, G. H. & Miziorko, H. M. (1980). *Biochemistry*, **19**, 5321–5328.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Monroe, N., Han, H., Gonciarz, M. D., Eckert, D. M., Karren, M. A., Whitby, F. G., Sundquist, W. I. & Hill, C. P. (2014). *J. Mol. Biol.* **426**, 510–525.
Mueller-Cajar, O., Stotz, M. & Bracher, A. (2014). *Photosynth. Res.* **119**, 191–201.
Mueller-Cajar, O., Stotz, M., Wendler, P., Hartl, F. U., Bracher, A. & Hayer-Hartl, M. (2011). *Nature (London)*, **479**, 194–199.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. (1999). *Genome Res.* **9**, 27–43.
Ott, C. M., Smith, B. D., Portis, A. R. Jr & Spreitzer, R. J. (2000). *J. Biol. Chem.* **275**, 26241–26244.
Pearce, F. G. & Andrews, T. J. (2003). *J. Biol. Chem.* **278**, 32526–32536.
Portis, A. R. Jr (2003). *Photosynth. Res.* **75**, 11–27.
Portis, A. R. Jr, Li, C., Wang, D. & Salvucci, M. E. (2008). *J. Exp. Bot.* **59**, 1597–1604.
Putnam, C. D., Clancy, S. B., Tsuruta, H., Gonzalez, S., Wetmur, J. G. & Tainer, J. A. (2001). *J. Mol. Biol.* **311**, 297–310.
Salvucci, M. E. (1992). *Arch. Biochem. Biophys.* **298**, 688–696.
Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., Levchenko, I., Neher, S. B., Oakes, E. S. C., Siddiqui, S. M., Wah, D. A. & Baker, T. A. (2004). *Cell*, **119**, 9–18.
Stotz, M., Mueller-Cajar, O., Ciniawsky, S., Wendler, P., Hartl, F. U., Bracher, A. & Hayer-Hartl, M. (2011). *Nature Struct. Mol. Biol.* **18**, 1366–1370.
Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D. & Eisenberg, D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8060–8065.
Tang, W. K., Li, D., Li, C.-C., Esser, L., Dai, R., Guo, L. & Xia, D. (2010). *EMBO J.* **29**, 2217–2229.
Wang, Z. Y., Ramage, R. T. & Portis, A. R. Jr (1993). *Biochim. Biophys. Acta*, **1202**, 47–55.

Wang, Z. Y., Snyder, G. W., Esau, B. D., Portis, A. R. & Ogren, W. L. (1992). *Plant Physiol.* **100**, 1858–1862.
Werneke, J. M., Chatfield, J. M. & Ogren, W. L. (1989). *Plant Cell*, **1**, 815–825.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zhang, N., Kallis, R. P., Ewy, R. G. & Portis, A. R. Jr (2002). *Proc. Natl Acad. Sci. USA*, **99**, 3330–3334.
Zhang, N. & Portis, A. R. Jr (1999). *Proc. Natl Acad. Sci. USA*, **96**, 9438–9443.
Zhu, G. & Jensen, R. G. (1991). *Plant Physiol.* **97**, 1354–1358.